
ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ



ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ

ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ & ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ

**ΤΕΧΝΙΚΕΣ ΕΠΕΞΕΡΓΑΣΙΑΣ ΚΑΙ ΔΙΑΧΕΙΡΙΣΗΣ
ΟΠΤΙΚΟΑΚΟΥΣΤΙΚΟΥ ΠΕΡΙΕΧΟΜΕΝΟΥ
ΣΤΗ
ΜΕΛΕΤΗ ΒΙΟΑΚΟΥΣΤΙΚΩΝ ΦΑΙΝΟΜΕΝΩΝ**

Διδακτορική Διατριβή

Δημούλα Α. Χαράλαμπου

Διπλ. Ηλεκτρολόγου Μηχανικού ΑΠΘ

ΘΕΣΣΑΛΟΝΙΚΗ

ΣΕΠΤΕΜΒΡΙΟΣ 2006

Αντί προλόγου

Η παρούσα διατριβή υλοποιήθηκε από το διπλωματούχο ηλεκτρολόγο μηχανικό Δημούλα Χαράλαμπο, στο πλαίσιο συμμετοχής του στο πρόγραμμα μεταπτυχιακών σπουδών του Τμήματος Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών που οδηγεί σε απόκτηση διδακτορικού διπλώματος, υπό την επιστημονική καθοδήγηση των κκ. Αν. Καθ. Παπανικολάου Γεώργιο, Καθ. Πανά Σταύρο και Αν. Καθ. Σεργιάδη Γεώργιο. Η διατριβή πραγματεύεται ζητήματα λήψης και επεξεργασίας ακουστικών και οπτικών καταγραφών, με στόχο την αξιοποίησή τους στη μελέτη βιοακουστικών φαινομένων, όπως είναι οι ήχοι που παράγονται εξαιτίας της γαστρεντερικής κινητικότητας.

Για τις ανάγκες καταγραφής, διαχείρισης και ανάλυσης των πειραματικών δεδομένων δοκιμάστηκαν και αξιολογήθηκαν διάφοροι αισθητήρες με τις αντίστοιχες συνοδευτικές κυκλωματικές διατάξεις, αναπτύχθηκαν αλγόριθμοι επεξεργασίας σήματος, υλοποιήθηκαν υπολογιστικά περιβάλλοντα λογισμικού. Οι πειραματικές διαδικασίες πραγματοποιήθηκαν σε συνεργασία με το Γαστρεντερικό Τμήμα του Γενικού Περιφερειακού Νοσοκομείου Παπαγεωργίου Θεσσαλονίκης, όπου καθ' όλη τη διάρκεια εκπόνησης της διατριβής υπήρξε μια ανοικτή διαδικασία ανατροφοδότησης με το ιατρικό και νοσηλευτικό προσωπικό. Στόχος της συνεργασίας αυτής είναι η μελλοντική θεμελίωση μιας κλινικά αποδεκτής και εύκολα εφαρμόσιμης διαγνωστικής μεθόδου, γύρω από τις λειτουργικές διαταραχές της κινητικότητας του γαστρεντερικού συστήματος.

Ως αποτέλεσμα, διαμορφώθηκε ένα νέο πλαίσιο μελέτης της κινητικότητας του γαστρεντερικού συστήματος, με τη βοήθεια των παραγομένων βιοακουστικών φαινομένων. Επίσης, αναπτύχθηκε μια μεθοδολογία που απαντά στο γενικότερο ζήτημα ψυχοφυσιολογικής παρακολούθησης παρατεταμένου χρόνου, μέσω πολύ-αισθητηριακών καταγραφών. Ο σχεδιασμός αυτός προτείνει ευέλικτες μεθόδους παρακολούθησης και ελέγχου της μη εποπτευόμενης πειραματικής διαδικασίας, τεχνικές αυτοματοποίησης και δομές διαχείρισης περιεχομένου. Τα αποτελέσματα συμβάλλουν στην επέκταση της ιατρικής έρευνας με ταχύτερους και αποδοτικότερους ρυθμούς. Παράλληλα, επιτρέπουν τη συνδυασμένη εφαρμογή μεθόδων βιοϊατρικής παρακολούθησης, με στόχο την πιο εμπειρισταωμένη μελέτη του φαινομένου, τη θεμελίωση της επιδιωκόμενης ιατρικής γνώσης και των αντίστοιχων διαγνωστικών εργαλείων.

Με την ευκαιρία σύνταξης αυτού του προλόγου, θα ήθελα να ευχαριστήσω θερμά τα μέλη της τριμελούς συμβουλευτικής επιτροπής, και ειδικότερα τον κύριο επιβλέποντα κ. Γ. Παπανικολάου, για την επιστημονική καθοδήγηση που μου παρείχαν. Θα ήθελα επίσης να ευχαριστήσω τα μέλη του Εργαστηρίου Ηλεκτρακουστικής και Τηλεοπτικών Συστημάτων ΑΠΘ, με ιδιαίτερη έμφαση στον Επικ. Καθ. κ. Καλλίρη Γεώργιο του οποίου η συνεργασία ήταν πολύτιμη· ακόμη, τον ειδικό γαστρεντερολόγο Δρ. Καλαμπάκα Αθανάσιο και τους συνεργάτες του από το Νοσοκομείο Παπαγεωργίου, τον Επικ. Καθ. κ. Χατζηλεοντιάδη Λεόντιο, τους Καθ. κκ. Τροχίδη Αθανάσιο και Πετρίδη Βασίλειο, τους φίλους μου και την οικογένειά μου, που μου συμπαραστάθηκαν σε όλη αυτή την ερευνητική διαδρομή.

Δημούλας Χαράλαμπος,
Σεπτέμβριος 2006

Πίνακας Περιεχομένων

Εισαγωγή.....	6
Δομή της εργασίας.....	6
Επιτομή της συμβολής της εργασίας.....	9
Λίστα κυριότερων ακρωνυμίων.....	11
Λίστα κυριότερων συμβόλων.....	15
1. Αντικείμενο της διατριβής.....	16
1.1 Περιγραφή του προβλήματος.....	16
1.2 Μελέτη του γαστρεντερικού συστήματος.....	17
1.2.1 Στοιχεία φυσιολογίας.....	18
1.2.2 Υπάρχουσα ιατρική κατάσταση.....	20
1.2.3 Τεχνικές διάγνωσης.....	21
1.2.3.1 Ραδιο-ισοτοπικές τεχνικές.....	21
1.2.3.2 Ακτινοσκοπικές τεχνικές.....	22
1.2.3.3 Ηλεκτρομυογραφικές τεχνικές.....	22
1.2.3.4 Τεχνικές καταγραφής μηχανικής γαστρεντερικής δραστηριότητας.....	24
1.2.3.5 Τεχνικές ενδοσκόπησης.....	26
1.2.4 Ακουστική ανάλυση γαστρεντερικής δραστηριότητας.....	27
1.2.4.1 Μηχανισμοί παραγωγής γαστρεντερικών ήχων.....	27
1.2.4.2 Γαστρεντερική ακρόαση και διαγνωστικές προσεγγίσεις.....	28
1.2.4.3 Αξιοποίηση γαστρεντερικών βιοακουστικών επεισοδίων στη διάγνωση παθολογιών.....	29
1.2.4.4 Αξιοποίηση γαστρεντερικών ήχων στη μελέτη της κινητικότητας.....	30
1.3 Προδιαγραφές ερευνητικής προσέγγισης.....	31
1.4 Σχεδιασμός του συστήματος βιοϊατρικής παρακολούθησης.....	34
1.5 Επισκόπηση 1ου κεφαλαίου.....	34
2. Περιγραφή του συστήματος καταγραφής.....	36
2.1 Συστήματα ήχου και εικόνας στη βιοϊατρική τεχνολογία.....	36
2.1.1 Συστήματα καταγραφής βιοακουστικών σημάτων.....	36
2.1.1.1 Ακροαστική διάγνωση μέσω στηθοσκοπίων.....	37
2.1.1.2 Βασικές αρχές ηλεκτρακουστικών μετατροπών λήψης.....	37
2.1.1.3 Μικρόφωνα.....	38
2.1.1.4 Επιταχυνσιόμετρα.....	40
2.1.1.5 Ηλεκτρακουστικές διατάξεις προ-ενίσχυσης.....	42
2.1.1.6 Αναλογικά και ψηφιακά συστήματα ηχητικής εγγραφής.....	42
2.1.1.7 Προδιαγραφές βιοακουστικών συστημάτων καταγραφής.....	43
2.1.2 Συστήματα οπτικοακουστικής παρακολούθησης.....	44
2.1.2.1 Βιοϊατρικές εφαρμογές οπτικοακουστικής παρακολούθησης.....	44
2.1.2.2 Συστήματα καταγραφής και επεξεργασίας κινούμενης εικόνας.....	45
2.2 Πειραματικό μέρος: ανάπτυξη συστήματος καταγραφής.....	46
2.2.1 Πολυκαναλικό σύστημα καταγραφής βιοακουστικών σημάτων.....	46
2.2.1.1 Αισθητήρες λήψης και κυκλώματα προενίσχυσης.....	47
2.2.1.2 Διαμόρφωση πολυκαναλικού συστήματος βιοακουστικής καταγραφής.....	50
2.2.1.3 Πρόσθετα μέσα «στέγασης» και προσαρμογής των αισθητήρων.....	51
2.2.1.4 Ψηφιοποίηση βιοακουστικών καταγραφών και μονάδες εγγραφής.....	54
2.2.2 Σύστημα οπτικοακουστικής παρακολούθησης & ελέγχου.....	55
2.2.2.1 Εξοπλισμός οπτικοακουστικής παρακολούθησης - καταγραφής.....	56
2.2.2.2 Ψηφιοποίηση και τεχνικές επεξεργασίας των οπτικοακουστικών καταγραφών.....	57
2.2.2.3 Διαχείριση οπτικοακουστικών καταγραφών και τεχνικές ανίχνευση κίνησης.....	59

2.3	Επισκόπηση 2ου κεφαλαίου.....	63
2.3.1	Αποτίμηση των παρεμβάσεων και ερευνητική συνεισφορά της εργασίας	63
2.3.2	Προτάσεις μελλοντικής εξέλιξης του συστήματος καταγραφής.....	64
3.	Τεχνικές καταστολής θορύβου σε ηχητικές καταγραφές.....	65
3.1	Περιγραφή του προβλήματος καταστολής θορύβου	65
3.1.1	Θόρυβος φόντου ευρείας ζώνης.....	65
3.1.2	Μέθοδοι καταστολής θορύβου φόντου.....	66
3.1.2.1	Τεχνικές φασματικής αφαίρεσης και παραμετρικό φίλτρο Wiener.....	67
3.1.2.2	Κυματιδιακή ανάλυση και τεχνικές κατωφλίωσης κυματιδιακών συντελεστών.....	71
3.2	Καταστολή θορύβου σε καταγραφές κοιλιακών ήχων.....	80
3.2.1	Παρουσία θορύβου & προβλήματα ανάλυσης κοιλιακών ήχων.....	80
3.2.2	Υπάρχουσες τεχνικές καταστολής θορύβου κοιλιακών ήχων.....	81
3.2.3	Η προτεινόμενη μέθοδος καταστολής θορύβου.....	84
3.2.4	Επεξεργασία κυματιδιακών συντελεστών με το παραμετρικό φίλτρο Wiener.....	85
3.2.5	Νέες τεχνικές “Wavelet Domain Wiener Filtering” (WDWF).....	86
3.2.5.1	Επεξεργασία WDWF σημείο προς σημείο, τύπος-I.....	86
3.2.5.2	Επεξεργασία WDWF σημείο προς σημείο, τύπος-II.....	91
3.2.5.3	Επεξεργασία WDWF μέσω διακριτού μετασχηματισμού κυματιδίων 6 ζωνών.....	97
3.2.5.4	Επεξεργασία WDWF μέσω πακέτων κυματιδιακής ανάλυσης 17 ζωνών.....	98
3.2.5.5	Επεξεργασία WDWF μέσω αυτόματου εντοπισμού κατωφλίων.....	99
3.2.6	Αξιολόγηση των τεχνικών Wavelet Domain Wiener Filtering (WDWF).....	100
3.2.6.1	Έλεγχος της αποδοτικότητας των επιλεγμένων παραμέτρων επεξεργασίας.....	101
3.2.6.2	Ποιοτική αξιολόγηση των προτεινόμενων τεχνικών αποθορυβοποίησης.....	105
3.2.6.3	Ποσοτική αξιολόγηση των προτεινόμενων τεχνικών αποθορυβοποίησης.....	116
3.3	Επισκόπηση 3ου κεφαλαίου.....	120
3.3.1	Αποτίμηση προτεινόμενων παρεμβάσεων.....	120
3.3.2	Στοιχεία πρωτοτυπίας & ερευνητική συνεισφορά.....	120
3.3.3	Προτεινόμενες μελλοντικές εξελίξεις.....	121
4.	Επεξεργασία παρατεταμένου χρόνου.....	122
4.1	Επεξεργασία ηχητικών καταγραφών παρατεταμένου χρόνου.....	122
4.1.1	Τεχνικές ανίχνευσης και εντοπισμού σήματος.....	123
4.1.2	Ανάλυση πολυπλοκότητας κυματομορφών “Fractal Dimension”.....	123
4.2	Ανάλυση γαστρεντερικών ακουστικών καταγραφών παρατεταμένου χρόνου.....	125
4.2.1	Υπάρχουσα κατάσταση και ιδιαιτερότητες του προβλήματος.....	125
4.2.2	Τεχνικές σάρωσης με χρήση παραθύρων μικρής διάρκειας.....	126
4.2.3	Μέθοδος πολύ-διακριτικής σάρωσης.....	128
4.2.3.1	Επεξεργασία αποθορυβοποίησης σε καταγραφές παρατεταμένου χρόνου.....	131
4.2.3.2	Ανίχνευση σημαντικών συμβάντων γαστρεντερικής δραστηριότητας.....	135
4.2.3.3	Οριοθέτηση και αυτόματος διαχωρισμός συμβάντων.....	139
4.2.3.4	Τοπογραφική ερμηνεία συμβάντων.....	142
4.2.3.5	Εξαγωγή ενεργειακών παραμέτρων συνοπτικής απεικόνισης.....	144
4.2.4	Πειραματικά αποτελέσματα αξιολόγησης της μεθόδου LT-WDSS.....	145
4.3	Επισκόπηση 4ου κεφαλαίου.....	151
4.3.1	Αποτίμηση προτεινόμενων παρεμβάσεων.....	152
4.3.2	Στοιχεία πρωτοτυπίας και ερευνητική συνεισφορά.....	152
4.3.3	Προτεινόμενες μελλοντικές εξελίξεις.....	153
5.	Τεχνικές αναγνώρισης προτύπων βιοακουστικών φαινομένων.....	154
5.1	Εφαρμογές ταξινόμησης και αναγνώρισης προτύπων.....	154
5.1.1	Η έννοια του προτύπου.....	155
5.1.2	Τεχνικές στατιστικής ταξινόμησης προτύπων.....	156

5.1.3	Αναγνώριση προτύπων με χρήση νευρωνικών δικτύων	158
5.1.4	Προσεγγίσεις συντακτικής – ιεραρχικής ταξινόμησης προτύπων	161
5.2	Ανάλυση προτύπων γαστρεντερικών ακουστικών φαινομένων.....	162
5.3	Σχήμα συνοπτικής ταξινόμησης κοιλιακών ήχων (ASCC).....	164
5.3.1	Κανόνες περιγραφής των προτύπων συνοπτικής ταξινόμησης ASCC	164
5.3.2	Αυτόματη αναγνώριση των προτύπων συνοπτικής ταξινόμησης ASCC.....	168
5.3.2.1	Τεχνικές προ-επεξεργασίας (pre-processing)	168
5.3.2.2	Εξαγωγή χαρακτηριστικών παραμέτρων ανάλυσης (feature extraction)	169
5.3.2.3	Τοπολογίες δικτύων MLP και αποτελέσματα εκπαίδευσης	173
5.4	Σχήμα αναλυτικής ταξινόμησης κοιλιακών ήχων (ASFC).....	177
5.4.1	Κανόνες περιγραφής των προτύπων αναλυτικής ταξινόμησης ASFC.....	178
5.4.2	Αυτόματη αναγνώριση των προτύπων αναλυτικής ταξινόμησης ASFC	182
5.5	Επισκόπηση 5ου κεφαλαίου.....	186
5.5.1	Αποτίμηση προτεινόμενων παρεμβάσεων	186
5.5.2	Στοιχεία πρωτοτυπίας & ερευνητική συνεισφορά	187
5.5.3	Προτεινόμενες μελλοντικές εξελίξεις	187
6.	Τεχνικές ακουστικής χαρτογράφησης.....	188
6.1	Χωρική ανάλυση ήχου: θεωρία και εφαρμογές.....	188
6.1.1	Τεχνικές εντοπισμού ηχητικών πηγών.....	189
6.1.2	Χωρική καταγραφή και απεικόνιση ηχητικού πεδίου.....	191
6.1.3	Τεχνικές ηχογράφησης και αναπαραγωγής στερεοφωνικού - περιφερικού ήχου	192
6.2	Τοπογραφική ανάλυση γαστρεντερικού βιοακουστικού πεδίου	195
6.2.1	Η προτεινόμενη μέθοδος βιοακουστικής χαρτογράφησης.....	197
6.2.1.1	Ηχητικός εντοπισμός στο «χωροχρόνο» και οπτικοακουστική απόδοση των συμβάντων.....	201
6.2.1.2	Αξιολόγηση της μεθόδου DEBL μέσω προσομοιώσεων σε περιβάλλον λογισμικού	203
6.2.1.3	Αξιολόγηση της μεθόδου DEBL μέσω φυσικών πειραματικών μοντέλων	206
6.3	Επισκόπηση 6ου κεφαλαίου.....	208
7.	Τεχνικές διαχείρισης και ανάλυσης περιεχομένου.....	209
7.1	Το πρότυπο MPEG-7	209
7.1.1	Τα μέρη του MPEG-7	210
7.1.2	Δομή του προτύπου MPEG-7	211
7.1.3	Περιγραφείς και σχήματα περιγραφής.....	211
7.1.4	Προτεινόμενες εφαρμογές MPEG-7 στη βιοϊατρική τεχνολογία.....	213
7.2	Διαχείριση περιεχομένου και ανάλυση των πειραματικών καταγραφών	214
7.2.1	Σχεδιασμός εννοιολογικής δομής περιγραφής περιεχομένου	215
7.2.2	Προτεινόμενα σχήματα περιγραφής των πειραματικών καταγραφών	216
7.2.3	Τεχνικές μετα-επεξεργασίας και εργαλεία ανάλυσης	221
7.3	Επισκόπηση 7ου κεφαλαίου.....	222
8.	Συμπεράσματα - Προοπτικές.....	223
8.1	Ερευνητική συνεισφορά & πρωτοτυπία της διατριβής	223
8.1.1	Συνεισφορά στην έρευνα «βιοϊατρικής τεχνολογίας».....	224
8.1.2	Συνεισφορά στην έρευνα «μηχανικής».....	226
9.	Βιβλιογραφικές αναφορές.....	227
10.	ABSTRACT.....	244

Εισαγωγή

Η παρούσα διατριβή πραγματεύεται τεχνικές επεξεργασίας και ανάλυσης βιοακουστικών καταγραφών παρατεταμένου χρόνου, με στόχο την υποστήριξη της ιατρικής έρευνας*. Η προσπάθεια επικεντρώνεται στη μελέτη της φυσιολογίας του γαστρεντερικού συστήματος μέσω της ανάλυσης των παραγόμενων βιοακουστικών φαινομένων. Στις καταγραφές συμπεριλήφθηκαν και οπτικοακουστικά σήματα παρακολούθησης, για τον έλεγχο της πειραματικής διαδικασίας σε πραγματικές - κλινικές συνθήκες και τη δυνατότητα εξαγωγής σχετικών συμπερασμάτων. Σε γενικές γραμμές, οι τεχνικές οπτικοακουστικής επεξεργασίας που αναπτύχθηκαν, αφορούν τους λαμβανόμενους κοιλιακούς ήχους, τα σήματα οπτικοακουστικής εποπτείας της πειραματικής διαδικασίας και τις εξαγόμενες παραμέτρους ανάλυσης και απεικόνισης των πειραματικών δεδομένων.

Στο πλαίσιο αυτό, διάφορες τεχνικές λήψης και επεξεργασίας ακουστικών και οπτικών καταγραφών έχουν προταθεί και υλοποιηθεί. Παράλληλα σχεδιάστηκαν νέες, ευέλικτες δομές αρχειοθέτησης και διαχείρισης περιεχομένου. Οι προτεινόμενες λύσεις αναζητήθηκαν σε διάφορα επιστημονικά πεδία των τεχνολογιών του ήχου και της εικόνας. Ως παραδείγματα μπορούμε να αναφέρουμε μεθόδους βιοϊατρικής τεχνολογίας για τη μελέτη του γαστρεντερικού συστήματος αλλά και άλλων ψυχο-φυσιολογικών παραμέτρων, τεχνικές βάσεων δεδομένων για τη δομημένη καταχώρηση, διαχείριση και πλοήγηση πολυμεσικού περιεχομένου, τεχνικές χαρτογράφησης ηχητικού πεδίου και ακουστικής διάγνωσης μηχανικών συστημάτων, συστήματα οπτικοακουστικής παρακολούθησης και εποπτικού ελέγχου (συστήματα ασφαλείας). Έτσι, πολλές από τις μεθόδους που προτείνονται για την αντιμετώπιση του συγκεκριμένου προβλήματος, μπορούν να χρησιμοποιηθούν σε διάφορες εφαρμογές των παραπάνω επιστημονικών πεδίων.

Δομή της εργασίας

Το σύγγραμμα δομείται σε οκτώ συνολικά ενότητες, καθεμιά από τις οποίες επιχειρεί να «φωτίσει» διαφορετικές πλευρές του προβλήματος προς επίλυση. Από πλευράς δομικής διάρθρωσης, κάθε κεφάλαιο περιλαμβάνει μια σύντομη αναφορά στην υπάρχουσα ερευνητική κατάσταση του ειδικού αντικείμενου με το οποίο ασχολείται (state of research), καθώς και συνοπτική παρουσίαση της αντίστοιχης θεωρίας. Ακολουθεί αναλυτική παρουσίαση των προτεινόμενων μεθόδων, όπου συμπεριλαμβάνονται τα κίνητρα, οι στόχοι και τα αναμενόμενα αποτελέσματα, καθώς και η αξιολόγηση της λειτουργίας και αποδοτικότητάς τους. Έτσι πραγματοποιείται μια συνολική αποτίμηση των προτεινόμενων παρεμβάσεων και της ερευνητικής τους συνεισφοράς.

Στο πρώτο κεφάλαιο παρουσιάζονται τα κύρια ερευνητικά αντικείμενα της διατριβής. Επιχειρείται μια σύντομη παρουσίαση των προβλημάτων που σχετίζονται με τη μελέτη της συσταλτικής δραστηριότητας του γαστρεντερικού συστήματος, παρατίθενται τα απαραίτητα στοιχεία φυσιολογίας και σχολιάζονται οι έως τώρα ερευνητικές προσπάθειες. Επίσης, αναλύονται οι τεχνικές προδιαγραφές της ιατρικής έρευνας, οι δυσκολίες διαχείρισης και ελέγχου των πειραματικών δεδομένων, και οι ανάγκες αυτοματοποίησης της μη εποπτευόμενης πειραματικής διαδικασίας, λόγω της ιδιαίτερα μεγάλης χρονικής διάρκειας των καταγραφών.

* δηλαδή της εξαγωγής ιατρικής γνώσης και της θεμελίωσης μεθόδων διάγνωσης.

Στο δεύτερο κεφάλαιο παρουσιάζονται τα κύρια τεχνικά ζητήματα που σχετίζονται με τη λήψη και ψηφιακή καταγραφή των οπτικοακουστικών σημάτων. Προηγείται μια σύντομη θεωρητική εισαγωγή που αφορά στους αισθητήρες λήψης, τις παραμέτρους ψηφιοποίησης και τις τεχνικές κωδικοποίησης - συμπίεσης. Κατόπιν, αναλύονται οι ιδιαιτερότητες του συγκεκριμένου προβλήματος, τόσο σε θέματα καταγραφής των βιοακουστικών συμβάντων και του αντίστοιχου ηχητικού πεδίου, όσο και σε ζητήματα παρακολούθησης της πειραματικής διαδικασίας. Τέλος, παρατίθενται οι λύσεις που δοκιμάστηκαν και τελικά επιλέχθηκαν, ενώ αξιολογείται η αποδοτικότητά τους στην αντιμετώπιση των διαφόρων προβλημάτων.

Το τρίτο κεφάλαιο αφορά στις τεχνικές επεξεργασίας ηχητικών σημάτων για τον εμπλουτισμό των πληροφοριακών συνιστωσών και την καταστολή των συνιστωσών θορύβου. Διαπραγματεύεται, δηλαδή, τα ζητήματα αποθορυβοποίησης και αποκατάστασης ενθόρυβων ηχητικών καταγραφών. Στο πλαίσιο αυτό, παρουσιάζονται τα προβλήματα που ανακύπτουν από την παρουσία προσθετικού θορύβου βάθους, καθώς και οι γενεσιουργές αιτίες του τελευταίου. Η ανάλυση αυτή εξειδικεύεται στα ζητήματα λήψης και καταγραφής των γαστρεντερικών ήχων*. Κατόπιν, αναλύονται οι βασικές θεωρητικές αρχές των τεχνικών καταστολής θορύβου, όπως είναι οι μέθοδοι φασματικής αφαίρεσης και το παραμετρικό φίλτρο Wiener. Επίσης, παρουσιάζεται ο μετασχηματισμός κυματιδίων (wavelet transform) και οι τεχνικές κυματιδιακής κατωφλίωσης, που έχουν χρησιμοποιηθεί ως επί το πλείστον για την αποθορυβοποίηση των κοιλιακών ήχων. Το κεφάλαιο ολοκληρώνεται με τη διεξοδική παρουσίαση μιας νέας μεθόδου, που επιχειρεί να δώσει απάντηση στα ανεπίλυτα, έως τώρα, προβλήματα. Εδώ, συμπεριλαμβάνονται διαδικασίες ποιοτικής και ποσοτικής αξιολόγησης της αποδοτικότητας της προτεινόμενης τεχνικής αποθορυβοποίησης, καθώς και συγκρίσεις με τις υπάρχουσες μεθόδους.

Το τέταρτο κεφάλαιο πραγματεύεται ζητήματα επεξεργασίας παρατεταμένου χρόνου, που αφορούν κυρίως στις βιοακουστικές καταγραφές των λαμβανόμενων κοιλιακών δονήσεων. Παρουσιάζονται οι έως τώρα ερευνητικές κατευθύνσεις ανάλυσης των γαστρεντερικών ακουστικών φαινομένων, σε καταγραφές σύντομης και παρατεταμένης διάρκειας. Στη συνέχεια διατυπώνεται μια νέα ερευνητική προσέγγιση, που συνδυάζει πλεονεκτήματα και των δύο προηγούμενων προσεγγίσεων. Έτσι, γίνεται αναφορά σε τεχνικές ανίχνευσης σήματος και αυτόματου διαχωρισμού συμβάντων, με χρήση χρονικών παραθύρων επιλογής. Ιδιαίτερη έμφαση δίνεται στις μεθόδους που χρησιμοποιήθηκαν στην παρούσα εφαρμογή, όπως είναι οι τελεστές ενεργειακής σύγκρισης και κατωφλίωσης κυματιδιακών συντελεστών, καθώς και οι αλγόριθμοι εντοπισμού μεταβατικών φαινομένων (transient detection), μέσω ανάλυσης πολυπλοκότητας κυματομορφών "fractal dimension". Ακολούθως, παρουσιάζεται το συνολικό πλαίσιο της προτεινόμενης μεθοδολογίας που περιλαμβάνει: την ανίχνευση των σημαντικών βιοακουστικών συμβάντων, την ενσωμάτωση των πληροφοριών σε ευέλικτα σχήματα περιγραφής και διαχείρισης περιεχομένου, την τοπογραφική ερμηνεία των απομονωμένων «επεισοδίων» και την εξαγωγή παραμέτρων συνοπτικής απεικόνισης του συνολικού φαινομένου. Το κεφάλαιο ολοκληρώνεται με την παράθεση ποιοτικών και ποσοτικών στοιχείων αξιολόγησης της νέας μεθόδου.

Το πέμπτο κεφάλαιο επικεντρώνεται στη διαδικασία προσδιορισμού προτύπων κοιλιακών ήχων. Πολλοί ερευνητές θεωρούν ότι η απουσία συγκεκριμένης ορολογίας και σχημάτων ταξινόμησης στο χώρο αυτό, αποτελούν και έναν από τους κυριότερους

* Είναι χαρακτηριστικό ότι η πλειονότητα των δημοσιεύσεων που έχουν γίνει σε τεχνικά συνέδρια και περιοδικά γύρω από το θέμα των γαστρεντερικών ήχων, αφορά αλγόριθμους αποθορυβοποίησης.

παράγοντες που δρουν ως τροχοπέδη στην εξέλιξη των σχετικών μεθόδων διάγνωσης. Στην προκειμένη περίπτωση, η απάντηση στο πρόβλημα δίνεται μέσα από μια συστηματική μελέτη, όπου επιχειρείται η ενσωμάτωση όλης της υπάρχουσας γνώσης και η τυποποίηση των πειραματικών παρατηρήσεων σε αντίστοιχα λεξικά κανόνων ταξινόμησης. Για το λόγο αυτό, το κεφάλαιο παραθέτει την απαραίτητη θεωρία γύρω από τις τεχνικές ταξινόμησης και αναγνώρισης προτύπων, με έμφαση στις μεθοδολογίες στατιστικού διαχωρισμού κλάσεων, τα νευρωνικά δίκτυα και τις τεχνικές συντακτικής - ιεραρχικής ταξινόμησης. Παράλληλα, παρουσιάζεται το συγκεκριμένο πρόβλημα, αναλύονται τα σχετικά ερευνητικά εγχειρήματα που πραγματοποιήθηκαν στο παρελθόν, τα κίνητρα, οι στόχοι και τα επιδιωκόμενα αποτελέσματα. Το κεφάλαιο επικεντρώνεται στα αποτελέσματα κατηγοριοποίησης και τους κανόνες ταξινόμησης των κοιλιακών ήχων, καθώς και στα εργαλεία αυτόματης αναγνώρισης προτύπων που αναπτύχθηκαν. Επίσης, πραγματοποιείται αξιολόγηση της αποδοτικότητας των μεθόδων αυτών, ενώ αναλύονται και τα πλεονεκτήματα ενσωμάτωσής τους στη μελέτη των μη εποπτευόμενων καταγραφών παρατεταμένου χρόνου.

Το έκτο κεφάλαιο αναφέρεται στις τεχνικές χωρικής ανάλυσης των καταγεγραμμένων βιοακουστικών φαινομένων. Έτσι, παρουσιάζονται οι θεωρητικές προσεγγίσεις των εφαρμογών εντοπισμού ηχητικής πηγής, των τεχνικών γραφικής απεικόνισης του ηχητικού πεδίου και των μεθόδων «τριδιάστατης» αναπαραγωγής στερεοφωνικού - περιφερικού ήχου. Παράλληλα, γίνεται αναφορά στις έως τώρα προσπάθειες χαρτογράφησης της εντερικής δραστηριότητας, στα κίνητρα ανάπτυξης τέτοιων μεθόδων, στις ιδιαιτερότητες του κοιλιακού βιοακουστικού πεδίου και στις σχετικές δυσκολίες επεξεργασίας. Ως απάντηση στα υπάρχοντα προβλήματα παρουσιάζεται μια νέα μέθοδος τοπογραφικής ανάλυσης, με βάση τον εντοπισμό της χωρο-χρονικής προέλευσης των διαφόρων βιοακουστικών διεγέρσεων. Παρατίθενται τα αποτελέσματα αξιολόγησης της μεθόδου με τη βοήθεια πειραματικών διαδικασιών και τεχνικών προσομοίωσης. Επίσης, παρουσιάζονται εργαλεία ενεργειακής χαρτογράφησης του γαστρεντερικού ηχητικού πεδίου, τεχνικές χωρο-χρονικής ολοκλήρωσης για τη συνοπτική αποτύπωση της εντερικής δραστηριότητας σε χρωματικούς χάρτες, καθώς και εργαλεία «στερεοφωνικής ακρόασης» των επιμέρους συμβάντων.

Το έβδομο κεφάλαιο, που φέρει τον τίτλο «*Τεχνικές διαχείρισης και ανάλυσης περιεχομένου*», ενσωματώνει όλα τα προηγούμενα εργαλεία επεξεργασίας σε μια ευέλικτη δομή αρχειοθέτησης και μια αντίστοιχη οντολογία ανάλυσης, με στόχο την αποτελεσματική διαχείριση και μελέτη όλων των πειραματικών δεδομένων. Αρχικά γίνεται μια σύντομη περιγραφή του πρωτοκόλλου MPEG-7, που ως τεχνολογία αιχμής, προτείνεται για απαιτητικές εφαρμογές διαχείρισης και πλοήγησης οπτικοακουστικού περιεχομένου. Κατόπιν παρουσιάζεται η προτεινόμενη οντολογία ανάλυσης της μη-εποπτευόμενης πειραματικής διαδικασίας παρατεταμένου χρόνου, η οποία σχεδιάστηκε στα πρότυπα του MPEG-7. Στο τέλος του κεφαλαίου διατυπώνονται συμπεράσματα σχετικά με τη χρησιμότητα της νέας μεθοδολογίας στις προσεγγίσεις μελέτης της γαστρεντερικής λειτουργίας με τη βοήθεια των παραγομένων κοιλιακών ήχων. Επίσης, παρουσιάζονται οι δυνατότητες μελλοντικής εξέλιξης, η επέκταση των παραμέτρων καταγραφής (συνδυασμένη ανάλυση με άλλες διαγνωστικές μεθόδους) και η εφαρμογή της προτεινόμενης οντολογίας σε εφαρμογές ψυχο-φυσιολογικής παρακολούθησης με παρόμοια χαρακτηριστικά.

Στο όγδοο και τελευταίο κεφάλαιο διατυπώνονται τα γενικά συμπεράσματα επί του συνόλου της διατριβής. Στην ουσία πρόκειται για μια ανακεφαλαίωση της ερευνητικής συνεισφοράς κάθε ενότητας και της αποτίμησης των σχετικών παρεμβάσεων στον τομέα της ιατρικής έρευνας, της βιοϊατρικής τεχνολογίας και γενικότερα της μηχανικής. Στο τέλος του

συγγράμματος παρατίθενται οι βιβλιογραφικές αναφορές της διατριβής, καθώς και μια σύντομη περίληψη στην Αγγλική γλώσσα.

Επιτομή της συμβολής της εργασίας

Τα σημαντικότερα στοιχεία ερευνητικής συνεισφοράς και πρωτότυπης συμβολής της διατριβής συνοψίζονται στα ακόλουθα:

- Ανάπτυξη ενός πολυκαναλικού συστήματος ακουστικής καταγραφής της γαστρεντερικής κινητικότητας, χρησιμοποιώντας πιεζοηλεκτρικούς αισθητήρες επαφής και (για πρώτη φορά) ένα επιταχυνσιόμετρο τριών αξόνων. Η προτεινόμενη διάταξη χαρακτηρίζεται από μεγάλη χωρική ευαισθησία και παρέχει νέες δυνατότητες τοπογραφικής ανάλυσης. Επίσης, περιλαμβάνει τις συμπληρωματικές διατάξεις στερέωσης, προσαρμογής και ακουστικής απομόνωσης των αισθητήρων, για τον περιορισμό των θορύβων περιβάλλοντος και την αύξηση της ευαισθησίας λήψης.
- Σχεδιασμός ενός επικουρικού συστήματος οπτικοακουστικής εποπτείας, με δυνατότητες αυτοματοποιημένης ανάλυσης, βάσει των αποτελεσμάτων ανίχνευσης κίνησης βίντεο. Το σύστημα αυτό επιτρέπει την παρακολούθηση των υποκειμένων κατά τη διεξαγωγή της πειραματικής διαδικασίας, την ανίχνευση σφαλμάτων καταγραφής λόγω κίνησής τους, και τη σύνδεση των αποτελεσμάτων κίνησης με ψυχο-φυσιολογικές παραμέτρους.
- Ανάπτυξη ενός νέου αλγορίθμου καταστολής του προσθετικού θορύβου φόντου ευρείας ζώνης, συνδυάζοντας τα πλεονεκτήματα του φίλτρου WIENER και των κυματιδιακών τεχνικών. Συνολικά, προτείνονται τέσσερις εναλλακτικές υλοποιήσεις, που είναι ιδιαίτερα αποτελεσματικές σε όλες τις κατηγορίες βιοακουστικών σημάτων, σε αντίθεση με τις προγενέστερες τεχνικές, που αναπτύχθηκαν αποκλειστικά για την επεξεργασία κρουστικών ήχων. Κύρια χαρακτηριστικά της νέας προσέγγισης είναι το μικρό, σχετικά, υπολογιστικό κόστος, η εύκολη εφαρμογή της σε καταγραφές παρατεταμένης διάρκειας και η εξίσου ικανοποιητική απόδοση, τόσο στις περιοχές παρουσίας σήματος, όσο και στα διαστήματα σιγής. Η χρήση των νέων μεθόδων εντοπίζεται στις τεχνικές αυτοματοποιημένης ανάλυσης (ανίχνευση σημαντικών συμβάντων, εξαγωγή σημάτων σύννοψης, αναγνώριση προτύπων, ακουστική χαρτογράφησης, κ.ά.), καθώς και στις ιατρικές διαδικασίες γαστρεντερικής ακρόασης, όπου είναι δυνατή και η εισαγωγή αντιληπτικών-ψυχοακουστικών κριτηρίων.
- Σχεδιασμός και υλοποίηση μιας νέας διαδικασίας πολυδιακριτικής σάρωσης, για την ανίχνευση των τμημάτων σημαντικής δραστηριότητας και την εξαγωγή ενεργειακών παραμέτρων σύννοψης. Η νέα μέθοδος χαρακτηρίζεται από μεγάλη διακριτική ικανότητα εντοπισμού, χωρίς αντίστοιχα μεγάλες υπολογιστικές απαιτήσεις, παρουσιάζοντας πολλά πλεονεκτήματα για την επεξεργασία πολύ-καναλικών καταγραφών παρατεταμένου χρόνου. Παράλληλα, επιτρέπει τον χαρακτηρισμό της χωρικής «σημαντικότητας» των καναλιών και την καταχώρηση των συμβάντων σε ευέλικτα αρχεία «αναφοράς», μαζί με τις υπόλοιπες παραμέτρους περιγραφής περιεχομένου. Κατά την ανάπτυξη της μεθόδου υιοθετήθηκαν τεχνικές κυματιδιακής ανάλυσης, αντιληπτικά κριτήρια και αλγόριθμοι γεωμετρικής πολυπλοκότητας των κυματομορφών.
- Καθορισμός των κοιλιακών ακουστικών προτύπων και ανάπτυξη έμπειρων συστημάτων αυτόματης αναγνώρισης. Η πειραματική αυτή διαδικασία βασίστηκε στην αξιοποίηση της υπάρχουσας γνώσης, τη συστηματική μελέτη και την ομαδοποίηση των χαρακτηριστικών ενός μεγάλου αριθμού δειγμάτων, που διεξήχθη σε συνεργασία με την ιατρική ερευνητική ομάδα. Το αποτέλεσμα ήταν η σύσταση δύο σχημάτων (λεξικών) ταξινόμησης, όπου εκτός από τους κανόνες προσδιορισμού και περιγραφής των κλάσεων, αναπτύχθηκαν και δύο

αντίστοιχα «δίκτυα» αναγνώρισης προτύπων, χρησιμοποιώντας νευρωνικές τεχνικές και προσεγγίσεις συντακτικής - ιεραρχικής ταξινόμησης. Η εργασία αυτή είναι κεφαλαιώδους σημασίας, καθώς είναι η πρώτη φορά που πραγματοποιείται μια τόσο συστηματική μελέτη κατάταξης των βιοακουστικών φαινομένων γαστρεντερικής δραστηριότητας. Τα νέα συστήματα αναγνώρισης προτύπων επιτρέπουν, αφενός την απομόνωση των ανεπιθύμητων σημάτων παρεμβολής, αφετέρου την αξιοποίηση νέων μεθοδολογιών ανάλυσης, με βάση τη χωρική κατανομή και εξέλιξη των διαφόρων προτύπων στο χρόνο.

- Ανάπτυξη μιας νέας μεθόδου ακουστικής χαρτογράφησης, αξιοποιώντας τις συνιστώσες 3-αξονικού επιταχυνσιομέτρου κατά αντιστοιχία με τα μικρόφωνα *ηχητικού πεδίου*, και σε συνδυασμό με αριθμητικές τεχνικές ενεργειακού εντοπισμού των ηχητικών θέσεων. Τα εργαλεία αυτά παρέχουν δυνατότητες παρακολούθησης των βιοακουστικών φαινομένων με τη βοήθεια χαρτών χωρικής - ενεργειακής κατανομής και κινούμενων εικόνων αναπαράστασης του ηχητικού πεδίου. Παράλληλα, υποβοηθείται και η ακροαστική διαδικασία μέσω προτύπων αναπαραγωγής τρισδιάστατου - περιφερικού ήχου. Η προτεινόμενη μέθοδος ακουστικού εντοπισμού αξιολογήθηκε με τη βοήθεια «φυσικών» πειραματικών διατάξεων, αλλά και μέσω προσομοιώσεων λογισμικού.
- Σχεδιασμός μιας ευέλικτης οντολογίας ανάλυσης όλων των πειραματικών δεδομένων, δηλαδή των πολυκάναλων βιοακουστικών καταγραφών, των σημάτων οπτικοακουστικής παρακολούθησης και των εξαγομένων οπτικοακουστικών πληροφοριών περιγραφής του παραγόμενου βιοακουστικού πεδίου. Η προτεινόμενη οντολογία έχει σχεδιαστεί με βάση το πρότυπο MPEG-7, παρέχοντας εξελιγμένες δυνατότητες διαχείρισης περιεχομένου, αναζήτησης και πλοήγησης. Πρόκειται για μια τυποποίηση που μπορεί να χρησιμοποιηθεί, τόσο για την ανάλυση των μεμονωμένων συμβάντων μικρής διάρκειας, όσο και για τις καταγραφές παρατεταμένου χρόνου, ευνοώντας όλες τις ιατρικές ερευνητικές προσεγγίσεις βιοακουστικής μελέτης της γαστρεντερικής κινητικότητας.

Λίστα κυριότερων ακρωνυμίων

IBS (<i>Irritable Bowel Syndrome</i>)	: σύνδρομο ευερέθιστου εντέρου.
MMC (<i>Motor Migrating Complex</i>)	: περιοδικό σύμπλεγμα γαστρεντερικής κινητικότητας.
BS / IS (<i>Bowel / Intestinal Sounds</i>)	: εντερικοί ήχοι.
AS (<i>Abdominal Sounds</i>)	: κοιλιακοί ήχοι.
GIS (<i>Gastro-Intestinal Sounds</i>)	: γαστρεντερικοί ήχοι.
GIP (<i>Gastro-Intestinal Phonography</i>)	: γαστρεντερική φωνογραφία (διαδικασία ακουστικής ανάλυσης της γαστρεντερικής κινητικότητας)
A-Soundfield 4+3 (<i>Abdominal Soundfield 4+3</i>)	: διάταξη καταγραφής του κοιλιακού βιοακουστικού πεδίου, αποτελούμενη από τέσσερις περιφερειακούς αισθητήρες και ένα επιταχυνσιόμετρο τριών αξόνων στο κέντρο συμμετρίας.
FSA (<i>Foil for Sensors' Attaching</i>)	: λεπτή μεταλλική πλάκα (αλουμινίου) για τη στερέωση των βιοακουστικών αισθητήρων.
WAAV (<i>Wearable Absorbing Abdominal Vest</i>)	: ηχοαπορροφητικό γιλέκο (ζώνη) προσαρμογής που φοριέται γύρω από την κοιλιακή περιοχή για την ηχητική απομόνωση των βιοακουστικών αισθητήρων και την εξασφάλιση καλύτερης μηχανικής προσαρμογής.
SHIELD (<i>Structure for Housing, Sensors insulation and Adaptation</i>)	: προσαρμοστική διάταξη εφαρμογής του βιοακουστικού μετρητικού συστήματος, αποτελούμενη από τις διατάξεις FSA και WAAV.
ExpMA (<i>Exponential Moving Average</i>)	: διαδικασία κινούμενου εκθετικού μέσου όρου.
ABN (<i>Additive Broadband Noise</i>)	: προσθετικός θόρυβος βάθους ευρείας ζώνης.
GWN (<i>Gaussian White Noise</i>)	: λευκός θόρυβος γκαουσιανής κατανομής.
FDPWF (<i>Frequency Depended Parametric Wiener Filter</i>)	: παραμετρικό φίλτρο Wiener (PWF), εξαρτώμενο από τη συχνότητα.
JTFA (<i>Joint Time Frequency Analysis</i>)	: συνδυασμένη ανάλυση χρόνου - συχνότητας.
DWT (<i>Discrete Wavelet Transform</i>)	: διακριτός μετασχηματισμός κυματιδίων.
WP (<i>Wavelet Packets</i>)	: ανάλυση κυματιδιακών πακέτων.
QMF (<i>Quadrature Mirror Filter-banks</i>)	: κατοπτρικές τράπεζες φίλτρων, δύο καναλιών.
FD (<i>Fractal Dimension</i>)	: ανάλυση γεωμετρικής πολυπλοκότητας κυματομορφών.

WTST-NST (<i>wavelet transform-based stationary-non stationary</i>)	: κομματιδιακή μέθοδος καταστολής θορύβου με αυτόματο εντοπισμό των κατωφλίων.
WT-FD (<i>Wavelet Transform - Fractal Dimension -based</i>)	: κομματιδιακή μέθοδος καταστολής θορύβου με αυτόματο εντοπισμό των κατωφλίων, αξιοποιώντας τεχνικές FD.
FDFPA (<i>Fractal Dimension Peak Peeling Algorithm</i>)	: αλγόριθμος εντοπισμού των σημείων αιχμής (κορυφές - <i>peaks</i>) των σημάτων, αξιοποιώντας τεχνικές FD.
IKD (<i>Iterative Kurtosis-based Detector</i>)	: επαναληπτικός αλγόριθμος εντοπισμού κρουστικών σημάτων, με βάση την ανάλυση κύρτωσης.
EBS (<i>Explosive Bowel Sounds</i>)	: κρουστικοί εντερικοί ήχοι.
WDWF (<i>Wavelet Domain Wiener Filtering</i>)	: κομματιδιακό φίλτρο Wiener με επεξεργασία "σημείο - προς - σημείο" (<i>point to point</i>).
DWT-6 (<i>6-band DWT</i>)	: τοπολογία διακριτού μετασχηματισμού κομματιδίων 6 ζωνών, που χρησιμοποιήθηκε στα φίλτρα WDWF.
WP-17 (<i>17-band WP</i>)	: τοπολογία κομματιδιακών πακέτων 17 ζωνών, που χρησιμοποιήθηκε στα φίλτρα WDWF.
CRFB (<i>Constant Rate Filter Banks</i>)	: τράπεζες φίλτρων σταθερού ρυθμού δειγματοληψίας.
CCI (<i>Cross Correlation Index</i>)	: δείκτης γραμμικής συσχέτισης σημάτων.
LCF (<i>Logarithmic Crest Factor</i>)	: λογαριθμική έκφραση του «παράγοντα κλίσης» (<i>Crest Factor</i>).
LSNR (<i>Localized Signal to Noise Ratio</i>)	: τοπικός λόγος σήματος προς θόρυβο.
ESNR (<i>Effective Signal to Noise Ratio</i>)	: «αποτελεσματικός» λόγος σήματος προς θόρυβο.
SDR (<i>Signal to Deviation Ratio</i>)	: λόγος σήματος προς απόκλιση.
SDM (<i>Spectral Distortion Measure</i>)	: μέτρηση φασματικής παραμόρφωσης.
SBS (<i>Short Block Scanning / Streaming Bowel Sounds</i>)	: διαδικασία σάρωσης των καταγραφών με χρήση παραθύρων μικρού μήκους, για τον εντοπισμό και τη μεταφορά των σημαντικών τμημάτων σε αντίστοιχα *.sbs αρχεία.
MRS (<i>Multi-Resolution Scanning / Multi-channel Recording Source indexing</i>)	: διαδικασία πολυδιακριτικής σάρωσης των καταγραφών με χρήση διπλών παραθύρων, για τον εντοπισμό και την καταχώρηση του χρονισμού των συμβάντων σε *.mrs αρχεία «διεύθυνσης».
LT-WDSS (<i>Long-Term Wavelet-based Detection, Segmentation and Summarization</i>)	: κομματιδιακή μέθοδος υλοποίησης της διαδικασίας MRS, που έχει τα στάδια εντοπισμού-διαχωρισμού των συμβάντων, και εξαγωγής σημάτων «σύντοψης» των καταγραφών παρατεταμένου χρόνου.

FDPD (Fractal Dimension Pause Detector)	: μέθοδος ανίχνευσης σιγής βάσει της ανάλυσης πολυπλοκότητας κυματομορφών FD.
TD-FDPD (Time Domain -based FDPD)	: υλοποίηση της μεθόδου FDPD στο πεδίο του χρόνου.
WT-FDPD (Wavelet Transform -based FDPD)	: υλοποίηση της μεθόδου FDPD στο πεδίο του μετασχηματισμού κυματιδίων.
TDDA (Time Duration Distribution Analysis)	: ανάλυση κατανομής «χρόνου - διάρκειας» (της γαστρεντερικής δραστηριότητας).
WSE (Wavelet Spectrographic Envelope)	: σπεκτρογραφική ανάλυση που προκύπτει από τις περιβάλλουσες ισχύος των κυματιδιακών ακολουθιών.
MLP (Multi-Layer Perceptron)	: πολυεπίπεδο δίκτυο τεχνητών νευρώνων.
RPROP (Resilient Propagation)	: αλγεβρικός αλγόριθμος εκπαίδευσης πίσω διάδοσης.
ASPA (Abdominal Sound Pattern Analysis)	: ανάλυση προτύπων κοιλιακών - γαστρεντερικών ήχων.
TPDA (Temporal Pattern Distribution Analysis)	: ανάλυση κατανομής των προτύπων (κοιλιακών ήχων) στο χρόνο.
STPDA (Spatio-Temporal Pattern Distribution Analysis)	: ανάλυση κατανομής των προτύπων (κοιλιακών ήχων) στο χωρο-χρόνο.
ASCC (Abdominal Sound Coarse Classification)	: σχήμα συνοπτικής (χοντρικής) ταξινόμησης των κοιλιακών - γαστρεντερικών ήχων.
IB (Intestinal Bursts)	: εντερικοί ήχοι με έντονες κρουστικές εξάρσεις (κρουστικοί ήχοι).
RS (Regularly Sustained)	: ομαλά παραμένοντα σήματα (γαστρεντερικών ήχων).
SP (Silent Periods)	: περίοδοι εφησυχασμού της γαστρεντερικής κινητικότητας.
RESN (Respiration and Snoring)	: θόρυβοι «έντονης» αναπνοής και ροχαλητών.
SN (Sliding Noises)	: θόρυβοι τριβής λόγω κίνησης των υποκειμένων (τριβο-ηλεκτρικοί θόρυβοι).
AIMAS (Autonomous Intestinal Motility Analysis System)	: αυτόνομο σύστημα ανάλυσης της γαστρεντερικής κινητικότητας, μέσω της αναγνώρισης προτύπων.
TDF (Time-Domain Features)	: παράμετροι ανάλυσης στο πεδίο του χρόνου.
FDF (Frequency-Domain Features)	: παράμετροι ανάλυσης στο πεδίο της συχνότητας.
WDF (Wavelet-Domain Features)	: κυματιδιακές παράμετροι ανάλυσης.

PTM (Peak Transition Measure)	: παράμετρος ανάλυσης βασισμένος στα σημεία αιχμής (peaks) του σήματος περιβάλλουσας.
ASFC (Abdominal Sound Fine Classification)	: σχήμα αναλυτικής (λεπτομερούς) ταξινόμησης των κοιλιακών - γαστρεντερικών ήχων.
SCL (Solitary CLicks)	: γαστρεντερικοί ήχοι απομονωμένων κρουστικών παλμών.
RCL (Repeated CLicks)	: γαστρεντερικοί ήχοι επαναλαμβανόμενων κρουστικών παλμών.
SICS (Sequences of Irregularly Concatenated Segments)	: τυχαίες ακολουθίες «μικτών» γαστρεντερικοί ήχων.
CRSW (CRepitating SWeeps)	: «τρίζοντες» γαστρεντερικοί ήχοι.
WSSW (WhiStling SWeeps)	: «συρίτονες» γαστρεντερικοί ήχοι.
CPA (Closest Point of Approach)	: μέθοδος τοπογραφικής ανάλυσης ακουστικού πεδίου, βάσει του πλησιέστερου αισθητήρα στη θέση ηχητικής διέγερσης.
ISL (Inverse Square Low)	: νόμος αντιστρόφου τετραγώνου που καθορίζει την ιστροπική - σφαιρική κατανομή της ηχητικής έντασης.
DEBL (Direction and Energy Based Localization)	: μέθοδος εντοπισμού της διεύθυνσης και της θέσης της ηχητικής πηγής βάσει της διάταξης A-Soundfield 4+3.
TW-DEBL (Time Windowed DEBL)	: μέθοδος εντοπισμού των ηχητικών διεγέρσεων με εφαρμογή της μεθόδου DEBL σε διαδοχικά χρονικά παράθυρα επιλογής.
ESM (Equivalent Sound Map)	: ισοδύναμος ακουστικός - ενεργειακός χάρτης των συμβάντων (προκύπτει από την ολοκλήρωση όλων των TW-DEBL εικόνων).
SSM (Summarization Sound Map)	: συνολτικός ακουστικός χάρτης που προκύπτει από τη χρονική ολοκλήρωση των ESM εικόνων, των αντίστοιχων συμβάντων.
ASF-AVI (Abdominal SoundField Audio-Visual Information)	: οπτικοακουστικά αρχεία περιγραφής των γαστρεντερικών βιοακουστικών συμβάντων στο χωρο-χρόνο.
ASPM (Abdominal Sound Propagation Media Models)	: τεχνητά ομοιώματα (μοντέλα) του μέσου μετάδοσης των γαστρεντερικών ήχων στα κοιλιακά τοιχώματα.
D(s) -(Descriptors)	: MPEG-7 περιγραφείς χαμηλού επιπέδου (παράμετροι ανάλυσης).
DS(s) -(Description Schemes)	: MPEG-7 σχήματα περιγραφής (συνδυασμός περιγραφέων Ds).
GIMM DS (Gastro-Intestinal Motility Monitoring Description Scheme)	: οντολογία μελέτης της γαστρεντερικής κινητικότητας, που σχεδιάστηκε με βάση MPEG-7 σχήματα περιγραφής (DS).

Λίστα κυριότερων συμβόλων

$x(i), s(i), s^-(i)$:	{ενθόρυβο, καθαρό, αποθορυβοποιημένο}, ψηφιακό σήμα στο πεδίο χρόνου.
$X(k), S(k), S^-(k)$:	{ενθόρυβο, καθαρό, αποθορυβοποιημένο} ψηφιακό σήμα στο πεδίο της συχνότητας (ανάλυση Fourier).
$X_k(w), S_k(w), S^-_k(w) \cong$ $X(k,w), S(k,w), S^-(k,w)$:	{ενθόρυβες, καθαρές, αποθορυβοποιημένες} κομματιδιακές ακολουθίες, με "k" τη φασματική ζώνη και "w" το δείκτη των κομματιδιακών συντελεστών.
$n(i), n_{FP}(i), n^-(i)$:	{προσθετικός θόρυβος, προφίλ θορύβου, αφαιρούμενος θόρυβος}, στο πεδίο του χρόνου.
$N(k), N_{FP}(k), N^-(k)$:	{προσθετικός θόρυβος, προφίλ θορύβου, αφαιρούμενος θόρυβος}, στο πεδίο της συχνότητας (ανάλυση Fourier).
$N_k(w), N_{FP;k}(w), N^-_k(w) \cong$ $N(k,w), N_{FP}(k,w), N^-(k,w)$:	{προσθετικός θόρυβος, προφίλ θορύβου, αφαιρούμενος θόρυβος} στο πεδίο της κομματιδιακής ανάλυσης, με "k" τη φασματική ζώνη και "w" το δείκτη των κομματιδιακών συντελεστών.
$P_{x;a}(k,w) \cong \langle X(k,w) \rangle^a$:	εκτίμηση ισχύος (σημείο προς σημείο) των ενθόρυβων κομματιδιακών ακολουθιών, στη δύναμη "a", για κάθε ζώνη "k" και κάθε συντελεστή "w".
$P_{n_{FP};a}(k) \cong \langle N_{FP}(k,w) \rangle^a$:	εκτίμηση μέσης κομματιδιακής ισχύος του προφίλ θορύβου στη δύναμη "a", σε κάθε φασματική ζώνη k.
$x_m(i)$:	τμήμα ενθόρυβου σήματος στο πεδίο του χρόνου, που έχει επιλεγεί από το m-χρομικό παράθυρο.
$X_{(m)}(k,w)$:	κομματιδιακές ακολουθίες του $x_m(i)$ σήματος (στο m-χρομικό παράθυρο).
A, A_x, A_y, A_z	:	διανύσματα των συνιστωσών επιτάχυνσης του 3-αξονικού επιταχυνσιόμετρο· τα x, y, z δηλώνουν τις 3 αξονικές συνιστώσες, ενώ A είναι η συνισταμένη.
θ	:	κατακόρυφη γωνία διεύθυνσης του ηχητικού κύματος.
Φ	:	οριζόντια γωνία διεύθυνσης του ηχητικού κύματος.
$p_\zeta(i)$:	σήμα ακουστικής πίεσης του πιεζοηλεκτρικού αισθητήρα "ζ".
I_ζ	:	διάνυσμα ηχητικής έντασης που φθάνει στη θέση του αισθητήρα "ζ".
P_{Ws}	:	ακουστική ισχύς εκπομπής της εκάστοτε βιοακουστικής διέγερσης.
(X_s, Y_s, Z_s)	:	θέση βιοακουστικής πηγής (συσταλτικής διέγερσης) στο χώρο.
L_{SSM}	:	ισοδύναμη εικόνα (χάρτης) χωρικής κατανομής της ηχητικής στάθμης.

ΚΕΦΑΛΑΙΟ 1^ο

1. Αντικείμενο της διατριβής

Στο κεφάλαιο αυτό παρουσιάζονται οι διάφορες πτυχές του αντικειμένου της διατριβής, που είναι η τεχνολογική υποστήριξη της ιατρικής έρευνας στη μελέτη βιοακουστικών φαινομένων. Αρχικά, γίνεται περιγραφή του γενικού προβλήματος (problem definition) που αφορά στις μη εποπτευόμενες εφαρμογές βιοϊατρικής παρακολούθησης παρατεταμένου χρόνου. Κατόπιν, παρουσιάζονται οι ιδιαιτερότητες στη μελέτη του μηχανισμού κινητικότητας, τα στοιχεία φυσιολογίας του γαστρεντερικού συστήματος και η υπάρχουσα ερευνητική κατάσταση στον τομέα αυτό. Επίσης, δίνονται οι προδιαγραφές, τα κίνητρα και οι επιδιωκόμενοι στόχοι της συγκεκριμένης ερευνητικής προσέγγισης. Το κεφάλαιο ολοκληρώνεται με το σχεδιασμό μιας νέας μεθοδολογίας καταγραφής, διαχείρισης και ανάλυσης όλων των πειραματικών δεδομένων ιατρικής παρακολούθησης.

1.1 Περιγραφή του προβλήματος

Σε εφαρμογές βιοϊατρικής παρακολούθησης παρατεταμένου χρόνου, ανακύπτουν δυσκολίες που σχετίζονται με το μη εποπτευόμενο χαρακτήρα της διαδικασίας, τη διαχείριση ενός τεράστιου όγκου δεδομένων και τις ανάγκες αυτοματοποίησης της ανάλυσης. Όλα αυτά έρχονται να προστεθούν στις γενικότερες απαιτήσεις για ελάχιστη ενόχληση στα υπό παρακολούθηση υποκείμενα, έτσι ώστε να περιοριστεί η επίδραση της μεθόδου, να αποφευχθεί η αλλοίωση της «απόκρισής» των, και να εξασφαλισθεί η αντικειμενικότητα και αξιοπιστία των παρατηρήσεων. Ως χαρακτηριστικά παραδείγματα των παραπάνω προδιαγραφών μπορούμε να αναφέρουμε διάφορες εφαρμογές ψυχοφυσιολογικής παρακολούθησης, συμπεριλαμβανομένης της ανάλυσης βιοϊατρικών σημάτων και βιοακουστικών φαινομένων.

Στις περιπτώσεις αυτές, οι ερευνητές καλούνται να εφαρμόσουν συνδυαστικές τεχνικές ανάλυσης, με σκοπό τη συνοπτική παρουσίαση των υπό παρακολούθηση φαινομένων, την ανίχνευση των σημαντικών γεγονότων κατά μήκος των καταγραφών και την εξαγωγή των κατάλληλων παραμέτρων στατιστικής ανάλυσης. Συχνά, δε, ενσωματώνονται και διαδικασίες οπτικοακουστικής παρακολούθησης της πειραματικής διαδικασίας, με τη βοήθεια εικονοληπτών χαμηλού φωτισμού, υπέρυθρων ή και θερμικών συστημάτων. Τα εργαλεία αυτά αποσκοπούν στην έμμεση εποπτεία της δραστηριότητας των υποκειμένων και την αξιολόγηση της κατάστασής τους, τον έλεγχο πειραματικών παραμέτρων, την εξαγωγή πρόσθετων στοιχείων ανάλυσης, την αποτίμηση των εμπλεκόμενων τεχνικών και την ανίχνευση ενδεχομένων σφαλμάτων καταγραφής [1] -[14].

Σε γενικές γραμμές, η παρούσα διατριβή επιχειρεί να εισάγει δύο βασικές καινοτομίες:

α) Την πρόταση μιας νέας μεθοδολογίας καταγραφής και επεξεργασίας βιοακουστικών σημάτων σε πειραματικές διαδικασίες ψυχοφυσιολογικής παρακολούθησης παρατεταμένου χρόνου.

β) Την εξειδίκευση της μεθοδολογίας αυτής στη μελέτη της φυσιολογίας του γαστρεντερικού συστήματος με τη βοήθεια των παραγομένων κοιλιακών ήχων.

Η προτεινόμενη μέθοδος επεξεργασίας και ανάλυσης κοιλιακών ήχων, που αποτελεί και το κύριο αντικείμενο της διατριβής, επιχειρεί να παράσχει νέες δυνατότητες στη μελέτη της φυσιολογίας της γαστρεντερικής κινητικότητας και στη μελλοντική διάγνωση των αντίστοιχων λειτουργικών διαταραχών. Συγκεντρώνει, δε, ως μέθοδος, πολλές ιδιαιτερότητες, που την καθιστούν ως μια απαιτητική «περίπτωση μελέτης», παρουσιάζοντας με ανάγλυφο τρόπο τις δυσκολίες και τις ανάγκες αντιμετώπισης των διαφόρων θεμάτων, που σχετίζονται με το γενικό μοντέλο, σύνθετων καταγραφών, παρατεταμένου χρόνου.

1.2 Μελέτη του γαστρεντερικού συστήματος

Κατά τη διάρκεια των προηγούμενων δεκαετιών, έρευνες στην επιστήμη της Γαστρεντερολογίας επισήμαναν τις δυσκολίες τεκμηριωμένης διάγνωσης των λειτουργικών διαταραχών, που σχετίζονται με την κινητικότητα του λεπτού εντέρου. Το γεγονός αυτό οφείλεται στην απουσία μιας αξιόπιστης και εύκολα εφαρμόσιμης μεθόδου, που να μπορεί να χρησιμοποιηθεί για την παρατεταμένη παρακολούθηση της εντερικής δραστηριότητας. Ως αποτέλεσμα, η μελέτη των μηχανισμών κινητικότητας του γαστρεντερικού συστήματος δεν έχει προχωρήσει όσο θα έπρεπε και η ιατρική γνώση στον τομέα αυτό κρίνεται μη επαρκής. Έτσι, οι περισσότερες λειτουργικές διαταραχές εντοπίζονται μέχρι σήμερα με έμμεσο τρόπο, δηλαδή διά του αποκλεισμού άλλων ασθενειών [15]-[17].

Το σύνδρομο ευερέθιστου εντέρου (Irritable Bowel Syndrome - IBS) αποτελεί την πιο αντιπροσωπευτική περίπτωση, ενώ άλλες σχετικές δυσλειτουργίες είναι τα περιστατικά τυμπανισμού και οι ενοχλήσεις «κοιλιακών φουσκωμάτων» (abdominal bloating), η λειτουργική δυσπεψία (functional dyspepsia), η διάρροια, η δυσκοιλιότητα και ο κοιλιακός πόνος* (abdominal pain). Τα περισσότερα από τα συμπτώματα αυτά επηρεάζονται έντονα από διάφορους παράγοντες όπως είναι η διατροφή, η φαρμακευτική αγωγή, το άγχος, η παρουσία άλλων ασθενειών και οι αντίστοιχες διαδικασίες ανάρρωσης, γεγονός που χρίζει περαιτέρω διερεύνησης [15]-[18]. Στατιστικές μελέτες αναφέρουν πως ένα ποσοστό 20-40% περίπου του συνολικού πληθυσμού εκτιμάται ότι υποφέρει από τέτοιου είδους διαταραχές [19]. Στις περισσότερες, δε, των περιπτώσεων, τα σχετικά συμπτώματα αποδίδονται στο σύγχρονο τρόπο ζωής, τις διατροφικές συνήθειες και το άγχος, χωρίς να υπάρχει δυνατότητα έγκαιρης και αντικειμενικής εκτίμησης.

Από τα παραπάνω προκύπτουν δύο βασικά ανεπιθύμητα επακόλουθα. Το πρώτο αφορά το γεγονός ότι, η πλήρως τεκμηριωμένη διάγνωση απαιτεί τον αποκλεισμό όλων των εν δυνάμει παθολογιών, μέσω κλινικών μεθόδων που είναι συνήθως επεμβατικές, επίπονες και επιβαρυντικές για τον ανθρώπινο οργανισμό. Ως αποτέλεσμα, η περαιτέρω διερεύνηση αποφεύγεται στην πλειονότητα των περιστατικών και ιδιαίτερα σε περιπτώσεις όχι επιβαρηνμένου ιατρικού ιστορικού ή για σχετικά μικρές ηλικίες (ομάδες περιορισμένης επικινδυνότητας). Το γεγονός αυτό έχει οδυνηρές συνέπειες στην περίπτωση που υποβόσκει μια πιο σοβαρή παθολογία, καθώς σε πολλά περιστατικά η καθυστέρηση ιατρικής φροντίδας μπορεί να αποβεί μοιραία (αυτό είναι και το δεύτερο αρνητικό επακόλουθο). Έτσι, υπάρχει σοβαρό πρόβλημα στην έγκαιρη διάγνωση και σωστή αγωγή αντιμετώπισης των λειτουργικών διαταραχών, ενώ δεν υπάρχει δυνατότητα αντικειμενικής καταγραφής και αξιολόγησης της γαστρεντερικής κινητικότητας. Ενώ, λοιπόν, έχει διαπιστωθεί σύνδεση των μηχανισμών γαστρεντερικής λειτουργίας με την ψυχολογική κατάσταση των υποκειμένων, δεν έχει γίνει ποσοτικοποίηση αυτών των αποτελεσμάτων για τη διάγνωση και

* Εννοείται ο πόνος που δεν οφείλεται σε κάποια παθολογική κατάσταση του πεπτικού συστήματος ή κάποιου άλλου γειτονικού οργάνου, αλλά σε λειτουργικά ζητήματα κινητικότητας.

αντιμετώπιση των σχετικών προβλημάτων. Ανάλογα, δεν είναι δυνατή η παρακολούθηση των επιδράσεων της διατροφής με στόχο τη σύσταση κατάλληλης, σε κάθε περίπτωση, διαίτας. Ακόμη πιο σημαντικό είναι το γεγονός ότι δεν υπάρχει δυνατότητα αντικειμενικής εκτίμησης της επίδρασης των διαφόρων φαρμάκων στην κινητικότητα*, όπως επίσης υπάρχουν δυσκολίες αξιολόγησης της λειτουργίας του εντέρου σε μετεγχειρητικό στάδιο†.

Τα παραπάνω προβλήματα έχουν καταγραφεί από τις αρχές του προηγούμενου κίολας αιώνα, όπου διάφοροι ερευνητές αναζήτησαν τις λύσεις αξιοποιώντας τεχνικές μανομετρίας και ηλεκτρομυογραφίας, αλλά και πιο επιβαρυντικές μεθόδους ακτινοσκόπησης [20]-[22]. Οι τεχνικές γαστρεντερικής ακρόασης, στις οποίες εμπίπτει και η προτεινόμενη μέθοδος, έκαναν επίσης την εμφάνισή τους πολύ νωρίς, με την πρωτοπόρα εργασία του Cannon το 1905 [23]. Οι προσεγγίσεις μελέτης μέσω της ακουστικής ανάλυσης χαρακτηρίζονται από το κυρίαρχο πλεονέκτημα της «μη επεμβατικότητας», αλλά και από πολλές δυσκολίες που σχετίζονται με προβλήματα επεξεργασίας, διαχείρισης και ανάλυσης των καταγεγραμμένων σημάτων. Τις τελευταίες δεκαετίες, όπου η τεράστια τεχνολογική ανάπτυξη και η εξέλιξη των υπολογιστικών και επικοινωνιακών συστημάτων έδωσε ώθηση στις εφαρμογές βιοϊατρικής τεχνολογίας, το συγκεκριμένο θέμα επανήλθε στο προσκήνιο, με στόχο την επίτευξη μιας αποδοτικότερης λύσης.

1.2.1 Στοιχεία φυσιολογίας

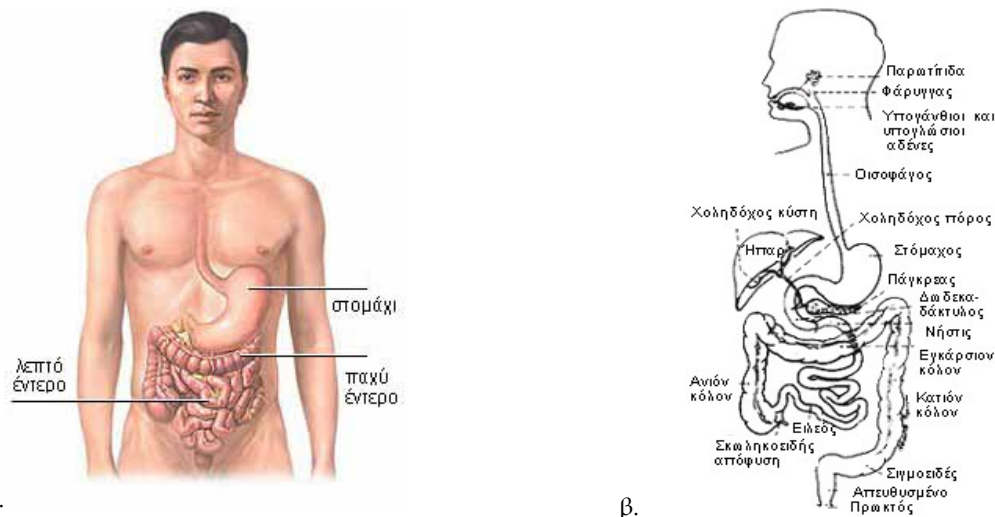
Το πεπτικό σύστημα (σχ. 1-1) αποτελείται από τον εντερικό ή πεπτικό σωλήνα (στοματική κοιλότητα, φάρυγγας, οισοφάγος, στομάχι, λεπτό και παχύ έντερο) και τους πεπτικούς αδένες. Στην ουσία πρόκειται για ένα μακρύ σωλήνα που διαπερνά το ανθρώπινο σώμα και έχει ως κύρια λειτουργία την πέψη, δηλαδή την απορρόφηση των χρήσιμων συστατικών των λαμβανομένων τροφών και την απόρριψη των αποβλήτων στο κατώτατο άκρο του παχέος εντέρου. Ολόκληρο το σύστημα βρίσκεται υπό ορμονικό έλεγχο, όπου η παρουσία τροφίμων στο στόμα προκαλεί μια αλυσίδα ορμονικών ενεργειών. Όταν υπάρχει τροφή στο στομάχι οι ορμόνες ενεργοποιούν την έκκριση οξέων, την αυξανόμενη κινητικότητα του εντέρου, την απελευθέρωση ενζύμων, κ.ά. Τα τρία βασικά τμήματα του πεπτικού συστήματος είναι το στομάχι, το λεπτό έντερο και το παχύ έντερο, ενώ επιπλέον συμπεριλαμβάνονται ο οισοφάγος, το ήπαρ και το πάγκρεας. Ο πεπτικός σωλήνας εκτείνεται από το στόμα μέχρι τον πρωκτό και έχει μήκος περί τα 15 μέτρα. Το 75 % περίπου του συνολικού του μήκους καταλαμβάνεται από το λεπτό έντερο, το οποίο θεωρείται και το κυρίως υπεύθυνο όργανο για την παραγωγή των γαστρεντερικών δονήσεων στην επιφάνεια των κοιλιακών τοιχωμάτων. Ωστόσο, σύμφωνα με τον Cannon [23], κοιλιακοί ήχοι δύναται επίσης να παραχθούν λόγω της δραστηριότητας του στομάχου, και σε πολύ πιο μικρά ποσοστά του παχέος εντέρου.

Η διαδικασία της πέψης περιλαμβάνει τα ακόλουθα στάδια. Αρχικά η τροφή τεμαχίζεται με τη διαδικασία της μάσησης και κατόπιν διαποτίζεται με τη σιέλο οπότε σχηματίζεται ο βλωμός (ημίρρευστη μάζα). Στη συνέχεια, κατά τη διαδικασία της κατάποσης, ο βλωμός μεταφέρεται μέσα από το φάρυγγα και τον οισοφάγο στο στομάχι. Το στομάχι επιτελεί πέντε σημαντικές λειτουργίες: α) προσωρινή αποθήκευση τροφής, β) έλεγχο του ρυθμού με τον οποίο η τροφή εισάγεται στο δωδεκαδάκτυλο, γ) όξινη έκκριση και

* Τα τεστ, που διεξάγονται σε δείγματα υποκειμένων (εθελοντές) πριν την έγκριση κυκλοφορίας των φαρμάκων, βασίζονται αποκλειστικά σε αξιολόγηση σχετικών ερωτηματολογίων.

† Η μετεγχειρητική λειτουργία του εντέρου κρίνεται ιδιαίτερα σημαντική στην περαιτέρω πορεία ανάρρωσης του ασθενή.

αντιβακτηριδιακή δράση, δ) ρευστοποίηση των περιεχομένων του στομάχου, και ε) προκαταρκτική πέψη. Εδώ συναντάται μία πρώτης μορφής κινητικότητα, που καλείται γαστρική κινητικότητα. Το στομάχι παρουσιάζει δύο είδη κυμάτων: τα κύματα μίξεως και τα περισταλτικά κύματα. Τα πρώτα έχουν ως στόχο τη μίξη της τροφής με τα γαστρικά υγρά για τη δημιουργία του χυμού. Τα δεύτερα έχουν ως στόχο την εκκένωση του στομάχου και την προώθηση του χυμού στο έντερο, στην περιοχή του δωδεκαδακτύλου. Στη διάρκεια των παραπάνω λειτουργιών το ήπαρ και το πάγκρεας διευκολύνουν τη μίξη των τροφών, εκκρίνοντας τη χολή και παγκρεατικό υγρό αντίστοιχα. [24]-[29]

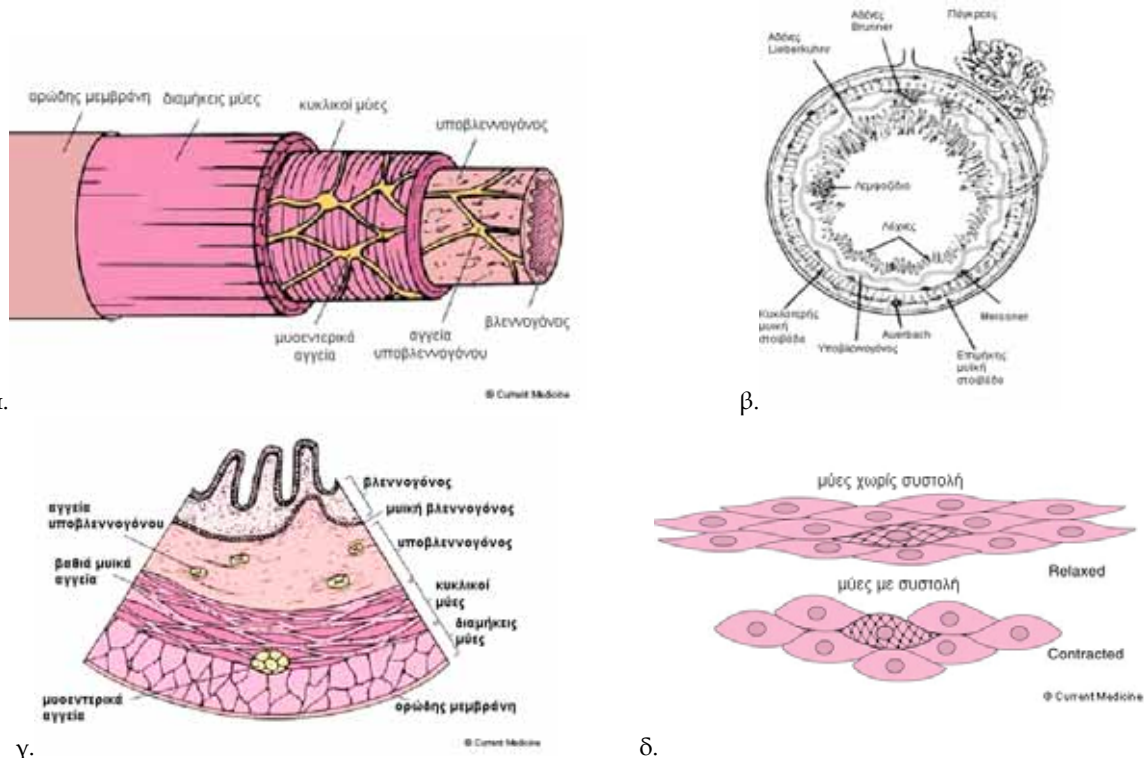


Σχήμα 1-1. Σχηματική αναπαράσταση της ανατομίας του πεπτικού συστήματος.

Στο λεπτό έντερο (σχ. 1-2) συντελείται το μεγαλύτερο μέρος της χημικής και μηχανικής πέψης, καθώς και όλη η απορρόφηση των χρήσιμων θρεπτικών συστατικών. Πρόκειται για ένα σωλήνα μικρής διαμέτρου, χαρακτηριστικό από το οποίο προήλθε και το όνομά του, και μήκους επτά μέτρων περίπου. Η ανατομία του λεπτού εντέρου έχει σαν βάση έναν απορροφητικό βλεννογόνο τύπο, με συγκεκριμένες τροποποιήσεις για κάθε διαφορετικό τμήμα. Το έντερο αποτελείται από ένα λείο μυϊκό τοίχωμα δύο στρωμάτων, όπου οι ρυθμικές συστολές πιέζουν και μεταφέρουν τα προϊόντα της πέψης (περισταλτισμός). Διακρίνονται τρία κύρια τμήματα στο λεπτό έντερο: ο δωδεκαδάκτυλος, η νήστιδα και ο ειλεός (σχ. 1-1β). Ο δωδεκαδάκτυλος είναι ένα όργανο διαμορφωμένο σε σχήμα "C" γύρω από το πάγκρεας. Η κύρια λειτουργία του είναι να εξουδετερώνει το όξινο γαστρικό περιεχόμενο και να ενεργοποιεί τη διαδικασία της περαιτέρω πέψης. Οι αδένες Brunner στον υποβλεννογόνο εκκρίνουν μια αλκαλική βλέννα που εξουδετερώνει τον όξινο χαρακτήρα και προστατεύει την επιφάνεια του δωδεκαδάκτυλου. Η νήστιδα και ο ειλεός αποτελούν μέρη του λεπτού εντέρου που είναι έντονα αναδιπλωμένα (κουλουριασμένα) με συνολικό μήκος περίπου 4-6 μέτρα. Το όριο σύνδεσης και διαχωρισμού τους δεν είναι καθορισμένο με σαφήνεια. Ο υποβλεννογόνος αυτών των τμημάτων έχει υφή έντονα διπλωμένων στρωμάτων με πτυχές (σχ. 12-β,γ), γεγονός που αυξάνει σε μεγάλο βαθμό την διαθέσιμη επιφάνεια για την απορρόφηση των θρεπτικών συστατικών.

Οι κινήσεις των τοιχωμάτων του λεπτού εντέρου διακρίνονται επίσης σε κινήσεις μίξεως και προωθήσεως. Οι κινήσεις μίξεως έχουν ως στόχο την ανάμειξη του χυμού με τα πεπτικά υγρά και την απορρόφηση των θρεπτικών ουσιών από τα τοιχώματά του. Οι κινήσεις προωθήσεως προέρχονται από περισταλτικά κύματα τα οποία προωθούνται μόνο προς την κατεύθυνση του παχέος εντέρου. Η περισταλτική δραστηριότητα του λεπτού εντέρου δεν

παρατηρείται κατά τη διάρκεια της πέψης, ενώ αντίθετα αυξάνει όταν ο γαστρικός χυμός εισέλθει στο έντερο οπότε και το διατείνει. Όσοι να φθάσουν τα προϊόντα πέψης στο παχύ έντερο, σχεδόν όλα τα χρήσιμα θρεπτικά συστατικά έχουν απορροφηθεί. Το παχύ έντερο αφαιρεί το υγρό από το περιεχόμενο που έχει απομείνει, οδηγώντας τα ημισταθερά περιττώματα στο ορθό του εντέρου για αποβολή. Λόγω της στερεάς κατάστασης του περιεχομένου του παχέος εντέρου, η παραγωγή βιοακουστικών φαινομένων στην περιοχή αυτή είναι αρκετά περιορισμένη, σε βαθμό που να θεωρείται αμελητέα [24]-[29].



Σχήμα 1-2. Σύσταση και μορφολογικά χαρακτηριστικά του λεπτού εντέρου.

1.2.2 Υπάρχουσα ιατρική κατάσταση

Πολλές μέθοδοι έχουν χρησιμοποιηθεί στο παρελθόν για τη μελέτη της γαστρεντερικής κινητικότητας, τόσο σε ανθρώπους όσο και σε πειραματόζωα. Από αυτές, ξεχωριστή θέση κατέχουν οι τεχνικές μανομετρίας και ηλεκτρομυογραφίας, βάσει των οποίων προέκυψαν και οι πρώτες ιατρικές γνώσεις γύρω από τη φυσιολογία της εντερικής κινητικότητας. Σύμφωνα με τα στοιχεία αυτά, οι περισταλτικές προωθητικές κινήσεις του λεπτού εντέρου εμφανίζονται σε κύκλους επανάληψης, δηλαδή παρουσιάζουν περιοδικότητα ως προς την εμφάνισή τους. Το σύμπλεγμα αυτών των συστατικών δραστηριοτήτων, που στη διεθνή ιατρική ορολογία αναφέρεται ως “*Motor Migrating Complex*” (MMC), έχει σαν αφετηρία την περιοχή του δωδεκαδάκτυλου και μετακινείται προς το μέρος του παχέος εντέρου μέχρι την ειλεοτυφλική βαλβίδα [20]-[22], [26]-[28], [30]-[34]. Για όσο διάστημα το στομάχι παραμένει άδειο (περίοδος νηστείας*), το περιοδικό αυτό φαινόμενο επαναλαμβάνεται σ’ ένα κύκλο διάρκειας 120-150 λεπτών, ενώ η εντερική δραστηριότητα εναλλάσσεται μεταξύ τριών βασικών φάσεων λειτουργίας. Η πρώτη φάση (phase 1) είναι αυτή του εφησυχασμού (motor quiescence). Χαρακτηρίζεται από μειωμένη εντερική δραστηριότητα (ή και πλήρη απουσία

* Πρόκειται για περίοδο που ξεκινά με την παρέλευση 2 περίπου ωρών, μετά τη λήψη γεύματος.

δράσης) και συνήθως διαρκεί 45-60 λεπτά. Καταλαμβάνει το 45-50% του συνολικού κύκλου του φαινομένου MMC [34]. Η δεύτερη φάση (phase 2) είναι γνωστή ως «ακανόνιστη» (irregular) φάση. Χαρακτηρίζεται από αλληπάλληλες εντερικές συστολές με «τυχαίο» τρόπο εμφάνισης και αρκετά διαφορετικά χαρακτηριστικά μεταξύ τους. Διαρκεί περί τα 45-84 λεπτά, καταλαμβάνοντας και πάλι ένα ποσοστό περίπου 45-50% του συνολικού κύκλου. Τέλος, η τρίτη φάση (phase 3) είναι γνωστή και ως «κανονική» φάση (regular) ή φάση ομαλής κατάστασης. Εδώ παρατηρούνται περιοδικές επαναλήψεις εντερικών συστολών με παρόμοια χαρακτηριστικά ενέργειας-διάρκειας, και με μια συχνότητα 11~13 επαναλήψεις στη διάρκεια του λεπτού. Η ομάδα αυτών των συσταλτικών κινήσεων διαρκεί περίπου 6 με 10 λεπτά σε κάθε τμήμα του εντέρου (5% του κύκλου). Κατόπιν «μεταναστεύει» προς την πλευρά του παχέος εντέρου με μια ταχύτητα 1-4 cm / λεπτό [20]-[22], [28], [34]. Αντίθετα, κατά την παρουσία γεύματος (μεταγευματική περίοδος) η παραπάνω περιοδικότητα παύει να υπάρχει). Η περίοδος αυτή χαρακτηρίζεται από την έντονη και ακανόνιστη συσταλτική δραστηριότητα της δεύτερης φάσης και μια σύντομη κατάσταση μετάβασης στην ηρεμία, η οποία από πολλούς ερευνητές αναφέρεται ως τέταρτη φάση -“phase 4” [35].

1.2.3 Τεχνικές διάγνωσης

Ραδιοϊσοτοπικές, ακτινοσκοπικές, ηλεκτρομυογραφικές και μανομετρικές τεχνικές, που χρησιμοποιούνται στη διάγνωση παθολογιών του πεπτικού συστήματος, και όχι μόνο, έχουν κατά καιρούς αξιοποιηθεί και για τη μελέτη της γαστρεντερικής κινητικότητας και των αντίστοιχων φυσιολογικών διαταραχών. Από αυτές, οι δύο τελευταίες (μανομετρία και ηλεκτρομυογραφία) παρουσιάζουν αρκετά πλεονεκτήματα, παρέχοντας τη δυνατότητα παράλληλης παρακολούθησης διαφόρων περιοχών (multi-site monitoring) και για μεγάλη χρονική διάρκεια (prolonged recordings).

1.2.3.1 Ραδιο-ισοτοπικές τεχνικές

Στις ραδιο-ισοτοπικές, ή ραδιογραφικές τεχνικές, ο εξεταζόμενος υποβάλλεται στην λήψη και κατάποση ραδιοϊσοτόπων που θα λειτουργήσουν ως δείκτες βιολογικής δράσης. Στόχος είναι ο προσδιορισμός του χρόνου, από τη στιγμή που λαμβάνεται η ουσία, μέχρι τη στιγμή που θα παρουσιαστεί σε κάποιο σημείο του εντέρου (εκτίμηση της ταχύτητας διόδου). Ο τρόπος με τον οποίο προσδιορίζεται ο χρόνος αυτός, καθώς επίσης και το είδος των «ουσιών ανιχνευτών» που χρησιμοποιούνται, διαφέρουν. Έτσι, σε ορισμένες περιπτώσεις χρησιμοποιείται ραδιοσεσυμασμένη ουσία η οποία ανιχνεύεται μετά από κάποιο χρονικό διάστημα στα κόπρανα του εξεταζόμενου. Σε άλλες πάλι τεχνικές λαμβάνεται ραδιενεργός άνθρακα (C13) που μόλις φθάσει στην περιοχή του τυφλού (αρχή του παχέος εντέρου) διασπάται και ελευθερώνεται υδρογόνο, στοιχείο που μπορεί να ανιχνευτεί στην αναπνοή του εξεταζόμενου. Έτσι, προσδιορίζεται ο χρόνος που πέρασε από τη στιγμή που λήφθηκε η ουσία, και εκτιμάται η ταχύτητα διόδου. Μια τρίτη περίπτωση είναι εκείνη στην οποία διάφορα πειραματόζωα οδηγούνται στην κατάποση ραδιοϊσοτόπων και αργότερα «θυσιάζονται» για να μελετηθεί η εμφάνιση του ισότοπου στα διάφορα τμήματα του εντέρου. Οποσδήποτε η τελευταία μέθοδος δεν είχε και δεν μπορούσε να έχει ευρύ πεδίο εφαρμογής. Εφαρμόστηκε κυρίως στα πρώτα πειράματα που διεξήχθησαν στις αρχές του προηγούμενου αιώνα [20], [24].

Κύριο μειονέκτημα των μεθόδων αυτών είναι ότι δεν μπορούν να δώσουν στοιχεία (τόσο ποσοτικά όσο και ποιοτικά) για τις μορφές συσταλτικής δραστηριότητας που παρουσιάζονται κατά μήκος του εντέρου, αφού βασίζονται σε γενικές εκτιμήσεις των

αποτελεσμάτων αυτής της δραστηριότητας. Εξ' άλλου, κάποιος μπορεί να εκτιμήσει μόνο την κίνηση των «ουσιών ανιχνευτών», χωρίς να είναι απόλυτα σίγουρος για το αν αυτή συμπίπτει με την κίνηση των τροφών, ή τον εντερικό περισταλτισμό. Η αμφιλεγόμενη ποιότητα των παρεχομένων πληροφοριών, σε συνδυασμό με το γεγονός ότι οι μέθοδοι αυτές είναι επιβαρυντικές για τον ανθρώπινο οργανισμό και επηρεάζουν σημαντικά τους μηχανισμούς εντερικής κινητικότητας, εξηγούν την περιορισμένη χρήση τους [20], [24], [25].

1.2.3.2 Ακτινοσκοπικές τεχνικές

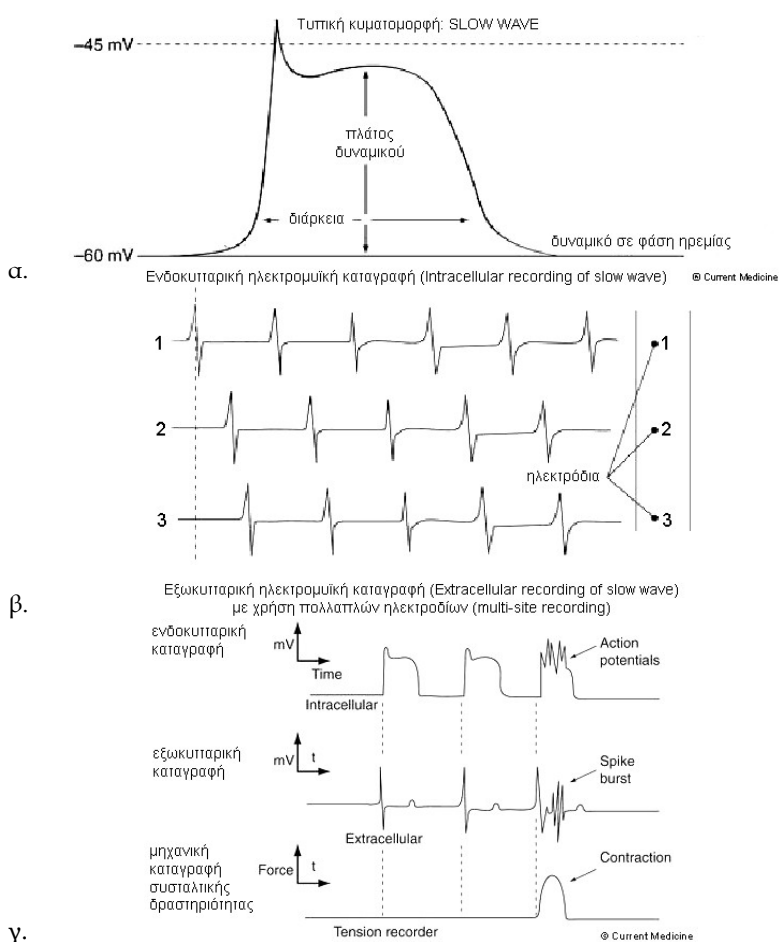
Η πρώτη εφαρμογή των ακτινοσκοπικών τεχνικών για την καταγραφή της γαστρεντερικής δραστηριότητας έγινε περίπου στις αρχές του αιώνα από τον Cannon (1902) [36]. Ο τελευταίος ανέμειξε κονσερβοποιημένο σολομό με βισμούθιο και το χρησιμοποίησε ως τροφή για γάτες. Έτσι, χωρίς να χρειασθεί να αναισθητοποιήσει τα πειραματόζωά του, μπορούσε με την βοήθεια ακτινοβολίας (γάμα κάμερα) να παρακολουθεί την κίνηση του βισμούθιου κατά μήκος του εντέρου. Μπορούσε δηλαδή να παρατηρεί την προώθηση του περιεχομένου του εντέρου, που είναι το αποτέλεσμα της συσταλτικής δραστηριότητας των τοιχωμάτων του, χωρίς, και πάλι, να μπορεί να παρακολουθεί τη δραστηριότητα αυτή κάθε-αυτή, και ακόμη περισσότερο, χωρίς να μπορεί να την εκφράσει ποσοτικά. Το γεγονός της περιορισμένης ποιοτικής και ποσοτικής αξίας στην παροχή «πληροφοριών κινητικότητας», σε συνδυασμό με τις βλαβερές συνέπειες της παρατεταμένης έκθεσης σε ακτινοβολία, καθιστούν τη μέθοδο ιδιαίτερα ασύμφορη. Η λήψη των ραδιενεργών ουσιών μπορεί να είναι επίσης επιβλαβής, αλλά και να επιδρά στους μηχανισμούς συσταλτικής δραστηριότητας. Έτσι, η συγκεκριμένη προσέγγιση δεν εφαρμόζεται σε καταγραφές παρατεταμένου χρόνου που έχουν αντικείμενο τη μελέτη της φυσιολογίας. Ωστόσο, αποτελεί τη βάση των μεθόδων φθοροσκοπίας, που χρησιμοποιούνται ευρύτατα για την αποτύπωση του εσωτερικού του εντέρου και την ανίχνευση τυχόν όγκων ή άλλων μορφωμάτων και διαταραχών ανατομικής φύσης [20], [24], [28], [37], [38].

1.2.3.3 Ηλεκτρομυογραφικές τεχνικές

Οι μέθοδοι καταγραφής της μυοηλεκτρικής γαστρεντερικής δραστηριότητας επιχειρούν να καταγράψουν τις μυϊκές συσπάσεις των εντερικών τοιχωμάτων, οι οποίες και αποτελούν τη γενεσιουργό αιτία της γαστρεντερικής συσταλτικής κινητικότητας (σχήμα 1-2δ). Πολλοί ερευνητές (κυρίως στο παρελθόν), διατύπωσαν κάποιες επιφυλάξεις για το αν η καταγραφόμενη μυοηλεκτρική δραστηριότητα είναι σε θέση να προσδιορίσει και να περιγράψει πλήρως τους μηχανισμούς περισταλτισμού. Σύμφωνα με κάποιες εργασίες [20]-[21], η σχέση αυτή δεν είναι απόλυτα γνωστή, αλλά προκύπτει κατόπιν υποθέσεων, διαδικασιών ανάλυσης και συμπερασμάτων. Επομένως η ποιότητα των παρεχομένων πληροφοριών είναι περιορισμένης αξίας και χρησιμότητας. Επιπρόσθετα, τα συμπεράσματα που προέκυψαν από τις ηλεκτρομυογραφικές μελέτες βασίστηκαν σε μεγάλο βαθμό σε πειραματόζωα, ενώ σύμφωνα με πολλούς ερευνητές [39]-[40], οι μηχανισμοί κινητικότητας και ο κύκλος MMC του ανθρώπου παρουσιάζουν σημαντικές διαφορές. Παρόλα αυτά, σήμερα είναι σε γενικές γραμμές αποδεκτό ότι η μυογενετική δραστηριότητα καθορίζει σε πολύ μεγάλο βαθμό την ένταση, τη συχνότητα, την ταχύτητα και την κατεύθυνση μετάδοσης των παραγομένων γαστρεντερικών συστολών [26], [28], [31].

Η αξιοποίηση των τεχνικών ηλεκτρομυογραφίας συνέβαλε σε μεγάλο βαθμό στη θεμελίωση της υπάρχουσας γνώσης, γύρω από τους μηχανισμούς εντερικής κινητικότητας. Έτσι, έχει διαπιστωθεί η παρουσία δύο βασικών μορφών μυϊκής δραστηριότητας, που

μεταφράζεται σε αντίστοιχες μορφές προτύπων ηλεκτρομυογραφικών σημάτων. Η πρώτη αφορά τη φάση ηρεμίας και είναι γνωστή ως Electrical Control Activity (ECA) ή «slow waves» (σχ.1-3α). Η δεύτερη αφορά φάσεις έξαρσης δραστηριότητας και χαρακτηρίζεται ως Electrical Response Activity (ERA) ή «spike bursts». Επιπρόσθετα, έχει μελετηθεί η σχέση των δύο αυτών μορφών με τις τέσσερις φάσεις του MMC κύκλου. Στην πρώτη φάση (P1) υπάρχει απουσία ή περιορισμένη εμφάνιση της μορφής ERA (little or no ERA). Στη δεύτερη (P2), παρουσιάζονται τυχαίες και εντελώς ακανόνιστες συσπάσεις ERA (irregular ERA bursts). Στην τρίτη φάση (P3) υπάρχει μία «τάξη» στην εμφάνιση της ηλεκτρομυϊκής μορφής ERA, η οποία έχει έντονο και επαναλαμβανόμενο χαρακτήρα (intense and repetitive ERA burst). Τέλος η τέταρτη φάση (P4) έχει παρόμοιο χαρακτήρα με αυτό της δεύτερης (P2, P4: irregular ERA bursts), με τη διαφορά ότι η P4 έχει μικρότερη διάρκεια από την P2, ενώ είναι πιθανό να απουσιάζει εντελώς [26], [28], [31]. Στο σχ. 1-3 απεικονίζονται χαρακτηριστικές ηλεκτρομυογραφικές καταγραφές της εντερικής δραστηριότητας, κάνοντας χρήση ενός ή περισσότερων ηλεκτροδίων λήψης, που τοποθετούνται ενδοκυτταρικά και εξωκυτταρικά. Από τη συνδυαστική ανάλυση των κυματομορφών μπορούν να εξαχθούν πληροφορίες για την ένταση και τη συχνότητα επανάληψης των συσταλτικών κινήσεων, ενώ από τις φασικές διαφορές προκύπτουν η κατεύθυνση και η ταχύτητα μετάδοσής των [28].



Σχήμα 1-3. Χαρακτηριστικά ηλεκτρομυογραφικά σήματα των εντερικών συσπάσεων

Με βάση τα παραπάνω, θα περίμενε κανείς ότι οι τεχνικές ηλεκτρομυογραφικής παρακολούθησης είναι κατάλληλες για τη μελέτη της φυσιολογίας της γαστρεντερικής κινητικότητας σε κλινικό επίπεδο. Ωστόσο, η πραγματικότητα διαφέρει, καθώς υπάρχουν

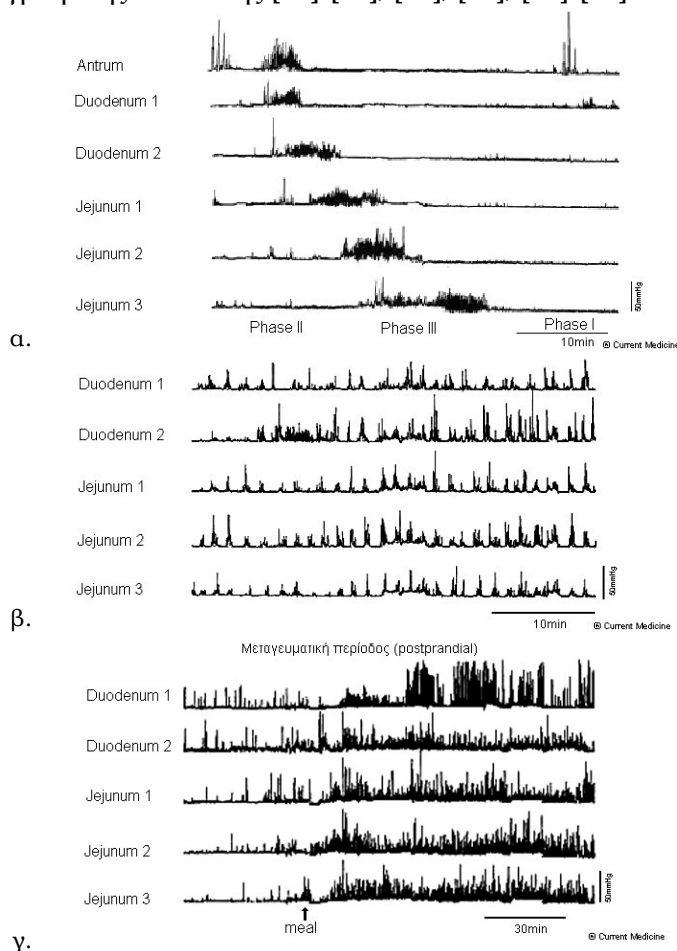
αρκετά προβλήματα κατά την τοποθέτηση των ηλεκτρομυογραφικών αισθητήρων, όπου υπάρχουν δύο βασικές επιλογές. Στην πρώτη, η τοποθέτηση των ηλεκτροδίων γίνεται στην ορογόνο επιφάνεια (εξωτερική μυϊκή στοιβάδα του λεπτού εντέρου) και συνοδεύεται από χειρουργική επέμβαση. Στη δεύτερη, τα ηλεκτρόδια εφαρμόζονται εξωτερικά, στη δερματική επιφάνεια των κοιλιακών τοιχωμάτων. Οι πληροφορίες που λαμβάνονται στην περίπτωση των επεμβατικών τεχνικών είναι αρκετά χρήσιμες και αντικειμενικές. Εντούτοις, δύσκολα εφαρμόζονται για τη διερεύνηση των λειτουργικών (μη παθολογικών) δυσλειτουργιών του ανθρώπου και ιδιαίτερα για μεγάλες χρονικές διάρκειες καταγραφής, λόγω του επεμβατικού τους χαρακτήρα. Τα στοιχεία επεμβατικότητας τις καθιστούν επώδυνες και ενοχλητικές για τα υποκείμενα, ενώ παράλληλα είναι πολύ πιθανό να επιδρούν και στους μηχανισμούς κινητικότητας, αλλοιώνοντας την αντικειμενικότητα των παρατηρήσεων.

Στην περίπτωση επιφανειακής καταγραφής, η τοποθέτηση των ηλεκτροδίων γίνεται εύκολα και ανώδυνα, γεγονός όμως που αντισταθμίζεται από πλήθος άλλων προβλημάτων. Σύμφωνα με κάποιους ερευνητές, οι αποκτούμενες πληροφορίες περιορίζονται στον προσδιορισμό της συχνότητας που αφορά τη δραστηριότητα ηλεκτρικού ελέγχου, του πυλωρικού άντρου του στομάχου [21]. Επιπρόσθετα, κατά την τοποθέτηση των ηλεκτροδίων στη δερματική περιοχή, τα σήματα πληροφορίας είναι πολύ ασθενικά κι έτσι η επίδραση των συνιστωσών θορύβου, λόγω ηλεκτρικών συζεύξεων στον περιβάλλοντα χώρο, ή λόγω τυχαίας κίνησης των υποκειμένων, είναι ιδιαίτερα έντονη. Οι ερευνητικές εργασίες επιφανειακής ηλεκτρομυογραφίας που πραγματοποιήθηκαν στο παρελθόν δεν είχαν ιδιαίτερη επιτυχία, λόγω των προβλημάτων που προαναφέρθηκαν. Ωστόσο, η εξέλιξη της τεχνολογίας επιτρέπει την αντιμετώπιση δυσκολιών που παλαιότερα θεωρούνταν ανυπέρβλητες. Για το λόγο αυτό, τα τελευταία χρόνια γίνονται προσπάθειες προς την κατεύθυνση εξέλιξης των μη επεμβατικών μεθόδων επιδερμικής ηλεκτρομυογραφίας, οι οποίες συνεχίζουν να αποτελούν σοβαρό αντικείμενο έρευνας [41].

1.2.3.4 Τεχνικές καταγραφής μηχανικής γαστρεντερικής δραστηριότητας

Οι τεχνικές καταγραφής της μηχανικής δραστηριότητας λαμβάνουν απευθείας τη μηχανική ενέργεια, που συνδέεται με τη γαστρεντερική συσταλτική κινητικότητα. Για το λόγο αυτό, πολλοί ερευνητές τις θεωρούν ως τις πιο αντικειμενικές [20], [21], καθώς επιχειρούν άμεσα να καταγράψουν και να αξιολογήσουν τα υπό μελέτη φαινόμενα. Αν εξαιρέσουμε τις τεχνικές ανάλυσης των γαστρεντερικών ήχων, που στην ουσία αποτελούν μια υποπερίπτωση αυτής της κατηγορίας, η καταγραφή της μηχανικής δραστηριότητας του εντέρου μπορεί να υλοποιηθεί με δύο τρόπους: α) χρησιμοποιώντας αισθητήρες πίεσης που καταγράφουν τη μηχανική ενέργεια των συστολών ενδοαυτικά, και β) χρησιμοποιώντας αισθητήρες που μετρούν την εκτόπιση των κοιλιακών μυϊκών τοιχωμάτων εξωτερικά (strain gauge transducers). Η δεύτερη προσέγγιση δεν έτυχε ευρείας εφαρμογής στη μελέτη της γαστρεντερικής κινητικότητας, καθώς είχε πολλά μειονεκτήματα (παρουσία έντονου εξωτερικού θορύβου λόγω κίνησης του υποκειμένου, προβλήματα στερέωσης, κ.ά.). Αντίθετα, η πρώτη προσέγγιση έτυχε ευρείας εφαρμογής. Στην περίπτωση αυτή, οι αισθητήρες τοποθετούνται στο εσωτερικό του εντέρου από την πλευρά του οισοφάγου, με κατάποσή τους από τους εξεταζομένους (ημι-επεμβατικές μέθοδοι). Για τη διευκόλυνση της τοποθέτησής τους χορηγείται συνήθως φάρμακο αναισθησίας στα υποκείμενα. Στο πλαίσιο αυτό έχουν χρησιμοποιηθεί μανομετρικοί σωλήνες καθώς και μετατροπείς στερεάς κατάστασης. Η μέθοδος που συναντάται περισσότερο, ή τουλάχιστον χρησιμοποιήθηκε στο παρελθόν, είναι αυτής της μανομετρίας (κυρίως λόγω του χαμηλού κόστους σε σχέση με τους

μετατροπείς στερεάς κατάστασης). Για το λόγο αυτό, οι συγκεκριμένες προσεγγίσεις συχνά αναφέρονται ως μέθοδοι μανομετρίας. Η τοποθέτηση πολλαπλών αισθητήρων πίεσης επιτρέπει την αναλυτικότερη καταγραφή της εντερικής δραστηριότητας, παρέχοντας δυνατότητες τοπογραφικής ανάλυσης [21]-[22], [24], [30], [32]-[34].



Σχήμα 1-4. Χαρακτηριστικές μανομετρικές καταγραφές της γαστρεντερικής δραστηριότητας: α. κύκλος MMC, β. φάση ακανόνιστης δραστηριότητας στην περίοδο νηστείας (phase II), γ. φάση ακανόνιστης δραστηριότητας στη μεταγευματική περίοδο (postprandial).

Στο σχήμα 1-4 αποτυπώνονται μανομετρικές καταγραφές, από περιοχές του πυλωρικού άνδρου (antrum), του δωδεκαδάκτυλου (duodenum 1, 2) και της νήστιδας (jejunum 1, 2, 3). Το πρώτο διάγραμμα (σχ. 1-4α) αναφέρεται σε περίοδο νηστείας, όπου είναι ευδιάκριτα τα διαφορετικά χαρακτηριστικά των τριών φάσεων του κύκλου MMC. Στη δεύτερη περίπτωση (σχ. 1-7β), παρουσιάζεται η φάση II, η οποία χαρακτηρίζεται από έντονη γενικά κινητικότητα και με σημαντικές διακυμάνσεις, χωρίς να εμφανίζεται κάποιος συγκεκριμένος κανόνας επανάληψης ή ομοιομορφίας (τυχαία-ακανόνιστη εμφάνιση). Στο σχ. 1-7γ, απεικονίζονται οι ίδιες μεταβολές ενδοαυλικής πίεσης κατά τη μεταγευματική περίοδο. Είναι χαρακτηριστική η απότομη έξαρση δραστηριότητας που λαμβάνει χώρα αμέσως μετά το γεύμα, ενώ δεν υφίσταται πλέον ο κύκλος MMC.

Στα θετικά των μανομετρικών μεθόδων συμπεριλαμβάνεται η αξιοπιστία στην καταγραφή των εσωτερικών πιέσεων, η αμεσότητα στη μελέτη του φαινομένου και η «καθαρότητα» των καταγραφών (περιορισμός θορύβου). Επίσης, τα καταγεγραμμένα σήματα, που αντιστοιχούν στις σταθμικές μεταβολές πίεσης, δεν παρουσιάζουν έντονες διακυμάνσεις στο χρόνο, κι έτσι, το περιορισμένο φασματικό τους περιεχόμενο επιτρέπει την

επιλογή μικρών ρυθμών δειγματοληψίας κατά την ψηφιοποίηση (ή την χρήση μικρών ταχυτήτων εγγραφής στα αντίστοιχα αναλογικά μέσα μαγνητικής ταινίας). Ο μικρός όγκος δεδομένων διευκολύνει τη διαχείριση των σημάτων κατά τη διαδικασία της ανάλυσης, ενώ παράλληλα ελαχιστοποιεί τις απαιτήσεις χωρητικότητας του συστήματος εγγραφής. Έτσι, σχετικά ιατρικά πρωτόκολλα διερεύνησης της εντερικής λειτουργίας έχουν εφαρμοστεί για πολύ μεγάλους χρόνους παρακολούθησης (π.χ. τρεις ή και περισσότερες μέρες), αξιοποιώντας φορητές συσκευές εγγραφής, ακόμη και με περιορισμένες τεχνολογικές δυνατότητες [21], [22], [30], [32]-[34].

Ένα άλλο πλεονέκτημα των μανομετρικών μεθόδων έχει να κάνει με την ευρεία χρησιμοποίησή τους σε ερευνητικές προσπάθειες, γεγονός στο οποίο συντέλεσαν η εύκολη χρήση, το μικρό κόστος και τα υπόλοιπα πλεονεκτήματα που αναφέρθηκαν προηγουμένως. Ως αποτέλεσμα, ένα βασικό κομμάτι της υπάρχουσας γνώσης έχει αποκτηθεί με αξιοποίηση της μανομετρίας, με κυριότερο εύρημα, αυτό του κύκλου MMC στην περίοδο νηστείας, και των αντίστοιχων φάσεων μεταγευματικής δραστηριότητας. Επίσης, έχουν γίνει απόπειρες παρακολούθησης της γαστρεντερικής κινητικότητας κατά τη μετεγχειρητική φάση [32], [34]. Παρά τα σημαντικά πλεονεκτήματα, η μελέτη του γαστρεντερικού περσταλτισμού, μέσω μανομετρικών τεχνικών, δεν προχώρησε όσο θα περίμενε κανείς. Ο κύριος λόγος είναι ο επεμβατικός τους χαρακτήρας, που προκαλεί ενοχλήσεις στα υποκείμενα, ειδικά σε καταγραφές μεγάλης χρονικής διάρκειας. Σύμφωνα με αρκετούς ερευνητές [42], μειονέκτημα αποτελεί και το γεγονός της άγνωστης επίδρασης του μανομετρικού καθετήρα στους μηχανισμούς της γαστρεντερικής λειτουργίας. Έτσι, οι τεχνικές μανομετρίας δεν εξελίχθηκαν σε κλινικά αποδεκτές διαγνωστικές μεθόδους, ενώ δεν υπήρξαν και περαιτέρω εξελίξεις στην αποκτούμενη ιατρική γνώση.

Στην προσπάθεια εξάλειψης των μειονεκτημάτων αυτών, προτάθηκαν εναλλακτικές τεχνικές καταγραφής της ενδοαυλικής πίεσης, με χρήση μικρών κάψουλων που λαμβάνονται με κατάποση από τα υποκείμενα, και στη συνέχεια εκπέμπουν τις λαμβανόμενες πληροφορίες κατά τη διαδρομή τους στον πεπτικό σωλήνα [43]. Πρόκειται για μικροσκοπικούς αισθητήρες στερεάς κατάστασης (integrated circuits), οι οποίοι συνήθως λαμβάνουν και πρόσθετες πληροφορίες εκτός από τις στάθμες πίεσης (pH, θερμοκρασία, κ.ά.) [44], ενώ ενδέχεται να ενσωματώνουν και αισθητήρες ενδοσκοπικής [45]-[47]. Ωστόσο, και οι τεχνικές αυτές δεν χρησιμοποιήθηκαν σε μεγάλο βαθμό για τη μελέτη της γαστρεντερικής κινητικότητας. Ο κύριος λόγος είναι το αυξημένο κόστος τους, καθώς κάθε κάψουλα μπορεί να αξιοποιηθεί για μία μόνο εξέταση κάθε φορά. Στις περισσότερες, δε, περιπτώσεις δεν είναι δυνατή η καταγραφή για παρατεταμένο χρονικό διάστημα [44], ενώ η μη ελεγχόμενη κίνηση των κάψουλων κατά μήκος του πεπτικού σωλήνα, αλλά και η αδυναμία παράλληλης καταγραφής διαφορετικών θέσεων, αποτελούν πρόσθετα μειονεκτήματα για τη συγκεκριμένη εφαρμογή.

Όπως έχει ήδη αναφερθεί, στην κατηγορία των μεθόδων καταγραφής της μηχανικής γαστρεντερικής δραστηριότητας, εντάσσονται και οι τεχνικές γαστρεντερικής ακουστικής ανάλυσης. Δεδομένου ότι σ' αυτό το πεδίο εντάσσεται και η προτεινόμενη μέθοδος, η πορεία εξέλιξης των σχετικών τεχνικών και η υπάρχουσα ερευνητική κατάσταση θα παρουσιαστούν πιο διεξοδικά, σε ξεχωριστή παράγραφο.

1.2.3.5 Τεχνικές ενδοσκοπικής

Οι τεχνικές ενδοσκοπικής αποβλέπουν στην αξιοποίηση κατάλληλων αισθητήρων εικονοσκοπικής, που διοχετεύονται στο εσωτερικό του ανθρωπίνου σώματος, έτσι ώστε να

αποτυπώσουν οπτικά την κατάσταση των υπό παρατήρηση οργάνων. Στην περίπτωση του πεπτικού συστήματος, χρησιμοποιούνται οι μέθοδοι γαστροσκόπησης και κολonosκόπησης, με την είσοδο κατάλληλων αγωγών - εικονοληπτών στα αντίστοιχα μέρη του ανθρώπινου σώματος. Κατά την αρχική εμφάνιση αυτών των τεχνολογιών, δεν ήταν δυνατή η παρακολούθηση μεγάλου τμήματος του λεπτού εντέρου, κυρίως λόγω ανατομικών ιδιαιτεροτήτων που δεν επέτρεπαν την πρόσβασή του από τους ενδοσκοπικούς αισθητήρες. Το κενό αυτό αρχικά καλύφθηκε με την αξιοποίηση των τεχνικών φθοροσκοπίας, αξονικής τομογραφίας και υπερηχογραφίας, όπου η καθεμιά είχε διαφορετικά μειονεκτήματα και πλεονεκτήματα [29], [37]-[38], [48]-[52]. Σήμερα, τα προβλήματα αυτά έχουν περιοριστεί σε μεγάλο βαθμό, με χρήση κατάλληλων τεχνικών οδήγησης και ελέγχου των ενδοσκοπιών, συμπεριλαμβανομένων λαπαροσκοπικών και ρομποτικών μεθόδων [38], [53]-[56]. Επίσης, έχουν προστεθεί και νέες μέθοδοι απεικόνισης - παρατήρησης, όπως οι μικροσκοπικές κάψουλες που προαναφέρθηκαν [45]-[47], καθώς και οι τεχνικές εικονικής ενδοσκόπησης μέσω συστημάτων αξονικής και μαγνητικής τομογραφίας [57]-[58].

Οι ενδοσκοπικές μέθοδοι προτείνονται κυρίως για διάγνωση παθολογιών που σχετίζονται με ανατομικά και ιστολογικά ευρήματα, και όχι για την παρατεταμένη μελέτη των μηχανισμών κινητικότητας και των αντίστοιχων λειτουργικών διαταραχών. Έτσι, ενώ η αξιοποίηση των ακουστικών γαστρεντερικών φαινομένων στη μελέτη παθολογικών καταστάσεων εκτοπίζεται σημαντικά από τις νέες μεθόδους, αυτό δε συμβαίνει με το πρόβλημα μελέτης του μηχανισμού γαστρεντερικής κινητικότητας, που παραμένει δυσεπίλυτο. Αναφορικά με την αξιοποίηση των ενδοσκοπικών τεχνικών στην παρούσα μελέτη, αυτές αξιοποιήθηκαν σε μικρό βαθμό. Συγκεκριμένα, χρησιμοποιήθηκαν, ενδεικτικά, τεχνικές φθοροσκόπησης, υπερηχογράφησης και ενδοσκόπησης, με στόχο τη διεξαγωγή παρατηρήσεων σχετικά με τους μηχανισμούς γένεσης των γαστρεντερικών ήχων και τη σύνδεσή τους με τη διαδικασία ταξινόμησης προτύπων.

1.2.4 Ακουστική ανάλυση γαστρεντερικής δραστηριότητας

Η αναζήτηση διαγνωστικών πληροφοριών στα χαρακτηριστικά των γαστρεντερικών ήχων ξεκίνησε με την πρωτοπόρα εργασία του Cannon το 1905 [23]. Ωστόσο, η αξιοποίηση των κοιλιακών ήχων στην κλινική πρακτική βρίσκεται ακόμη σε νηπιακό στάδιο, έτσι που οι αρχικές παρατηρήσεις του Cannon εξακολουθούν να είναι, εδώ και εκατό περίπου χρόνια, επίκαιρες, ή ακόμη και αξιόπεραστες. Πολλοί ερευνητές υποστηρίζουν ότι απουσία σημαντικής προόδου οφείλεται αποκλειστικά και μόνο στις δυσκολίες επιστημονικής υποστήριξης των σχετικών ιατρικών ερευνών, και σίγουρα όχι στη διαγνωστική αξία των κοιλιακών ήχων [59], [60]. Σε κάθε περίπτωση, η περιορισμένη πρόοδος μαρτυρά την πολυπλοκότητα του θέματος, αλλά και την πληθώρα των τεχνολογικών ζητημάτων που πρέπει να διευθετηθούν, έτσι ώστε οι προτεινόμενες διαγνωστικές μέθοδοι να καταστούν κλινικά εφαρμόσιμες.

1.2.4.1 Μηχανισμοί παραγωγής γαστρεντερικών ήχων

Σύμφωνα με τον Cannon [23], οι γαστρεντερικοί ήχοι παράγονται ως αποτέλεσμα των μυϊκών συσπάσεων κατά μήκος του πεπτικού σωλήνα, που στόχο έχει τον τεμαχισμό και την προώθηση του πεπτικού περιεχομένου. Οι περισταλτικές κινήσεις δημιουργούν ηχητικά φαινόμενα, που παράγονται από την τυρβώδη ροή, αλλά και τους παφλασμούς του πεπτικού χυμού πάνω στα τοιχώματα του πεπτικού σωλήνα. Έτσι, τα ακουστικά χαρακτηριστικά των κοιλιακών δονήσεων εξαρτώνται από την ένταση της συσταλτικής

δραστηριότητας, το ιξώδες του περιεχομένου, τα ανατομικά - ιστολογικά χαρακτηριστικά κατά μήκος του πεπτικού σωλήνα και τις «ελαστικές» ιδιότητες των μυϊκών ιστών στην κοιλιακή περιοχή, που αποτελούν το μέσο διάδοσης του ηχητικού κύματος.

Όσον αφορά την προέλευση των ηχητικών διεγέρσεων, αυτές δημιουργούνται στην πλωρική κατάληξη του στομάχου, κυρίως κατά μήκος του λεπτού εντέρου, που αποτελεί τη σημαντικότερη πηγή μηχανικής δραστηριότητας, καθώς επίσης και στο παχύ έντερο [23]. Η παραγωγή ακουστών δονήσεων στην τελευταία περίπτωση είναι περιορισμένη, καθώς το πεπτικό περιεχόμενο βρίσκεται ήδη σε στερεά κατάσταση. Οι τεχνικές «κοιλιακής» ακρόασης στοχεύουν κυρίως στη μελέτη της συσταλτικής δραστηριότητας του γαστρεντερικού. Αυτό αιτιολογείται με βάση τα ανατομικά χαρακτηριστικά του ανθρώπινου σώματος, όπου το μεγαλύτερο μέρος της κοιλιακής επιφάνειας βρίσκεται πλησιέστερα στο λεπτό έντερο. Ενισχύεται, δε, από το γεγονός ότι, στις περιόδους νηστείας η συσταλτική δραστηριότητα του στομάχου δεν υφίσταται. Για το λόγο αυτό, τα αντίστοιχα ακουστικά φαινόμενα συχνά αναφέρονται ως εντερικοί ήχοι (Bowel Sounds -BS, Intestinal Sounds -IS), χαρακτηρίζοντας με τον όρο αυτό και τις αντίστοιχες μεθόδους ανάλυσης. Ωστόσο, πολλοί ερευνητές προτιμούν τους πιο γενικούς όρους «κοιλιακοί» ή «γαστρεντερικοί» ήχοι (Abdominal Sounds -AS, Gastro-Intestinal Sounds -GIS), καθώς τα παραγόμενα βιοακουστικά φαινόμενα ενδέχεται να συνδέονται και με τις λειτουργίες των υπόλοιπων οργάνων του γαστρεντερικού συστήματος [61].

1.2.4.2 Γαστρεντερική ακρόαση και διαγνωστικές προσεγγίσεις

Από τη σύντομη παρουσίαση των μεθόδων διάγνωσης, γίνεται φανερό ότι όλες οι εν δυνάμει μέθοδοι παρακολούθησης παρουσιάζουν αρκετά μειονεκτήματα, είτε ως προς τις επιπτώσεις της εφαρμογής τους, είτε ως προς τα αποτελέσματα που παρέχουν. Από τις μεθόδους που δοκιμάστηκαν, η μανομετρία συγκέντρωσε τον καλύτερο συνδυασμό πλεονεκτημάτων - μειονεκτημάτων. Χαρακτηρίζεται, ωστόσο, από το κυρίαρχο πρόβλημα της επεμβατικότητας και των προκαλούμενων ενοχλήσεων στα εξεταζόμενα υποκείμενα. Στην προσπάθεια να ξεπεραστούν οι σχετικές δυσκολίες, προτάθηκαν οι τεχνικές «κοιλιακής ακουστικής ανάλυσης», οι οποίες είναι μη επεμβατικές και εντελώς ανώδυνες, προκαλώντας ελάχιστη ενόχληση στα υπό εξέταση υποκείμενα. Το στοιχείο αυτό είναι πάρα πολύ σημαντικό, καθώς οι μηχανισμοί της γαστρεντερικής λειτουργίας είναι ιδιαίτερα ευαίσθητοι σε ζητήματα επίδρασης της ψυχοφυσιολογικής κατάσταση των υποκειμένων [24], [25], [42].

Επιπρόσθετα, οι βιοακουστικές τεχνικές εξακολουθούν να διατηρούν το πλεονέκτημα της αμεσότητας στην καταγραφή της γαστρεντερικής μηχανικής δραστηριότητας, όπως παρουσιάστηκε στην περίπτωση των μανομετρικών μεθόδων. Συγκρίνοντας δε, τις δύο αυτές προσεγγίσεις μεταξύ τους, παρατηρούμε ότι οι «ακουστικές» μέθοδοι, εκτός από την παρακολούθηση των ποσοτικών παραμέτρων κινητικότητας (ενεργειακές μεταβολές των λαμβανομένων σημάτων στο χρόνο), επιτρέπουν την εξαγωγή πρόσθετων ποιοτικών δεικτών, που σχετίζονται με διάφορα ακουστικά χαρακτηριστικά (πλάτος, συχνότητα, στοιχεία κρουστικότητας, περιοδικότητα - τονικότητα, κ.ά.). Δηλαδή, οι ακουστικές πληροφορίες υπερκαλύπτουν αυτές των αντίστοιχων μανομετρικών καταγραφών, παρέχοντας νέες δυνατότητες ανάλυσης και παρακολούθησης των υπό εξέταση φαινομένων.

Σήμερα, η ακρόαση των κοιλιακών ήχων είναι ένα στοιχείο της ιατρικής πρακτικής για τη διερεύνηση συμπτωμάτων παθολογίας, που χρησιμοποιείται σε πολύ περιορισμένο βαθμό. Η χρήση της ακροαστικής εξέτασης συνίσταται, αποκλειστικά και μόνο, στην διερεύνηση της παρουσίας (ή απουσίας) ακουστικών φαινομένων, με στόχο την εξαγωγή

γενικών συμπερασμάτων που αφορούν τη λειτουργία ή όχι του λεπτού εντέρου. Στο πλαίσιο αυτό, η περιορισμένη παρουσία ήχων συνδέεται με ενδείξεις «παραλυτικού ειλεού», παθολογία στην οποία τμήμα του λεπτού εντέρου δεν λειτουργεί. Παραπλήσια ερμηνεία δίνεται και στην απουσία ήχων κατά τη μετεγχειρητική φάση, φαινόμενο που συνδέεται με τον «μετεγχειρητικό παραλυτικό ειλεό». Αντίστοιχα, η ακρόαση «έντονου περισταλισμού» δημιουργεί υποψίες για τυχόν περιστατικά απόφραξης του εντέρου (intestinal obstruction). Σε καθεμιά από τις προηγούμενες περιπτώσεις, η «κοιλιακή ακρόαση» έχει διερευνητικό και συμβουλευτικό χαρακτήρα, αδυνατώντας, από μόνη της, να παράσχει περισσότερες πληροφορίες [62], [63].

Είναι σημαντικό να αποσαφηνιστεί ότι, οι προσεγγίσεις γαστρεντερικής ακρόασης διαφέρουν σημαντικά από τις αντίστοιχες των καρδιακών και αναπνευστικών ήχων. Στις δύο τελευταίες κατηγορίες, η ακροαστική εξέταση υπήρξε για μεγάλο χρονικό διάστημα το κύριο εργαλείο διάγνωσης, γεγονός που οδήγησε σε σημαντικά ευρήματα γύρω από τη διαγνωστική αξία των βιοακουστικών αυτών φαινομένων. Δυστυχώς, δεν υπάρχει αντίστοιχα θεμελιωμένη ιατρική γνώση γύρω από τα χαρακτηριστικά των κοιλιακών ήχων. Μάλιστα, σε πολλές περιπτώσεις ενδέχεται η ακροαστική εξέταση να συμπέσει με περίοδο εφησυχασμού ή περιορισμένης συσταλτικής δραστηριότητας, όπως συμβαίνει έντονα στις περιόδους νηστείας. Στο πλαίσιο αυτό, ξεκίνησαν οι πρώτες ερευνητικές προσπάθειες για τη μελέτη των μηχανισμών παραγωγής γαστρεντερικών ακουστικών φαινομένων, και τη διερεύνηση μιας νέας ακροαστικής-διαγνωστικής προσέγγισης [23], [64]-[69].

1.2.4.3 Αξιοποίηση γαστρεντερικών βιοακουστικών επεισοδίων στη διάγνωση παθολογιών

Σε γενικές γραμμές μπορούμε να διακρίνουμε δύο ερευνητικές προσεγγίσεις στην ανάλυση κοιλιακών ήχων. Η πρώτη, που αποτελεί αντικείμενο της παρούσας παραγράφου, περιλαμβάνει διαδικασίες καταγραφής και ανάλυσης σύντομου χρονικού διαστήματος (μεμονωμένων βιοακουστικών συμβάντων) για τη μελέτη οργανικών παθολογιών, ενώ η δεύτερη αφορά σε αναλύσεις παρατεταμένου χρόνου. Στην πρώτη περίπτωση, συνήθως αξιολογούνται ακουστικά χαρακτηριστικά και ιδιότητες των σημάτων, ώστε να χρησιμοποιηθούν ως διαγνωστικά εργαλεία, για το διαχωρισμό των «υγιών κλάσεων» από τις «παθολογικές». Η προσέγγιση αυτή παρουσιάζει ομοιότητες με την ακροαστική εξέταση των πνευμονικών και καρδιακών ήχων. Αντιλαμβανόμαστε ότι η επίτευξη ενός τέτοιου στόχου έχει πάρα πολλές δυσκολίες, για διάφορους λόγους. Πρώτον, η ιατρική εμπειρία, που έχει αποκτηθεί εδώ και δεκαετίες με τις εξετάσεις γαστρεντερικής ακρόασης, δεν έχει καταδείξει αντίστοιχους μηχανισμούς διαχωρισμού. Επομένως, στις προσεγγίσεις αυτές αναζητούνται κανόνες που, κατά βάση, ξεπερνούν τις αντιληπτικές ικανότητες της ανθρώπινης εμπειρίας. Δεύτερον, τα χαρακτηριστικά των κοιλιακών ήχων παρουσιάζουν μεγάλες διακυμάνσεις ακόμη και σε καταγραφές του ίδιου υποκειμένου [24], [25]. Έτσι, δεδομένου του πολύ μικρού χρόνου παρατήρησης, υπάρχουν πρόσθετες δυσκολίες στην αναζήτηση των κατάλληλων χαρακτηριστικών παραμέτρων (features) διαχωρισμού και ταξινόμησης. Ένας τρίτος βασικός λόγος, που περιορίζει τη διεξαγωγή ερευνητικών προσπαθειών προς αυτή την κατεύθυνση, έχει να κάνει με το γεγονός ότι σήμερα είναι διαθέσιμες πολλές διαγνωστικές μέθοδοι, που χρησιμοποιούνται με επιτυχία στη διάγνωση των παθολογιών του γαστρεντερικού συστήματος.

Έτσι, περιορισμένος αριθμός εργασιών αυτής της προσέγγισης έχει καταγραφεί στη διεθνή επιστημονική βιβλιογραφία. Χαρακτηριστική είναι η εργασία διάγνωσης των περιστατικών απόφραξης του εντέρου [70], που είχε ως βάση τη φασματική ανάλυση των

παραγόμενων εντερικών ήχων. Η εργασία κατέληξε στην διαπίστωση τριών διαφορετικών τύπων φάσματος (type I, II, III), χωρίς ωστόσο να σημειωθεί περαιτέρω αξιοποίηση των ευρημάτων αυτών (αυτό προκόπτει και από τη βιβλιογραφική αναζήτηση που έγινε στο πλαίσιο της διατριβής). Πολλοί γιατροί ερμηνεύουν την εξέλιξη αυτή αναμενόμενη, καθώς τα περιστατικά απόφραξης μπορούν εύκολα να εντοπιστούν με τις «απεικονιστικές» διαγνωστικές μεθόδους (medical imaging), ενώ σε καμιά περίπτωση τα αποτελέσματα ακουστικής - φασματικής ανάλυσης δεν συνιστούν τεκμηριωμένη διάγνωση. Πρόσφατα, αναπτύχθηκε και μια νέα μέθοδος ακουστικής ανάλυσης για τη διάγνωση του «ασκήτη» (ascites) [71], δηλαδή της παθολογίας κατά την οποία παρουσιάζεται υγρό στο εσωτερικό των κοιλιακών τοιχωμάτων. Η μέθοδος βασίζεται στην καταγραφή και ανάλυση των εντερικών ήχων, αξιοποιώντας τεχνικές κυματιδιακής επεξεργασίας, και μεθόδους στατιστικής ανάλυσης ανώτερης τάξης. Πρόκειται για μια προσέγγιση αξιοποίησης των λαμβανόμενων ηχητικών κυμάτων με στόχο τον χαρακτηρισμό του μέσου διάδοσης. Σύμφωνα με τους συγγραφείς της, η προτεινόμενη μέθοδος μπορεί να αξιοποιηθεί, εναλλακτικά ή συμπληρωματικά με τις μεθόδους υπερηχογραφίας, κυρίως στα πρώτα στάδια εμφάνισης της ασθένειας [71].

1.2.4.4 Αξιοποίηση γαστρεντερικών ήχων στη μελέτη της κινητικότητας

Η δεύτερη κατηγορία ανάλυσης, όπου εντοπίζεται και η συντριπτική πλειοψηφία των ερευνητικών προσπαθειών, αποσκοπεί στην αξιολόγηση του γενικότερου μηχανισμού συσταλτικής δραστηριότητας, μέσω των κοιλιακών ήχων. Επιδιώκει, δηλαδή, τον καθορισμό των επιπέδων περισταλισμού στα διάφορα χρονικά διαστήματα. Η προσέγγιση αυτή έχει περισσότερες ομοιότητες με την κλινική εφαρμογή της γαστρεντερικής ακρόασης, όπου οι παθολόγοι διερευνούν την παρουσία (ή απουσία) των κοιλιακών ακουστικών φαινομένων. Με βάση τις βιβλιογραφικές αναφορές, οι ερευνητές της κατηγορίας αυτής τείνουν να εφαρμόζουν δύο διαφορετικές στρατηγικές. Στην πρώτη, η καταγραφή της γαστρεντερικής δραστηριότητας γίνεται για παρατεταμένο χρονικό διάστημα (συνήθως περισσότερο από 2 ώρες). Η σημαντικότερη δυσκολία αυτών των προσεγγίσεων είναι η διαχείριση ενός τεράστιου όγκου δεδομένων, στοιχείο που συχνά οδηγεί σε υιοθέτηση πολύ μικρών ρυθμών δειγματοληψίας [72], [73]. Έτσι, κατά τη διαδικασία της ανάλυσης χρησιμοποιούνται τεχνικές ολοκλήρωσης στο χρόνο και απεικονίσεις συνοπτικής παρουσίασης των καταγεγραμμένων βιοακουστικών φαινομένων. Χαρακτηριστικό παράδειγμα αποτελούν τα σήματα περιβάλλουσας ισχύος, που αναπαριστούν τις ενεργειακές μεταβολές κατά μήκος του χρόνου παρατήρησης, κατ' αντιστοιχία με τις σταθμικές μεταβολές της ενδοαυλικής πίεσης στη μανομετρία. Συχνά, δε, ενσωματώνονται και πρόσθετα εργαλεία στατιστικής ανάλυσης [25], [64], [69], [72]-[73]. Οι προσεγγίσεις αυτές παρουσιάζουν αρκετές ομοιότητες με την μανομετρία, όπου η κύρια διαφορά των δύο μεθόδων έγκειται στη φύση των λαμβανομένων «μηχανικών» σημάτων και την αντίστοιχη διαδικασία καταγραφής τους. Επίσης υπάρχουν κοινά στοιχεία με τις κλινικές ακροαστικές εξετάσεις, μόνο που η πειραματική διαδικασία επιμηκώνεται χρονικά και υποβοηθείται, πλέον, από υπολογιστικές τεχνικές (computerized auscultation).

Στη δεύτερη υποκατηγορία, χρησιμοποιούνται καταγραφές περιορισμένης διάρκειας, αλλά συγκεκριμένου χρονισμού στη φάση της εντερικής λειτουργίας (π.χ. σήματα διάρκειας δύο λεπτών που λαμβάνονται δύο ώρες μετά από το γεύμα των υποκειμένων) [42],[74]-[75], οι οποίες αναλύονται ως αντιπροσωπευτικά δείγματα του συνολικού μηχανισμού κινητικότητας. Οι σχετικές εργασίες που πραγματοποιήθηκαν με το μοντέλο αυτό [42], [60],

[74], [75], είχαν ως αντικείμενο τη μελέτη λειτουργικών διαταραχών, τη διάγνωση του συνδρόμου IBS και το διαχωρισμό του από την πάθηση του Chron. Συγκρίνοντας την προσέγγιση αυτή με τη στρατηγική που παρουσιάστηκε στην προηγούμενη παράγραφο, μπορούμε να διακρίνουμε διαφορετικά πλεονεκτήματα και μειονεκτήματα σε καθεμιά. Για παράδειγμα, τα σήματα συνοπτικής απεικόνισης των καταγραφών παρατεταμένου χρόνου, έχουν περιορισμένη περιγραφική αξία, ενώ οι μικροί ρυθμοί δειγματοληψίας και οι τεχνικές ολοκλήρωσης είναι επισφαλείς σε τυχόν σφάλματα (artifacts) καταγραφής. Στον αντίποδα, οι αναλύσεις περιορισμένης χρονικής διάρκειας ευνοούν την αποδοτικότερη διαχείριση περιεχομένου και την εφαρμογή πιο εξελιγμένων αλγορίθμων επεξεργασίας και ανάλυσης. Βέβαια, ο μικρός χρόνος παρατήρησης έχει αρνητικές επιπτώσεις στην στατιστική αξιοπιστία της μεθόδου.

Εκτός από τις παραπάνω βασικές κατευθύνσεις, οι εντερικοί ήχοι έχουν χρησιμοποιηθεί, σε μικρότερο βαθμό, σε άλλες εφαρμογές, όπως είναι οι τεχνικές βιοανατροφοδότησης με στόχο τον ψυχοφυσιολογικό έλεγχο της γαστρεντερικής κινητικότητας [76], [77], και οι μελέτες επίδρασης φαρμάκων στη συστατική δραστηριότητα [72]. Παράλληλα, υπάρχει ένας σημαντικός αριθμός δημοσιεύσεων τεχνικής φύσης, που έχουν ως κύριο αντικείμενο την ανάπτυξη αλγορίθμων ανίχνευσης - εντοπισμού των γαστρεντερικών βιοακουστικών συμβάντων και τον εμπλουτισμό (enhancement) των χρήσιμων πληροφοριακών συνιστωσών έναντι του θορύβου [59], [78]-[89]. Ωστόσο, υπάρχουν αρκετά τεχνικά ζητήματα που δεν έχουν ακόμη αντιμετωπιστεί ή δεν έχουν διευθετηθεί πλήρως. Το γεγονός αυτό οφείλεται στις δυσκολίες εξωτερικής καταγραφής των κοιλιακών δονήσεων, που σχετίζονται με την ασθενική φύση των σημάτων και τα συσχετιζόμενα προβλήματα παρεμβολών - θορύβου, την επίδραση των κοιλιακών τοιχωμάτων (ακουστικό μέσο) στα χαρακτηριστικά λήψης, τις δυσκολίες τοπογραφικής ερμηνείας και τα σχετικά φαινόμενα «συνακρόασης» (στην περίπτωση χρήσης πολλαπλών αισθητήρων), τα σφάλματα καταγραφής λόγω κίνησης των υποκειμένων ή/και άλλων παραγόντων, και γενικότερα τα προβλήματα διαχείρισης και ανάλυσης των δεδομένων, που προκύπτουν από τις μη εποπτευόμενες πειραματικές διαδικασίες παρατεταμένου χρόνου. Τα θέματα αυτά αποτελούν αντικείμενο συστηματικής έρευνας και θα καθορίσουν σε μεγάλο βαθμό την πορεία των ακουστικών μεθόδων γαστρεντερικής ανάλυσης.

1.3 Προδιαγραφές ερευνητικής προσέγγισης

Με βάση την ανάλυση που προηγήθηκε, μπορούμε πλέον να διατυπώσουμε τις βασικές προδιαγραφές της προτεινόμενης ιατρικής ερευνητικής προσέγγισης. Αυτή αφορά τη μελέτη των μηχανισμών κινητικότητας, εντάσσεται δηλαδή στη δεύτερη κατηγορία διαγνωστικών μεθόδων. Επιδιώκει, ωστόσο, την ανάπτυξη ενός γενικότερου περιβάλλοντος ιατρικής μελέτης, που έχει ως στόχο την εξαγωγή γνώσης γύρω από τους μηχανισμούς γαστρεντερικής κινητικότητας και τις διαγνωστικές δυνατότητες των κοιλιακών ήχων. Πρόκειται, δηλαδή, για μια μακρόπνη προσπάθεια ιατρική έρευνας, όπου η παρούσα διατριβή επιδιώκει να αντιμετωπίσει τα ζητήματα τεχνικής φύσεως και να υλοποιήσει τον αρχικό σχεδιασμό του συστήματος βιοϊατρικής τεχνολογίας.

Η προτεινόμενη μεθοδολογία επιχειρεί να αξιοποιήσει τα πλεονεκτήματα όλων των προσεγγίσεων επεξεργασίας κοιλιακών ήχων, παρέχοντας νέες δυνατότητες διαχείρισης και ανάδειξης της υποβόσκουσας ιατρικής γνώσης. Δηλαδή, το υιοθετούμενο πρωτόκολλο συνδυάζει τεχνικές ανάλυσης παρατεταμένου χρόνου, που εφαρμόζονται στο συνολικό μήκος των καταγραφών, αλλά και ποιο εξελιγμένους αλγόριθμους επεξεργασίας, για τα

μεμονωμένα ηχητικά στιγμιότυπα μικρότερης χρονικής διάρκειας. Στο πλαίσιο αυτό συμπεριλαμβάνονται τεχνικές προ-επεξεργασίας (για τον εμπλουτισμό των σημάτων και την απομάκρυνση των συνιστωσών θορύβου), διαδικασίες χρονικής ολοκλήρωσης - περιληψης (για την εξαγωγή παραμέτρων σύνοψης), εργαλεία στατιστικής ανάλυσης και αλγόριθμοι εντοπισμού - οριοθέτησης των επιμέρους συμβάντων γαστρεντερικής δραστηριότητας, που αποτελούν τις δομικές μονάδες περιγραφής του συνολικού φαινομένου. Παράλληλα, εισάγεται η χρήση πολλαπλών αισθητήρων καταγραφής, με στόχο την αντιμετώπιση δυσκολιών που σχετίζονται με τη λήψη των «αδύναμων» κοιλιακών ήχων, την εξάλειψη των φαινομένων συνακρόασης και τη δυνατότητα παροχής στοιχείων τοπογραφικής ανάλυσης. Εν συντομία, η προτεινόμενη μεθοδολογία καταγραφής, επεξεργασίας και ανάλυσης περιλαμβάνει τα ακόλουθα στάδια:

α) πολυ-καναλική λήψη των υπό μελέτη βιοακουστικών φαινομένων, ικανοποιώντας τις απαραίτητες προϋποθέσεις ασφάλειας των υποκειμένων και τις απαιτήσεις «ποιότητας» των ηχογραφήσεων,

β) εξαγωγή παραμέτρων συνοπτικής παρουσίασης των πληροφοριών, μέσα από σήματα περιβάλλουσας παρατεταμένου χρόνου, που παρουσιάζουν τις μεταβολές ενέργειας-χρόνου, και ενέργειας-χρόνου-συχνότητας,

γ) εφαρμογή τεχνικών σάρωσης και ανίχνευσης σήματος, για τον εντοπισμό των σημαντικών συμβάντων γαστρεντερικής δραστηριότητας, και την απομόνωσή τους από τις περιόδους εφησυχασμού,

δ) ανάλυση ηχητικών προτύπων (pattern analysis - classification) στα καταγεγραμμένα συμβάντα γαστρεντερικής δραστηριότητας, με στόχο την απομάκρυνση των ανεπιθύμητων κλάσεων (σήματα παρεμβολών) αλλά και την παροχή πιο λειτουργικών και περιγραφικών δομών ανάλυσης (time pattern distribution analysis, spatio-temporal distribution analysis),

ε) εφαρμογή τεχνικών χωρικής και χωρο-χρονικής ανάλυσης του παραγόμενου ηχητικού πεδίου, με την ενσωμάτωση εργαλείων ακουστικής χαρτογράφησης - απεικόνισης (sound field visualization) και στερεοφωνικής ηχητικής αναπαραγωγής (stereo - surround sound),

στ) ενοποίηση όλων των προηγούμενων παραμέτρων σε ευέλικτες δομές διαχείρισης περιεχομένου και ανάλυσης, παρέχοντας δυνατότητες πλοήγησης, αναζήτησης και σύγκρισης μεταξύ «στιγμιότυπων» διαφορετικών πειραματικών καταγραφών.

Στα παραπάνω θα πρέπει να προστεθούν οι δομημένες περιγραφές και οι «σχολιασμοί» (structural annotations) των πειραματικών διαδικασιών και των υποκειμένων. Αυτό γίνεται έτσι ώστε να υπάρχει δυνατότητα σύνδεσης των διαφόρων πειραματικών στοιχείων (συμπτώματα, ιατρικό ιστορικό, δημογραφικά στοιχεία, ψυχοφυσιολογική κατάσταση, συνθήκες πειράματος, ειδικά ιατρικά πρωτοκόλλα, κ.ά.) με τα αποτελέσματα ανάλυσης των κοιλιακών ήχων. Επίσης, αποφασίστηκε η ενσωμάτωση συστήματος οπτικοακουστικής παρακολούθησης (video surveillance) της πειραματικής διαδικασίας, για τους ακόλουθους λόγους:

- δυνατότητα παρακολούθησης της «κατάστασης» του υποκειμένου, ιδιαίτερα κατά τις νυκτερινές καταγραφές σε περίοδο ύπνου,

- ανάλυση των μηχανισμών παραγωγής μη επιθυμητών σημάτων, όπως είναι οι ήχοι αναπνοής - ροχαλητών και τα σήματα τριβής λόγω κίνησης, με στόχο την αντιμετώπισή τους μέσω συστημάτων αυτόματης αναγνώρισης,
- δυνατότητα ελέγχου της πειραματικής διαδικασίας και αποτίμηση της επίδρασης των σφαλμάτων (artifacts) και των αντίστοιχων εργαλείων αυτόματης αναγνώρισης (artifact/motion detection),
- διευθέτηση όλων των τεχνικών ζητημάτων και σχεδιασμός του αντίστοιχου περιβάλλοντος παρακολούθησης, που θα επιτρέψει τη μελλοντική αξιοποίηση του συστήματος σε συνδυαστικές μελέτες με άλλες μεθόδους, (π.χ. εργαστήριο ύπνου, καρδιο-αναπνευστικές καταγραφές, γνωστικές δοκιμασίες, ψυχοφυσιολογικές εξετάσεις, κ.ά.).

Υγιείς εθελοντές, δηλαδή άτομα χωρίς ιατρικό ιστορικό δυσλειτουργιών του λεπτού εντέρου, επιλέχθηκαν αρχικά ως η καταλληλότερη ομάδα για την πραγματοποίηση των πειραμάτων. Παράλληλα, συμπεριλήφθηκαν και ενδεικτικές καταγραφές μη υγιών υποκειμένων (περιστατικά ασθένειας Chron, γαστρεντερίτιδας, αποφρακτικού ειλεού, παραλυτικού ειλεού, ολικής γαστρεκτομής), για λόγους πληρότητας της πειραματικής διαδικασίας, και κυρίως κατά τη φάση καθορισμού των προτύπων κοιλιακών ήχων. Οι καταγραφές διεξήχθησαν, στον κύριο όγκο τους, στις εγκαταστάσεις του Γενικού Περιφερειακού Νοσοκομείου Παπαγεωργίου, όπου για τις ανάγκες των πειραμάτων διαμορφώθηκε κατάλληλα μία αίθουσα εξέτασης. Εναλλακτικά, χρησιμοποιήθηκαν και φορητά συστήματα για διεξαγωγή πειραμάτων σε άλλους χώρους, ανάλογα με τη διαθεσιμότητα υποκειμένων (υγιών και μη).

Οι πρώτες καταγραφές πραγματοποιήθηκαν σε περιόδους «νηστείας» (fasting state), δηλαδή μετά την πάροδο αρκετού χρονικού από τη λήψη γεύματος. Πρακτικά αυτό σημαίνει ότι θα πρέπει να έχουν παρέλθει τουλάχιστον δύο ώρες από τη στιγμή που τα υποκείμενα έλαβαν το τελευταίο γεύμα. Έτσι, τα πειράματα διεξήχθησαν συνήθως στη διάρκεια νυκτερινού ύπνου, ή πρωινές ώρες με επαρκή χρονική απόσταση από τη στιγμή του πρωινού γεύματος. Οι καταγραφές στη φάση ύπνου παρουσιάζουν το πλεονέκτημα της ελάχιστης ενόχλησης των υποκειμένων, ενώ δεν απαιτείται περαιτέρω δέσμευση στον ελεύθερο χρόνο τους. Ωστόσο, οι ανεξέλεγκτες σωματικές κινήσεις δημιουργούν αρκετά σφάλματα καταγραφής λόγω θορύβων τριβής. Η επιλογή της περιόδου νηστείας προτάθηκε από τους ιατρούς της ερευνητικής ομάδας ως η καταλληλότερη για την παρακολούθηση της εντερικής συστατικής δραστηριότητας. Επιπρόσθετα, συμπεριλήφθησαν και καταγραφές μεταγευματικής περιόδου (postprandial state) κυρίως για λόγους σύγκρισης και πληρότητας της ερευνητικής διαδικασίας. Η διάρκεια των διαθέσιμων καταγραφών κυμαίνεται από μία έως και έξι ώρες, με πιο συνηθισμένη αυτή των δύο ωρών, καθώς συμπεριλαμβάνει έναν πλήρη κύκλο (MMC) εντερικής δραστηριότητας. Παράλληλα, κρίθηκε χρήσιμη και η πραγματοποίηση πολλών δειγματοληπτικών καταγραφών μικρότερης χρονικής διάρκειας (μερικών λεπτών), κυρίως για την εξασφάλιση μεγάλου στατιστικού δείγματος, όσον αφορά τα ποιοτικά χαρακτηριστικά των κοιλιακών ήχων και την ανάλυση προτύπων.

Στο πλαίσιο αυτό, η Επιτροπή Ηθικής του νοσοκομείου ενέκρινε σχετικό πρωτόκολλο για τη διεξαγωγή των πειραμάτων (αριθμός 22Δ / 21-11-2000), βάσει του οποίου ενημερώνονται όλοι οι εθελοντές και υπογράφουν τη σχετική έντυπη φόρμα οικειοθελούς συμμετοχής στη διαδικασία.

1.4 Σχεδιασμός του συστήματος βιοϊατρικής παρακολούθησης

Η ικανοποίηση των παραπάνω προδιαγραφών απαντήθηκε με το σχεδιασμό ενός συστήματος βιοϊατρικής παρακολούθησης, που περιλαμβάνει τις υπομονάδες καταγραφής, επεξεργασίας και ανάλυσης. Όσον αφορά τα δύο πρώτα στοιχεία, οι σχετικές υλοποιήσεις επικεντρώθηκαν στη λήψη και επεξεργασία των γαστρεντερικών βιοακουστικών καταγραφών, αλλά και των σημάτων οπτικοακουστικής εποπτείας. Τα εργαλεία που αναπτύχθηκαν ικανοποιούν τις βασικές απαιτήσεις του συγκεκριμένου προβλήματος, ενώ πολλά στοιχεία μπορούν να ενσωματωθούν και να αξιοποιηθούν σε παραπλήσιες εφαρμογές βιοακουστικής και γενικότερα ψυχοφυσιολογικής παρακολούθησης.

Σχετικά με τις διαδικασίες μελέτης και «ερμηνείας» των πειραματικών δεδομένων, δόθηκε ιδιαίτερη έμφαση στην ανάπτυξη ενός ευέλικτου περιβάλλοντος διαχείρισης των πληροφοριών και μιας αντίστοιχης οντολογίας ανάλυσης. Τα ζητήματα σχεδιασμού και αξιοποίησης οντολογιών ιατρικής μελέτης, αρχίζουν να προβάλλονται ως ιδιαίτερα σημαντικά σε όλο και περισσότερες εφαρμογές ψυχο-φυσιολογικής παρακολούθησης του ανθρώπου, και κυρίως σε δι-επιστημονικές συνεργασίες [90]-[93]. Η διαδικασία αυτή συμπεριλαμβάνει τη σύνταξη κατάλληλων κανόνων περιγραφής, το σχεδιασμό ευέλικτων δομών αρχειοθέτησης και διαχείρισης των πληροφοριών, την αξιοποίηση στοιχείων ελέγχου των πειραματικών δοκιμασιών, καθώς και την ανάδειξη των βέλτιστων στοιχείων επεξεργασίας και ανάλυσης, με στόχο την εξαγωγή γνώσης [8], [90]-[93].

Στην προκειμένη περίπτωση, οι μηχανισμοί γαστρεντερικής κινητικότητας συνδέονται με πολλές παραμέτρους της ψυχοφυσιολογίας του ανθρώπου, ενώ η αντίστοιχη γνώση στον τομέα αυτό, βρίσκεται σε πρωταρχικό επίπεδο. Έτσι, δόθηκε ιδιαίτερα μέριμνα στις δυνατότητες μελλοντικής επέκτασης του περιβάλλοντος ανάλυσης, και τη συνδυαστική μελέτη μέσω διαφόρων μεθόδων βιοϊατρικής παρακολούθησης. Ως παραδείγματα μπορούμε να αναφέρουμε τη σύνδεση των κύκλων γαστρεντερικής κινητικότητας με τους κύκλους του ύπνου και τα σχετικά επεισόδια «άπνοιας» (ή άλλων ψυχοσωματικών διαταραχών). Επίσης, την επίδραση των διαφόρων στρεσο-παραγόντων, μέσω δοκιμασιών που εφαρμόζονται στην πειραματική ψυχολογία [94]-[95]. Την ανάπτυξη πρωτοκόλλων για τον έλεγχο της ποιοτικής και ποσοτικής επίδρασης των φαρμάκων, αλλά και των διατροφικών στοιχείων, στους μηχανισμούς της ανθρώπινης πέψης. Την προσαρμογή των παραπάνω μεθοδολογιών στις ανάγκες παιδιατρικών εφαρμογών, καθώς και σε άλλες ειδικές περιπτώσεις ανάρρωσης και «αποκατάστασης» ασθενών.

1.5 Επισκόπηση 1ου κεφαλαίου

Στο κεφάλαιο αυτό έγινε ανάλυση του αντικειμένου της παρούσης διατριβής, που αφορά στην τεχνολογική υποστήριξη της ιατρικής έρευνας για την καταγραφή και ανάλυση της γαστρεντερικής μηχανικής δραστηριότητας. Αρχικά παρουσιάστηκαν τα στοιχεία φυσιολογίας του πεπτικού συστήματος. Επίσης, έγινε μια αναλυτική καταγραφή των ιατρικών διαγνωστικών μεθόδων, που αξιοποιούνται στην ανίχνευση των διαφόρων παθολογιών. Ιδιαίτερη έμφαση δόθηκε στα προβλήματα λειτουργικών διαταραχών, δηλαδή των δυσλειτουργιών εκείνων που δεν προέρχονται από οργανικές παθήσεις και για τις οποίες δεν έχουν, έως και σήμερα, αναπτυχθεί τεχνικές άμεσης διάγνωσης - αντιμετώπισης. Στο πλαίσιο αυτό, παρουσιάστηκαν οι σχετικές ερευνητικές μέθοδοι που έχουν δοκιμασθεί μέχρι σήμερα, οι οποίες αφορούν κυρίως στις τεχνικές ηλεκτρομυογραφικής, μανομετρικής

και ακουστικής καταγραφής. Επίσης, αναλύθηκαν τα πλεονεκτήματα και μειονεκτήματα της κάθε προσέγγισης.

Με βάση την υπάρχουσα ερευνητική κατάσταση (state of research), παρουσιάστηκε ο προτεινόμενος σχεδιασμός βιοϊατρικής καταγραφής και ανάλυσης, ο οποίος περιλαμβάνει την πολυκαναλική λήψη και εγγραφή των παραγόμενων βιοακουστικών φαινομένων, την οπτικοακουστική παρακολούθηση της πειραματικής διαδικασίας και την ενσωμάτωση όλων των παρατηρήσεων σε μια ευέλικτη δομή επεξεργασίας, διαχείρισης περιεχομένου και ανάλυσης. Η νέα μεθοδολογία επιδιώκει την άμεση αντιμετώπιση των προαναφερθέντων προβλημάτων, τη μαζική εφαρμογή του ιατρικού πρωτοκόλλου και την αποτελεσματική διαχείριση των πειραματικών στοιχείων, με στόχο την επιτάχυνση της ιατρικής έρευνας.

ΚΕΦΑΛΑΙΟ 2^ο

2. Περιγραφή του συστήματος καταγραφής

Στο κεφάλαιο αυτό παρουσιάζονται τα ζητήματα που αφορούν στη λήψη και ψηφιακή εγγραφή των πειραματικών δεδομένων, δηλαδή των βιοακουστικών καταγραφών (κυρίως), αλλά και των σημάτων οπτικοακουστικής παρακολούθησης. Αρχικά, γίνεται μια σύντομη θεωρητική εισαγωγή, που αφορά τις τεχνολογίες ήχου και εικόνας, με ιδιαίτερη έμφαση στις σχετικές εφαρμογές βιοϊατρικής τεχνολογίας και τα θέματα που απασχόλησαν την τρέχουσα εργασία. Κατόπιν, αναλύονται οι ιδιαιτερότητες του συγκεκριμένου προβλήματος, περιγράφονται οι τεχνικές προδιαγραφές και ο σχεδιασμός των μονάδων καταγραφής, παρατίθενται οι λύσεις που δοκιμάστηκαν και, τελικά, επιλέχθηκαν. Τέλος, διατυπώνονται συμπεράσματα για τα στοιχεία καινοτομίας, αποδοτικότητας και λειτουργικότητας του αναπτυχθέντος συστήματος καταγραφής.

2.1 Συστήματα ήχου και εικόνας στη βιοϊατρική τεχνολογία

Τα συστήματα οπτικοακουστικής καταγραφής και επεξεργασίας αποτελούν, σήμερα, κοινό τόπο στις περισσότερες εφαρμογές βιοϊατρικής τεχνολογίας. Το γεγονός αυτό οφείλεται στην πληθώρα των πληροφοριών που παρέχουν τα σήματα ήχου και εικόνας, γεγονός που οδήγησε στην αξιοποίησή τους ως ιατρικά εργαλεία διάγνωσης, από πολύ νωρίς. Έτσι, οι ηλεκτρακουστικοί μετατροπείς και οι εικονολήπτες αποτέλεσαν προεκτάσεις των αισθητηρίων του ανθρώπου, εξελίσσοντας παράλληλα τις διαδικασίες φυσικής εξέτασης στην σημερινή μορφή των προηγμένων μεθόδων διάγνωσης.

Τα συστήματα βιοακουστικής μελέτης προέκυψαν ως εξέλιξη των αντίστοιχων ακροαστικών μεθόδων διάγνωσης. Το πρώτο διαγνωστικό εργαλείο στον τομέα αυτό ήταν το στηθοσκόπιο, το οποίο καθόρισε σε μεγάλο βαθμό τη μετέπειτα πορεία των μεθόδων ακουστικής διάγνωσης. Κατά αναλογία, τα συστήματα ιατρικής απεικόνισης (medical imaging) επιχείρησαν να καταγράψουν πληροφορίες από το εσωτερικό του ανθρώπινου οργανισμού, όπου δεν υπήρχε δυνατότητα παρατήρησης με τις διαδικασίες φυσικής εξέτασης. Σήμερα, στην καθημερινή ιατρική πρακτική χρησιμοποιούνται πλήθος εργαλείων εικονογράφησης και ενδοσκόπησης, που στόχο έχουν αποτύπωση της κατάστασης των ανθρώπινων οργάνων, και την εξέτασή τους μέσα από ακίνητες και κινούμενες εικόνες. Παράλληλα, σε πολλές εφαρμογές χρησιμοποιούνται συστήματα παρακολούθησης (video surveillance), που στόχο έχουν τον έλεγχο της κατάστασης των υποκειμένων και των αντίστοιχων διαδικασιών εξέτασης [1]-[14], [90]-[93].

2.1.1 Συστήματα καταγραφής βιοακουστικών σημάτων

Είναι γνωστό ότι ένα μεγάλο μέρος των αναγκών ιατρικής διάγνωσης καλύπτεται από συστήματα βιοακουστικής μελέτης, δηλαδή μεθόδων ιατρικής εξέτασης μέσω ακουστικών σημάτων. Χαρακτηριστικά παραδείγματα αποτελούν οι παγιωμένες μέθοδοι ανάλυσης των καρδιακών και πνευμονικών ήχων [96]-[98], οι τεχνικές ακρόασης των εμβρύων [99]-[100] και οι μέθοδοι ακουστικής καταγραφής των μεταβολών πίεσης του αίματος (blood pressure

auscultation) [98], [101]. Παράλληλα, υπάρχει έντονη ερευνητική δραστηριότητα σε παραπλήσια πεδία όπως, στην αξιοποίηση των βιοακουστικών σημάτων σε ψυχοφυσιολογικές δοκιμασίες και εφαρμογές βιοανατροφοδότησης [76]-[77], [94]-[95], [98], [102], στην αξιολόγηση και διάγνωση των φωνητικών διαταραχών (voice disorders) [102]-[104], στην ανάλυση των μυϊκών και αρθρικών ήχων [98], κ.ά. Στην ίδια κατηγορία εντάσσονται και οι ερευνητικές προσπάθειες μελέτης των κουλιακών ήχων [24].

2.1.1.1 Ακροαστική διάγνωση μέσω στηθοσκοπιών

Ο όρος ακρόαση (auscultation) αναφέρεται στο άκουσμα των εσωτερικών ήχων του σώματος, που συνήθως χρησιμοποιούνται για τους σκοπούς της εξέτασης του καρδιακού, αναπνευστικού και γαστρεντερικού συστήματος. Ο τεχνικός αυτός όρος εφευρέθηκε από το Γάλλο παθολόγο Laennec και προέρχεται από το λατινικό ρήμα auscultare, που σημαίνει ακροάζομαι. Το πρώτο μέσο που χρησιμοποιήθηκε για τους σκοπούς της ακροαστικής διάγνωσης, ήταν το στηθοσκόπιο. Τα ακουστικά στηθοσκόπια είναι μέσα μηχανικής σύζευξης με τις δονούμενες επιφάνειες του ανθρωπίνου σώματος, με στόχο τη λήψη και ακροαστική ερμηνεία των αντίστοιχων βιοακουστικών φαινομένων. Το πρόβλημα με τα ακουστικά στηθοσκόπια εντοπίζεται στο γεγονός ότι οι παραγόμενοι ήχοι έχουν εξαιρετικά χαμηλά επίπεδα στάθμης, καθιστώντας τη διαδικασία διάγνωσης δύσκολη. Για το λόγο αυτό εφευρέθηκαν τα ηλεκτρονικά στηθοσκόπια, τα οποία αξιοποιούν τις βασικές αρχές και τους μηχανισμούς λειτουργίας των μικροφώνων. Δηλαδή, περιέχουν έναν ηλεκτρακουστικό μετατροπέα που λαμβάνει τις μηχανικές δονήσεις και τις μετατρέπει σε ηλεκτρικές. Έτσι, υπάρχει δυνατότητα πρόσθετων διεργασιών όπως είναι η ενίσχυση, η απομόνωση μη επιθυμητών συνιστωσών (φιλτράρισμα), η εφαρμογή πρόσθετων τεχνικών επεξεργασίας και η αποθήκευση των σημάτων σε μονάδες εγγραφής. Οι συσκευές αυτές συχνά αναφέρονται και ως «στηθόφωνα». [96], [98], [105].

Σύμφωνα με πολλούς ερευνητές [98], το στηθοσκόπιο ως αισθητήρας λήψης ακουστικών σημάτων, ποτέ δεν είχε τα επιθυμητά χαρακτηριστικά απόκρισης, όπως αυτά διατυπώνονται σε σύγχρονες εφαρμογές μηχανικής και ακουστικών μετρήσεων. Το γεγονός αυτό έχει να κάνει με τις συνθήκες εφεύρεσης του στηθοσκοπίου και τα πρώτα βήματα της χρήσης του, όπου το ζήτημα της σταθερής ευαισθησίας καθ' όλο το μήκος των ακουστών συχνοτήτων (επίπεδη απόκριση συχνοτήτων) δεν αποτέλεσε βασικό αντικείμενο μελέτης. Αυτό συνέβη τόσο για λόγους περιορισμένων τεχνολογικών δυνατοτήτων, όσο και για λόγους μικρού ερευνητικού ενδιαφέροντος. Με την εξέλιξη της τεχνολογίας και των μεθόδων διεξαγωγής ακουστικών μετρήσεων, πολλοί ερευνητές-μηχανικοί, έθεσαν το παραπάνω ζήτημα. Ωστόσο, η ιατρική κοινότητα «αποφάσισε» τη διατήρηση των βασικών χαρακτηριστικών των στηθοσκοπιών για λόγους που σχετίζονται με την αποκτηθείσα ιατρική γνώση και εμπειρία. Σήμερα, πολλά ηλεκτρονικά στηθοσκόπια παρέχουν δυνατότητες επιλογής μεταξύ πολλών συναρτήσεων απόκρισης, ενώ αρκετοί ερευνητές προτιμούν να χρησιμοποιούν απευθείας διάφορους ηλεκτρακουστικούς αισθητήρες (μικρόφωνα, επιταχυνσιόμετρα) [98], [105], [106].

2.1.1.2 Βασικές αρχές ηλεκτρακουστικών μετατροπέων λήψης

Οι ηλεκτρακουστικοί μετατροπείς δέχονται στην είσοδο τους ακουστικά κύματα και παράγουν στην έξοδο ηλεκτρικά σήματα. Πρόκειται δηλαδή για «μετασχηματιστές», που λαμβάνουν μηχανική ενέργεια και τη μετατρέπουν σε ηλεκτρική. Στην κατηγορία αυτή εντάσσονται τα μικρόφωνα και τα επιταχυνσιόμετρα, που αποτελούν τα κύρια μέσα λήψης των ακουστικών πιέσεων και των αντίστοιχων δονήσεων [107]-[109]. Κατά την «πρόσληψη»

ηχητικών σημάτων στις διάφορες εφαρμογές ηχοληψίας και ακουστικών μετρήσεων, υπάρχουν κάποιες βασικές προδιαγραφές που θα πρέπει να ικανοποιούνται. Έτσι, για να επιτευχθεί υψηλή πιστότητα (high fidelity) κατά την ηχογράφηση, τα μικρόφωνα και οι αισθητήρες καταγραφής δονήσεων πρέπει να χαρακτηρίζονται από τα ακόλουθα κύρια χαρακτηριστικά: α) ελάχιστη επίδραση στο ηχητικό πεδίο, β) πιστότητα στην απόδοση, και γ) ικανοποιητικό εύρος λειτουργίας [107]-[109].

Το παραγόμενο ηλεκτρικό σήμα εξόδου θα πρέπει να είναι ανεξάρτητο από το μέγεθος και το σχήμα του ηλεκτρακουστικού μετατροπέα, ελαχιστοποιώντας τα τυχόν κυματικά φαινόμενα απόσβεσης - σκέδασης, που αυτός προκαλεί δρώντας ως εμπόδιο (λόγω των μη μηδενικών διαστάσεων του). Αυτό έχει ιδιαίτερη σημασία στις υψηλές συχνότητες, όπου τα μήκη κύματος είναι συγκρίσιμα με τις φυσικές διαστάσεις των αισθητήρων. Αντίστοιχο είναι το φαινόμενο της «βαρυτικής φόρτισης» των δονούμενων επιφανειών, από τη μάζα του επιταχυνσιομέτρου. Επιπλέον, η απόκριση του μετατροπέα θα πρέπει να είναι σταθερή, ανεξάρτητα από τη στάθμη ή το φασματικό περιεχόμενο του ακουστικού σήματος εισόδου, ενώ σε κάθε περίπτωση η περιοχή λειτουργίας του αισθητήρα θα πρέπει να είναι συγκρίσιμη με τα αντίστοιχα χαρακτηριστικά των προς παρακολούθηση ακουστικών φαινομένων. Το χαρακτηριστικό αυτό αποδίδεται με το εύρος δυναμικής περιοχής, δηλαδή το λόγο της μέγιστης προς την ελάχιστη επιτρεπτή τιμή εισόδου [107]-[111].

2.1.1.3 Μικρόφωνα

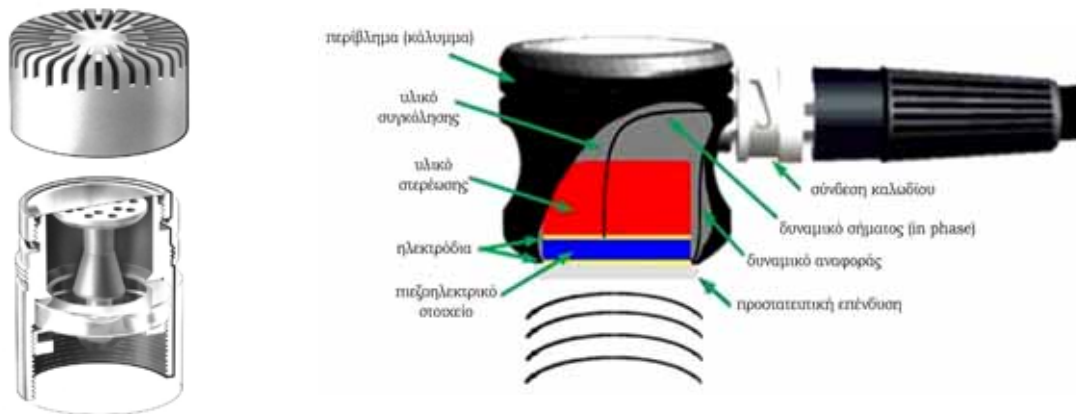
Στα μικρόφωνα συναντούμε τριών ειδών παραμέτρους: ακουστικές, μηχανικές και ηλεκτρικές. Για να επιτευχθεί πιστή σύλληψη και ηλεκτρική μετατροπή θα πρέπει: α) το ηχητικό κύμα που θα φτάσει στο διάφραγμα του μικροφώνου να είναι πιστό αντίγραφο αυτού που θέλουμε να καταγράψουμε, β) το διάφραγμα να ταλαντωθεί με τα ίδια ακριβώς χαρακτηριστικά που έχει η κυματομορφή του προσπίπτοντος κύματος, γ) οι κινήσεις του διαφράγματος να μετατραπούν σε ακριβές αντίγραφο μεταβολών τάσης ή ρεύματος [109], [112]. Υπάρχουν διάφοροι μηχανισμοί λειτουργίας, όπως είναι οι ηλεκτρομαγνητικοί και ηλεκτροδυναμικοί (μικρόφωνα ταινίας -Ribbon, κινούμενου πηνίου), η μεταβολή της ηλεκτρικής αντίστασης (μικρόφωνο άνθρακα), η αξιοποίηση του πιεζοηλεκτρικού και του ηλεκτροστατικού φαινομένου (κρυσταλλικά και πυκνωτικά μικρόφωνα, αντίστοιχα). Η παραγόμενη τάση εξόδου μπορεί να είναι ανάλογη, είτε προς την ταχύτητα, είτε προς το πλάτος της απομακρύνσεως της ταλάντωσης. Έτσι, διακρίνονται οι κατηγορίες των μικροφώνων σταθερής ταχύτητας και των μικροφώνων σταθερού πλάτους [107]-[111].

Τα κύρια χαρακτηριστικά των μικροφώνων είναι η ευαισθησία, η απόκριση συχνοτήτων, η σύνθετη αντίσταση, η δυναμική περιοχή και η κατευθυντικότητα. Η ευαισθησία ορίζεται ως η πραγματική τάση που παράγεται στην έξοδο του μικροφώνου, με διέγερση ακουστικό τόνο συχνότητας 1kHz και στάθμης πίεσης 74dB (0,1Pa). Έτσι, η ευαισθησία εκφράζεται σε μονάδες mV/Pa ή V/dB. Στην πράξη, τα ηχητικά σήματα είναι συνήθως πιο σύνθετα από τους καθαρούς τόνους, γεγονός που καθιστά αναγκαία την απόδοση της καμπύλης, που θα καθορίζει την ευαισθησία σε όλο το μήκος των ακουστικών συχνοτήτων. Πρόκειται για την απόκριση συχνοτήτων, δηλαδή την καμπύλη μεταβολών της στάθμης εξόδου του μικροφώνου (σε dB) ως προς τη συχνότητα, η οποία πρέπει να παραμένει επίπεδη (με μικρές αποκλίσεις της τάξης των ± 3 dB). Η κατευθυντικότητα, αντίστοιχα, καθορίζει την ευαισθησία του μικροφώνου σε συνάρτηση με τη γωνία πρόσπτωσης των ηχητικών κυμάτων. Όσον αφορά τη δυναμική περιοχή, το κατώτερο όριο λειτουργίας προκύπτει από την ισοδύναμη στάθμη εσωτερικού θορύβου του μικροφώνου,

που οφείλεται, κατά κύριο λόγο, σε θερμικά φαινόμενα. Έτσι, εάν η στάθμη του λαμβανομένου σήματος βρίσκεται στα ίδια επίπεδα με αυτή του θορύβου, το αποτέλεσμα είναι μια κακή σηματοθορυβική σχέση στην έξοδο. Αντίστοιχα, η μέγιστη επιτρεπτή είσοδος καθορίζεται από το όριο γραμμικής λειτουργίας του αισθητήρα, σημείο στο οποίο το μικρόφωνο παράγει αρμονική παραμόρφωση 3%. Αν η τιμή αυτή ξεπεραστεί, τότε έχουμε υπεροδήγηση του μετατροπέα, με έντονα φαινόμενα αρμονικής παραμόρφωσης (Total Harmonic Distortion - THD). Η σύνθετη αντίσταση εκφράζει το μέτρο της εμπέδησης που παρουσιάζεται στην έξοδο του μικροφώνου, και οφείλεται στο σύνολο των ωμικών, επαγωγικών και ηλεκτροστατικών φαινομένων που λαμβάνουν χώρα κατά τη λειτουργία του. Πρόκειται για μέγεθος που μεταβάλλεται με τη συχνότητα (η ονομαστική του τιμή ορίζεται στη συχνότητα 1kHz), το οποίο παίζει καθοριστικό ρόλο στον τρόπο σύνδεσης του μικροφώνου με τις επόμενες κυκλωματικές βαθμίδες, και τα σχετικά ζητήματα προσαρμογής [107]-[112].

Ανάμεσα στις διάφορες κατηγορίες μικροφώνων, σε εφαρμογές βιοακουστικής έχουν χρησιμοποιηθεί περισσότερο τα πυκνωτικά και πιεζοηλεκτρικά (σχ. 2-1) ενώ σε μικρότερο βαθμό αξιοποιήθηκαν και δυναμικά μικρόφωνα [96], [98], [106], [113]. Οι δύο αυτές βασικές κατηγορίες δοκιμάστηκαν και στο πλαίσιο της παρούσης διατριβής. Το κύριο πλεονέκτημα των πυκνωτικών μικροφώνων είναι η μικρή μάζα του διαφράγματος (σχ. 2-1α), γεγονός που επιτρέπει την άμεση ανταπόκριση στις απότομες ηχητικές μεταβολές. Έτσι, χαρακτηρίζονται από πολύ καλή απόκριση συχνοτήτων, ενώ παράλληλα μπορούν να έχουν τα επιθυμητά (ανά περίπτωση) κατευθυντικά χαρακτηριστικά [107]-[113]. Το γεγονός αυτό τα κάνει ιδανικά για ηχογραφήσεις υψηλής πιστότητας (φωνητικά-στούντιο), καθώς και για απαιτητικές εφαρμογές ακουστικών μετρήσεων. Τα πυκνωτικά μικρόφωνα μπορούν να κατασκευαστούν σε πολύ μικρά μεγέθη, παρέχοντας νέες λειτουργικές δυνατότητες που είναι ιδιαίτερα επιθυμητές σε πολλές εφαρμογές, συμπεριλαμβανομένων και εφαρμογών βιοακουστικής [96], [98], [106]. Ως αρνητικά στοιχεία μπορούν να θεωρηθούν το σχετικά υψηλό κόστος, η υψηλή αντίσταση εξόδου (της τάξης των MΩ), που καθιστά απαραίτητη τη χρήση προενισχυτικών βαθμίδων και η ανάγκη τροφοδοσίας για την πόλωση του πυκνωτή του μικροφώνου [107]-[113].

Τα κρυσταλλικά μικρόφωνα (σχ. 2-1β) διαθέτουν έναν πιεζοηλεκτρικό κρύσταλλο, που έχει την ιδιότητα να παράγει ηλεκτρικά φορτία όταν πάνω του εξασκούνται μηχανικές δυνάμεις (πιεζοηλεκτρικό φαινόμενο). Το κύριο πλεονέκτημα αυτών των στοιχείων είναι ότι είναι ιδιαίτερα ευαίσθητα, παρέχοντας πολύ μεγάλη (συγκριτικά) τάση στην έξοδό τους. Χρησιμοποιούνται κυρίως ως αισθητήρες επαφής σε εφαρμογές ηχογράφησης μουσικών οργάνων, σε συστήματα ελέγχου μεγαφώνων - ηχείων, αλλά και σε εφαρμογές βιοϊατρικής τεχνολογίας (ηλεκτρονικά στηθοσκόπια, συστήματα υπερήχων, κ.ά.). Αναφορικά με την τελευταία κατηγορία, ιδιαίτερα θετικό είναι το γεγονός ότι αποτελούν παθητικά στοιχεία που δεν χρειάζονται τροφοδοσία για να λειτουργήσουν, ενώ επίσης βολικό είναι το μικρό τους μέγεθος και το χαμηλό τους κόστος. Η σύνθετη αντίσταση των πιεζοηλεκτρικών μετατροπέων κυμαίνεται μεταξύ 10 και 40 kΩ, ανάλογα με τη συχνότητα, γεγονός που καθιστά απαγορευτική τη χρήση αγωγών μεγάλου μήκους, καθώς δημιουργούνται προβλήματα θορύβου. Στα αρνητικά στοιχεία συμπεριλαμβάνονται οι όχι τόσο καλοί παράμετροι απόδοσης, σε σχέση με αυτές του πυκνωτικού μικροφώνου. Έτσι, η καμπύλη απόκρισης συχνοτήτων δεν είναι τόσο επίπεδη, κυρίως για τις χαμηλές συχνότητες (<100Hz), όπου η ευαισθησία λήψης μειώνεται σημαντικά. Επίσης, το πιεζοηλεκτρικό μικρόφωνο έχει σχετικά υψηλά επίπεδα εσωτερικού θορύβου [24], [112], [114], [115].



α.

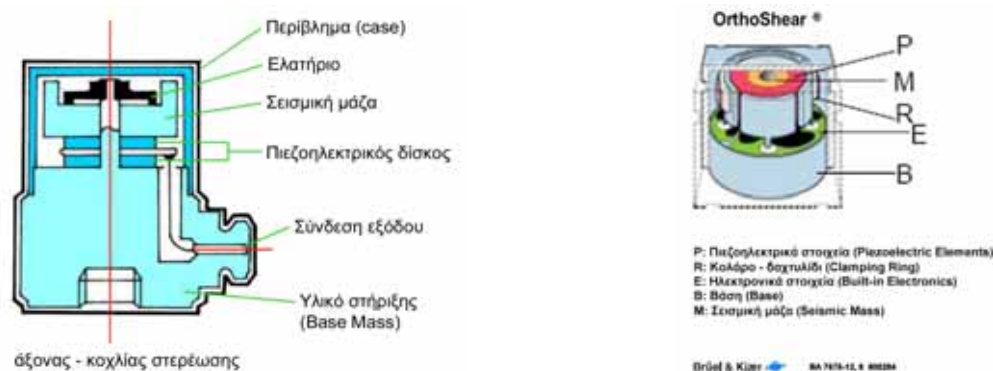
β.

Σχήμα 2-1. Σχηματική αναπαράσταση ηλεκτρακουστικών μετατροπέων: α. πυκνωτικό μικρόφωνο, β. πιεζοηλεκτρικό μικρόφωνο επαφής.

2.1.1.4 Επιταχυνσιόμετρα

Τα επιταχυνσιόμετρα αποτελούν ηλεκτρακουστικούς μετατροπείς που σχεδιάστηκαν για τη λήψη «στερεοφερτών» ηχητικών κυμάτων, δηλαδή δονήσεων. Παρουσιάζουν πολλές λειτουργικές ομοιότητες με τα κρυσταλλικά μικρόφωνα επαφής. Διακρίνουμε τα πιεζοηλεκτρικά, τα πιεζο-ωμικά (piezoresistive), τα πυκνωτικά και τα σερβο-δυναμικά (servo accelerometers - force balance) επιταχυνσιόμετρα [114], [116], [117]. Το βασικό μοντέλο λειτουργίας τους (σχ. 2-2α) περιλαμβάνει μία σεισμική μάζα, η οποία εφάπτεται και βρίσκεται σε μηχανική σύζευξη με το περίβλημα, όπου εξασκούνται οι εξωτερικές δονήσεις. Έτσι, η διάταξη του ελατηρίου μεταφέρει τη δόνηση στη σεισμική μάζα, ασκώντας, τελικά, δυνάμεις στο στοιχείο του ηλεκτρομηχανικού μετατροπέα (πιεζοηλεκτρικός δίσκος εν προκειμένω). Το ηλεκτρικό σήμα που παράγεται στην έξοδο είναι ανάλογο προς την επιβαλλόμενη μηχανική διέγερση. Με τον τρόπο αυτό καταγράφονται τα μεγέθη (πλάτη ή ενεργές τιμές) της μετατόπισης, της ταχύτητας και της επιτάχυνσης της ταλάντωσης. Το σήμα της ταχύτητας προηγείται φασικά αυτό της μετατόπισης κατά $\pi/2$, ενώ η ταχύτητα, με τη σειρά της, υπολείπεται της επιτάχυνσης κατά $\pi/2$. Η επιλογή των παραμέτρων μετατόπισης, ταχύτητας ή επιτάχυνσης εξαρτάται από το είδος της εφαρμογής, και σχετίζεται κυρίως με τις ανάγκες φασματικής ανάλυσης [117]-[118]. Πέρα από τις φασικές διαφορές των παραπάνω μεγεθών, η φασική πληροφορία (πολικότητα) αντιπροσωπεύει τη φορά της επιβαλλόμενης δύναμης [114]-[115], [119]-[120].

Τα πιεζοηλεκτρικά επιταχυνσιόμετρα (σχ. 2-2α) έχουν παρόμοια λειτουργικά στοιχεία με τα κρυσταλλικά μικρόφωνα επαφής. Η κύρια διαφορά τους έγκειται στα χαρακτηριστικά του μετατροπέα αλλά και στον τρόπο προσαρμογής - στερέωσης του κρυστάλλου (housing), στοιχείο που καθορίζει σε μεγάλο βαθμό την απόδοση του αισθητήρα. Έτσι, τα επιταχυνσιόμετρα έχουν καλύτερες προδιαγραφές (ευαισθησία, συχνотική απόκριση, δυναμική περιοχή, κλπ), επιτρέποντας την ικανοποιητική λήψη ιδιαίτερα ασθενικών δονήσεων, αλλά και σε πολύ χαμηλές συχνότητες (υπόηχοι). Επίσης, υπάρχουν διαφορές σχετικά με τη διαμόρφωση του σήματος εξόδου, καθώς το ζητούμενο δεν είναι πλέον η καταγραφή της ακουστικής πίεσης. Σήμερα, τα πιεζοηλεκτρικά επιταχυνσιόμετρα καλύπτουν το μεγαλύτερο ποσοστό των εφαρμογών ανάλυσης δονήσεων, ενώ από πολλούς ερευνητές θεωρούνται ως η καλύτερη κατηγορία επιταχυνσιόμετρων [120]-[121].



α.

β.

Σχήμα 2-2. Βασικό μοντέλο λειτουργίας επιταχυνσιόμετρων: α. πιεζοηλεκτρικό επιταχυνσιόμετρο συμπίεσης, β. πιεζοηλεκτρικό επιταχυνσιόμετρο διάτμησης, τύπος OrthoShear τριών αξόνων.

Υπάρχουν δύο κατηγορίες πιεζοηλεκτρικών επιταχυνσιόμετρων, τα συμπίεσης – “compression type” (σχ. 2-2α) και τα διάτμησης – “shear type” (σχ. 2-2β). Στην πρώτη περίπτωση ασκούνται δυνάμεις συμπίεσης, ενώ στη δεύτερη δυνάμεις διάτμησης. Τα επιταχυνσιόμετρα συμπίεσης χαρακτηρίζονται από σχετικά υψηλό λόγο ευαισθησίας προς μάζα (sensitivity to mass ratio). Τα επιταχυνσιόμετρα διάτμησης είναι λιγότερο ευαίσθητα σε επιδράσεις του περιβάλλοντος, ενώ είναι δυνατός ο σχεδιασμός μικροσκοπικών διατάξεων (miniature accelerometers) παρέχοντας επίσης υψηλούς λόγους ευαισθησίας προς μάζα. Σε πολλές εφαρμογές, όπου απαιτείται ταυτόχρονη καταγραφή των δονήσεων στους τρεις βασικούς άξονες (tri-axial accelerometer), γίνεται χρήση πολλαπλών πιεζοηλεκτρικών στοιχείων, που τοποθετούνται σε κατάλληλες διατάξεις στους «πολύ-αξονικούς» αισθητήρες. Στην κατηγορία αυτή ανήκει και ο πιεζοηλεκτρικός τύπος OrthoShear (σχήμα 2-2β) της εταιρίας Brüel & Kjær [121], τεχνολογία που αξιοποιήθηκε για τις ανάγκες της διατριβής. Αντίθετα με τους κοινούς τύπους πολλαπλών αξόνων, ο τύπος OrthoShear περιλαμβάνει μία ενιαία σεισμική μάζα, η οποία βρίσκεται στο κέντρο συμμετρίας (μάζας) και περιβάλλεται από έναν πιεζοηλεκτρικό δακτύλιο. Έτσι, παρέχεται η δυνατότητα εξαγωγής τριών διαφορετικών σημάτων (X, Y και Z έξοδοι, ένα σήμα για κάθε άξονα), που λαμβάνονται στις αντίστοιχες τερματικές συνδέσεις (κόμβους) του δακτυλίου [121]-[122].

Όπως και στην περίπτωση των μικροφώνων, η επιλογή του κατάλληλου επιταχυνσιόμετρου γίνεται με βάση τα χαρακτηριστικά απόδοσης (απόκριση συχνοτήτων, λόγος ευαισθησίας-μάζας, δυναμική περιοχή) και τις απαιτήσεις της εκάστοτε εφαρμογής. Η καμπύλη απόκρισης εκφράζει την ευαισθησία του αισθητήρα σε συνάρτηση με τη συχνότητα. Η διαφορά, σε σχέση με τα μικρόφωνα, είναι ότι εμφανίζεται μια χαρακτηριστική απότομη κορύφωση (συχνότητα συντονισμού), η οποία καθορίζει το ανώτερο συχνотικό όριο λειτουργίας του αισθητήρα (~1/3 της συχνότητας συντονισμού), έτσι ώστε να μην έχουν σφάλμα μεγαλύτερο από 12%. Στα επιταχυνσιόμετρα είναι δυνατή η καταγραφή πολύ χαμηλών - μη ακουστών - συχνοτήτων (υπόηχοι), ενώ σε ορισμένους τύπους αυτό το όριο αντιστοιχεί στη μηδενική (DC) συχνότητα [114], [119], [121]. Γενικά, οι υψηλές τιμές ευαισθησίας είναι επιθυμητές, καθώς επιτρέπουν την καταγραφή πολύ ασθενικών δονήσεων. Στην πράξη, η μεγάλη ευαισθησία μεταφράζεται (συνήθως) σε μεγάλη μάζα του επιταχυνσιόμετρου, στοιχείο που δεν είναι και τόσο θετικό, καθώς αύξηση του βάρους συνεπάγεται μείωση της συχνότητας συντονισμού. Επίσης, ιδιαίτερη προσοχή χρειάζεται όταν η δονούμενη επιφάνεια έχει μικρό μέγεθος, ώστε να μην αλλοιώνονται οι

αναπτυσσόμενες δονήσεις, λόγω «βαρυτικής» φόρτισης του αντικειμένου (mass – loading). Ένας πρακτικός κανόνας είναι ότι, η μάζα του επιταχυνσιομέτρου δεν θα πρέπει να ξεπερνά το 1/10 της μάζας του δονούμενου σώματος [119], [121]. Σχετικά με τη δυναμική περιοχή, το κατώτερο όριο λειτουργίας καθορίζεται και πάλι από τον εσωτερικό ηλεκτρικό θόρυβο, ενώ το ανώτερο όριο καθορίζεται από την περιοχή γραμμικής λειτουργίας του κρυστάλλου και από τη μηχανική αντοχή του επιταχυνσιομέτρου [114], [117], [119], [121].

2.1.1.5 Ηλεκτρακουστικές διατάξεις προ-ενίσχυσης

Στην πράξη, τα ηλεκτρικά και μηχανικά χαρακτηριστικά των ηλεκτρακουστικών αισθητήρων παρουσιάζουν αποκλίσεις από το ιδεατό μοντέλο γραμμικότητας. Δηλαδή, οι εμπλεκόμενοι αισθητήρες έχουν περιοχές μη γραμμικής λειτουργίας, ενώ κατά την ηλεκτρακουστική μετατροπή υπάρχουν ενεργειακές απώλειες. Έτσι, παρεμβάλλονται κυκλωματικές βαθμίδες ενίσχυσης και επεξεργασίας (signal conditioning), που στόχο έχουν την εξιδανίκευση των χαρακτηριστικών απόκρισης, και γενικότερα τη βελτίωση της ηχητικής λήψης. Οι διατάξεις αυτές, που συνήθως κατασκευάζονται από ολοκληρωμένα κυκλώματα τελεστικών ενισχυτών, χαρακτηρίζονται από υψηλή αντίσταση εισόδου (της τάξης των MΩ ή GΩ), μηδενική σχεδόν αντίσταση εξόδου και υψηλό κέρδος απολαβής (της τάξης των 60 dB ή περισσότερο). Παράλληλα είναι δυνατή η διαμόρφωση της επιθυμητής φασματικής απόκρισης (ενεργά φίλτρα). Συνολικά, οι προενισχυτικές βαθμίδες επιτελούν τις ακόλουθες λειτουργίες:

- προσαρμογή σύνθετης αντίστασης ανάμεσα στον μετατροπέα και την επόμενη κυκλωματική βαθμίδα, με στόχο τη μέγιστη μεταφορά τάσης και τον περιορισμό των ρευμάτων φόρτισης της τελευταίας,
- ηλεκτρική «απομόνωση» του αισθητήρα από το υπόλοιπο κύκλωμα,
- ενίσχυση του παραγόμενου ηλεκτρικού σήματος,
- μεγαλύτερη ανοχή στο θόρυβο λόγω της χαμηλής αντίστασης εξόδου και της υψηλότερης στάθμης,
- ισοστάθμιση και επιλεκτικό φιλτράρισμα για τη βελτίωση της σηματοθορυβικής σχέσης.

[24], [107], [112], [121], [123]-[124]

2.1.1.6 Αναλογικά και ψηφιακά συστήματα ηχητικής εγγραφής

Αν εξαιρέσουμε τις αρχικές «τεχνικές ακουστικής ηχογράφησης» (γραμμόφωνο, δίσκος βινυλίου), το μέσο που χρησιμοποιήθηκε ευρύτατα και εξακολουθεί να βρίσκεται εφαρμογή σε συστήματα αναλογικής εγγραφής και αναπαραγωγής ήχου, είναι η μαγνητική ταινία. Η αντιμετώπιση των τεχνικών και λειτουργικών προβλημάτων της μαγνητικής εγγραφής (πόλωση της ταινίας στη γραμμική περιοχή, μαγνητικός θόρυβος - hiss noise, περιορισμένη δυναμική περιοχή), αποτέλεσε αντικείμενο έρευνας για πολλές δεκαετίες. Έτσι, αναπτύχθηκαν μαγνητόφωνα εξαιρετικής ποιότητας από πλευράς ηλεκτρικής και μηχανικής συμπεριφοράς, ενσωματώνοντας εξελιγμένες τεχνικές διαμόρφωσης σήματος και μαγνητικής εγγραφής. Σήμερα, οι απαιτήσεις μηχανικής και ηλεκτρικής «τελειότητας» των

αναλογικών συστημάτων ηχητικής εγγραφής, έχει περιοριστεί σε μεγάλο βαθμό με την είσοδο της ψηφιακής τεχνολογίας [125]-[127].

Τα ψηφιακά ηχητικά σήματα προέρχονται από την ψηφιοποίηση των αντίστοιχων αναλογικών σημάτων. Όπως είναι γνωστό, η ψηφιοποίηση εισάγει δύο βασικές διεργασίες: α) τη μεταβολή της συνεχής ροής του χρόνου σε διακριτές χρονικές στιγμές, δηλαδή τη χρονική δειγματοληψία, και β) τη μετατροπή των συνεχών μεταβολών τάσης σε διακριτές αριθμητικές τιμές, δηλαδή την κβάντιση πλάτους. Η συχνότητα δειγματοληψίας (sampling frequency- f_s), ο ρυθμός δηλαδή λήψης δειγμάτων στο χρόνο, θα πρέπει να υπακούει στο κριτήριο του Nyquist. Αντίστοιχα, ο αριθμός των ψηφίων (bits) κβάντισης, θα πρέπει να εξασφαλίζει μια δυναμική περιοχή ισοδύναμη με αυτή του αρχικού αναλογικού σήματος [125], [127]-[128]. Για την αποτελεσματικότερη μεταφορά των ψηφιακών σημάτων έχουν αναπτυχθεί διάφορα πρωτόκολλα επικοινωνίας, που επιτρέπουν τη διασύνδεση των ψηφιακών συστημάτων χωρίς περαιτέρω απώλειες (π.χ. AES/EBU, SPDIF, MADI). Για τις ανάγκες ψηφιακής εγγραφής ήχου χρησιμοποιούνται αντίστοιχα μέσα αποθήκευσης, όπως είναι τα γραμμικά μέσα μαγνητικής εγγραφής σε ταινία (ψηφιακό κασετόφωνο, "Digital Audio Tape" -DAT), καθώς επίσης τα μαγνητικά και οπτικά μέσα τυχαίας προσπέλασης (σκληρός δίσκος - "Hard Disk", οπτικοί δίσκοι CD-DVD, μνήμες RAM - flash memories) [125], [127]-[129].

2.1.1.7 Προδιαγραφές βιοακουστικών συστημάτων καταγραφής

Σε γενικές γραμμές, οι προδιαγραφές των βιοακουστικών συστημάτων καταγραφής δεν παρουσιάζουν μεγάλες διαφορές με τα αντίστοιχα ηλεκτρακουστικά συστήματα μετρήσεων, αλλά αποτελούν υποκατηγορία των γενικότερων τεχνικών «ακουστικής διάγνωσης» [98], [130]-[131]. Ωστόσο, υπάρχουν και κάποιες ιδιαιτερότητες που σχετίζονται με τη φύση και τα χαρακτηριστικά των βιοακουστικών φαινομένων, καθώς και με τα ζητήματα ασφάλειας των υποκειμένων. Το τελευταίο στοιχείο συνδέεται με τις τεχνικές εφαρμογής - τοποθέτησης των ηλεκτρακουστικών αισθητήρων, αλλά και τις εμπλεκόμενες τεχνολογίες καταγραφής και μετάδοσης των λαμβανομένων σημάτων [24]. Υπάρχουν δύο κατηγορίες μικροφώνων που χρησιμοποιούνται σε εφαρμογές βιοακουστικής μελέτης: τα μικρόφωνα επαφής (contact mics) και τα μικρόφωνα αέρινης σύζευξης (air-coupled). Τα πρώτα έχουν το πλεονέκτημα ότι είναι λιγότερο ευαίσθητα σε ακουστικούς θορύβους περιβάλλοντος (ambient noise). Στον αντίποδα, τα χαρακτηριστικά τους εξαρτώνται σε μεγάλο βαθμό από τον τρόπο στερέωσης στην επιφάνεια ακρόασης και τις στατικές δυνάμεις που ασκούνται εκεί [98], [113].

Όσον αφορά τα λειτουργικά χαρακτηριστικά των βιοακουστικών συστημάτων, έχουν γενικά διατυπωθεί οι «ενδεδειγμένες» παράμετροι λειτουργίας για τις περιπτώσεις των αναπνευστικών και καρδιακών ήχων, ενώ πιο σπάνιες είναι οι αναφορές για την περίπτωση των κοιλιακών ήχων [24], [113], [132]-[134]. Έτσι, σε εφαρμογές καταγραφής αναπνευστικών σημάτων η συχνοτική απόκριση θα πρέπει να είναι επίπεδη (με μικρές διακυμάνσεις της τάξης των ± 6 dB ή λιγότερο) στην περιοχή από 60-100 Hz μέχρι 2-4 kHz, ενώ η δυναμική περιοχή του αισθητήρα δεν θα πρέπει να είναι μικρότερη από 50 dB. Επιθυμητοί είναι, επίσης, οι υψηλοί λόγοι σήματος προς θόρυβο (της τάξης των 60 dB και περισσότερο), ενώ η αρμονική παραμόρφωση δεν θα πρέπει να υπερβαίνει το 10% [113]. Στην περίπτωση των καρδιακών ήχων η «χρήσιμη» φασματική περιοχή περιορίζεται μέχρι τις συχνότητες 1-1,5kHz (περίπου) [132], [134]. Όσον αφορά του γαστρεντερικούς ήχους, το φασματικό του περιεχόμενο ξεκινά από τα 100-200Hz και εκτείνεται μέχρι τα 2-3 kHz, ενώ η δυναμική τους περιοχή κυμαίνεται στα επίπεδα 70-90 dB [24]-[25], [64], [70].

Κατά τη χρήση προενισχυτικών βαθμίδων, η σηματοθορυβική τους σχέση θα πρέπει να είναι μεγαλύτερη από εκείνη του «βιοακουστικού αισθητήρα», έτσι ώστε ο τελικός λόγος σήματος προς θόρυβο να είναι της τάξης των 60 dB (ή και περισσότερο), στις συχνότητες ενδιαφέροντος [113]. Για τον περιορισμό των ηλεκτρομαγνητικών παρεμβολών και των σχετικών βόμβων θα πρέπει να ακολουθούνται οι ενδεδειγμένες λύσεις διασύνδεσης (θωρακισμένα καλώδια, φίλτρα, κυκλώματα απομόνωσης / προσαρμογής –“buffers”). Όσον αφορά το μείζον θέμα της ασφάλειας των υποκειμένων, συνήθως προτιμώνται παθητικοί ηλεκτρακουστικοί μετατροπείς (πιεζοηλεκτρικά στοιχεία), ενώ οι ανάγκες τροφοδοσίας (κυκλωμάτων και πυκνωτικών αισθητήρων, όταν χρησιμοποιούνται) καλύπτονται συνήθως με τη χρήση συσσωρευτών (μπαταρίες). Τα παραπάνω, εκτός από την παροχή ασφάλειας, περιορίζουν τις πιθανότητες εμφάνισης «αρμονικού» θορύβου. Σε περίπτωση τροφοδοσίας από το δίκτυο παροχής ηλεκτρισμού, χρειάζεται ιδιαίτερη προσοχή στη σύνδεση και «προστασία» των σχετικών κυκλωματικών βαθμίδων, τόσο για αποφυγή βόμβων, όσο και για την ασφάλεια των υποκειμένων από ρεύματα διαρροής [113], [135]-[136].

Σήμερα, τα περισσότερα συστήματα βιοακουστικής ανάλυσης κάνουν χρήση των συστημάτων ψηφιακής επεξεργασίας, που παρέχουν δυνατότητες εφαρμογής εξελιγμένων αλγορίθμων επεξεργασίας [128]. Σε εφαρμογές αναπνευστικών και καρδιακών ήχων, κρίνονται επαρκείς ρυθμοί δειγματοληψίας των 2.500 Hz σε συνδυασμό με 12 bit κβάντιση [132]-[133]. Αντίστοιχα διαμορφώνονται και οι απαιτήσεις των γαστρεντερικών ήχων, όπου έχουν χρησιμοποιηθεί ρυθμοί δειγματοληψίας από 1 έως και 11 kHz [24]-[25], [59]-[61], [64]-[89]. Πρακτικά, η εξέλιξη της τεχνολογίας επιτρέπει τη χρήση «ποιοτικότερων» παραμέτρων ψηφιοποίησης, χωρίς τα ζητήματα υπολογιστικής ισχύος και χωρητικότητας να αποτελούν πλέον σημαντικό πρόβλημα. Σε κάθε περίπτωση, ο σχεδιασμός του βιοακουστικού συστήματος θα πρέπει να γίνεται προσεκτικά, λαμβάνοντας υπόψη τις πραγματικές απαιτήσεις των εφαρμογών, σε συνδυασμό με τις διαθέσιμες τεχνολογικές και λειτουργικές δυνατότητες.

2.1.2 Συστήματα οπτικοακουστικής παρακολούθησης

Τα συστήματα οπτικοακουστικής εποπτείας (video monitoring – surveillance) έχουν χρησιμοποιηθεί σε εφαρμογές βιοϊατρικής - ψυχοφυσιολογικής παρακολούθησης, με στόχο τον έλεγχο των υποκειμένων και των συνθηκών διεξαγωγής των αντίστοιχων πειραμάτων. Στην ενότητα αυτή παρουσιάζονται οι πιο χαρακτηριστικές εφαρμογές, καθώς και οι αντίστοιχες τεχνικές επεξεργασίας και αυτοματοποίησης των διαδικασιών παρακολούθησης, έτσι ώστε να γίνουν κατανοητές οι αντίστοιχες προδιαγραφές της παρούσης εφαρμογής.

2.1.2.1 Βιοϊατρικές εφαρμογές οπτικοακουστικής παρακολούθησης

Η αξιοποίηση των τεχνικών οπτικοακουστικής παρακολούθησης συνίσταται, κυρίως, σε ψυχοφυσιολογικές καταγραφές παρατεταμένου χρόνου, όπου ο φυσικός έλεγχος της πειραματικής διαδικασίας είναι προβληματικός, παρέχοντας χρήσιμες πληροφορίες στις περιπτώσεις όπου οι κινήσεις των υποκειμένων (ενήλικες, νήπια ή/και νεογνά) συνδέονται με τα υπό μελέτη φαινόμενα, ή όταν υπάρχουν καταστάσεις αβεβαιότητας - κινδύνου. Η συντριπτική πλειοψηφία των εφαρμογών αυτών έχει παρατηρηθεί στα «εργαστήρια ύπνου». Πρόκειται για διαδικασίες πολύ-αισθητηριακών καταγραφών (polysomnography), που στόχο έχουν τη διερεύνηση των αναπνευστικών και εγκεφαλικών λειτουργιών, σε συνδυασμό με τους κύκλους ύπνου. Χαρακτηριστικές είναι οι εφαρμογές ανίχνευσης των επεισοδίων άπνοιας [137], εντοπισμού και επαλήθευσης των συμπτωμάτων βήχα [12], οι

τεχνικές εκτίμησης του αναπνευστικού ρυθμού και η σύνδεσή του με τις κινήσεις του υποκειμένου και τις εναλλαγές των φάσεων ύπνου [138]. Σημαντική, είναι και η παρακολούθηση κατά τη φάση ηλεκτροεγκεφαλικών καταγραφών, για την ανίχνευση κρίσεων επιληψίας ή/και άλλων χαρακτηριστικών ευρημάτων της εγκεφαλικής λειτουργίας [1], [6], [139]-[143]. Όπως έχει ήδη επισημανθεί, η δυνατότητα παράλληλης διερεύνησης (με τα παραπάνω) των φάσεων γαστρεντερικής λειτουργίας, αποτέλεσε και ένα από τα πιο σημαντικά κίνητρα για την ενσωμάτωση εικονοληπτών παρακολούθησης στις τεχνικές προδιαγραφές του προτεινόμενου συστήματος καταγραφής.

Στις περισσότερες από τις προηγούμενες εφαρμογές γίνεται χρήση εικονοληπτών χαμηλού φωτισμού και νυχτερινής όρασης (night shot), όπως επίσης συστημάτων υπέρυθρης ακτινοβολίας και θερμικής απεικόνισης [7], [9], [10], [14]. Στο πλαίσιο αυτό έχουν αναπτυχθεί διάφορες τεχνικές για την ανίχνευση της κίνησης βίντεο [11], [139], [142]-[147], το συγχρονισμό των οπτικοακουστικών σημάτων με τις υπόλοιπες καταγραφές [2], [4]-[5], την εξέλιξη και προσαρμογή των αλγορίθμων συμπίεσης στις ειδικές ανάγκες των ιατρικών εφαρμογών [141], [148]-[149], την αξιοποίηση σχημάτων περιγραφής του οπτικοακουστικού περιεχομένου [150]. Σήμερα, οι τεχνικές καταγραφής και ψηφιακής επεξεργασίας εικόνας (βίντεο) και ήχου, καθιστούν πιο ευέλικτη και αποδοτική τη χρήση των οπτικοακουστικών συστημάτων παρακολούθησης, σε συνδυασμό με καταγραφές άλλων ψυχο-φυσιολογικών παραμέτρων, συμπεριλαμβανομένων και των γαστρεντερικών ήχων. Οι νέες λειτουργικές δυνατότητες σχετίζονται με τις εξελιγμένες μεθόδους λήψης, συμπίεσης, μεταφοράς, επεξεργασίας και διαχείρισης του οπτικοακουστικού - πολυμεσικού περιεχομένου.

2.1.2.2 Συστήματα καταγραφής και επεξεργασίας κινούμενης εικόνας

Τα συστήματα καταγραφής εικόνων, περιλαμβάνουν τους οπτικούς αισθητήρες, τα συνοδευτικά κυκλώματα επεξεργασίας (video signal conditioning) και τις μονάδες εγγραφής (video recording units). Αυτά είναι τα βασικά μέρη που συναντά κανείς στους εικονολήπτες εγγραφής (camcorders), όπου συνήθως είναι ενσωματωμένα και συστήματα ηχοληψίας [151]-[153]. Η δημιουργία ηλεκτρικού τηλεοπτικού σήματος προϋποθέτει τον κατάλληλο φωτισμό του οπτικού πεδίου παρατήρησης. Στην περίπτωση συνθηκών χαμηλού φωτισμού, όπως συμβαίνει συχνά με τις βιοϊατρικές εφαρμογές παρακολούθησης, χρησιμοποιούνται συστήματα νυχτερινής όρασης, παθητικά ή ενεργά. Στην πρώτη κατηγορία, γίνεται χρήση οπτικο-ηλεκτρονικών κυκλωμάτων συγκέντρωσης και ενίσχυσης της λαμβανόμενης φωτεινότητας, ενώ στη δεύτερη χρησιμοποιούνται προβολείς υπέρυθρης ακτινοβολίας. Σε κάθε περίπτωση το παραγόμενο τηλεοπτικό σήμα έχει «πρασινωπή» απόχρωση. Η συγκεκριμένη χρωματική συνιστώσα προτιμάται εξαιτίας της μεγαλύτερης ευαισθησίας της ανθρώπινης όρασης [105], [152]. Υπάρχουν και συστήματα θερμικής εικονοσκόπησης, όπου αντί των φακών και των οπτικών εξαρτημάτων, χρησιμοποιούνται συστοιχίες αισθητήρων θερμότητας [7], [9], [10], [105], [152].

Στην περίπτωση έγχρωμης εικόνας χρησιμοποιούνται οι συνιστώσες των τριών βασικών χρωμάτων (RGB components, δηλαδή το κόκκινο "R", το πράσινο "G" και το μπλε "B"). Η προσθετική μίξη αυτών των συνιστωσών δίνει το σύνθετο (έγχρωμο) τηλεοπτικό σήμα (Composite Video Signal - CVS), ενώ υπάρχει δυνατότητα χρήσης και άλλων συνιστωσών, ανάλογα με το τηλεοπτικό πρότυπο και το πρωτόκολλο μεταφοράς - εγγραφής. Έτσι διακρίνουμε το σύστημα φωτεινότητας - χρωμικότητας (Y-C components) και το σύστημα φωτεινότητας - χρωματοδιαφορών (YCrCb components, Cr=R-Y, Cb=B-Y). Το τελευταίο χρησιμοποιείται ως αφητηρία σε πολλές τεχνικές συμπίεσης (JPEG, MPEG, κ.ά.), ενώ υπάρχει

δυνατότητα διαφορετικού ορισμού των χρωματικών συνιστωσών, όπως για παράδειγμα στο σύστημα “YUV components”, που χρησιμοποιείται στο πρότυπο PAL [153]-[154]. Στην περίπτωση που δεν μας ενδιαφέρει η χρωματική πληροφορία, όπως συμβαίνει κατά τον οπτικοακουστικό έλεγχο των βιοϊατρικών πειραματικών διαδικασιών, επιλέγεται μόνο η συνιστώσα φωτεινότητας («ασπρόμαυρο σήμα»). Με βάση τα παραπάνω, διαμορφώθηκαν διάφορα αναλογικά πρότυπα, κάνοντας χρήση διαφορετικών κάθε φορά βίντεο συνιστωσών. Αντίστοιχα, προέκυψαν τα ψηφιακά συστήματα κινούμενης εικόνας, εφαρμόζοντας κανόνες ψηφιοποίησης στις εμπλεκόμενες βίντεο συνιστώσες [153]-[154].

Τα ψηφιακά συστήματα κινούμενης εικόνας συγκεντρώνουν όλα τα πλεονεκτήματα της ψηφιακής τεχνολογίας και κυρίως τη δυνατότητα συμπίεσης, δηλαδή μείωσης του όγκου των πληροφοριών. Οι τεχνικές συμπίεσης κινούμενης εικόνας επιδιώκουν τον περιορισμό της «πλεονάζουσας» πληροφορίας με όσο το δυνατόν λιγότερες επιπτώσεις στην ποιότητα των αρχικών εικόνων. Για το σκοπό αυτό συνήθως αξιοποιούν χωρικές ομοιότητες εντός των ακίνητων πλαισίων εικόνας, τεχνικές αναγνώρισης-κωδικοποίησης αντικειμένων (objects-based coding) και κυρίως αλγορίθμους εντοπισμού και αντιστάθμισης κίνησης (motion detection - compensation). Κατά την εγγραφή και αναπαραγωγή οπτικοακουστικού περιεχομένου σε περιβάλλον Η/Υ, όπως συμβαίνει και με την παρούσα εφαρμογή, οι διαδικασίες συμπίεσης και αποσυμπίεσης πραγματοποιούνται με «κωδικοποιητές» (codecs - compressor/ decompressor), που είναι εφαρμογές λογισμικού χαμηλού επιπέδου [153]-[157]. Παράλληλα, έχουν τυποποιηθεί πρωτόκολλα επικοινωνίας, τα οποία επιτρέπουν την εύκολη μεταφορά ψηφιακού βίντεο χωρίς απώλειες, τόσο στα γραμμικά μέσα (ψηφιακή εγγραφή σε μαγνητική ταινία), όσο και στα μέσα τυχαίας προσπέλασης (σκληροί δίσκοι, οπτικοί δίσκοι DVD, μνήμες στερεάς κατάστασης) [153]-[157].

2.2 Πειραματικό μέρος: ανάπτυξη συστήματος καταγραφής

Το σύστημα καταγραφής που αναπτύχθηκε για τις ανάγκες της παρούσης εφαρμογής περιλαμβάνει δύο υπο-συστήματα: τις διατάξεις καταγραφής των βιοακουστικών σημάτων και τα μέσα οπτικοακουστικής παρακολούθησης των υποκειμένων και της πειραματικής διαδικασίας. Και τα δύο σχεδιάστηκαν λαμβάνοντας υπόψη τις ιδιαιτερότητες του συγκεκριμένου προβλήματος, τις υπάρχουσες τεχνολογικές δυνατότητες και τη διεθνή εμπειρία και πρακτική, κατά την ανάπτυξη παρόμοιων εφαρμογών.

2.2.1 Πολυκαναλικό σύστημα καταγραφής βιοακουστικών σημάτων

Για την καταγραφή και επεξεργασία των γαστρεντερικών ήχων, αναπτύχθηκε πολυκαναλικό σύστημα, το οποίο αποτελείται από συστοιχία αισθητήρων, μαζί με τα απαραίτητα συνοδευτικά κυκλώματα και τις μονάδες εγγραφής. Η χρήση περισσότερων του ενός αισθητήρων προτείνεται για δύο λόγους. Πρώτον, για την εξασφάλιση υψηλής χωρικής ευαισθησίας, επιλέγοντας, κάθε φορά τον ηλεκτρακουστικό μετατροπέα, που βρίσκεται πλησιέστερα στο σημείο μηχανικής διέγερσης (Closest Point of Approach - CPA) [87], [158]. Δεύτερον, για τη δυνατότητα παράλληλης παρακολούθησης πολλαπλών θέσεων (multi-site monitoring), καθώς και στην ενσωμάτωση πρόσθετων λειτουργιών τοπογραφικής ανάλυσης και ακουστικής χαρτογράφησης (sound field visualization).

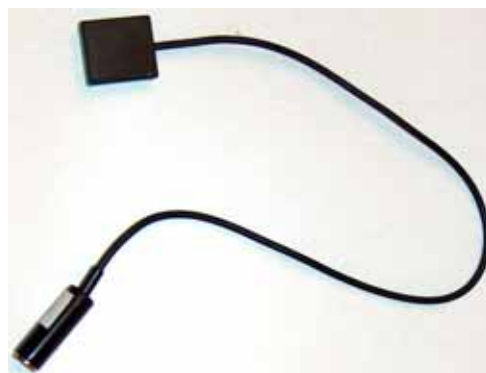
2.2.1.1 Αισθητήρες λήψης και κυκλώματα προενίσχυσης

Για τις ανάγκες καταγραφής των κοιλιακών ήχων έχουν χρησιμοποιηθεί διάφοροι τύποι αισθητήρων, με πιο συχνή τη χρήση ηλεκτρονικών στηθοσκοπίων, καθώς και μικροφώνων που προσαρμόζονται σε ακουστικά στηθοσκόπια [59], [106]. Τα στηθοσκόπια έχουν συνήθως μεγάλο μέγεθος και δεν προσφέρονται για καταγραφές παρατεταμένου χρόνου, ιδιαίτερα στη διάρκεια νυχτερινού ύπνου. Για τις ανάγκες της παρούσης εφαρμογής δοκιμάστηκαν πυκνωτικά μικρόφωνα επαφής (σχήματος ταινίας) CDUCER (Capacitance transDUCER), καθώς και πιεζοηλεκτρικοί κρύσταλλοι επαφής [24]. Οι τελευταίοι αποδείχθηκαν πολύ καλή λύση, καθώς έχουν βολικό μέγεθος και σχήμα, μεγάλη ευαισθησία στις «στερεόφερτες» δονήσεις της κοιλιακής επιφάνειας, περιορισμένη ευαισθησία στους «αερόφερτους» θορύβους περιβάλλοντος (ambient noise) και χαμηλό κόστος. Ιδιαίτερα θετικό είναι και το γεγονός ότι αποτελούν παθητικά στοιχεία, που δεν χρειάζονται τροφοδοσία, παρέχοντας ασφάλεια στα υπό εξέταση υποκείμενα κατά τη διαδικασία καταγραφής [24].

Εκτός από τα πιεζοηλεκτρικά μικρόφωνα επαφής (contact piezoelectric transducers - contact microphone pick-ups), χρησιμοποιήθηκαν και μονάδες επιταχυνσιόμετρων (accelerometers). Η τελευταία επιλογή έγινε για να υπάρξει ένα μέτρο αναφοράς και σύγκρισης με διατάξεις υψηλής πιστότητας και προδιαγραφών, όπως είναι τα μετρητικά επιταχυνσιόμετρα (instrumentation sensors). Έτσι, κατέστη δυνατή η διερεύνηση του φασματικού περιεχομένου των γαστρεντερικών ήχων σε περιοχές κάτω από τα 150 Hz, όπου η απόκριση των κλασσικών πιεζοηλεκτρικών μικροφώνων επαφής δεν είναι ικανοποιητική. Παράλληλα, η ενσωμάτωση επιταχυνσιόμετρου τριών αξόνων στο μετρητικό σύστημα παρείχε πρόσθετες δυνατότητες τοπογραφικής ανάλυσης του υπό παρακολούθηση ηχητικού πεδίου (η σχετική ανάλυση γίνεται στο 6^ο κεφάλαιο του παρόντος).



α.

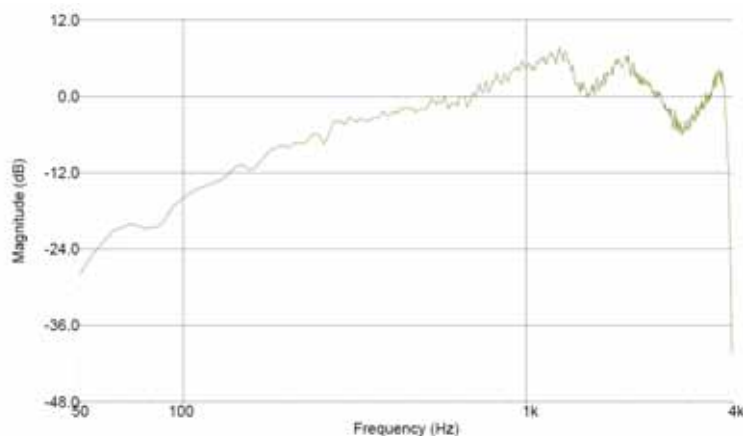


β.

Σχήμα 2-3. Ο πιεζοηλεκτρικός μετατροπέας PowerBeat PU-35.

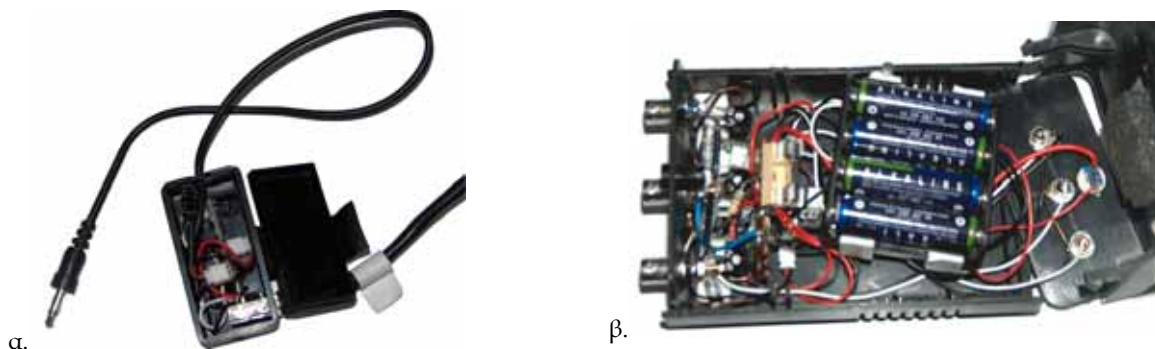
Ειδικότερα, χρησιμοποιήθηκε ο πιεζοηλεκτρικός μετατροπέας επαφής PowerBeat PU-35 (σχ. 2-3). Ο μετατροπέας αυτός περιλαμβάνει τον πιεζοηλεκτρικό δίσκο, που προσαρμόζεται σε πλαστικό περίβλημα-πλαίσιο ορθογωνίου σχήματος, πολύ μικρών διαστάσεων και βάρους (15x25x4mm / 5,5 grams). Χαρακτηρίζεται από υψηλή σύνθετη αντίσταση (high impedance), μεγάλη ευαισθησία και γραμμική συμπεριφορά, με αποκλίσεις της τάξης των $\pm 5\text{dB}$ στην περιοχή 120Hz-6kHz. Για την περαιτέρω διερεύνηση των παραμέτρων απόδοσης πραγματοποιήθηκαν μετρήσεις απόκρισης συχνότητας (σχ. 2-4), όπου χρησιμοποιήθηκε η γεννήτρια δονήσεων Vibration Exciter - B&K Type 4809, με λευκό και ροζ θόρυβο ως διέγερση. Παρά τη σχετικά υψηλή στάθμη εσωτερικού θορύβου και την όχι επίπεδη απόκριση στις χαμηλές συχνότητες (<200 Hz), οι προδιαγραφές του πιεζοηλεκτρικού μικροφώνου

PU-35 υπερκαλύπτουν τις απαιτήσεις των βιοακουστικών συστημάτων καταγραφής, όπως αυτές παρουσιάστηκαν στην παράγραφο 2.1.1.7.



Σχήμα 2-4. Η μετρούμενη καμπύλη απόκρισης συχνοτήτων του μετατροπέα PU-35.

Από το σχ. 2-4, επιβεβαιώνεται ότι, η καμπύλη απόκρισης συχνοτήτων του μετατροπέα παρουσιάζει μικρές διακυμάνσεις, μικρότερες των ± 6 dB στην περιοχή 150Hz έως 4kHz. Τα παραπάνω χαρακτηριστικά είναι ιδιαίτερα ικανοποιητικά, συγκρινόμενα με τις αντίστοιχες παραμέτρους απόδοσης των ακουστικών και ηλεκτρονικών στηθοσκοπίων. Ωστόσο, απέχουν αρκετά από τις «ιδεατές» καμπύλες απόκρισης των μετρητικών διατάξεων. Το γεγονός αυτό μπορεί να αντισταθμιστεί με τη χρήση φίλτρων ισοστάθμισης, τα οποία μπορεί να είναι διατάξεις προενίσχυσης υλικού (signal conditioning), αλλά και υλοποιήσεις λογισμικού (software). Στην τελευταία περίπτωση, που αποτέλεσε και αντικείμενο διερεύνησης στην παρούσα προσέγγιση, είναι δυνατή η εφαρμογή «αντίστροφου φιλτραρίσματος» με τη βοήθεια του αλγορίθμου FFT ή με τη χρήση τραπεζών φίλτρων [24], [25].



Σχήμα 2-5. α. Κύκλωμα προενίσχυσης και, β. διάταξη τροφοδοσίας των προενισχυτών.

Για την επίτευξη υψηλών τιμών ενίσχυσης και την προσαρμογή αντίστασης (matching impedance) του μετατροπέα, χρησιμοποιήθηκε προενισχυτής υψηλής απολαβής (σχ. 2-5α). Το κύκλωμα αυτό διαθέτει υψηλό κέρδος (>50 dB), ιδιαίτερα χαμηλή στάθμη θορύβου, πολύ μεγάλη αντίσταση εισόδου (πρακτικά άπειρη, της τάξης των $G\Omega$) και σχεδόν μηδενική αντίσταση εξόδου. Επίσης, χαρακτηρίζεται από επίπεδη καμπύλη απόκρισης συχνοτήτων με αποκλίσεις της τάξης των ± 1 dB, στην περιοχή από 40Hz έως 50kHz. Η τροφοδοσία του προενισχυτή έγινε με συστοιχία συσσωρευτών (σχ. 2-5β), έτσι ώστε το σύστημα καταγραφής να είναι στο σύνολό του παθητικό (για λόγους ασφάλειας των υποκειμένων), ενώ παράλληλα να μην εμφανίζονται προβλήματα βόμβων λόγω «βρόχων γείωσης» με το

δίκτυο παροχής ηλεκτρισμού. Οι προενισχυτικές βαθμίδες τοποθετήθηκαν σε μικρές αποστάσεις από τους αισθητήρες, με στόχο τον περιορισμό του προσθετικού θορύβου και την επίτευξη υψηλών σηματοθορυβικών σχέσεων.

Εκτός από το μικρόφωνο επαφής PU-35, χρησιμοποιήθηκε και το πιεζοηλεκτρικό επιταχυνσιόμετρο τριών αξόνων “Type 4506, Brüel & Kjær” (σχ. 2-6α). Πρόκειται για αισθητήρα μικρών διαστάσεων, τύπου OrthoShear που ανήκει στην οικογένεια DeltaTron της εταιρίας Brüel & Kjær (miniature triaxial DeltaTron accelerometer, OrthoShear design). Χαρακτηρίζεται από επίπεδη καμπύλη απόκριση συχνότητων (με αποκλίσεις $\pm 10\%$) στην περιοχή 0,3Hz - 6kHz για τον άξονα X, και 0,6Hz - 3kHz για τις εξόδους Y, Z, στοιχεία που κρίνονται ιδιαίτερα ικανοποιητικά για τις ανάγκες της παρούσης εφαρμογής [122], [159]. Το επιταχυνσιόμετρο 4506 έχει μεγάλη δυναμική περιοχή (>100 dB), διαθέτει ενσωματωμένο προενισχυτή και υψηλή ευαισθησία (10 mV/ms^{-2} at 159,2Hz), μικρό εσωτερικό θόρυβο σε όλους τους άξονες (X: $< 40\mu\text{VRMS}$; Y, Z: $< 20\mu\text{VRMS}$) και πολύ μικρό ποσοστό συνακρόασης μεταξύ των αξόνων (cross-talk): “*Transverse Sensitivity $< 5\%$ of the sensitivity of the axis in question*” [159]. Για τις ανάγκες τροφοδοσίας και προενίσχυσης χρησιμοποιήθηκε η διάταξη WB 1453 (σχ. 2-6β) [159].

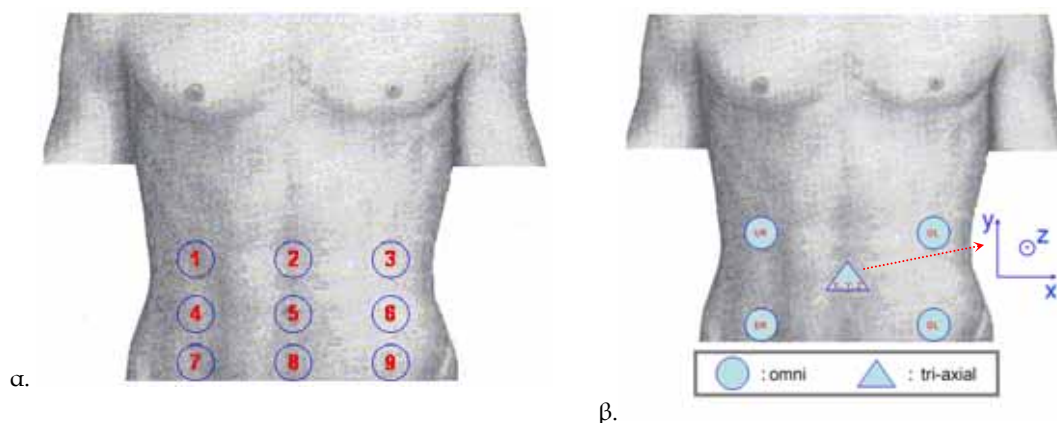


Σχήμα 2-6 α. Το επιταχυνσιόμετρο τριών αξόνων B&K 4506, β. η διάταξη τροφοδοσίας τριών καναλιών WB 1453.

Ιδιαίτερα σημαντική υπήρξε και η διαδικασία βαθμονόμησης (calibration) της μετρητικής διάταξης. Οι ρυθμίσεις κέρδους των προενισχυτικών βαθμίδων έγιναν, έτσι ώστε, σε κάθε περίπτωση, να παράγεται η ίδια στάθμη εξόδου για την ίδια ακουστική - μηχανική διέγερση. Η βαθμονόμηση του επιταχυνσιόμετρου τριών αξόνων, σε σχέση με τα πιεζοηλεκτρικά μικρόφωνα, έγινε με βάση τον κατακόρυφο άξονα (Z), που λειτούργησε ως πιλότος και για τη ρύθμιση των υπόλοιπων αξόνων (X, Y). Έτσι, αν θεωρήσουμε ότι τα μικρόφωνα επαφής PU-35 παρουσιάζουν πανκατευθυντικό (omni-directional) πολικό διάγραμμα, τότε η μηχανική διέγερση στον κατακόρυφο άξονα θα πρέπει να δίνει ίδια έξοδο με τον άξονα Z του επιταχυνσιόμετρου (για την ίδια «ποσότητα» διέγερσης). Λόγω του χαμηλού ενεργειακού περιεχομένου των γαστρεντερικών βιοακουστικών κυμάτων, επιλέχθηκαν υψηλές τιμές κέρδους στα κυκλώματα προενίσχυσης, της τάξης των 60-65 dB. Έτσι, οι κανονικοποιημένες στάθμες εγγραφής (recording level) των πιο έντονων κοιλιακών ήχων διαμορφώθηκαν γύρω στα -3 dB, κανόνας που αποτελεί πάγια τακτική στις εφαρμογές ηχοληψίας και ακουστικών μετρήσεων. Οι αντίστοιχες στάθμες του προσθετικού θορύβου βάθους (θόρυβος ευρείας ζώνης, Additive Broadband Noise - ABN), περιορίστηκαν στις τιμές μεταξύ -50dB και -40dB.

2.2.1.2 Διαμόρφωση πολυκαναλικού συστήματος βιοακουστικής καταγραφής

Σχετικά με τη χωρική διάταξη του πολυκαναλικού συστήματος ηχητικής λήψης, αρχικά χρησιμοποιήθηκαν δύο αισθητήρες παράλληλης καταγραφής [24]-[25], που τοποθετήθηκαν διαγωνίως - συμμετρικά στα δύο από τα τέσσερα τεταρτημόρια της κοιλιακής χώρας (θέσεις 3 και 7 του σχ. 2-7α), δηλαδή στο πάνω-αριστερό (LU) και κάτω δεξιό (DR) μέρος (διάταξη "LURD-stereo"). Γενικά, έχουν αναφερθεί εφαρμογές με τρεις [42], τέσσερις [72], πέντε [69] και έξι [87] αισθητήρες παράλληλης καταγραφής των κοιλιακών ήχων, αποσκοπώντας στην ενσωμάτωση πολλαπλών «οπτικών γωνιών» παρακολούθησης (multi-site monitoring), αλλά και στην αξιοποίηση πρόσθετων τεχνικών χαρτογράφησης - απεικόνισης του ηχητικού πεδίου [42], [87]. Από τις παραπάνω προσεγγίσεις, η συμμετρική - περιφερειακή διάταξη των 4 αισθητήρων (θέσεις 1, 3, 7, 9 του σχ. 2-7α) [72], είχε θεωρηθεί ως η πιο ενδεδειγμένη, από τις πρώτες κιόλας πειραματικές προσεγγίσεις [24], [25]. Το σύστημα αυτό, το οποίο στο εξής θα καλείται *A-Quadrants* (Abdominal Quadrants), επιλέχθηκε, αφού πρωτίστως εξετάστηκαν μέχρι και εννέα συνολικά θέσεις καταγραφής (σχ. 2-7α). Για τις ανάγκες αυτής της διερεύνησης πραγματοποιήθηκαν δειγματοληπτικές μετρήσεις, τόσο σε υποκείμενα, όσο και με τη βοήθεια πειραματικών διατάξεων (λεπτομέρειες δίνονται στο έκτο κεφάλαιο).



Σχήμα 2-7. α. Οι θέσεις τοποθέτησης των αισθητήρων που εξετάστηκαν, και β. οι τελικές θέσεις των ηλεκτρακουστικών μετατροπέων που επιλέχθηκαν για το πολύ-καναλικό σύστημα βιοακουστικής καταγραφής (UR: Up Right, UL: Up Left, DR: Down Right, DL: Down Left, X: x-axis, Y: y-axis, Z: z-axis).

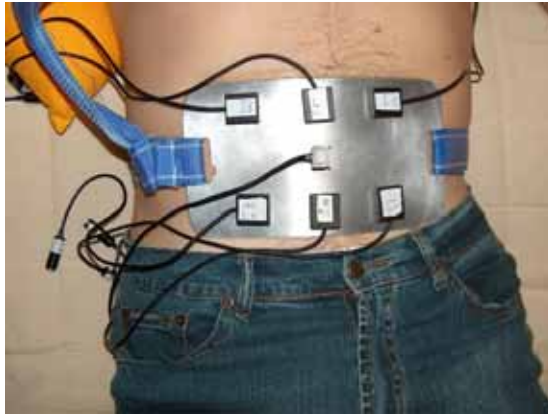
Η διάταξη *A-Quadrants* αποδείχθηκε ιδιαίτερα λειτουργική, εξασφαλίζοντας την ικανοποιητική λήψη ακουστικών σημάτων από ολόκληρη την κοιλιακή περιοχή με περιορισμένο, σχετικά, αριθμό αισθητήρων. Η χρήση περισσότερων των 4 αισθητήρων αποδείχθηκε μάλλον υπερβολική και χωρίς πρόσθετα ουσιαστικά πλεονεκτήματα. Στο σύστημα τεσσάρων καναλιών προστέθηκε και το επιταχυνσιόμετρο τριών αξόνων B&K 4506, το οποίο εισήχθη στο κέντρο συμμετρίας των περιφερειακών, «πανκατευθυντικών» (omni-directional), αισθητήρων (θέση 5 του σχ. 2-7α). Η διάταξη αυτή, που στο εξής θα αναφέρεται ως *A-Soundfield 4+3* (σχ. 2-7β), διαμορφώθηκε με στόχο την ανάπτυξη τεχνικών ακουστικής χαρτογράφησης. Ο όρος "*Soundfield*" υιοθετήθηκε από την τεχνολογία των αντίστοιχων μικροφώνων ηχητικού πεδίου [160], που αποτέλεσαν αφορμή και έμπνευση για την παραπάνω υλοποίηση. Όλες οι σχετικές λεπτομέρειες παρουσιάζονται στο έκτο κεφάλαιο, μαζί με την ανάλυση των μεθόδων τοπογραφικής ανάλυσης και χωρικής απεικόνισης - αναπαραγωγής του ηχητικού πεδίου. Αυτό που μας ενδιαφέρει στην παρούσα φάση, είναι το γεγονός ότι το προτεινόμενο σύστημα καταγραφής περιλαμβάνει επτά (7), συνολικά, κανάλια ηχητικής εγγραφής.

2.2.1.3 Πρόσθετα μέσα «στεγασής» και προσαρμογής των αισθητήρων

Για την τοποθέτηση των πιεζοηλεκτρικών αισθητήρων στην κοιλιακή επιφάνεια των υποκειμένων, αρχικά, έγινε χρήση χάρτινων κολλητικών ταινιών και αυτοκόλλητων ταινιών διπλής όψης. Παράλληλα, χρησιμοποιήθηκαν ελαστικοί επίδεσμοι για την καλύτερη στερέωση και συγκράτηση των αισθητήρων, καθώς και σπογγοειδή απορροφητικά υλικά κάλυψης για τον περιορισμό των ακουστικών θορύβων περιβάλλοντος. Η παρεμβολή παχύρρευστου GEL υπερήχων βοήθησε στην καλύτερη προσαρμογή και την επίτευξη μεγαλύτερης ευαισθησίας [24], [25]. Οι προηγούμενες διευθετήσεις δίνουν ικανοποιητικά αποτελέσματα εφόσον οι πειραματικές διαδικασίες διεξάγονται σε σχετικά «ήσυχους» χώρους (ηχητικά μονωμένους) και ο αριθμός των εμπλεκόμενων αισθητήρων καταγραφής δεν είναι μεγάλος (συνήθως, όχι περισσότεροι από δύο). Σε διαφορετική περίπτωση αρχίζουν να εμφανίζονται διάφορα προβλήματα, όπως είναι η αδυναμία των απορροφητικών υλικών να παράσχουν ικανοποιητικό βαθμό ηχομείωσης, οι λειτουργικές δυσκολίες και οι ενοχλήσεις που σχετίζονται με την τοποθέτηση πολλαπλών επιδέσμων ή τη χρήση ισχυρών αυτοκόλλητων ταινιών. Έτσι, αν και οι αισθητήρες που επιλέχθηκαν δεν είναι ευαίσθητοι σε αερόφερτους ήχους, οι θόρυβοι περιβάλλοντος υπεισέρχονται στις βιοακουστικές καταγραφές, ακόμη κι όταν είναι πολύ μικρής στάθμης (ήχοι από γειτονικούς χώρους, ομιλίες υποκειμένων και ιατρών, ροχαλητά, κ.ά.). Αυτό συμβαίνει καθώς, ο ακουστικός θόρυβος διεγείρει εξωτερικά τα κοιλιακά τοιχώματα, και οι αναπτυσσόμενες, εκεί, δονήσεις λαμβάνονται από τους αισθητήρες καταγραφής.

Η ενσωμάτωση των εικονοληπτών ελέγχου, έδωσε τη δυνατότητα καταγραφής τέτοιων φαινομένων σε πραγματικές συνθήκες διεξαγωγής της πειραματικής διαδικασίας. Έτσι, διαπιστώθηκε ότι, συχνά, οι αισθητήρες δεν παραμένουν σωστά στερεωμένοι καθ' όλη τη διάρκεια της καταγραφής, ενώ σε πολλές περιπτώσεις, η επικόλληση των ταινιών στερέωσης μπορεί να προσθέσει πρόσθετους ήχους τριβής. Οι τελευταίοι, έχουν συνήθως περιοδικό χαρακτήρα, αντίστοιχο με αυτό της αναπνοής, καθώς παράγονται από τις αναπνευστικές θωρακικές κινήσεις. Η χαλάρωση της επαφής των ηλεκτρακουστικών μετατροπέων έχει άμεσες επιπτώσεις στην ευαισθησία, αλλά και στο σύνολο των χαρακτηριστικών του μετρητικού συστήματος. Είναι αυτονόητο ότι τέτοιου είδους διαφοροποιήσεις θα πρέπει να είναι περιορισμένες, τόσο κατά μήκος της ίδιας καταγραφής, όσο και μεταξύ καταγραφών διαφορετικών υποκειμένων. Έτσι, η διαδικασία προσαρμογής ενός μεγάλου αριθμού αισθητήρων καθίσταται ιδιαίτερα απαιτητική και χρονοβόρα. Έχει παρατηρηθεί ότι, για τη σωστή τοποθέτηση και των πέντε αισθητήρων του μετρητικού βιοακουστικού συστήματος, απαιτείται χρονική διάρκεια μισής τουλάχιστον ώρας, χρόνος που είναι αρκετά μεγάλος συγκρινόμενος με την τυπική χρονική διάρκεια δύο ωρών της πειραματικής διαδικασίας.

Για την αντιμετώπιση των παραπάνω προβλημάτων προτάθηκαν οι ακόλουθες λύσεις: α) η μόνιμη στερέωση των αισθητήρων καταγραφής σε μια επιφάνεια προσαρμογής, η οποία θα εφάπτεται ενιαία στην κοιλιακή περιοχή, και β) η κατασκευή ενός ηχομονωτικού γιλέκου προσαρμογής («κοιλιακή ζώνη»), που θα φοριέται από τα υποκείμενα κατά την πειραματική διαδικασία. Για την πρώτη περίπτωση, χρησιμοποιήθηκε μια πολύ λεπτή επιφάνεια αλουμινίου, πάχους περίπου 0,5mm (σχ. 2-8), όπου στερεώνονται οι αισθητήρες με «ισχυρές» αυτοκόλλητες ταινίες διπλής όψης, χωρίς να υπάρχει κίνδυνος ερεθισμών ή ενοχλήσεων στην επιδερμίδα των υποκειμένων. Κατόπιν, η επιφάνεια στερέωσης (FSA - Foil for Sensors' Attaching) εφάπτεται στην κοιλιακή περιοχή παρεμβάλλοντας GEL υπερήχων, ενώ η συγκράτησή της γίνεται με τη χρήση ζώνης ρυθμιζόμενου μήκους (σχ. 2-8).



Σχήμα 2-8. Η μεταλλική επιφάνεια στερέωσης των αισθητήρων καταγραφής FSA (Foil for Sensors' Attaching) κατά την εφαρμογή της σε υποκείμενο.

Η εισαγωγή της επιφάνειας FSA προσέφερε δύο σημαντικές βελτιώσεις. Το ένα αφορά τη λειτουργική πλευρά της πειραματικής διαδικασίας, όπου η τοποθέτηση της μετρητικής διάταξης στα υποκείμενα γίνεται πλέον σε πολύ μικρότερο χρονικό διάστημα (λιγότερο από 10 λεπτά). Το άλλο αφορά στην καλύτερη προσαρμογή των αισθητήρων, προκαλώντας μικρότερες ενοχλήσεις στα υποκείμενα και εξασφαλίζοντας μεγαλύτερη ευαισθησία λήψης. Προς την κατεύθυνση αυτή συμβάλλει και η παρεμβολή του GEL υπερήχων. Παράλληλα, η καλύτερη συγκράτηση των αισθητήρων περιορίζει αισθητά τα σφάλματα καταγραφής λόγω τριβών (εξαιρούνται οι περιπτώσεις έντονης κίνησης των υποκειμένων, για τις οποίες γίνεται λόγος σε επόμενο κεφάλαιο). Ένα άλλο θετικό στοιχείο είναι ότι η επιφάνεια FSA έχει πάρα πολύ μικρό βάρος και μέγεθος. Συνεκτιμώντας τις διαστάσεις της, την ταχύτητα του ήχου στο μυϊκό ιστό των κοιλιακών τοιχωμάτων ($\approx 1.500 \text{ m/sec}$) και τις συχνότητες των παραγομένων κοιλιακών ήχων (200Hz - 2kHz, περίπου), παρατηρούμε ότι η επίδραση στο μετρούμενο ηχητικό πεδίο είναι αμελητέα, καθώς τα διαμορφούμενα μήκη κύματος (7,5m - 750cm, περίπου) είναι πολύ μεγαλύτερα από τις διαστάσεις του «εμποδίου». Με βάση τις κρίσεις των υποκειμένων, στα οποία χρησιμοποιήθηκε η διάταξη FSA, η λύση αυτή είναι προτιμητέα από την απευθείας επικόλληση των αισθητήρων στην κοιλιακή περιοχή. Στην τελευταία άποψη συνηγορεί και το ιατρικό - νοσηλευτικό προσωπικό, που επιμελούνταν και επέβλεπε τις πειραματικές διαδικασίες.

Όσον αφορά τη δεύτερη διεύθυνση, δηλαδή το ηχομονωτικό γιλέκο προσαρμογής WAAV - "Wearable Absorbing Abdominal Vest" (σχ. 2-9), αυτό καλείται να επιτελέσει δύο λειτουργίες: α) να απομονώσει (μηχανικά - ηχητικά) την επιφάνεια FSA, αλλά και ολόκληρη την κοιλιακή περιοχή, από τους τυχόν θορύβους περιβάλλοντος, και β) να συγκρατήσει τη διάταξη FSA σε στενή επαφή με την κοιλιακή επιφάνεια, ώστε να εξασφαλιστεί μεγαλύτερη ευαισθησία λόγω καλύτερης μηχανικής προσαρμογής. Για την επίτευξη ικανοποιητικής ηχομείωσης αναζητήθηκαν λύσεις σε «αδρανείς μάζες», δηλαδή υλικά με ισχυρή ηχομονωτική ικανότητα [108]-[109]. Έτσι, χρησιμοποιήθηκε φύλλο μολύβδου πάχους 1mm, αντίστοιχο μ' εκείνο των ακτινοσκοπικών ιατρικών ποδιών, το οποίο προσαρμόστηκε στη θήκη της ανθεκτικής υφασμάτινης ζώνης του σχ. 2-9. Η μολύβδινη επιφάνεια απομονώθηκε με πλήρη επένδυση από απορροφητικό υλικό πάχους 1cm. Με την προσθήκη αυτή ενισχύθηκε η ηχοαπορροφητική ικανότητα του γιλέκου, ενώ έγιναν και πιο μαλακές οι επιφάνειές του, ώστε να μην προκαλείται ενόχληση στα υποκείμενα. Το μήκος της ηχομονωτικής επένδυσης είναι περίπου 30 cm, ενώ το βάρος του γιλέκου WAAV μετρήθηκε στα 860 gr. Η ζώνη WAAV στερεώνεται περιμετρικά γύρω από την κοιλιακή περιοχή (σχ. 2-10), όπου «κουμπώνει» στο πίσω μέρος με αυτοκόλλητες ταινίες τύπου "Velcro". Κατά τη

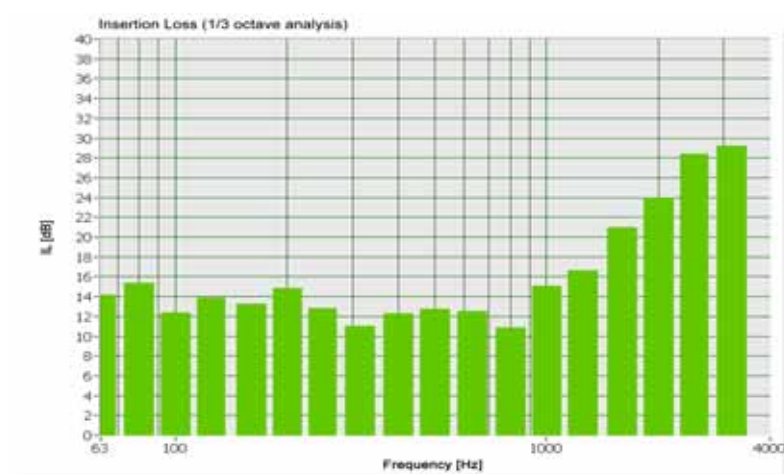
χρήση του ηχομονωτικού γιλέκου WAAV, παρατηρήθηκε ελαχιστοποίηση των ακουστικών θορύβων περιβάλλοντος (σχ. 2-11), ενώ δεν αναφέρθηκαν προβλήματα ενοχλήσεων από τα υποκείμενα. Στο γεγονός αυτό συντελεί το σχετικά μικρό μέγεθος και βάρος της διάταξης, καθώς και το γεγονός ότι τα υποκείμενα είναι ξαπλωμένα, κι έτσι δεν επωμίζονται το σύνολο του «βαρύτερου φορτίου». Ιδιαίτερα θετικές είναι οι εκτιμήσεις του ιατρικού - νοσηλευτικού προσωπικού, σχετικά με την ευκολία εφαρμογής του γιλέκου WAAV αλλά και την πιο αποτελεσματική προσαρμογή της μετρητικής διάταξης.



Σχήμα 2-9. Το γιλέκο WAAV (Wearable Absorbing Abdominal Vest) που χρησιμοποιήθηκε για την καλύτερη μηχανική προσαρμογή του μετρητικού βιοακουστικού συστήματος.



Σχήμα 2-10. Εφαρμογή του ηχομονωτικού γιλέκου WAAV (Wearable Absorbing Abdominal Vest) σε υποκείμενο, κατά την πειραματική διαδικασία.



Σχήμα 2-11. Η μετρούμενη ηχομονωτική ικανότητα (IL - Insertion Loss) του γιλέκου WAAV, εκφρασμένη σε dB ακουστικής πίεσης, για τις διάφορες τρίτο-οκταβικές περιοχές συχνότητας.

Εκτός από τις εμπειρικές παρατηρήσεις, η ηχομονωτική ικανότητα της διάταξης WAAV αξιολογήθηκε βάσει πειραματικών μετρήσεων. Τα σχετικά πειράματα διεξήχθησαν σε ένα τυπικό δωμάτιο εξέτασης, το οποίο διεγέρθηκε ακουστικά (98dB SPL) με τη βοήθεια συστήματος ηχοτροφοδοσίας και γεννήτριας ροζ θορύβου. Κατόπιν, εφαρμόστηκε το σύστημα βιοακουστικής καταγραφής σε υποκείμενο, σαν να επρόκειτο για διεξαγωγή βιοακουστικών καταγραφών σε περιβάλλον έντονου θορύβου. Από τις συγκρίσεις των καταγραφών, με και χωρίς τη χρήση της διάταξης WAAV, υπολογίστηκε η ηχομονωτική ικανότητα του γιλέκου, δηλαδή η τριτο-οκταβική καμπύλη του δείκτη ηχομείωσης IL (Insertion Loss). Τα σχετικά αποτελέσματα δίνονται στο σχήμα 2-11.

Στο σχήμα 2-12 παρουσιάζεται το σύνολο των διατάξεων προσαρμογής της μετρητικής διάταξης, η οποία θα περιγράφεται πλέον με το ακρωνύμιο SHIELD (Structure for HousIng, sEnsors insuLation and aDaptation). Εκτός από την επιφάνεια FSA και το γιλέκο WAAV, διακρίνουμε το στρώμα επάλειψης με GEL υπερήχων, τις θέσεις των αισθητήρων της διάταξης A-Soundfield 4+3 και τις θέσεις του ιμάντα ανάρτησης.



Σχήμα 2-12. Σχηματική περιγραφή του συστήματος προσαρμογής των αισθητήρων καταγραφής (SHIELD - Structure for HousIng, sEnsors insuLation and aDaptation).

2.2.1.4 Ψηφιοποίηση βιοακουστικών καταγραφών και μονάδες εγγραφής

Για την καταγραφή των λαμβανομένων βιοακουστικών σημάτων χρησιμοποιήθηκαν μονάδες ψηφιακής εγγραφής, γραμμικής και τυχαίας προσπέλασης. Μερικά από τα επιθυμητά χαρακτηριστικά του συστήματος εγγραφής είναι η δυνατότητα αποθήκευσης καταγραφών παρατεταμένης διάρκειας, η δυνατότητα ενσωμάτωσης πολλαπλών καναλιών, η εύκολη λήψη και μεταφορά των σημάτων σε υπολογιστικό περιβάλλον και φυσικά, η εξασφάλιση των απαιτήσεων εύρους φάσματος και δυναμικής περιοχής, που σχετίζονται με το πληροφοριακό περιεχόμενο των γαστρεντερικών ήχων. Επίσης, σε πολλές περιπτώσεις κρίθηκε αναγκαία η αξιοποίηση φορητών συστημάτων [24]-[25], [161]-[162].

Όπως έχει αναφερθεί, σε εφαρμογές επεξεργασίας γαστρεντερικών ήχων έχουν χρησιμοποιηθεί διάφοροι ρυθμοί δειγματοληψίας από 1 έως και 11.025 kHz [24]-[25], [59]-[61], [64]-[89], ανάλογα με τη διάρκεια των καταγραφών και τις δυνατότητες των μέσων αποθήκευσης. Πολλοί ερευνητές υποστηρίζουν ότι συχνότητες δειγματοληψίας της τάξης των 4 με 5 kHz, υπερκαλύπτουν τις ανάγκες παρακολούθησης του φαινομένου [59], [64], [71], [80]-[89]. Αντίστοιχα, έχουν δοκιμαστεί 8, 12 και 16 bit, όπου οι δύο πρώτες επιλογές παρουσιάζουν πιο έντονα προβλήματα θορύβου κβάντισης [24], [25], [89], ενώ η τελευταία κρίνεται καταλληλότερη, καθώς παρέχει μεγαλύτερη δυναμική περιοχή [25], [161]-[162]. Με βάση τα παραπάνω και λαμβάνοντας υπόψη ότι το φασματικό περιεχόμενο των κοιλιακών ήχων δεν ξεπερνά τα 2-2,5 kHz [59], [70], επιλέχθηκαν οι παράμετροι ψηφιοποίησης 8kHz,

16bit, καθώς κρίθηκαν υπεραρκετοί για τις ανάγκες της παρούσης εφαρμογής. Έτσι, για την αποθήκευση των δεδομένων χρησιμοποιήθηκε το πρότυπο αρχείων ήχου “WAV” (Wav, PCM, 16bit, 8kHz). [161]-[162].

Όσον αφορά τις μονάδες εγγραφής, για τις ανάγκες δι-καναλικής ηχογράφησης της διάταξης *LURD-stereo* χρησιμοποιήθηκε το ψηφιακό μαγνητόφωνο ήχου DAT, DA-R100, CASIO (σχ. 2-13α). Η φορητή αυτή μονάδα υποστηρίζει ψηφιακή εγγραφή 16bit/(32, 44,1 και 48 kHz), χαρακτηρίζεται από μεγάλη δυναμική περιοχή (της τάξης των 95 dB) και εύρος φάσματος από 10Hz έως 20kHz. Διαθέτει ενσωματωμένο προενισχυτή μικροφώνου υψηλής απολαβής, με δυνατότητα ρύθμισης κέρδους και πολύ καλή σηματοθορυβική σχέση [25], ενώ είναι δυνατή η απευθείας μεταφορά των ψηφιακών δεδομένων, μέσω του ψηφιακού πρωτοκόλλου SPDIF, [25], [162]. Όσον αφορά τις πολυκαναλικές καταγραφές των διατάξεων *A-Quadrants* και *A-soundfield 4+3*, αποφασίστηκε η απευθείας ηχογράφηση σε περιβάλλον Η/Υ με χρήση καρτών ήχου πολλών καναλιών, αξιοποιώντας τα πλεονεκτήματα των μέσων τυχαίας προσπέλασης. Το στοιχείο αυτό παρείχε επιπρόσθετα τη δυνατότητα αντιγράφων ασφαλείας ολόκληρων των καταγραφών, πριν την εφαρμογή τεχνικών κατωφλίωσης για τον εντοπισμό και το διαχωρισμό των επιμέρους συμβάντων [161]-[162]. Ειδικότερα, χρησιμοποιήθηκε η κάρτα “PreSonus FIREPOD - 24-bit/96K FireWire Recording Studio” (σχ. 2-13β), η οποία συνδέεται στον Η/Υ εξωτερικά, μέσω των πρωτοκόλλων USB-2 και IEEE-1394 (firewire). Οι προδιαγραφές της υπερκαλύπτουν τις ανάγκες της παρούσης εφαρμογής, καθώς υποστηρίζει παραμέτρους ψηφιοποίησης μέχρι και 24bit/96kHz, ενώ παρέχει τη δυνατότητα παράλληλης ηχογράφησης έως και 10 καναλιών. Η κάρτα FIREPOD αποδείχθηκε ιδιαίτερα εύχρηστη, καθώς διαθέτει εξωτερικές ρυθμίσεις κέρδους για τη βαθμονόμηση της μετρητικής διάταξης, ενώ η όλη διαδικασία ηχογράφησης ελέγχεται από περιβάλλον λογισμικού.



Σχήμα 2-13. α. Φορητό σύστημα στέρεο-ηχογράφησης DAT, DA-R100, CASIO, β. πολυκαναλική κάρτα ήχου PreSonus FIREPOD - 24-bit/96K FireWire Recording Studio, σε στιγμιότυπο πειραματικής εφαρμογής.

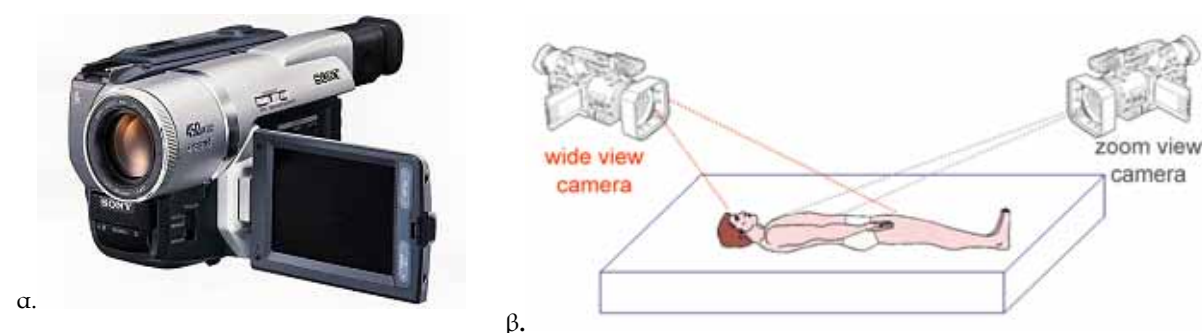
2.2.2 Σύστημα οπτικοακουστικής παρακολούθησης & ελέγχου

Όπως έχει ήδη αναφερθεί, το σύστημα οπτικοακουστικής παρακολούθησης χρησιμοποιήθηκε με στόχο την διεξαγωγή παρατηρήσεων σχετικά με τις ιδιαιτερότητες του πειραματικού πρωτοκόλλου (μη εποπτευόμενη διαδικασία παρατεταμένου χρόνου), αλλά και τη διευθέτηση όλων των τεχνικών ζητημάτων που θα επιτρέψουν τη μελλοντική συνδυαστική του χρήση με άλλες μεθόδους ψυχοφυσιολογικής παρακολούθησης. Η αξιοποίηση του συστήματος αυτού έδωσε τη δυνατότητα εντοπισμού συγκεκριμένων ιδιαιτεροτήτων - ατελειών, κατά την εφαρμογή της βιοακουστικής μετρητικής διάταξης (αναλύθηκαν στην §2.2.1.2), συμβάλλοντας στη βελτίωση των χαρακτηριστικών απόδοσης και λειτουργικότητας.

Η διάταξη οπτικοακουστικής παρακολούθησης αποτελείται από τους αισθητήρες λήψης εικόνας - ήχου, τις μονάδες εγγραφής και τις εφαρμογές πλοήγησης - πρόσβασης του καταγεγραμμένου υλικού. Στο πλαίσιο αυτό, συμπεριλαμβάνονται τεχνικές επεξεργασίας κινούμενης εικόνας (video) και ήχου, που στόχο έχουν τη βελτίωση της ποιότητας των πληροφοριών, τη συμπίεση και αποδοτικότερη αρχειοθέτησή τους, και την αξιοποίηση ευέλικτων δομών διαχείρισης και ανάλυσης του πολυμεσικού περιεχομένου. Ειδικά για την τελευταία περίπτωση, εφαρμόστηκαν τεχνικές ανίχνευσης «κίνησης εικόνας» (video motion detection), οι οποίες ενδείκνυνται για την εξαγωγή παραμέτρων σύνοψης (video abstracting -summarization) και τη δυνατότητα γρήγορης πλοήγησης (video browsing). Ως αποτέλεσμα της πιλοτικής αυτής διερεύνησης, διευθετήθηκε το σύνολο των τεχνικών ζητημάτων και προδιαγραφών που αφορούν στο εποπτικό σύστημα παρακολούθησης, ώστε να είναι δυνατή η άμεση εφαρμογή του στις πειραματικές διαδικασίες βιοϊατρικής έρευνας.

2.2.2.1 Εξοπλισμός οπτικοακουστικής παρακολούθησης - καταγραφής

Από πλευράς εξοπλισμού, το σύστημα οπτικοακουστικής καταγραφής περιλαμβάνει τους εικονολήπτες με τις μονάδες εγγραφής εικόνας (cameras & video recording), τα μικρόφωνα λήψης με τα κυκλώματα προενίσχυσης και τις αντίστοιχες μονάδες εγγραφής ήχου (sound recording). Στην προκειμένη περίπτωση, και δεδομένου ότι η υψηλή ποιότητα λήψης δεν αποτελεί κρίσιμη παράμετρο για τις ανάγκες της εφαρμογής, επιλέχθηκε ο ψηφιακός εικονολήπτης χειρός (handy-cam) Sony DCR-TRV520E DIGITAL8 (σχ. 2-14α). Υποστηρίζει ψηφιακή εγγραφή βάσει του πρωτοκόλλου Digital-8, διαθέτει ενσωματωμένο πυκνωτικό μικρόφωνο και ψηφιακό εγγραφέα ήχου-εικόνας (camcorder). Μπορεί να εικονοληπτεί ικανοποιητικά σε πολύ χαμηλές συνθήκες φωτισμού, αξιοποιώντας τα συστήματα νυχτερινής όρασης "Night Shot" και "Super Night Shot" (*minimum illumination: 3 lux, 0 lux with Night-Shot*). Τα καταγεγραμμένα οπτικοακουστικά σήματα μπορούν να μεταφερθούν με ψηφιακό τρόπο (απευθείας) στο υπολογιστικό περιβάλλον επεξεργασίας, κάνοντας χρήση του πρωτοκόλλου IEEE 1394 (γνωστό και ως DV). Έτσι, οι πληροφορίες ήχου-εικόνας αποθηκεύονται και καταχωρούνται ως αρχεία βίντεο τύπου "avi" (audio video interlaced) και κωδικοποίησης DVPAL.



Σχήμα 2-14. Η διάταξη οπτικοακουστικής παρακολούθησης της πειραματικής διαδικασίας: **α)** ο εικονολήπτης Sony DCR-TRV520E DIGITAL8, **β)** το «σχήμα» διάταξης των δύο εικονοληπτών.

Κατά την πιλοτική χρήση του συστήματος οπτικοακουστικής εποπτείας, αποφασίστηκε η ενσωμάτωση δύο εικονοληπτών (σχ. 2-14β), με στόχο τη «σχολαστική» ανάλυση της κίνησης των υποκειμένων. Οι εικονολήπτες τοποθετήθηκαν εκατέρωθεν του υποκειμένου, ενώ για την καταγραφή του ακουστικού θορύβου περιβάλλοντος αξιοποιήθηκαν τα αντίστοιχα εσωτερικά (build-in) μικρόφωνα. Επιπρόσθετα, ο ένας από τους εικονολήπτες ρυθμίστηκε έτσι ώστε να εστιάζει στην κοιλιακή περιοχή όπου στερεώνεται η βιοακουστική

διάταξη (zoom-view camera) με στόχο τη διεξαγωγή παρατηρήσεων που αναφέρθηκαν στην §2.2.1.2, ενώ ο άλλος να παρακολουθεί την πειραματική διαδικασία συνολικά (wide-view camera). Στο σχ. 2-15 δίνονται χαρακτηριστικά στιγμιότυπα εικόνων παρακολούθησης όπου φαίνεται και η ενσωμάτωση οπτικών σημείων ελέγχου (οι μαύρες βούλες στην κοιλιακή περιοχή), που προστέθηκαν για τη διευκόλυνση των παρατηρήσεων. Με δεδομένες τις πειραματικές παρατηρήσεις της §2.2.1.2, είναι φανερό ότι τα «σημάδια» αυτά μπορούν να παραλειφθούν κατά την παρακολούθηση των υποκειμένων σε επίπεδο κλινικής εφαρμογής. Επίσης, το σύστημα των δύο εικονοληπτών μπορεί να αντικατασταθεί από την τοποθέτηση μια πανοραμικής κάμερας ευρυγώνιας λήψης επάνω ακριβώς από το υποκείμενο, λύση που συχνά ακολουθείται σε αντίστοιχες βιοϊατρικές εφαρμογές [6].



Σχήμα 2-15. Παραδείγματα οπτικοακουστικής παρακολούθησης.

2.2.2.2 Ψηφιοποίηση και τεχνικές επεξεργασίας των οπτικοακουστικών καταγραφών

Το επόμενο βήμα μετά την οπτικοακουστική καταγραφή, αφορά στη μεταφορά των λαμβανομένων σημάτων σε ψηφιακό υπολογιστικό περιβάλλον, με στόχο την περαιτέρω επεξεργασία και ανάλυση. Έτσι, δεδομένης της απευθείας ψηφιακής διασύνδεσης μέσω IEEE 1394, οι πληροφορίες ήχου λαμβάνονται ως δεδομένα ασυμπίεστης PCM κωδικοποίησης (16bit, 48 kHz, stereo), ενώ οι πληροφορίες εικόνας είναι συμπιεσμένες με τον κωδικοποιητή DV PAL (720x576, 25fps, 24bit RGB, DV compression). Τα παραπάνω στοιχεία αντιστοιχούν σε ρυθμούς ροής δεδομένων 1.536 kbps και 25 Mbps (για τον ήχο και την εικόνα αντίστοιχα), όπου αν συμπεριληφθούν και τα δεδομένα ελέγχου, έχουμε συνολικά ένα ρυθμό της τάξης των 30 Mbps. Γίνεται εύκολα αντιληπτό, ότι τα παραπάνω μεγέθη δημιουργούν προβλήματα αποθήκευσης, επεξεργασίας και ανάλυσης, καθώς μια τυπική καταγραφή διάρκειας δύο ωρών έχει μέγεθος περίπου 26 GB. Στον αντίποδα, παρατηρείται πλεονασμός στην κωδικοποίηση των οπτικοακουστικών πληροφοριών, τόσο γιατί η ποιότητα των καταγραφών αυτών είναι δευτερευούσης σημασίας, όσο και γιατί μπορούν να χρησιμοποιηθούν πολύ αποδοτικότεροι παράμετροι κωδικοποίησης - συμπίεσης. Έτσι, αποφασίστηκε ο περιορισμός των καναλιών ήχου (μίξη καναλιών) και η χρήση μικρότερης συχνότητα δειγματοληψίας (υποδειγματοληψία). Με τα νέα χαρακτηριστικά των ηχητικών καταγραφών ελέγχου (PCM, 8kHz, 16 bit, mono) γίνεται σημαντική εξοικονόμηση αποθηκευτικού χώρου (περίπου 55MB/ώρα) κι έτσι δεν υπάρχει λόγος για ενσωμάτωση τεχνικών συμπίεσης.

Όσον αφορά τα ζητήματα εικόνας, εδώ τα πράγματα είναι πιο σύνθετα, καθώς αυξάνει σημαντικά η ποσότητα των πληροφοριών, ενώ παράλληλα υπάρχουν προβλήματα θερμικού θορύβου λόγω του συστήματος νυχτερινής όρασης. Έτσι, αποφασίστηκε ο υποδιπλασιασμός της αρχικής ανάλυσης (από 720x576 σε 360x288), λύση που συχνά ακολουθείται στα

συστήματα παρακολούθησης - ασφαλείας. Επίσης, αποφασίστηκε η απόρριψη των χρωματικών συνιστωσών, δεδομένου ότι οι λαμβανόμενες εικόνες δεν είναι κατά βάση έγχρωμες (παρουσιάζουν διακυμάνσεις αποχρώσεων του «πρασίνου»). Έτσι, επιλέχθηκε το σήμα φωτεινότητας και η αναπαράσταση των εικόνων στην πιο «φυσική» κλίμακα του γκριζου (grayscale). Για τον περιορισμό του θερμικού θορύβου «νυχτερινή λήψης» γίνεται επεξεργασία με χρήση φίλτρων τάξης (order filters: size 3x3, order 4). Σε κάθε περίπτωση, οι τεχνικές επεξεργασίας κινούμενης εικόνας δεν αποτέλεσαν κύριο αντικείμενο της παρούσης εργασίας, αλλά κατευθύνθηκαν από κλασσικές βιβλιογραφικές προσεγγίσεις. Για την υλοποίηση των σχετικών εφαρμογών χρησιμοποιήθηκε το λογισμικό πακέτο National Instruments LabVIEW 7.1™ και ειδικότερα το πρόσθετο πακέτο εφαρμογών επεξεργασίας εικόνας NI-IMAQ VISION 7.1™ (add-on toolkit)[156]-[157].

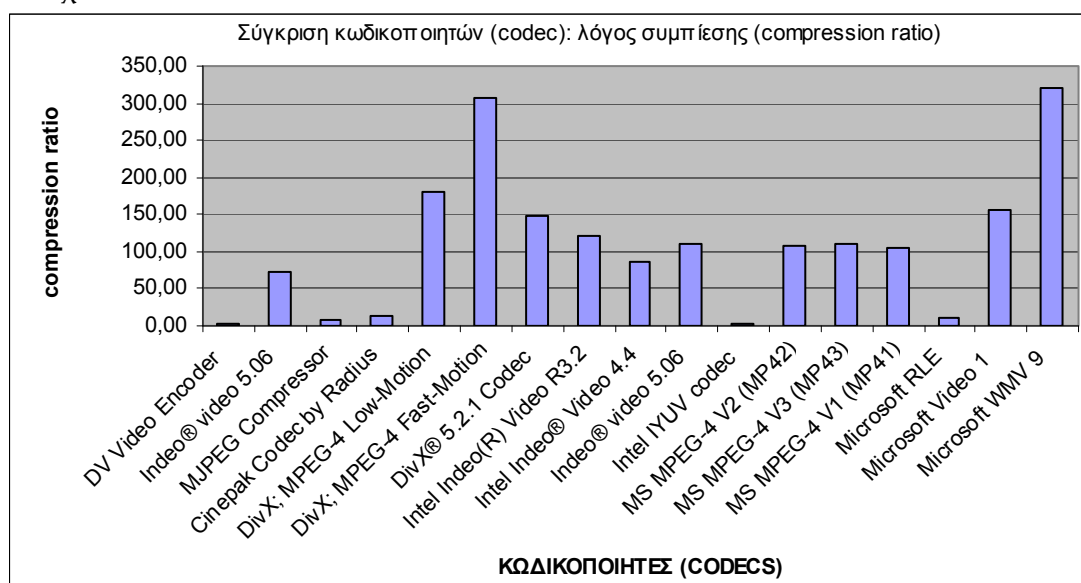
Για την επιλογή της μεθόδου συμπίεσης, δοκιμάστηκαν μια σειρά από κωδικοποιητές (CODECs) κινούμενης εικόνας, οι οποίοι εφαρμόστηκαν σε αποσπάσματα των διαθέσιμων καταγραφών παρακολούθησης (ευρείας λήψης -wide, και χωρικής εστίασης -zoom). Η αξιολόγηση έγινε με βάση τα ποιοτικά χαρακτηριστικά των συμπιεσμένων εικόνων (ποιοτική αξιολόγηση), καθώς και με την αξιοποίηση τριών ποσοτικών δεικτών που παρουσιάζονται στον πίνακα Π2-1: α) το λόγο συμπίεσης, εκφρασμένο από το μέγεθος του αρχικού - ασυμπίεστου προς το μέγεθος του συμπιεσμένου αρχείου, β) την πολυπλοκότητα του αλγορίθμου, εκφρασμένη από το μέσο χρόνο εγγραφής κάθε συμπιεσμένου πλαισίου (average write time), και γ) την ποιότητα (quality) συμπίεσης, υπολογισμένη ως η μέση απόλυτη διαφορά (mean absolute difference) της αρχικής ασυμπίεστης εικόνας από τη συμπιεσμένη (κανονικοποιημένη στην περιοχή 0-1000).

Compressor	Λόγος συμπίεσης	Avg Write Time (ms)	Quality (0 - 1000)
DV Video Encoder	3,44	1,35	987,22
Indeo® video 5.06	73,59	8,91	977,20
MJPEG Compressor	8,13	1,21	985,00
Cinepak Codec by Radius	13,60	160,19	989,68
DivX; MPEG-4 Low-Motion	181,53	1,17	977,62
DivX; MPEG-4 Fast-Motion	306,73	1,26	975,62
DivX® 5.2.1 Codec	147,31	1,89	975,97
Intel Indeo(R) Video R3.2	120,04	1,75	959,01
Intel Indeo® Video 4.4	86,19	87,59	974,42
Indeo® video 5.06	110,14	17,53	976,74
Intel IYUV codec	2,66	1,27	660,37
MS MPEG-4 V2 (MP42)	106,49	1,33	982,53
MS MPEG-4 V3 (MP43)	109,48	1,82	982,54
MS MPEG-4 V1 (MP41)	105,60	1,24	982,53
Microsoft RLE	11,01	1,86	844,22
Microsoft Video 1	156,59	1,75	953,10
Microsoft WMV 9	319,26	1,88	980,58

Πίνακας Π2-1. Αξιολόγηση αλγορίθμων συμπίεσης με βάση τους δείκτες α) λόγο συμπίεσης, β) μέσος χρόνος εγγραφής (Avg Write Time - ms), και γ) ποιότητα (Quality 0-1000).

Στο διάγραμμα του σχ. 2-16 αποδίδεται γραφικά ο λόγο συμπίεσης κάθε μεθόδου, που αποτελεί και τον πιο ουσιαστικό δείκτη αξιολόγησης (με την προϋπόθεση ότι η ποιότητα συμπίεσης είναι σε σχετικά αποδεκτά επίπεδα). Με βάση τα αποτελέσματα ποιοτικής και ποσοτικής αξιολόγησης αποφασίστηκε η επιλογή του κωδικοποιητή συμπίεσης Microsoft

Windows Media Video 9 (MS-WMV9), καθώς συνδυάζει πολύ καλή ποιότητα εικόνας, με υψηλά ποσοστά συμπίεσης και μικρό, συγκριτικά, υπολογιστικό κόστος. Ο αλγόριθμος συμπίεσης WMV9 αναπτύχθηκε από τη Microsoft και στη συνέχεια τυποποιήθηκε από τον οργανισμό SMPTE, ως SMPTE 421M. Επιτυγχάνει υψηλούς λόγους συμπίεσης ενσωματώνοντας τεχνικής εντοπισμού και αντιστάθμισης κίνησης με προσαρμοστικό μέγεθος μπλοκ, ενώ γενικά χαρακτηρίζεται από καλύτερη απόδοση σε σχέση με τους κωδικοποιητές MPEG-2 και MPEG-4, που προηγήθηκαν χρονικά [163]. Όσον αφορά τη συγκεκριμένη εφαρμογή, η συμπίεση των καταγραφών παρακολούθησης με τον εν λόγω κωδικοποιητή, είχε ως αποτέλεσμα αρχεία μεγέθους μόλις 60 MB/ώρα και 50 MB/ώρα, για τις καταγραφές ευρείας λήψης (wide-view video) και εστίασης (zoom-view video), αντίστοιχα.



Σχήμα 2-16. Σχηματική αποτύπωση των λόγων συμπίεσης, για τους διάφορους κωδικοποιητές (codecs) που δοκιμάστηκαν στις δεδομένες καταγραφές παρακολούθησης.

2.2.2.3 Διαχείριση οπτικοακουστικών καταγραφών και τεχνικές ανίχνευση κίνησης

Ένα από τα σημαντικότερα ζητήματα που σχετίζονται με τη διαχείριση οπτικοακουστικών καταγραφών παρατεταμένου χρόνου, αφορά στην δυνατότητα εξαγωγής πληροφοριών «περίληψης» και ευέλικτων παραμέτρων σύνοψης (abstracting and summarization). Τα στοιχεία αυτά απασχόλησαν την ερευνητική κοινότητα κατά την τυποποίηση του πρωτοκόλλου περιγραφής πολυμεσικού περιεχομένου (Multimedia Content Description Interface), που είναι πλέον γνωστό ως MPEG-7 [164]-[165]. Στην προκειμένη περίπτωση αποφασίστηκε ο υπολογισμός των «φακέλων» ηχητικής στάθμης (Sound Level Envelopes) και η χρήση παραμέτρων κίνησης (motion parameters), για τα σήματα ήχου και εικόνας, αντίστοιχα, του συστήματος οπτικοακουστικής παρακολούθησης. Για την εξαγωγή της «περιβάλλουσας ήχου», υιοθετήθηκε η χρήση μη επικαλυπτόμενων παράθυρων (non-overlapping windows) και ο υπολογισμός της ηχητικής στάθμης (Sound Level - dB) για καθένα από αυτά. Το μήκος των παραθύρων τέθηκε ίσο με 1 sec, τιμή που αντιστοιχεί στη σταθερά χρόνου «αργών μεταβολών» (slow - time constant) των συστημάτων ηχομέτρησης [166]. Με δεδομένη τη κβάντιση των 16 bit, οι υπολογιζόμενες στάθμες κυμαίνονται, θεωρητικά τουλάχιστον, από 0dB μέχρι και 90 dB. Για την εξομάλυνση τυχόν απότομων μεταβολών στις παραπάνω καμπύλες, χρησιμοποιείται φίλτρο εκθετικά κινούμενου μέσου

όρου (Exponential Moving Average, -ExpMA). Με τον τρόπο αυτό εξασφαλίζεται η σωστή εκτίμηση των σημάτων περιβάλλουσας, περιορίζοντας, τόσο το μήκος των εξαγομένων ακολουθιών, όσο και το υπολογιστικό κόστος της διαδικασίας.

Όσον αφορά τους περιγραφείς εικόνας, αποφασίστηκε η αξιοποίηση παραμέτρων κίνησης (visual motion descriptors), οι οποίοι συνδέονται στενά με τις μεταβολές δραστηριότητας του οπτικού πεδίου παρατήρησης [164]-[165], [167]-[170]. Γενικά, υπάρχουν διάφορες προσεγγίσεις εντοπισμού κίνησης σε καταγραφές βίντεο [154], [171]-[174]. Σ' αυτές συμπεριλαμβάνονται οι μέθοδοι σύγκρισης μακρο-μπλοκ [154], [171]-[172], οι μέθοδοι αξιοποίησης χρονικών διαφορών (time differencing), οι μέθοδοι εκτίμησης και αφαίρεσης φόντου (background subtraction) και οι τεχνικές «οπτικής ροής» (optical flow), με τις τελευταίες να χρησιμοποιούνται (συνήθως) σε περιπτώσεις κίνησης των εικονοληπτών [173]-[174]. Από τις παραπάνω, πιο δημοφιλείς είναι οι τεχνικές σύγκρισης μακρο-μπλοκ (Block-Matching Algorithms -BMA), οι οποίες προσδιορίζουν τα διανύσματα κίνησης (motion vectors), παρακολουθώντας τη μετατόπιση των εσωτερικών περιοχών (μπλοκ) της εικόνας στο χρόνο [154], [171]-[172]. Η πιο απλή προσέγγιση των τεχνικών BMA είναι η μέθοδος πλήρους αναζήτησης (full search -FS), η οποία χαρακτηρίζεται από πολύ μεγάλο υπολογιστικό κόστος. Για το λόγο αυτό έχουν προταθεί διάφοροι αλγόριθμοι γρήγορης αναζήτησης, οι οποίοι περιορίζουν σε μεγάλο βαθμό το υπολογιστικό κόστος, αυξάνοντας, ωστόσο, την πολυπλοκότητα [171]-[172]. Λόγω των παραπάνω χαρακτηριστικών, καθώς και του γεγονότος ότι οι παρεχόμενες πληροφορίες κίνησης είναι πλεοναστικές για τις ανάγκες της παρούσης εφαρμογής, η χρήση των τεχνικών BMA ήταν περιορισμένη.

Έτσι, το ενδιαφέρον στράφηκε σε παραμετρικές μεθόδους εντοπισμού κίνησης, κάνοντας χρήση των τεχνικών «χρονικών διαφορών» και «αφαίρεσης φόντου», που συχνά χρησιμοποιούνται σε συστήματα παρακολούθησης - ελέγχου (surveillance) [173]-[176]. Αρχικά, υλοποιήθηκε μια μέθοδος χρονικών διαφορών [175], όπου οι πίνακες φωτεινότητας των εικόνων κλιμακώνονται σε ένα μικρό αριθμό επιπέδων του γκριζου (gray-level band, GLB), ενώ η «κίνηση» προκύπτει από τη σύγκριση των εικόνων GLB των διαδοχικών καρέ:

$$M_n(x, y) = \begin{cases} 0, & GLB_n(x, y) \neq GLB_{n-1}(x, y) \\ 1, & GLB_n(x, y) = GLB_{n-1}(x, y) \end{cases} \quad (2-1)$$

όπου $M_n(x, y)$ είναι η «εικόνα κίνησης», x, y οι μεταβλητές που προσδιορίζουν τις χωρικές συντεταγμένες κάθε εικόνας και n ο δείκτης που χρησιμοποιείται για την αρίθμηση των πλαισίων. Στην προκειμένη περίπτωση, χρησιμοποιήθηκαν τρία επίπεδα για τη διαμόρφωση των εικόνων GLB, όπως προτείνεται για λήψεις «νυχτερινής όρασης» στην αυθεντική εργασία [175]. Η προηγούμενη μέθοδος είναι πολύ απλή στη σύλληψη και έχει μικρές υπολογιστικές απαιτήσεις. Ωστόσο, απορρίφθηκε καθώς αποδείχθηκε ιδιαίτερα ευαίσθητη στις μεταβολές φωτισμού. Στην απόρριψή της συντέλεσε και η ανάπτυξη μιας πιο αποδοτικής μεθόδου, συνδυάζοντας τεχνικές χρονικών διαφορών και αφαίρεσης φόντου.

Η τελική προσέγγιση που υιοθετήθηκε, αφορά μια προσαρμοστική, επαναληπτική τεχνική, όπου σε κάθε χρονική στιγμή γίνεται εκτίμηση της εικόνας παρασκηνίου (φόντου), από την οποία εκτιμάται εάν υπάρχει κίνηση, κατόπιν συγκρίσεων με τους πίνακες φωτεινότητας κάθε πλαισίου [173]-[174]. Για την υλοποίηση των σχετικών υπολογισμών γίνεται χρήση κατωφλίων σύγκρισης, ενώ όλα τα εμπλεκόμενα μεγέθη ανανεώνονται χρονικά, κάνοντας χρήση εκθετικά κινούμενων μέσων όρων. Έτσι, αν $I_n(x, y)$ είναι η ένταση

(φωτεινότητα) του εικονοστοιχείου με συντεταγμένες (x,y) στο n -οστό πλαίσιο της αλληλουχίας κινούμενων εικόνων (video), η εκτίμηση (πρόβλεψη) της έντασης παρασκηνίου $B_{n+1}(x,y)$ γίνεται ως εξής:

$$B_{n+1}(x,y) = \begin{cases} a_m \cdot B_n(x,y) + (1-a_m) \cdot I_n(x,y), & \text{αν δεν υπάρχει κίνηση στη θέση } (x,y) \\ B_n(x,y) & , \text{αν υπάρχει κίνηση στη θέση } (x,y) \end{cases} \quad (2-2)$$

όπου η παράμετρος ανανέωσης a_m παίρνει θετικές τιμές κοντά στη μονάδα ($0 < a_m < 1$). Η αρχική εκτίμηση της εικόνας φόντου $B_0(x,y)$ τίθεται ίση με τον αντίστοιχο πίνακα φωτεινότητας $I_0(x,y)$, ενώ οι πληροφορίες «κίνησης των εικονοστοιχείων» υπολογίζονται, συγκρίνοντας τις εντάσεις φωτεινότητας των διαδοχικών εικόνων $I_n(x,y)$ και $I_{n-1}(x,y)$:

$$|I_n(x,y) - I_{n-1}(x,y)| > T_n(x,y) \Leftrightarrow \text{υπάρχει κίνηση στη θέση } (x,y) \quad (2-3)$$

Οι τιμές σύγκρισης $T_n(x,y)$ ανανεώνονται βάσει του ακόλουθου αναδρομικού τύπου:

$$T_{n+1}(x,y) = \begin{cases} a_m T_n(x,y) + (1-a) \cdot c_m \cdot |I_n(x,y) - B_n(x,y)|, & \text{αν δεν υπάρχει κίνηση στη θέση } (x,y) \\ T_n(x,y) & , \text{αν υπάρχει κίνηση στη θέση } (x,y) \end{cases} \quad (2-4)$$

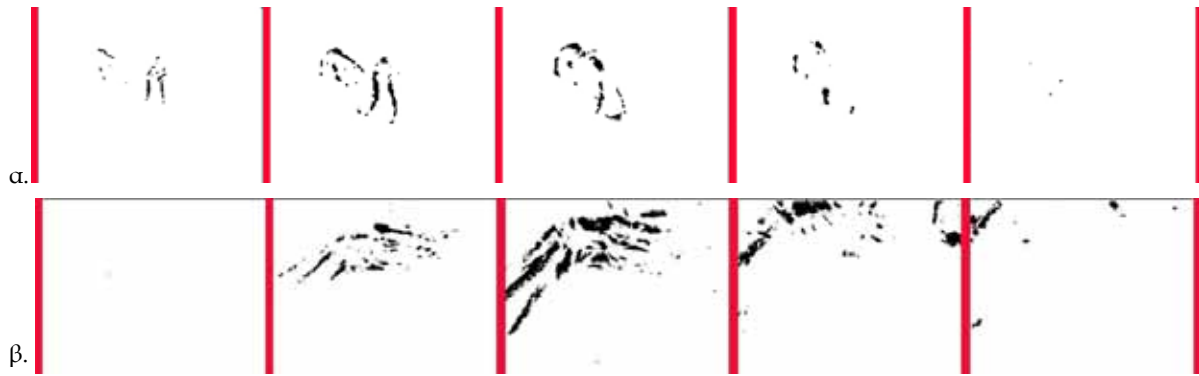
όπου η σταθερά c_m (πραγματικός αριθμός, $c_m > 1$) χρησιμοποιείται για τον έλεγχο της ευαισθησίας ανίχνευσης κίνησης (όσο αυξάνει η τιμή της παραμέτρου c_m , τόσο μειώνεται η ευαισθησία ανίχνευσης κίνησης, και αντίστροφα), ενώ η αρχική τιμή κατωφλίσωσης τίθεται εμπειρικά ($T_0(x,y) = Th_{mot}$). Έτσι, οι «εικόνες κίνησης» $M_n(x,y)$ προκύπτουν ως η διαφορά των τιμών φωτεινότητας από τις τιμές φόντου σε κάθε πλαίσιο (εξ. 2-5), ενώ η ολοκλήρωση των «κινούμενων εικονοστοιχείων» δίνει την καμπύλη «έντασης κίνησης» $I_{mot}(n)$ (motion intensity curve), βάσει της σχέσης 2-6.

$$M_n(x,y) = \begin{cases} 1, & \text{αν } |I_n(x,y) - B_n(x,y)| > T_n(x,y) \\ 0, & \text{διαφορετικά} \end{cases} \quad (2-5)$$

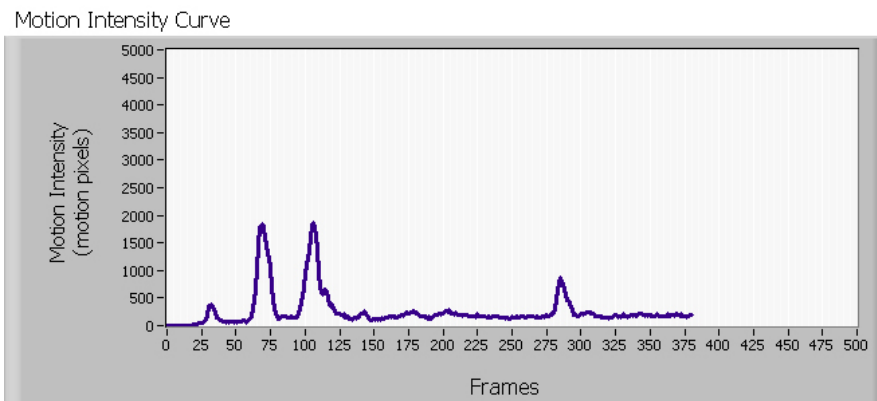
$$I_{mot}(n) = \sum_{x,y} M_n(x,y) \quad (2-6)$$

Η προηγούμενη μέθοδος ανίχνευσης κίνησης αποδείχθηκε ιδιαίτερα αποδοτική, καθώς συνδυάζει πλεονεκτήματα όπως: πολύ καλά αποτελέσματα ανίχνευση κίνησης, μικρό, σχετικά, υπολογιστικό κόστος και δυνατότητα ρύθμισης της ευαισθησίας, με την επιλογή των κατάλληλων παραμέτρων. Στην προκειμένη περίπτωση επιλέχθηκαν οι τιμές $a_m = 0,85$ / $c_m = 5$ / $Th_{mot} = 30$. Στο σχ. 2-17 παρουσιάζονται τα διαδοχικά πλαίσια κίνησης, για τις εικόνες ευρείας λήψης (wide) και χωρικής εστίασης (zoom), ενώ στο σχ. 2-18 παρουσιάζεται μια τυπική καμπύλη «έντασης κίνησης», που εκφράζει τον συνολικό αριθμό των κινούμενων εικονοστοιχείων, για κάθε χρονική στιγμή (για κάθε πλαίσιο εικόνας).

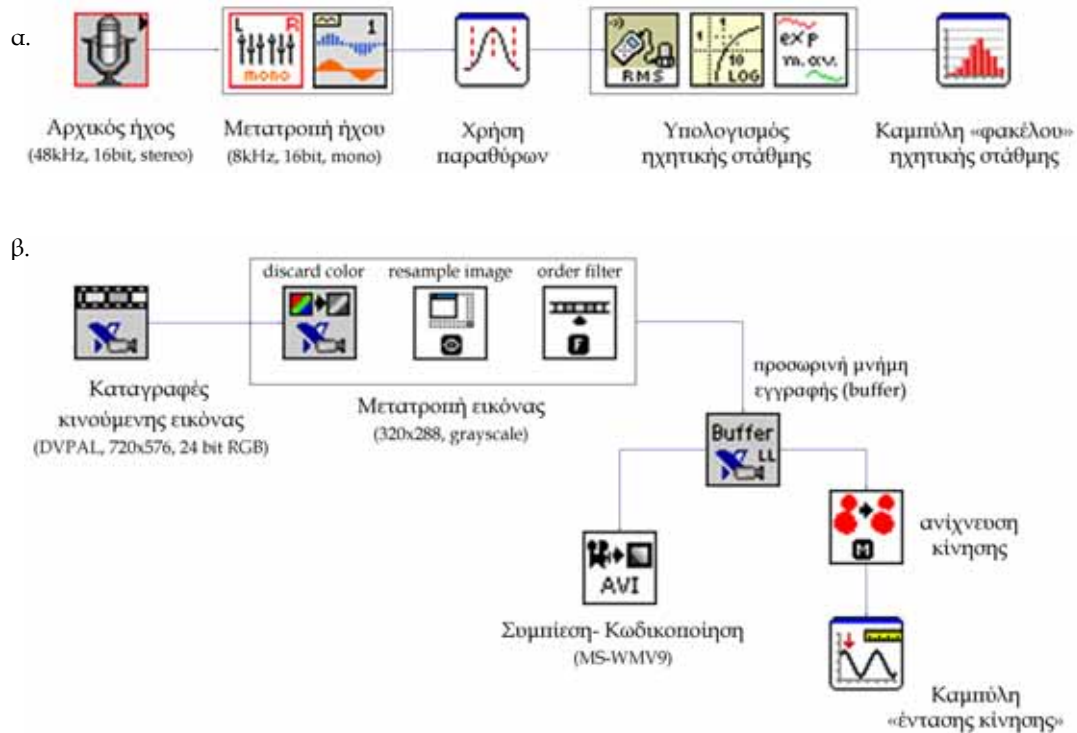
Συνοψίζοντας, οι τεχνικές επεξεργασίας που εφαρμόστηκαν στα οπτικοακουστικά σήματα παρακολούθησης, είχαν στόχο: α) τη βελτίωση της ποιότητας των πληροφοριών, β) τη συμπίεση και αποδοτικότερη αποθήκευση των οπτικοακουστικών δεδομένων, και γ) την εξαγωγή παραμέτρων σύνοψης - πλοήγησης των καταγραφών οπτικοακουστικής εποπτείας. Στο σχ. 2-19 δίνεται ένα συνοπτικό μπλοκ διάγραμμα των εμπλεκόμενων εφαρμογών.



Σχήμα 2-17. Διαδοχικά πλαίσια εικόνων κίνησης στους εικονολήπτες ευρείας γωνίας λήψης (α), και εστίασης (β).



Σχήμα 2-18. Τοπική καμπύλη έντασης κίνησης $I_{mot}(n)$ για την σύνοψη και παρακολούθηση των αλλαγών περιεχομένου, στις οπτικοακουστικές καταγραφές εποπτείας.



Σχήμα 2-19. Μπλοκ διάγραμμα του συστήματος οπτικοακουστικής παρακολούθησης και ελέγχου της πειραματικής διαδικασίας: α. επεξεργασία σημάτων ήχου, β. επεξεργασία καταγραφών κινούμενης εικόνας.

2.3 Επισκόπηση 2ου κεφαλαίου

Το κεφάλαιο αυτό επικεντρώνεται στην παρουσίαση του συστήματος καταγραφής, τόσο για τη λήψη των βιοακουστικών φαινομένων, όσο και για την παρακολούθηση της πειραματικής διαδικασίας. Είναι, επομένως, κεφαλαιώδους σημασίας για τα συνολικά αποτελέσματα της διατριβής, αφού το σύστημα καταγραφής αποτελεί το πρώτο και σημαντικότερο κομμάτι της «πειραματικής αλυσίδας». Στο πλαίσιο αυτό, παρουσιάστηκε η βασική θεωρία των οπτικοακουστικών τεχνολογιών, με έμφαση στις μετρητικές διατάξεις και τις σχετικές εφαρμογές βιοϊατρικής τεχνολογίας. Τα στοιχεία αυτά αξιοποιήθηκαν σε μεγάλο βαθμό, κατά το σχεδιασμό και τις διάφορες πειραματικές διαδικασίες αξιολόγησης των επιμέρους διατάξεων, που αναπτύχθηκαν. Το νέο σύστημα καταγραφής παρουσιάζει αρκετές καινοτομίες, ενώ ανταποκρίνεται πλήρως στις τεχνικές προδιαγραφές της ιατρικής έρευνας, όπως παρουσιάστηκαν στο προηγούμενο κεφάλαιο.

2.3.1 Αποτίμηση των παρεμβάσεων και ερευνητική συνεισφορά της εργασίας

Γενικά, δοκιμάστηκαν μια σειρά από ηλεκτρακουστικούς αισθητήρες λήψης, με στόχο τη διερεύνηση αξιοποίησής τους στο σύστημα βιοακουστικής καταγραφής. Από τις δοκιμές, αποφασίστηκε η χρήση πιεζοηλεκτρικών μετατροπέων (μικρόφωνα επαφής και επιταχυνσιόμετρα), καθώς χαρακτηρίζονται από ικανοποιητική απόδοση, έχουν ιδιαίτερα βολικά φυσικά χαρακτηριστικά και ικανοποιούν τα κριτήρια ασφάλειας της πειραματικής διαδικασίας. Τα στοιχεία αυτά πλαισιώθηκαν από τις κατάλληλες κυκλωματικές διατάξεις προενίσχυσης και τις αντίστοιχες μονάδες εγγραφής. Παράλληλα, αξιολογήθηκαν πολλές διαφορετικές διατάξεις χωρικής κατανομής των αισθητήρων, απ' όπου προέκυψε ένα νέο πολυκαναλικό σύστημα ηχητικής καταγραφής. Το σύστημα αυτό, που αποτελεί ερευνητική καινοτομία της παρούσης εργασίας, περιλαμβάνει πέντε αισθητήρες και επτά, συνολικά, κανάλια εγγραφής, παρέχοντας νέες δυνατότητες τοπογραφικής ερμηνείας και ακουστικής χαρτογράφησης (σχετικές λεπτομέρειες δίνονται στα κεφ. 4, 6).

Παράλληλα με το κομμάτι ηλεκτρακουστικού σχεδιασμού, ιδιαίτερη βαρύτητα δόθηκε στα ζητήματα προσαρμογής του μετρητικού συστήματος στα υποκείμενα. Έτσι, προέκυψε η προσαρμοστική διάταξη *SHIELD*, η οποία περιλαμβάνει την επιφάνεια στερέωσης των αισθητήρων (*FSA*) και το ηχομονωτικό γιλέκο προσαρμογής (*WAAV*). Η καινοτομία, αυτή, παρέμβαση, παρουσιάζει αρκετά πλεονεκτήματα, αντιμετωπίζοντας αποτελεσματικά τα σχετικά προβλήματα. Για να γίνει κατανοητή η λειτουργικότητα της διάταξης *SHIELD*, αρκεί να την συγκρίνουμε με την προσέγγιση των Gamer και Ehrenreich [69], οι οποίοι πρότειναν την αντιμετώπιση των προβλημάτων θορύβου με χρήση πολλών «σκεπασμάτων» (κουβέρτες), αλλά και αυτή των Tomomasa κ.ά. [72], [73], οι οποίοι συνιστούν την απόλυτη ακινησία και σιγή των υποκειμένων, σε ολόκληρη τη διάρκεια των καταγραφών. Με βάση τα αποτελέσματα των πειραματικών διαδικασιών αξιολόγησης, η προσθήκη της διάταξης *SHIELD* δικαίωσε απόλυτα τις προσδοκίες του αρχικού σχεδιασμού. Έτσι, εκτός από τη βελτίωση των λειτουργικών χαρακτηριστικών και την επιτάχυνση των διαδικασιών προετοιμασίας του πειραματικού πρωτοκόλλου, παρατηρήθηκε ελαχιστοποίηση των ακουστικών θορύβων περιβάλλοντος, σημαντικός περιορισμός των τριβο-ηλεκτρικών θορύβων (κυρίως αυτών που παράγονται από μικρές κινήσεις των υποκειμένων) και βελτίωση της μηχανικής προσαρμογής της μετρητικής διάταξης στη δονούμενη επιφάνεια. Στα συμπεράσματα αυτά συγκλίνουν και οι αυτοματοποιημένες παρατηρήσεις, που προέκυψαν από τη χρήση του συστήματος οπτικοακουστικής παρακολούθησης.

Ειδικότερα για το σύστημα εποπτείας, αξίζει να σημειωθεί ότι αποτελεί, επίσης, πρωτοτυπία της εργασίας, καθώς είναι πρώτη φορά που χρησιμοποιείται σε παρατεταμένες καταγραφές της γαστρεντερικής ακουστικής δραστηριότητας. Όπως έχει ήδη επισημανθεί, η μελλοντική του αξιοποίηση, σε συνδυασμό με την παράλληλη διενέργεια πρόσθετων ψυχοφυσιολογικών καταγραφών, αναμένεται με ιδιαίτερο ενδιαφέρον. Στην παρούσα φάση έγινε διευθέτηση όλων των τεχνικών ζητημάτων, που συνδέονται με την επιλογή των παραμέτρων ψηφιοποίησης - συμπίεσης, αλλά και τις δομές περιγραφής και διαχείρισης του οπτικοακουστικού περιεχομένου. Έτσι, διαμορφώθηκε μια νέα τεχνική ανίχνευσης κίνησης, η οποία προέκυψε από τροποποιήσεις υπάρχοντων μεθόδων. Η προσέγγιση αυτή δίνει ιδιαίτερα ικανοποιητικά αποτελέσματα με μικρό, σχετικά, υπολογιστικό κόστος, παρέχοντας παράλληλα τη δυνατότητα ρύθμισης της ευαισθησίας, ανάλογα με τις απαιτήσεις κάθε εφαρμογής. Τα σχετικά συμπεράσματα μπορούν να αξιοποιηθούν σε πολλές παραπλήσιες εφαρμογές ψυχο-φυσιολογικής παρακολούθησης και, γενικά, στο σύνολο των συστημάτων οπτικοακουστικής επιτήρησης και ελέγχου, που παρουσιάζουν κοινά χαρακτηριστικά.

2.3.2 Προτάσεις μελλοντικής εξέλιξης του συστήματος καταγραφής

Το σύνολο των μετρητικών διατάξεων, στην παρούσα τους μορφή, χρησιμοποιείται στις τρέχουσες πειραματικές καταγραφές που διεξάγονται από την ιατρική ερευνητική ομάδα. Η εφαρμογή τους στην πράξη θα αναδείξει και τις όποιες αδυναμίες, που προς το παρόν τουλάχιστον δεν είναι ορατές. Ωστόσο, από την εμπειρία των πρώτων μετρήσεων και των αντίστοιχων εργαστηριακών δοκιμών, είναι δυνατό να προταθούν κάποιες μελλοντικές εξελίξεις, οι οποίες μπορούν να δώσουν ακόμη μεγαλύτερη ευελιξία και λειτουργικότητα.

Μια πρόταση εξέλιξης αφορά στη χρήση ασύρματων συστημάτων βιοακουστικής καταγραφής, ώστε να μην υπάρχει ανάγκη «καλωδίωσης» των υποκειμένων. Στο πλαίσιο αυτό θα μπορούσε να αξιοποιηθεί η σύγχρονη τεχνολογία blue-tooth, καθώς συγκεντρώνει πολλά επιθυμητά χαρακτηριστικά, όπως είναι η εύκολη διασύνδεση σε ψηφιακά δίκτυα δεδομένων, η συμβατότητα πολλών ψηφιακών τερματικών με το πρωτόκολλο αυτό (μεταξύ των οποίων και πολλές βιοϊατρικές διατάξεις) και η μικρή ισχύς εκπομπής. Εναλλακτικά, θα μπορούσαν να αξιοποιηθούν ολοκληρωμένα κυκλώματα εγγραφής τυχαίας προσπέλασης (μνήμες), τα οποία χαρακτηρίζονται από πολύ μικρό μέγεθος και ολοένα αυξανόμενες αποθηκευτικές δυνατότητες. Στην περίπτωση αυτή θα μπορούσαν να συμπεριληφθούν και μονάδες ψηφιακής επεξεργασίας πραγματικού χρόνου (DSPs) για τις διάφορες ανάγκες, κατά τη λήψη των σημάτων (signal conditioning), καθώς και για την ενσωμάτωση πολλών από τις τεχνικές επεξεργασίας που περιγράφονται σε επόμενα κεφάλαια. Τέλος, προτείνεται η δοκιμή νέων πολυκαναλικών διατάξεων χωρικής καταγραφής με χρήση περισσότερων 3-αξονικών επιταχυνσιομέτρων (σχετικές λεπτομέρειες δίνονται στο 6ο κεφάλαιο).

Παράλληλα, όλα τα επιμέρους συστήματα λήψης - επεξεργασίας θα μπορούσαν να ενσωματωθούν σε μια ενιαία «βιοϊατρική» συσκευή καταγραφής, με τα χαρακτηριστικά εύκολης χρήσης και εφαρμογής της διάταξης προσαρμογής SHIELD. Μάλιστα, ενδιαφέρον παρουσιάζει και η εξέλιξη των πρωτότυπων πειραματικών κατασκευών FSA και WAAV, στο πλαίσιο ενός εγχειρήματος «βιομηχανικού σχεδιασμού». Έτσι, η διατήρηση της αρχικής φιλοσοφίας και η εκ νέου υλοποίηση των δύο κατασκευών σε επίπεδο «βιομηχανικής παραγωγής», θα επιτρέψει την αξιοποίηση νέων τεχνολογιών και υλικών, συμβάλλοντας στη βελτίωση φυσικών και λειτουργικών χαρακτηριστικών. Η πιστοποίηση, δε, των νέων παραμέτρων λειτουργίας, βάσει προτύπων διασφάλισης ποιότητας, θα οδηγήσει στην τυποποίηση μιας νέας βιοϊατρικής διάταξης, για χρήση σε επίπεδο κλινικής εφαρμογής.

ΚΕΦΑΛΑΙΟ 3^ο

3. Τεχνικές καταστολής θορύβου σε ηχητικές καταγραφές

Στο κεφάλαιο αυτό αναλύονται τα ζητήματα εμφάνισης του προσθετικού θορύβου βάθους ευρείας ζώνης (Additive Broadband Noise -ABN) και οι επιδράσεις του τελευταίου σε εφαρμογές επεξεργασίας - ανάλυσης ηχητικών καταγραφών. Αρχικά παρατίθεται το γενικό θεωρητικό μοντέλο περιγραφής και οι αντίστοιχες τεχνικές καταστολής του ανεπιθύμητου αυτού φαινομένου. Έτσι, γίνεται αναφορά στις κλασσικές μεθόδους αποθορυβοποίησης, με ιδιαίτερη έμφαση στις προσεγγίσεις κυματιδιακής ανάλυσης, οι οποίες χρησιμοποιήθηκαν εκτενώς στο πλαίσιο της παρούσης διατριβής. Ακολουθεί η βιβλιογραφική αναφορά των τεχνικών αποθορυβοποίησης που έχουν εφαρμοστεί στην επεξεργασία κουλιακών ήχων. Το μεγαλύτερο μέρος του κεφαλαίου αφιερώνεται στην παρουσίαση μιας νέας παραμετρικής μεθόδου, που αναπτύχθηκε ειδικά για τις ανάγκες της παρούσης εφαρμογής, επιδιώκοντας να ξεπεράσει μειονεκτήματα και αδυναμίες των τεχνικών που προηγήθηκαν. Στο πλαίσιο αυτό συμπεριλαμβάνονται διαδικασίες ποιοτικής και ποσοτικής αξιολόγησης της νέας μεθόδου, συγκρίσεις με κλασσικούς αλγόριθμους καταστολής θορύβου, ενώ παρατίθενται και προτάσεις μελλοντικής επέκτασης.

3.1 Περιγραφή του προβλήματος καταστολής θορύβου

Γενικά, υπάρχουν διάφοροι παράγοντες στους οποίους οφείλεται η παρουσία θορύβου, οι οποίοι χαρακτηρίζουν τη φύση και ονομασία του αντίστοιχου προβλήματος «μόλυνσης». Ανάμεσα στα συνηθισμένα είδη αλλοιώσεων μπορούμε να διακρίνουμε: α) κρουστικούς θορύβους (*impulsive noises, clicks, pops*), που παράγονται από ηλεκτροστατικές εκκενώσεις, οπτικές παρεμβολές ηλεκτρομαγνητικής φύσης και σφάλματα στα μέσα εγγραφής, (π.χ. δίσκοι βινυλίου), β) τυχαίους προσθετικούς θορύβους που οφείλονται σε παρεμβολές λόγω παρουσίας ανεπιθύμητων πηγών ή πολλαπλών οδεύσεων του ίδιου σήματος (*ambient noise, multi-path fading*), γ) θορύβους παραμόρφωσης που οφείλονται στη μη ιδεατή συμπεριφορά των μέσων εγγραφής και των καναλιών μετάδοσης (μη γραμμικές μεταβολές, φασικές παραμορφώσεις), σε φαινόμενα ψαλιδισμού (*clips, harmonic distortion*), καθώς και σε σφάλματα κατά τη διαδικασία ψηφιοποίησης, δ) θορύβους βάθους ευρείας ζώνης, οι οποίοι οφείλουν την ύπαρξή τους σε θερμικά, κυρίως, φαινόμενα, καθώς και σε συνδυασμούς των προαναφερθέντων παραγόντων (σε μικρότερο βαθμό) [177]-[178]. Από όλα τα παραπάνω, θα επικεντρώσουμε την προσοχή μας στην τελευταία κατηγορία, δηλαδή τον προσθετικό θόρυβο φόντου (ή βάθους) ευρείας ζώνης.

3.1.1 Θόρυβος φόντου ευρείας ζώνης

Ο θόρυβος *ABN* είναι συνηθισμένο πρόβλημα σε όλες σχεδόν τις μετρητικές εφαρμογές, ειδικά αυτές που σχετίζονται με λήψη ασθενικών σημάτων και επιβάλλουν υψηλά «κέρδη» ενίσχυσης [179]. Συνήθως μοντελοποιείται ως στάσιμη Γκαουσιανή διαδικασία (*Stationary Gaussian Process*), όπου στην απλούστερη περίπτωση έχει ομοιόμορφο φασματικό περιεχόμενο σε όλες τις συχνοτικές περιοχές (λευκός θόρυβος, *Gaussian White Noise -GWN*).

Όταν η καμπύλη φασματικής ανάλυσης δεν είναι επίπεδη, τότε μιλάμε πλέον για έγχρωμο θόρυβο (colored noise), όπως είναι ο ροζ και ο καφέ θόρυβος (Pink Noise –PN, Brown Noise) [108]-[109]. Η παρουσία του θορύβου φόντου οφείλεται σε διάφορους παράγοντες, όπως είναι: α) ο θερμικός θόρυβος που δημιουργείται κατά μήκος της «μετρητικής αλυσίδας» (στους ηλεκτρακουστικούς μετατροπείς και στις μονάδες προ-ενίσχυσης), β) ο θόρυβος περιβάλλοντος χαμηλής στάθμης λόγω της παρουσίας πρόσθετων (μη επιθυμητών) ηχητικών πηγών και των αναπόφευκτων προβλημάτων συνακρόασης (ambient noise), γ) τα προβλήματα εμφάνισης θερμικού θορύβου κατά την αναλογική εγγραφή των σημάτων, κυρίως σε μέσα μαγνητικής ταινίας (hiss noise), δ) ο θόρυβος κβάντισης (quantization noise), ειδικά στις περιπτώσεις που οι στάθμες εγγραφής δεν έχουν ρυθμιστεί σωστά και βρίσκονται σε χαμηλά επίπεδα [179]. Σε διάφορα προφίλ θορύβου βάθους ενδέχεται να συμπεριλαμβάνονται και αρμονικές συνιστώσες (σήματα ανεπιθύμητων βόμβων), ειδικά όταν δεν ικανοποιούνται οι προδιαγραφές προσαρμογής και ορθής διασύνδεσης των κυκλωματικών βαθμίδων, όπως περιγράφηκε στο προηγούμενο κεφάλαιο.

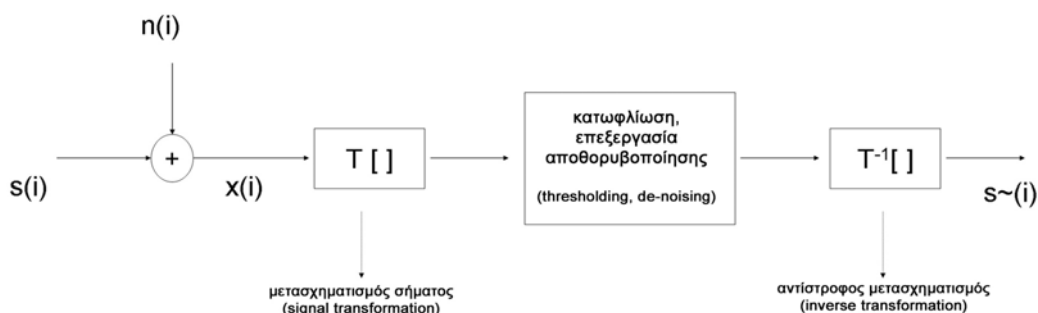
Αν συμβολίσουμε με $x(i)$ την ψηφιοποιημένη ακολουθία του ενθόρυβου σήματος $x(t)$, με τα $s(i)$ και $n(i)$ να περιγράφουν τις αντίστοιχες ακολουθίες του «καθαρού» σήματος $s(t)$ και του ασυσχέτιστου προσθετικού θορύβου $n(t)$, τότε ισχύει:

$$x(i) = s(i) + n(i) \quad (3-1)$$

Ο σκοπός των τεχνικών αποθορυβοποίησης είναι να ανακτηθεί μια προσέγγιση $\tilde{s}(i)$ του αρχικού σήματος, η οποία είναι απαλλαγμένη από τις συνιστώσες θορύβου και πλησιάζει στο αυθεντικό αρχικό σήμα $s(i)$ με το μικρότερο δυνατό σφάλμα. Η διαδικασία «μόλυνσης» ενός σήματος με τυχαίο θόρυβο απεικονίζεται σχηματικά στο σχ. 3-1, όπου παρουσιάζεται το βασικό μοντέλο των τεχνικών καταστολής θορύβου [179].

3.1.2 Μέθοδοι καταστολής θορύβου φόντου

Το «γενικό μοντέλο καταστολής θορύβου» (σχ. 3-1) περιλαμβάνει, σχεδόν κατά κανόνα, τα ακόλουθα βήματα: α) μετασχηματισμό (με την ευρύτερη έννοια του όρου) του αρχικού (ενθόρυβου) σήματος σε ένα κατάλληλο πεδίο επεξεργασίας, που διαχωρίζει κατά τον καλύτερο δυνατό τρόπο τις συνιστώσες θορύβου από το καθαρό σήμα, β) επεξεργασία κατωφλίωσης των ενθόρυβων συντελεστών, με στόχο την απομάκρυνση των συνιστωσών θορύβου, γ) αντίστροφος μετασχηματισμός των επεξεργασμένων, απαλλαγμένων από θόρυβο, συντελεστών για την απόκτηση του καθαρού σήματος [179].



Σχήμα 3-1. Σχηματική αναπαράσταση της «παρεμβολής» τυχαίου προσθετικού θορύβου και του βασικού μοντέλου αποθορυβοποίησης.

Διάφοροι μετασχηματισμοί και αντίστοιχα σχήματα ανάλυσης - σύνθεσης έχουν χρησιμοποιηθεί με βάση το παραπάνω μοντέλο αποθορυβοποίησης. Ιδιαίτερα δημοφιλείς είναι οι μέθοδοι φασματικής αφαίρεσης, που συχνά αναφέρονται ως γενικευμένο Wiener φίλτρο, και οι κυματιδιακές τεχνικές [179]. Οι προσεγγίσεις αυτές αποτέλεσαν τη βάση για την ανάπτυξη της προτεινόμενης μεθόδου, γι' αυτό και θα ακολουθηθεί μια πιο αναλυτική παρουσίασή τους. Εξάλλου, ο μετασχηματισμός κυματιδίων έχει αξιοποιηθεί και σε άλλες τεχνικές επεξεργασίας στο πλαίσιο της παρούσης διατριβής (περισσότερες πληροφορίες δίνονται στα κεφάλαια 4 και 5). Για το λόγο αυτό, και δεδομένου ότι πρόκειται για ένα νέο σχετικά επιστημονικό πεδίο, κρίνεται αναγκαία η παρουσίαση της βασικής θεωρίας που διέπει την κυματιδιακή ανάλυση.

3.1.2.1 Τεχνικές φασματικής αφαίρεσης και παραμετρικό φίλτρο Wiener

Η μέθοδος της φασματικής αφαίρεσης παρουσιάστηκε για πρώτη φορά από τον Boll το 1979 [180] και παραμένει μέχρι σήμερα μια από τις πιο χρησιμοποιούμενες μεθόδους για την αφαίρεση του θορύβου φόντου [181]-[183]. Αν συμβολίσουμε με $X(k)$, $S(k)$ και $N(k)$ τις φασματικές ακολουθίες του ενθόρυβου σήματος $x(i)$, του αρχικού «καθαρού» σήματος $s(i)$ και του θορύβου $n(i)$, αντίστοιχα, τότε, με βάση το μοντέλο του σχ. 3-1, ισχύει:

$$X(k) = S(k) + N(k) \Rightarrow S(k) = X(k) - N(k) \quad (3-2)$$

Η επίλυση στο πρόβλημα αποθορυβοποίησης είναι η εφαρμογή ενός φίλτρου με συνάρτηση μεταφοράς $H(k)$, η οποία προσεγγίζει κατά τον καλύτερο δυνατό τρόπο την πράξη φασματικής αφαίρεσης της εξίσωσης (3-2). Θεωρώντας ότι «μολυσματικός» θόρυβος αποτελεί μια στάσιμη τυχαία μεταβλητή, τότε μπορούμε να αποκτήσουμε μια φασματική του εκτίμηση $N_{FP}(k)$ από ένα αποτύπωμα (noise footprint) θορύβου $n_{FP}(i)$. Εφαρμόζοντας τον ταχύ μετασχηματισμό Fourier (Fast Fourier Transform -FFT) στην τελευταία ακολουθία (η οποία μπορεί να εντοπιστεί εύκολα στις περιόδους σιγής του ενθόρυβου σήματος) προκύπτει το φασματικό προφίλ $N_{FP}(k)$, οπότε η εξίσωση της φασματικής αφαίρεσης γίνεται:

$$S^{\sim}(k) = X(k) - N_{FP}(k) = H(k) \cdot X(k) \quad (3-3)$$

Φασματική αφαίρεση πλάτους και ισχύος

Οι δύο πιο βασικές τεχνικές φασματικής αφαίρεσης, είναι η φασματική αφαίρεση πλάτους (magnitude spectral subtraction):

$$|S^{\sim}(k)| = \begin{cases} |X(k)| - |N_{FP}(k)|, & |X(k)| > |N_{FP}(k)| \\ 0, & \text{διαφορετικά} \end{cases} \quad (3-4)$$

και η φασματική αφαίρεση ισχύος (power spectral subtraction):

$$|S^{\sim}(k)|^2 = \begin{cases} P_X(k) - a \cdot P_{N_{FP}}(k), & P_X(k) - a \cdot P_{N_{FP}}(k) > \beta \cdot P_{N_{FP}}(k) \\ \beta \cdot P_{N_{FP}}(k), & \text{διαφορετικά} \end{cases} \quad (3-5)$$

όπου:

$$P_X(k) \triangleq E\{|X(k)|^2\}, \quad P_{N_{FP}}(k) \triangleq E\{|N_{FP}(k)|^2\} \quad (3-6)$$

Οι παράμετροι a, β παίρνουν θετικές - πραγματικές τιμές και χρησιμοποιούνται για τον έλεγχο της «βαθμού αφαιρέσεως», ενώ παράλληλα είναι υπεύθυνες για τη διαμόρφωση του εναπομειναντος θορύβου (remaining noise floor) στα επιθυμητά ενεργειακά επίπεδα [184]-[185]. Έτσι, και χωρίς να γίνει καμιά φασική επεξεργασία (θεωρείται ότι κύρια επίδραση του θορύβου εντοπίζεται στο πλάτος του σήματος), το αποθορυβοποιημένο σήμα προκύπτει από τον αντίστροφο μετασχηματισμό Fourier (Inverse Fourier Transform -IFT) [185]-[186]:

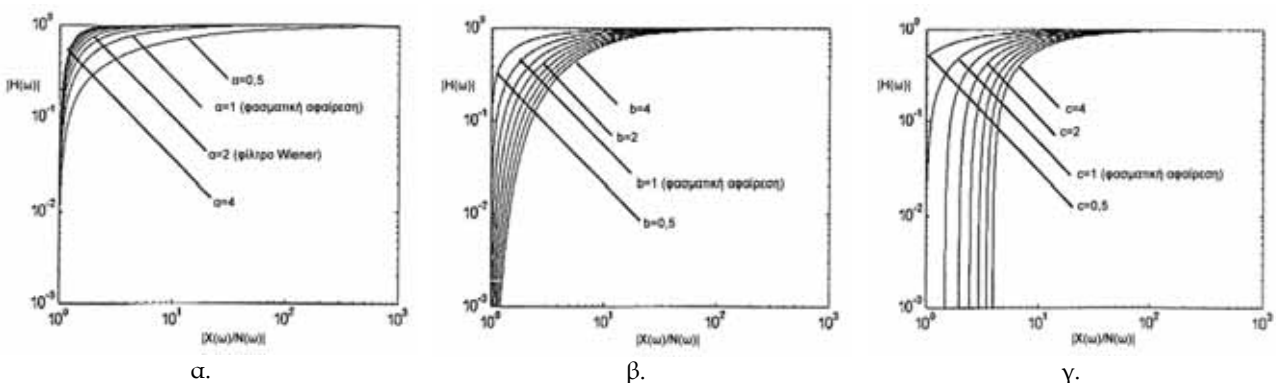
$$s^{\sim}(i) = IFT \left\{ \left| S^{\sim}(k) \right|, \mathcal{A} \left(X(k) \right) \right\} \quad (3-7)$$

Παραμετρικό φίλτρο Wiener (Parametric Wiener Filter -PWF)

Ένας πιο γενικευμένος κανόνας αποθορυβοποίησης περιγράφεται από το παραμετρικό φίλτρο Wiener, το οποίο έχει αποδεχθεί ότι είναι ισοδύναμο με την πράξη της φασματικής αφαιρέσεως, όταν εφαρμόζεται σε σήματα περιορισμένης χρονικής διάρκειας [177], [187]. Το φίλτρο Wiener ελαχιστοποιεί το μέσο τετραγωνικό σφάλμα ανάμεσα στο αρχικό και το ανακατασκευασμένο σήμα στο πεδίο του χρόνου, για την περίπτωση ασυσχέτιστου προσθετικού θορύβου μηδενικής μέσης τιμής (*uncorrelated, zero-mean, additive noise*) [185]. Η μαθηματική έκφραση της συνάρτησης μεταφοράς H_{PWF} του παραμετρικού φίλτρου Wiener, δίνεται από την ακόλουθη σχέση:

$$H_{PWF}(k) = \begin{cases} \left\{ 1 - c \left[\frac{N_{FP}(k)}{X_T(k)} \right]^a \right\}^b, & \text{αν } c \cdot \left[\frac{N_{FP}(k)}{X_T(k)} \right]^a \leq 1 \\ 0, & \text{διαφορετικά} \end{cases} \quad (3-8)$$

όπου οι παράμετροι a, b και c παίρνουν και πάλι θετικές-πραγματικές τιμές και χρησιμοποιούνται για τη ρύθμιση του φίλτρου (σχ. 3-2), $X_T(k)$ είναι το εκτιμώμενο φάσμα του επιλεγμένου ενθόρυβου σήματος διάρκειας T (windowed signal), ενώ $N_{FP}(k)$ είναι, όπως και προηγουμένως, η διαθέσιμη φασματική εκτίμηση του θορύβου [177]-[178], [184]-[187].



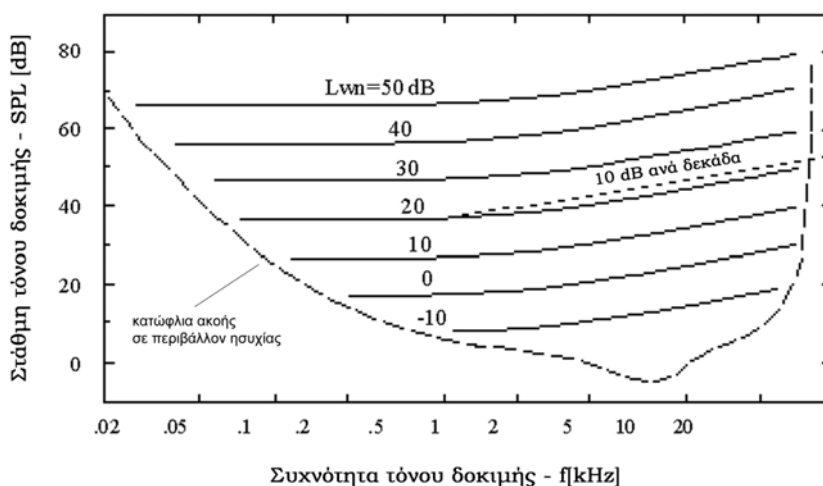
Σχήμα 3-2. Επίδραση των παραμέτρων a (σχήμα 3-2α), b (σχήμα 3-2β) και c (σχήμα 3-2γ) στην καμπύλη αποκοπής του φίλτρου $|H(\omega)|$, για διάφορες τιμές του λόγου $|X(\omega)/N(\omega)|$.

Στα διαγράμματα του σχήματος 3-2 παρατηρούμε ότι, για διαφορετικές τιμές των παραμέτρων a, b, c , πετυχαίνουμε διαφορετικές καμπύλες «συμπίεσης» του φίλτρου. Οι καμπύλες αυτές διαφοροποιούνται ανάλογα με το «βαθμό παρουσίας του θορύβου», δηλαδή τη σηματοθορυβική σχέση που εκφράζεται από το λόγο $|X(\omega)/N(\omega)|$. Όπως είναι φυσικό, όσο ο λόγος αυτός μεγαλώνει τόσο το κέρδος πλησιάζει τη μονάδα. Αν σε μια

φασματική συνιστώσα παρουσιάζεται πολύ μεγάλη σηματοθορυβική σχέση ($\gg 30\text{dB}$), τότε η συγκεκριμένη συνιστώσα παραμένει αμετάβλητη. Διαφορετικά, αν η σηματοθορυβική σχέση είναι πολύ μικρή ($< 0\text{dB}$), η συνιστώσα συμπιέζεται δραματικά. Παρατηρούμε ότι η αύξηση της παραμέτρου a έχει ως αποτέλεσμα πιο «οξύ» φιλτράρισμα, στις χαμηλές (κυρίως) σηματοθορυβικές σχέσεις. Η ακριβώς αντίστροφη συμπεριφορά παρατηρείται όταν μεταβάλλονται οι παράμετροι b , c , όπου αύξηση των τιμών τους οδηγεί σε «ηπιότερο» φιλτράρισμα για τις χαμηλές σηματοθορυβικές σχέσεις [177], [179], [184], [187]. Στην περίπτωση που $a=b=c=1$ η εξίσωση (3-8) περιγράφει την κλασική φασματική αφαίρεση πλάτους και είναι ισοδύναμη με τη σχέση (3-4). Ανάλογα, στην περίπτωση που $a=2$ και $b=c=1$ έχουμε την εξίσωση της φασματικής αφαίρεσης ισχύος, δηλαδή η σχέση (3-8) ισοδυναμεί με την (3-5). Σύμφωνα με την ορολογία που υιοθετούν πολλές εργασίες [177], [184], [187], η προηγούμενη ρύθμιση αντιστοιχεί στο κλασικό φίλτρο Wiener. Σε κάθε περίπτωση, η εξίσωση (3-8) είναι πιο γενική και εμπεριέχει τις υποπεριπτώσεις φασματικής αφαίρεσης, των εξισώσεων (3-4) και (3-5) [177], [179], [184], [187].

Εξαρτώμενο από τη συχνότητα παραμετρικό φίλτρο Wiener (Frequency Depended Parametric Wiener Filter - FDPWF)

Η διαδικασία εκτίμησης του θορύβου και η αντίστοιχη διαμόρφωση του γενικευμένου φίλτρου αποθορυβοποίησης (εξίσωση 3-8) όπως παρουσιάστηκε έως τώρα, δεν περιλαμβάνει καθόλου κριτήρια που να σχετίζονται με τις αντιληπτικές ικανότητες του ανθρώπου. Είναι γνωστό ότι η «απόκριση του ανθρώπου» στα ακουστικά ερεθίσματα δεν είναι φασματικά ομοιόμορφη. Ένα σχετικό ψυχο-ακουστικό φαινόμενο είναι αυτό του «θορύβου ομοιόμορφης απόκρυψης» (equivalent masking noise), σύμφωνα με το οποίο, ο βαθμός ενόχλησης που προκαλούν ηχητικά ερεθίσματα της ίδιας ενέργειας αλλά διαφορετικού φασματικού περιεχομένου, δεν είναι ίδιος. Η ενσωμάτωση ενός τέτοιου ψυχο-ακουστικού παράγοντα στην εξίσωση (3-8), θα μπορούσε να επεκτείνει ακόμη περισσότερο τις δυνατότητες του αποκαλούμενου «εξαρτώμενου από τη συχνότητα παραμετρικού φίλτρου Wiener» (Frequency Depended Parametric Wiener Filter) [177], [187]-[188].



Σχήμα 3-3. Απεικόνιση των κατωφλίων της ακοής, για τόνους δοκιμής στάθμης SPL [dB], παρουσία λευκού θορύβου σε διάφορες στάθμες L_{wn} ([188], p. 57).

Η έννοια του θορύβου ομοιόμορφης απόκρυψης γίνεται ευκολότερα κατανοητή με τη βοήθεια του σχ. 3-3. Στο διάγραμμα αυτό, το οποίο έχει προκύψει από στατιστικά ψυχοακουστικά πειράματα [188], απεικονίζονται τα κατώφλια της ακοής, όπως

διαμορφώνονται με την παρουσία λευκού θορύβου σε διάφορες στάθμες L_{wn} . Η μεγάλη διακεκομμένη γραμμή απεικονίζει το κατώφλι σε περιβάλλον «ησυχίας». Είναι σαφές ότι ο λευκός θόρυβος, παρόλο που είναι ομοιόμορφα κατανεμημένος μέσα στο ακουστό φάσμα, δεν προκαλεί ομοιόμορφα κατώφλια απόκρυψης (απεικονίζονται με συνεχείς γραμμές), με αποτέλεσμα να προκαλείται διαφορετική ακουστότητα σε τόνους ίδιας στάθμης αλλά διαφορετικής συχνότητας. Παρατηρούμε ότι τα κατώφλια απόκρυψης είναι οριζόντια μόνο στις χαμηλές συχνότητες (μέχρι τα 500Hz περίπου), ενώ για μεγαλύτερες συχνότητες παρουσιάζουν άνοδο, με μια κλίση της τάξης των 10 dB ανά δεκάδα (απεικονίζεται από τη μικρή διακεκομμένη γραμμή). Έτσι, προέκυψε η έννοια του θορύβου ομοιόμορφης απόκρυψης, ο οποίος παράγεται όταν ένα σήμα λευκού θορύβου περάσει από φίλτρο με την ακόλουθη συνάρτηση μεταφοράς [177], [187]- [188]:

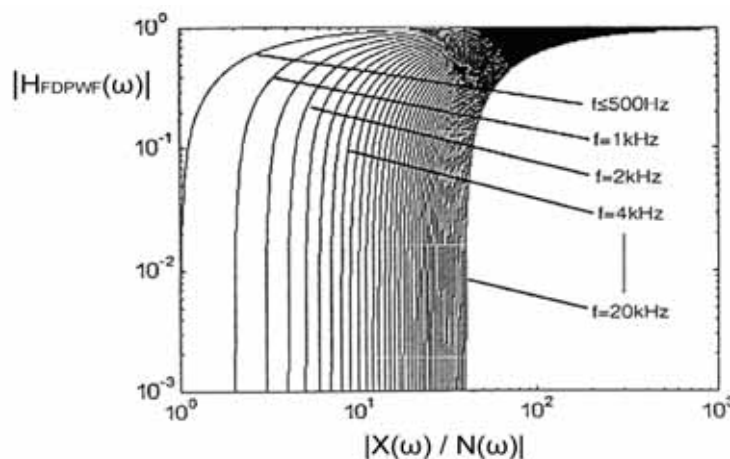
$$A_k \triangleq A(k) = \begin{cases} 1, & k \cdot df \leq 500 \\ \frac{k \cdot df}{500}, & k \cdot df > 500 \end{cases} \quad (3-9)$$

όπου $df=f_s/N$ είναι το διάστημα δειγματοληψίας, f_s η συχνότητα δειγματοληψίας και N το μήκος της ψηφιοποιημένης ακολουθίας.

Σύμφωνα με τη θεωρία του θορύβου ομοιόμορφης απόκρυψης, είναι δυνατόν να προσαρμόσουμε το φάσμα του θορύβου με βάση την ενόχληση που προκαλείται στο όργανο της ακοής. Αν λοιπόν εισάγουμε τη συνάρτηση μεταφοράς της προηγούμενης σχέσης (3-9) στη γενική εξίσωση (3-8) του παραμετρικού φίλτρου Wiener, τότε προκύπτει το παραμετρικό και εξαρτώμενο από τη συχνότητα φίλτρο Wiener (Frequency Dependent Parametric Wiener Filter - FDPWF) H_{FDPWF} :

$$H_{FDPWF}(k) = \begin{cases} \left\{ 1 - c \cdot A(k) \cdot \left[\frac{|N_{FP}(k)|}{|X_T(k)|} \right]^a \right\}^b, & \text{αν } c \cdot A(k) \cdot \left[\frac{|N_{FP}(k)|}{|X_T(k)|} \right]^a \leq 1 \\ 0, & \text{διαφορετικά} \end{cases} \quad (3-10)$$

Όπως προκύπτει από το παρακάτω διάγραμμα (σχ. 3-4) η ενσωμάτωση του όρου $A(k)$ διαφοροποιεί σημαντικά το φίλτρο H_{FDPWF} από την απλή φασματική αφαίρεση [177], [187].



Σχήμα 3-4. Συνάρτηση μεταφοράς του φίλτρου H_{FDPWF} για διαφορετικές συχνότητες.

Γενικά, το παραμετρικό φίλτρο Wiener παρέχει πολύ καλά αποτελέσματα αποθορυβοποίησης, με μικρή επίδραση στις χρήσιμες συνιστώσες του πληροφοριακού σήματος και χωρίς έντονα προβλήματα παραμορφώσεων. Το κύριο μειονέκτημά του οφείλεται στο γεγονός ότι αφήνει υπολείμματα θορύβου (noise residuals) στις περιόδους σιγής. Χαρακτηριστικά είναι και τα υπολείμματα θορύβου τύπου “birdy noise”, τα οποία εμφανίζονται κυρίως σε χαμηλές σηματοθορυβικές σχέσεις, δηλαδή όταν οι συνιστώσες του σήματος μεταβάλλονται πολύ κοντά στα επίπεδα κατωφλίωσης. Οι σχετικές υλοποιήσεις, μέσω βραχυπρόθεσμου μετασχηματισμού Fourier (Short Term Fourier Transform -STFT), τείνουν να παράγουν πρόσθετα «σφάλματα αποθορυβοποίησης» (de-noising artifacts), με ποιο χαρακτηριστική την εμφάνιση «μουσικών ήχων» [177]-[179], [184]-[187], [189]-[190]. Όσον αφορά την ενσωμάτωση του φίλτρου $A(k)$, αυτό παρέχει τη δυνατότητα περαιτέρω διαμόρφωσης της «συνολικής» κατωφλίωσης “ $c \cdot A(k)$ ”, ανάλογα με το φασματικό περιεχόμενο του ενθόρυβου σήματος και τα χαρακτηριστικά του θορύβου. Πειραματικές παρατηρήσεις [177], [179], [187] έχουν δείξει ότι δίνει πολύ καλά ακουστικά αποτελέσματα (audible results), περιορίζοντας τις μη αντιληπτά συνιστώσες σήματος και παρουσιάζοντας μεγαλύτερη ανοχή στα σφάλματα αποθορυβοποίησης (de-noising artifacts, noise residuals).

3.1.2.2 Κυματιδιακή ανάλυση και τεχνικές κατωφλίωσης κυματιδιακών συντελεστών

Στην ενότητα αυτή αναλύονται οι τεχνικές κυματιδιακής κατωφλίωσης, οι οποίες έχουν χρησιμοποιηθεί ευρύτατα σε εφαρμογές αποθορυβοποίησης. Πριν την περιγραφή των σχετικών τεχνικών, κρίνεται σκόπιμη η συνοπτική παρουσίαση της κυματιδιακής θεωρίας και η σύγκρισή της με τις υπόλοιπες μεθόδους «συνδυασμένης ανάλυσης χρόνου-συχνότητας» (Short Term Spectral Analysis -STSA), όπως είναι ο μετασχηματισμός STFT.

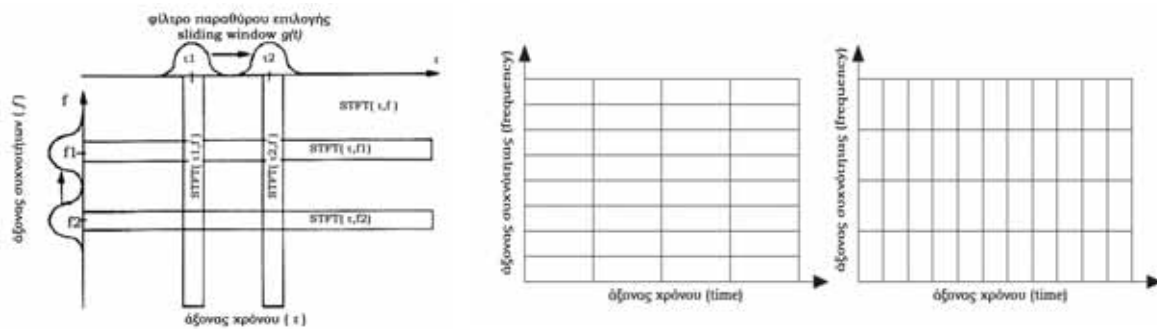
A. Συνδυασμένη ανάλυση χρόνου-συχνότητας (Joint Time Frequency Analysis -JTFA)

Δεδομένου ότι οι τεχνικές καταστολής θορύβου αφορούν ως επί το πλείστον μη στάσιμα σήματα, όλες οι προηγούμενες σχέσεις του παραμετρικού φίλτρου Wiener υλοποιούνται συνήθως με τεχνικές STSA. Στην περίπτωση του μετασχηματισμού STFT, χρησιμοποιούνται «χρονικά παράθυρα επιλογής» (συχνά επικαλυπτόμενα, overlapped windows), και στη συνέχεια προκύπτουν οι φασματικές εκτιμήσεις με χρήση του αλγορίθμου FFT σε κάθε τμήμα του σήματος. Η συνδυασμένη ανάλυση χρόνου - συχνότητας του STFT, έδωσε μεγαλύτερη προσαρμοστικότητα σε σχέση με την κλασική ανάλυση Fourier, καθώς πλέον επιτυγχάνεται καλύτερη διακριτική ικανότητα στα πεδία χρόνου και συχνότητας (time - frequency resolution). Δημιουργείται, δηλαδή, μια απεικόνιση δύο διαστάσεων (σχ. 3-5) που εκτιμά τα φασματικά χαρακτηριστικά του σήματος σαν συνάρτηση του χρόνου [191]-[193]:

$$STFT(\tau, f) = X_w(\tau, f) = \int x(t) \cdot g_w^*(t - \tau) \cdot e^{-2j\pi ft} dt \quad (3-11)$$

όπου $x(t)$ είναι το μη στάσιμο σήμα προς επεξεργασία, $g_w(t)$ είναι το παράθυρο (φίλτρο) επιλογής, ενώ το ολοκλήρωμα στο δεξιό μέρος της (3-11) αντιστοιχεί στο μετασχηματισμό Fourier του επιλεγμένου - «παραθυροποιημένου» - σήματος [191], [195]. Τα παραπάνω απεικονίζονται γραφικά στο σχ. 3-5, όπου μπορεί κανείς να επιλέξει όλες τις φασματικές συνιστώσες γύρω από μια χρονική στιγμή t (« t -κεντραρισμένο παράθυρο χρόνου»), ή αντίστοιχα να παρατηρήσει τις συνιστώσες γύρω από μια κεντρική συχνότητα f (« f -κεντραρισμένο φίλτρο») καθ' όλο το μήκος του άξονα χρόνου. Έτσι, ο STFT μπορεί να

θεωρηθεί ως μία διαμορφούμενη σειρά φίλτρων, η οποία δημιουργεί το γραμμικό πλέγμα ανάλυσης χρόνου - συχνότητας του σχήματος 3-5.



α.

β.

Σχήμα 3-5. α. Διοδιάστατο πλέγμα (grid) ανάλυσης που προκύπτει από την εφαρμογή του βραχυπρόθεσμου μετασχηματισμού Fourier, STFT, β. παραδείγματα διακριτικότητας χρόνου-συχνότητας (β).

Αν $G_w(f)$ είναι ο μετασχηματισμός Fourier της συνάρτησης του παραθύρου $g_w(t)$, τότε η φασματική διακριτικότητα Δf του φίλτρου και η αντίστοιχη διακριτικότητα Δt των παραθύρων στο πεδίο του χρόνου δίνονται από την ακόλουθη σχέση:

$$\Delta f = \sqrt{\frac{\int f^2 \cdot |G_w(f)|^2 df}{\int |G_w(f)|^2 df}}, \quad \Delta t = \sqrt{\frac{\int t^2 \cdot |g_w(t)|^2 dt}{\int |g_w(t)|^2 dt}} \quad (3-12)$$

όπου N είναι το μήκος του χρονικού παραθύρου, ενώ τα διαστήματα δειγματοληψίας df και dt προκύπτουν από τη συχνότητα δειγματοληψίας f_s ($df=f_s/N$, $dt=1/f_s$). Αυτό σημαίνει ότι δύο παλμοί που απέχουν χρονικά περισσότερο από Δt , μπορούν να διακριθούν με χρήση του STFT. Εντελώς ανάλογα, δύο φασματικές συνιστώσες στην ίδια χρονική στιγμή είναι διακριτές, αν οι συχνότητές τους απέχουν περισσότερο από Δf . Η διακριτικότητα που επιτυγχάνεται με εφαρμογή του STFT δεν μπορεί να είναι όσο μεγάλη θέλουμε και στα δύο πεδία. Μπορούμε να θυσιάσουμε τη διακριτικότητα στο πεδίο του χρόνου ώστε να εξασφαλίσουμε καλύτερη συχνотική ανάλυση (αριστερό διάγραμμα του σχ. 3-5β), και αντίστροφα να αυξήσουμε τη χρονική διακριτικότητα σε βάρος της φασματικής (δεξιό διάγραμμα του σχ. 3-5β). Σε καμιά περίπτωση όμως δεν μπορούμε να πετύχουμε υψηλή ανάλυση και στα δύο πεδία. Η παραπάνω διατύπωση αποτελεί την αρχή της αβεβαιότητας (αλλιώς λέγεται και ανισότητα του Heisenberg), σύμφωνα με την οποία το γινόμενο της χρονικής και συχνотικής διακριτικότητας παρουσιάζει ένα κάτω φράγμα:

$$\Delta t \cdot \Delta f \geq \frac{1}{4\pi} \quad (3-13)$$

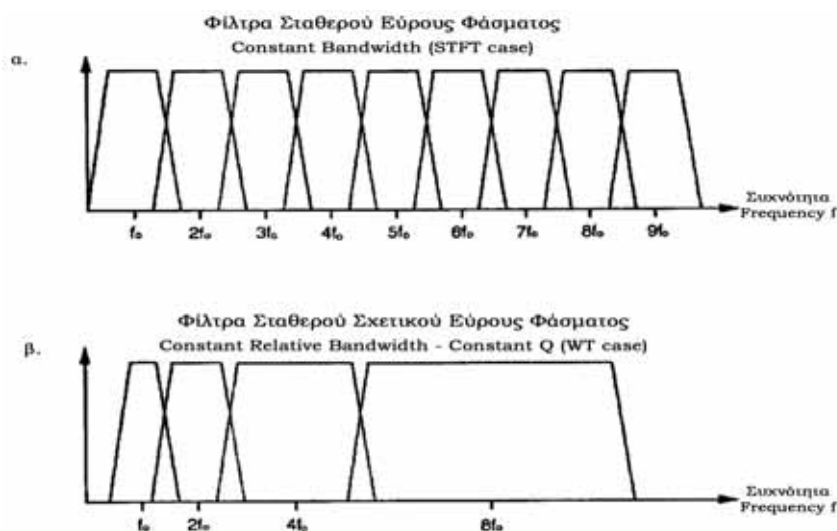
Από τη στιγμή που επιλεγούν οι παράμετροι f_s , N , αυτόματα ορίζονται και οι τιμές διακριτικότητας χρόνου (Δt) - συχνότητας (Δf), οι οποίες παραμένουν πλέον σταθερές και μετατοπίζονται γραμμικά κατά μήκος των δύο αξόνων. Αυτό σημαίνει ότι, ένα σήμα που αποτελείται από μικρές απότομες χρονικές μεταβολές (bursts) και συνδυασμό μακρόχρονων στάσιμων συνιστωσών μπορεί να αναλυθεί με πολύ καλή ανάλυση, είτε στο πεδίο του χρόνου, είτε της συχνότητας. Όχι όμως συγχρόνως και στα δύο πεδία [97], [191], [195]-[197].

Η αντιμετώπιση της προηγούμενης αδυναμίας, έγινε με την εισαγωγή της ανάλυσης σταθερού σχετικού εύρους φάσματος. Η ιδέα που κρύβεται πίσω από αυτή τη νέα προσέγγιση είναι απλή: τα Δt και Δf δεν είναι πλέον σταθερά αλλά μεταβάλλονται κατά

μήκος των δύο αξόνων. Έτσι, αν θεωρήσουμε την ανάλυση μέσα από μια σειρά ζωνοδιαβατών φίλτρων, διαπιστώνουμε ότι συμφέρει η διακριτικότητα στο πεδίο του χρόνου Δt να αυξάνει με την κεντρική συχνότητα f_c του εκάστοτε φίλτρου. Μπορούμε, λοιπόν, να επιβάλλουμε μια γραμμική σχέση ανάμεσα στα Δf και f_c έτσι που το εύρος Δf κάθε φίλτρου να είναι ανάλογο με την κεντρική του συχνότητα f_c :

$$\frac{\Delta f}{f_c} = const \quad (3-14)$$

όπου $const$ είναι σταθερά. Τα παραπάνω φίλτρα χαρακτηρίζονται από σταθερό σχετικό εύρος (= λόγος εύρους φάσματος προς κεντρική συχνότητα) και για το λόγο αυτό η ανάλυση αυτού του είδους χαρακτηρίζεται "constant-Q analysis". Η παράμετρος Q ισούται με τον ακριβώς αντίστροφο λόγο, δηλαδή της κεντρικής συχνότητας του φίλτρου f_c προς το εύρος φάσματος Δf ή BW ($Q = f_c / BW$). Αυτό σημαίνει ότι (σε αντίθεση με τον STFT) οι συχνотικές αποκρίσεις των φίλτρων ανάλυσης παρουσιάζονται σε λογαριθμική (οκταβική) και όχι γραμμική κλίμακα (σχ. 3-6). Το νέο αυτό πλέγμα ανάλυσης χρόνου - συχνότητας υιοθετείται και στην περίπτωση του μετασχηματισμού κυματιδίων (Wavelet Transform - WT).



Σχήμα 3-6. Τράπεζες φίλτρων ανάλυσης χρόνου - συχνότητας: α) φίλτρα σταθερού εύρους (γραμμική συχνотική ανάλυση, περίπτωση STFT) και β) φίλτρα σταθερού σχετικού εύρους (λογαριθμική συχνотική ανάλυση, περίπτωση οκταβικής constant-Q ανάλυσης / WT).

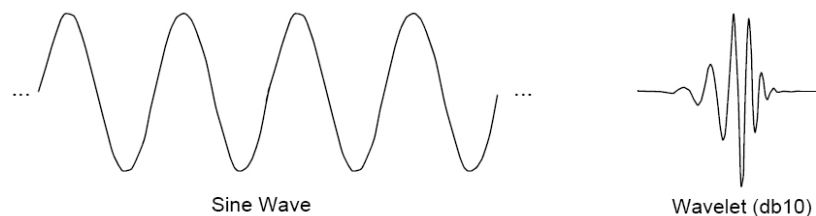
Κατά την ανάλυση σταθερού σχετικού εύρους εξακολουθεί και πάλι να ισχύει η αρχή της αβεβαιότητας (σχέση 3-13), μόνο που αυτή τη φορά μπορούμε να εξασφαλίζουμε πολύ μικρές τιμές χρονικής διακριτικότητας Δt στις υψηλές συχνότητες, κι αντίστοιχα μικρές τιμές φασματικής διακριτικότητας Δf στις χαμηλές συχνότητες. Αυτού του είδους η ανάλυση δουλεύει καλύτερα σε περιπτώσεις σημάτων που αποτελούνται από συνδυασμό συνιστωσών υψηλών συχνοτήτων με μικρές χρονικές διάρκειες και χαμηλών συχνοτήτων με μεγάλες χρονικές διάρκειες, δηλαδή των σημάτων που συναντώνται στην πράξη [191]-[198].

Ένα χαρακτηριστικό παράδειγμα του παραπάνω «φασματικού πλέγματος», αποτελεί η οκταβική και η τρίτο-οκταβική ανάλυση, που χρησιμοποιούνται ευρέως σε εφαρμογές ακουστικής. Οι μέθοδοι αυτές, αν και άρχισαν να χρησιμοποιούνται πολύ πριν από τη θεμελίωση της θεωρίας κυματιδίων, συναντούν πολλές από τις βασικές αρχές της τελευταίας. Γι' αυτό και μπορούν να θεωρηθούν ως προπομποί της ανάλυσης κυματιδίων. Το κύριο πλεονέκτημά τους εντοπίζεται στη λογαριθμική κλίμακα φασματικής ανάλυσης (οκτάβες),

που παρουσιάζει αρκετές ομοιότητες με τα χαρακτηριστικά της ανθρώπινης ακοής. Εκμεταλλευόμαστε δηλαδή το γεγονός ότι το ίδιο το αυτί αντιλαμβάνεται τους ήχους με «οκταβικό» τρόπο, στοιχείο που αποτελεί ένα πρόσθετο πλεονέκτημα των κυματιδιακών και οκταβικών προσεγγίσεων [108]-[109], [191], [199]. Όσον αφορά το τελευταίο ζήτημα, σχετική με τις αντιληπτικές ιδιότητες της ανθρώπινης ακοής είναι η θεωρία των κρίσιμων ζωνών (“critical bands” ή “bark scales”, όπως αναφέρονται διεθνώς) [188], [201]-[205]. Σύμφωνα με την παραπάνω θεωρία, η συχνοτική διακριτικότητα της ανθρώπινης ακοής είναι γραμμική μέχρι τα 500Hz, παρουσιάζοντας σταθερό εύρος φάσματος ίσο με 100Hz ($\Delta f = BW = 100\text{Hz}$, για $f < 500\text{Hz}$), ενώ σε μεγαλύτερες συχνότητες παραμένει σταθερό το σχετικό εύρος φάσματος, το οποίο διαμορφώνεται στο 20% της κεντρικής συχνότητας f_c ($\Delta f = BW = 20\%f_c$, για $f > 500\text{Hz}$). Το σχήμα αυτό ημι-γραμμικής / ημι-λογαριθμικής ανάλυσης παρουσιάζει πολλές ομοιότητες με τα χαρακτηριστικά του θορύβου ομοιόμορφης απόκρουσης [188], κι έχει εφαρμοστεί επιτυχώς σε πολλές τεχνικές συμπίεσης και αποθορυβοποίησης ήχου [181][200]-[205].

B. Μετασχηματισμός Κυματιδίων (Wavelets Transform -WT)

Ο μετασχηματισμός κυματιδίων (Wavelet Transform -WT) αναπτύχθηκε σε μια προσπάθεια να διευθετηθούν τα ζητήματα βέλτιστης διακριτικότητας χρόνου - συχνότητας, με ένα πιο «κομψό» και αποδοτικό τρόπο. Η κυματιδιακή επεξεργασία διατηρεί τα λογαριθμικά χαρακτηριστικά συχνοτικής διακριτικότητας, ενώ παράλληλα προσαρμόζει τη χρονική διακριτικότητα Δt , επιβάλλοντας μικρότερους ρυθμούς δειγματοληψίας σε χαμηλότερες συχνότητες (πολυδιακριτική ανάλυση, multi-rate scheme). Με τον τρόπο αυτό επιτυγχάνεται διατήρηση του αρχικού μεγέθους των δεδομένων, ευκολότερη υλοποίηση των σχετικών εφαρμογών και μικρότερο υπολογιστικό κόστος. Μια άλλη βασική διαφορά σε σχέση με το μετασχηματισμό Fourier (Fourier Transform -FT), αφορά στη χρήση διαφορετικών συναρτήσεων «βάσης» για τον υπολογισμό των φασματικών συνιστωσών. Όπως στην ανάλυση Fourier κάθε σήμα αναλύεται σε επιμέρους ημιτονικές συνιστώσες διαφορετικής συχνότητας, κατά αντίστοιχο τρόπο αναλύεται και σε σειρά διαφορετικών κυματιδίων. Τα τελευταία αποτελούν μετατοπίσεις και κλιμακώσεις μιας συνάρτησης ψ , που ονομάζεται μητρική συνάρτηση ή μητρικό κυματίδιο (mother wavelet). Τα κυματίδια έχουν περιορισμένη διάρκεια και μηδενική μέση τιμή (σχ. 3-7), χαρακτηριστικά που είναι συνηθισμένα στα φυσικά σήματα του «πραγματικού κόσμου», σε αντίθεση με τις ιδεατές περιοδικές κυματομορφές των ημιτονικών σημάτων [96]-[97], [191]-[198].



Σχήμα 3-7. Χαρακτηριστικές κυματομορφές **α.** συνημιτόνου, και **β.** κυματιδίου.

Το άθροισμα του εσωτερικού γινομένου (inner product) ενός σήματος $x(t)$ με τις αντίστοιχες συναρτήσεις $\psi()$, αποτελεί το συνεχή μετασχηματισμό κυματιδίων (Continuous Wavelet Transform - CWT). Αν, λοιπόν, συμβολίσουμε με $L^2(\mathcal{R})$ το χώρο των τετραγωνικά αθροιζόμενων μονοδιάστατων συναρτήσεων, δηλαδή των συναρτήσεων των οποίων το τετράγωνο του μέτρου τους είναι πεπερασμένος αριθμός, τότε, σε ένα χώρο Hilbert, ισχύουν οι ακόλουθες εξισώσεις [97], [191]-[198]:

$$\psi_{u,\tau}(t) = \frac{1}{\sqrt{|u|}} \cdot \psi\left(\frac{t-\tau}{u}\right), \quad u, \tau \in R, u \neq 0 \quad (3-15)$$

$$CWT_x(u, \tau) = W_{u,\tau} = \langle x, \psi_{u,\tau} \rangle = \int x(t) \cdot \psi_{u,\tau}^*(t) dt \quad (3-16)$$

όπου οι παράμετροι u, τ αποτελούν τους συντελεστές κλιμάκωσης και μετατόπισης αντίστοιχα, $\psi(t)$ είναι η μητρική συνάρτηση κυματιδίου, $\psi_{u,\tau}(t)$ είναι η οικογένεια όλων των παραγόμενων κυματιδίων, ενώ το σύμβολο * δηλώνει το συζυγή μιγαδικό. Τα επιμέρους γινόμενα του παραπάνω αθροίσματος είναι γνωστά ως συντελεστές κυματιδίων (wavelet coefficients), κατά αντιστοιχία με τους συντελεστές Fourier, και δίνονται από τη σχέση:

$$C_{u,\tau} = x(t) \cdot \psi_{u,\tau}^*(t) \quad (3-17)$$

Ο συνεχής μετασχηματισμός κυματιδίων είναι αντιστρέψιμος στην περίπτωση που το κυματίδιο ψ ικανοποιεί τη συνθήκη αποδεκτότητας (admissibility condition):

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{\omega} d\omega < \infty, \quad (\text{συνθήκη αποδεκτότητας}) \quad (3-18)$$

$$x(t) = C_\psi^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} CWT_x(u, \tau) \cdot \psi_{u,\tau}(t) \cdot \frac{du d\tau}{u^2}, \quad (\text{αντίστροφος μετασχηματισμός}) \quad (3-19)$$

όπου $\Psi(\omega)$ είναι ο μετασχηματισμός Fourier της $\psi(t)$. Η συνθήκη αποδεκτότητας αποτελεί απαραίτητη προϋπόθεση ώστε μία συνάρτηση $\psi(t)$ να μπορεί να χρησιμοποιηθεί ως μητρικό κυματίδιο. Από την ίδια συνθήκη προκύπτει ο όρος $\Psi(0)$ πρέπει να μηδενίζεται, προκειμένου να μην απειρίζεται το ολοκλήρωμα C_ψ , δηλαδή το κυματίδιο έχει συχνότητες μόνο διάφορες του μηδενός. Στην ιδιότητα αυτή οφείλεται και η ονομασία κυματίδιο, που σημαίνει μικρό κύμα [195]-[197].

Εκτός από το συνεχή μετασχηματισμό CWT , ορίζεται και ο διακριτός μετασχηματισμός κυματιδίων (Discrete Wavelet Transform -DWT), όπου οι παράμετροι κλιμάκωσης και μετατόπισης δεν μεταβάλλονται συνεχώς αλλά παίρνουν διακριτές τιμές σε ένα δυαδικό πλέγμα της μορφής $u=2^j, \tau=k \cdot 2^j, k, j \in \mathbb{Z}$:

$$\psi_{j,k}(t) = 2^{\frac{j}{2}} \cdot \psi(2^j t - k), \quad j, k \in \mathbb{Z} \quad (3-20)$$

Με αντικατάσταση της (3-20) στην (3-16) προκύπτει η έκφραση του DWT:

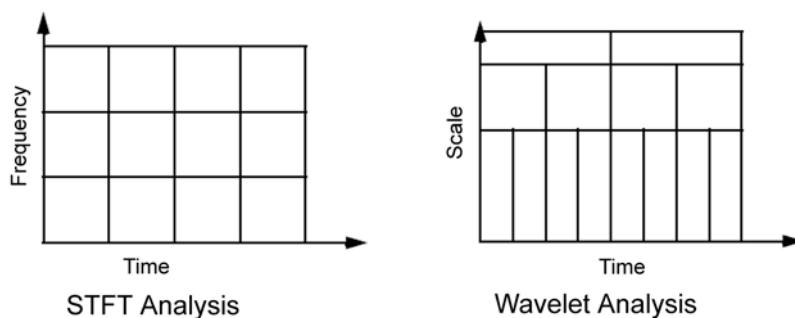
$$d_{j,k} = \langle x, \psi_{j,k} \rangle = \int x(t) \cdot \psi_{j,k}^*(t) dt, \quad j, k \in \mathbb{Z} \quad (3-21)$$

ενώ ο αντίστροφος μετασχηματισμός (Inverse DWT - IDWT), διαμορφώνεται ως ακολούθως:

$$x(t) = \sum_{j,k \in \mathbb{Z}} d_{j,k} \cdot \psi_{j,k}(t), \quad j, k \in \mathbb{Z} \quad (3-22)$$

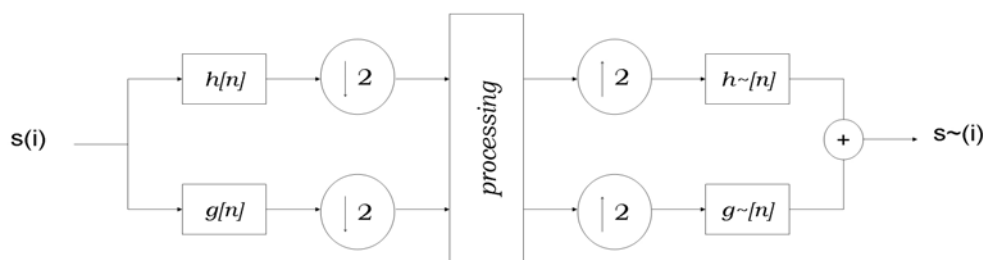
Στο σχ. 3-8 αποδίδεται γραφικά το πλέγμα ανάλυσης χρόνου - κλιμάκωσης (ή χρόνου - συχνότητας) που προκύπτει από την εφαρμογή του DWT , ενώ γίνεται και σύγκριση με την ανάλυση χρόνου - συχνότητας του μετασχηματισμού $STFT$. Δεδομένου ότι η κλιμάκωση μεταβάλλεται αντιστρόφως ανάλογα με τη συχνότητα, παρατηρούμε ότι η κυματιδιακή

ανάλυση ενσωματώνει τα χαρακτηριστικά των οκταβικών σχημάτων. Οι συναρτήσεις $\psi_{j,k}(t)$ αποτελούν βάση του χώρου $L^2(\mathcal{R})$, δηλαδή είναι ορθογώνιες μεταξύ τους και κάθε συνάρτηση $x \in L^2(\mathcal{R})$, μπορεί να γραφτεί ως γραμμικός τους συνδυασμός. Μια διαφορετική μαθηματική ερμηνεία των παραπάνω είναι ότι, κατά την κυματιδιακή ανάλυση το σήμα αναλύεται διαδοχικά σε συντελεστές «προσεγγίσεων» (approximations) και «λεπτομερειών» (details), όπως συμβαίνει με τη διαδοχική του προβολή σε «φωλιασμένους» υποχώρους [97], [191]-[198].



Σχήμα 3-8. Πλέγματα ανάλυσης χρόνου-συχνότητας (αριστερά) και χρόνου-κλιμάκωσης (δεξιά), για την περίπτωση των STFT και WT, αντίστοιχα ([192], p. 1-11).

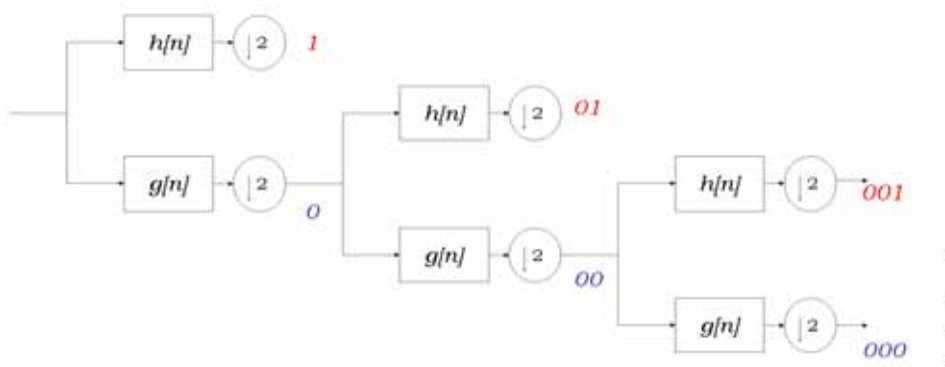
Στην πράξη, κατά την εφαρμογή των παραπάνω σχέσεων κυματιδιακής ανάλυσης, οι αναλογικές εκφράσεις των συναρτήσεων $x(t)$ και $\psi(t)$ αντικαθίστανται από τις αντίστοιχες ψηφιακές $x(i)$, $\psi(i)$. Μια τεχνική που επιτρέπει το γρήγορο υπολογισμό των κυματιδιακών συντελεστών, είναι η χρήση ζεύγους κατοπτρικών φίλτρων τέλει ανακατασκευής (“two-channel perfect reconstruction filter banks” ή “Quadrature Mirror Filter-banks” -QMF) [192]-[193], [196]. Πρόκειται για διαδικασία πολλαπλών «φιλτραρισμάτων» (πολυεπίπεδη ανάλυση), όπου σε κάθε κόμβο ανάλυσης (σχ. 3-9) το σήμα χωρίζεται φασματικά σε δύο συμμετρικές υπο-περιοχές (band splitting). Αν $h(n)$ και $g(n)$ είναι οι κρουστικές αποκρίσεις του υψηλερατού (High Pass Filter -HPF) και χαμηλοπερατού (Low Pass Filter -LPF) φίλτρου ανάλυσης, αντίστοιχα, τότε η επιλογή κατάλληλων φίλτρων $h\sim(n)$, $g\sim(n)$ κατά τη φάση της σύνθεσης, οδηγεί στην τέλεια ανακατασκευή του αρχικού σήματος [97], [191]-[198]. Η παραπάνω τεχνική συχνά αναφέρεται ως «ταχύς μετασχηματισμός κυματιδίων» (Fast Wavelet Transform -FWT) [179], [191]-[196], [206].



Σχήμα 3-9. Κατοπτρικά φίλτρα τέλει ανακατασκευής (QMF): $s(i)$ και $s\sim(i)$ είναι τα σήματα εισόδου - εξόδου, $h[n] - g[n]$ και $h\sim[n] - g\sim[n]$ τα φίλτρα ανάλυσης και σύνθεσης (HPF και LPF αντίστοιχα).

Στην πράξη η επιλογή του φίλτρου δεν καθορίζει μόνο το αν θα είναι εφικτή η τέλεια ανακατασκευή, αλλά και την μορφή του κυματιδίου ανάλυσης. Έτσι, η επαναλαμβανόμενη διαδικασία φασματικής διαίρεσης (band splitting), που εφαρμόζεται κάθε φορά στη χαμηλότερη από τις δύο φασματικές ζώνες (low frequency band), είναι ισοδύναμη με τον διακριτό μετασχηματισμό DWT (σχ. 3-10). Οι κυματιδιακές συναρτήσεις έχουν τη μορφή της εξίσωσης (3-20), ενώ το σήμα αναλύεται διαδοχικά σε συντελεστές «προσέγγισης» και

«λεπτομερειών» [97], [191]-[198]. Για το συμβολισμό των κόμβων ανάλυσης και των αντίστοιχων διαμορφούμενων κυματιδιακών ζωνών, χρησιμοποιούνται συνδυασμοί των ψηφίων “0” και “1”. Έτσι, ο συμβολισμός “0” αντιστοιχεί στη χαμηλότερη φασματική περιοχή (low frequency half-band splitting) κι επομένως περιγράφει τους συντελεστές προσέγγισης (approximation coefficients), ενώ το “1” αντιστοιχεί στην υψηλότερη φασματική περιοχή (high frequency half-band portion), δηλαδή στους συντελεστές λεπτομερειών (detail coefficients). Επιπρόσθετα, σε κάθε κόμβο ανάλυσης λαμβάνει χώρα υποδειγματοληψία κατά ένα παράγοντα ίσο με 2, με αποτέλεσμα, ο συνολικός αριθμός των δειγμάτων να διατηρείται ίσος με το μήκος της αρχικής ακολουθίας. Πρόκειται δηλαδή για μια διαδικασία πολύ-διακριτικής ανάλυσης (multi-resolution analysis), όπου όλα τα παραγόμενα φαινόμενα φασματικής αλλοίωσης εξαλείφονται κατά τη διαδικασία της ανασύνθεσης, χάρη στα χαρακτηριστικά των κατοπτρικών φίλτρων τέλει ανακατασκευής [97], [179], [191]-[198].



Σχήμα 3-10. Υλοποίηση του Διακριτού Μετασχηματισμού Κυματιδίων (DWT) με κατοπτρικά φίλτρα τέλει ανακατασκευής (QMFs).

Σύμφωνα με το σχήμα 3-10 κάθε σήμα s μπορεί να εκφραστεί ως άθροισμα όλων των «λεπτομερειών» (D_j) και της τελευταίας «προσέγγισης» (A_j) [192]:

$$s = A_j + \sum_{j \leq J} D_j, \quad j = 1, 2, \dots, J \quad (3-23)$$

Η παραπάνω σχέση προκύπτει από την επαναληπτική εφαρμογή της φασματικής διαίρεσης σε κάθε επίπεδο ανάλυσης:

$$A_{j-1} = A_j + D_j, \quad j = 1, 2, \dots, J \quad (3-24)$$

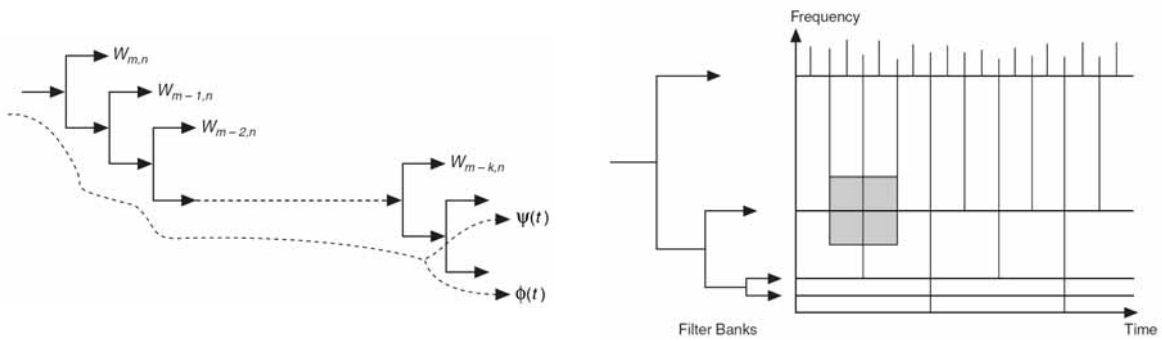
$$s \triangleq A_0 = A_1 + D_1$$

όπου η ακολουθία προσεγγίσεων του «μηδενικού επιπέδου ανάλυσης» A_0 είναι ταυτόσημη με την αρχική ακολουθία s . Αν συμβολίσουμε με $d_{j,k}$ τους συντελεστές των λεπτομερειών και $a_{j,k}$ τους συντελεστές προσέγγισης στο τελευταίο επίπεδο ανάλυσης J , τότε η εξ. (3-24) γίνεται:

$$s = \sum_{k=-\infty}^{\infty} a_{J,k} \cdot \phi_k + \sum_{k=-\infty}^{\infty} \sum_{j=1}^J d_{j,k} \cdot \psi_{j,k} \quad (3-25)$$

Η μητρική κυματιδιακή συνάρτηση ψ προκύπτει από τη συνάρτηση μεταφοράς h του υπερπαρατό φίλτρου, ενώ η συνάρτηση «κλιμάκωσης» ϕ καθορίζεται από το χαμηλοπερατό

φίλτρο g . Το σχ. 3-11 δείχνει τη σχέση μεταξύ των κατοπτρικών τραπεζών φίλτρων με το σχήμα χρόνου-συχνότητας της κυματιδιακής ανάλυσης [193].

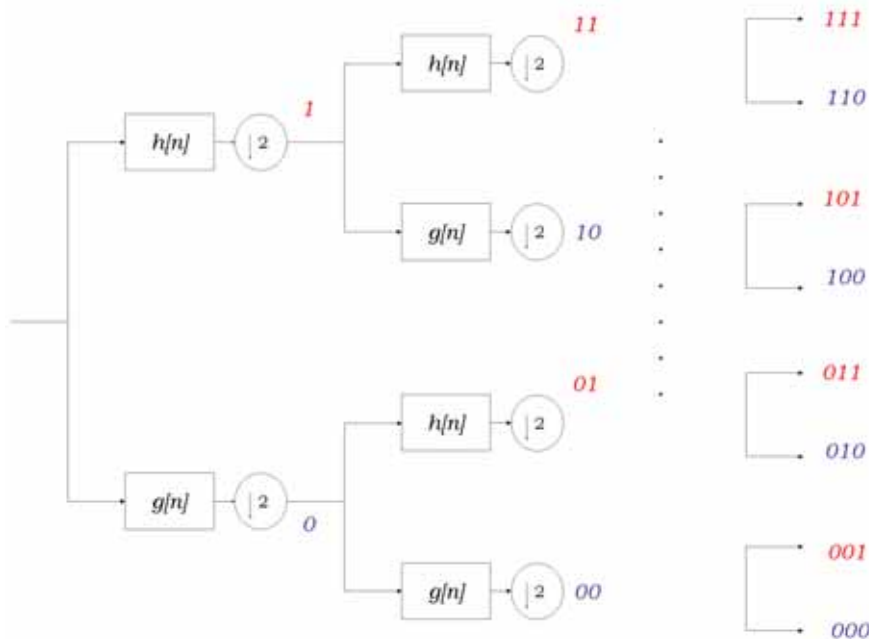


Σχήμα 3-11. Σχέση μεταξύ των τραπεζών φίλτρων με το μετασχηματισμό κυματιδίων και το σχήμα χρόνου - συχνότητας της κυματιδιακής ανάλυσης ([193], pp. 10-3, 10-4).

Όπως και στην περίπτωση του μετασχηματισμού Fourier, έτσι και στους μετασχηματισμούς *CWT* και *DWT* ισχύει η ισότητα του Parseval, που εκφράζει την ισότητα της ενέργειας του αρχικού σήματος με το ενεργειακό άθροισμα όλων των κυματιδιακών συντελεστών. Χρησιμοποιώντας τις ακολουθίες a_j και d_j ($j=1, 2, \dots, J$), και δεδομένου ότι $a_0 \equiv s$ (σχέση 3-24β), η διατήρηση της ενέργειας εκφράζεται από την ακόλουθη εξίσωση [207]:

$$\|a_0\|^2 = \|a_J\|^2 + \sum_{j=1}^J \|d_j\|^2 \tag{3-26}$$

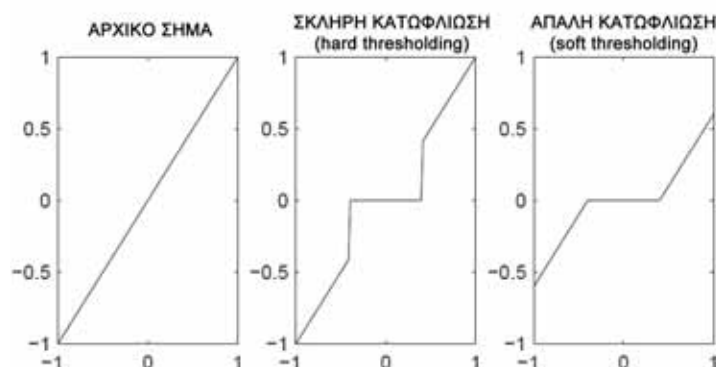
Εκτός από τη δομή του σχ. 3-10, συχνά χρησιμοποιείται και η πιο σύνθετη τοπολογία κυματιδιακής ανάλυσης του σχ. 3-12. Στην περίπτωση αυτή, και οι δύο φασματικές περιοχές (low and high frequency bands) μπορούν να αναλυθούν περαιτέρω με χρήση των φίλτρων φασματικού διαχωρισμού (band splitting). Η συγκεκριμένη τοπολογία είναι γνωστή ως ανάλυση κυματιδιακών πακέτων (Wavelet Packets -WP) [97], [191]-[198], [208]-[210].



Σχήμα 3-12. Υλοποίηση της ανάλυσης Κυματιδιακών Πακέτων (WP) με κατοπτρικά φίλτρα τέλει ανακατασκευής (QMFs).

Γ. Τεχνικές καταστολής θορύβου με κατωφλίωση των κυματιδιακών συντελεστών

Αν χρησιμοποιήσουμε τους μετασχηματισμούς DWT και WP στη θέση του τελεστή $T[\]$ στο σχ. 3-1, το πρόβλημα της αποθορυβοποίησης ανάγεται στην εύρεση των κατάλληλων τιμών κατωφλίωσης (threshold - th) για την επεξεργασία των κυματιδιακών συντελεστών. Σε πολλές περιπτώσεις τα επίπεδα κατωφλίωσης ορίζονται εμπειρικά (empirical - manual), ενώ παράλληλα έχουν αναπτυχθεί και αλγόριθμοι αυτόματου υπολογισμού [197], [211]. Δύο από τις πιο διαδεδομένες τεχνικές αυτόματου εντοπισμού, είναι η μέθοδος SURE (Stein's Unbiased Risk Estimate) και η "SQTWOLOG", που ονομάστηκε έτσι από την έκφραση υπολογισμού των κυματιδιακών κατωφλίων " $\sqrt{2 \cdot \log(\text{Length}\{x\{i\}\})}$ " [192]-[193], [197], [211]. Και οι δύο προσεγγίσεις αποτελούν κλασικά εργαλεία κυματιδιακής καταστολής θορύβου [192]-[193], και αξιοποιήθηκαν ως μέθοδοι αναφοράς κατά την αξιολόγηση της νέας μεθόδου [179]. Σημαντική είναι και η επιλογή της κατάλληλης, κάθε φορά, τοπολογίας κυματιδιακής ανάλυσης - σύνθεσης. Πολλές ερευνητικές μελέτες έχουν πραγματοποιηθεί για τη διευθέτηση των δύο προηγούμενων ζητημάτων, συμπεριλαμβανομένων και εφαρμογών επεξεργασίας βιοακουστικών σημάτων. Έτσι έχουν αναπτυχθεί τεχνικές για την ανάδειξη των βέλτιστων τοπολογιών κυματιδιακής προβολής - ανακατασκευής, την ενσωμάτωση αντιληπτικών - ψυχοακουστικών κριτηρίων, την επιλογή των κατάλληλων μητρικών κυματιδίων και την εφαρμογή επαναληπτικών (iterative) δομών επεξεργασίας [59], [83], [133], [203]-[205], [208]-[210].



Σχήμα 3-13. Γραφική αναπαράσταση της επεξεργασίας απαλής κατωφλίωσης (soft thresholding) και σκληρής κατωφλίωσης (hard thresholding), αντίστοιχα ([192], p. 6-101).

Γενικά υπάρχουν δύο βασικές «στρατηγικές» κατωφλίωσης κυματιδιακών συντελεστών (σχ. 3-13): η «σκληρή κατωφλίωση» (hard thresholding) και η «απαλή κατωφλίωση» (soft thresholding). Η πρώτη αφορά στην απόρριψη (μηδενισμό) των συντελεστών που βρίσκονται κάτω από το επίπεδο κατωφλίωσης και τη διατήρηση εκείνων που το ξεπερνούν, ενώ στη δεύτερη γίνεται συμπίεση των κυματιδιακών συντελεστών (wavelet shrinkage) για τιμές μεγαλύτερες του κατωφλίου [192]-[193], [211]. Τα αποτελέσματα των δύο τελεστών ($x_{hard-th}(i)$ και $x_{soft-th}(i)$ αντίστοιχα) με δεδομένα το σήμα εισόδου $x(i)$ και την τιμή σύγκρισης th δίνονται στην εξίσωση (3-27):

$$x_{hard-th}(i) = \begin{cases} x(i), & \text{αν } |x(i)| > th \\ 0, & \text{διαφορετικά} \end{cases} \quad (3-27)$$

$$x_{soft-th}(i) = \begin{cases} \text{sign}(x(i))(|x(i)| - th), & \text{αν } |x(i)| > th \\ 0, & \text{διαφορετικά} \end{cases}$$

3.2 Καταστολή θορύβου σε καταγραφές κοιλιακών ήχων

Εκτός από το γενικό πλαίσιο αντιμετώπισης των προβλημάτων θορύβου φόντου ευρείας ζώνης, υπάρχουν αρκετές ιδιαιτερότητες που σχετίζονται με την ανάλυση των κοιλιακών ήχων και τις αντίστοιχες απαιτήσεις αποθορυβοποίησης. Στις επόμενες παραγράφους, που στην ουσία αποτελούν το «πειραματικό μέρος» του «τρέχοντος κεφαλαίου καταστολής θορύβου», γίνεται ανάλυση όλων των σχετικών ζητημάτων, με έμφαση στα χαρακτηριστικά σχεδιασμού της νέας προτεινόμενης μεθόδου. Επίσης συμπεριλαμβάνονται διαδικασίες αξιολόγησης και σύγκρισης με τις προγενέστερες ερευνητικές προσεγγίσεις.

3.2.1 Παρουσία θορύβου & προβλήματα ανάλυσης κοιλιακών ήχων

Μια από τις βασικότερες δυσκολίες στις προσεγγίσεις ακουστικής ανάλυσης της γαστρεντερικής δραστηριότητας, είναι η αδύναμη φύση των παραγόμενων βιοακουστικών φαινομένων, σε συνδυασμό με τα ιδιαίτερα χαρακτηριστικά του μέσου διάδοσης. Έτσι, η ικανοποιητική καταγραφή των γαστρεντερικών δονήσεων στην επιφάνεια των κοιλιακών τοιχωμάτων απαιτεί τη χρήση αισθητήρων μεγάλης ευαισθησίας και κυκλωματικών βαθμίδων υψηλής απολαβής, ενώ το ζήτημα της ασφάλειας των υποκειμένων παραμένει ιδιαίτερα κρίσιμο. Ένα αρνητικό συνεπακόλουθο των παραπάνω στοιχείων είναι και η έντονη παρουσία θορύβου [59], [179], [215]-[218] του οποίου οι γενεσιουργίες αιτίες εντοπίζονται στους γενικότερους παράγοντες που παρουσιάστηκαν στην παράγραφο 3.1.1 (θερμικός ηλεκτρονικός θόρυβος, θόρυβος κβάντισης, σήματα βόμβων, ακουστικός θόρυβος περιβάλλοντος, κ.ά.). Όπως διαπιστώθηκε από τις πειραματικές καταγραφές που πραγματοποιήθηκαν στο πλαίσιο της παρούσης διατριβής [179], [217] το πρόβλημα εξακολουθεί να υπάρχει σε αρκετά έντονο βαθμό, ακόμη και στην περίπτωση χρήσης μετρητικού εξοπλισμού υψηλών προδιαγραφών και κόστους (όπως συμβαίνει με το επιταχυνσιόμετρο τριών αξόνων B&K 4506).

Η επίδραση του θορύβου φόντου μπορεί να εντοπιστεί σε διάφορες εκφάνσεις της γαστρεντερικής ακουστικής ανάλυσης, με αρνητικές επιπτώσεις σε όλες τις ιατρικές προσεγγίσεις που παρουσιάστηκαν στο πρώτο κεφάλαιο. Καταρχήν, ο θόρυβος δυσχεραίνει την ακροαστική ανάλυση των κοιλιακών ήχων, η οποία καθίσταται προβληματική και κουραστική για το ιατρικό προσωπικό. Κατά δεύτερο λόγο, δημιουργεί προβλήματα εσφαλμένης ανίχνευσης κατά τη διαδικασία εντοπισμού των χρήσιμων βιοακουστικών συμβάντων. Οι παραπάνω παράγοντες μπορεί να οδηγήσουν σε σημαντικά υπολογιστικά λάθη, καθώς α) η ανάλυση των σημάτων δεν είναι δυνατόν να είναι συνεχώς επιτηρούμενη από το ιατρικό προσωπικό, ιδιαίτερα σε καταγραφές παρατεταμένου χρόνου, β) οι καμπύλες σύνοψης της γαστρεντερικής δραστηριότητας (summarization curves: time-energy & time-energy-frequency envelopes) εμπεριέχουν λάθη, καθώς συνεκτιμών ως σήμα και τις συνιστώσες του θορύβου, γ) ο υπολογισμός των χαρακτηριστικών παραμέτρων ανάλυσης (features) είναι επίσης επισφαλής, με αρνητικές επιπτώσεις στις εμπλεκόμενες τεχνικές αυτοματοποιημένης ανίχνευσης και ταξινόμησης των διαφόρων συμβάντων [179].

Επιπρόσθετα, σε προσεγγίσεις ακουστικής χαρτογράφησης της γαστρεντερικής δραστηριότητας [42], [87], είναι απαραίτητη η αξιοποίηση περισσότερων του ενός αισθητήρων λήψης, κάποιιοι από τους οποίους (οι πιο απομακρυσμένοι από το σημείο διέγερσης) θα έχουν ιδιαίτερα χαμηλές σηματοθορυβικές σχέσεις. Έτσι, η ανάγκη αποθορυβοποίησης γίνεται ακόμη πιο επιτακτική στις περιπτώσεις αυτές (περισσότερες λεπτομέρειες δίνονται στο έκτο κεφάλαιο). Σύμφωνα, λοιπόν, με όλες τις παραπάνω

αιτιάσεις, το ζήτημα καταστολής του προσθετικού θορύβου ευρείας ζώνης εξακολουθεί να είναι ερευνητικά επίκαιρο, σε όλες τις προσεγγίσεις διαγνωστικής αξιοποίησης των γαστρεντερικών ήχων. Αυτός είναι και ο λόγος που ένας σημαντικός αριθμός ερευνητικών εργασιών έχουν ως αντικείμενο την αντιμετώπιση του συγκεκριμένου προβλήματος.

3.2.2 Υπάρχουσες τεχνικές καταστολής θορύβου κοιλιακών ήχων

Τα τελευταία χρόνια έχουν πραγματοποιηθεί πολλές ερευνητικές εργασίες με αντικείμενο την αποθορυβοποίηση βιοακουστικών σημάτων και ειδικότερα κοιλιακών ήχων. Αν εξαιρέσουμε την εργασία των Mansy και Sandler [78], η οποία είχε ως αντικείμενο το διαχωρισμό των γαστρεντερικών ακουστικών φαινομένων από τους καρδιακούς ήχους σε μικρά πειραματόζωα (αρουραίους), οι υπόλοιπες τεχνικές εστιάζουν στο πρόβλημα του ακουστικού θορύβου βάθους. Εξάλλου, οι καρδιακοί ήχοι είναι σπάνιοι σε ακουστικές γαστρεντερικές καταγραφές ενηλίκων (εμφανίζονται κυρίως σε πειραματόζωα μικρού μεγέθους και σπανιότερα σε νήπια [59], [179]). Όσον αφορά στις υπόλοιπες εργασίες αποθορυβοποίησης κοιλιακών ήχων, αυτές συνήθως περιλαμβάνουν τεχνικές αυτόματης κατωφλίωσης, αξιοποιώντας κυματιδιακές μεθόδους σε συνδυασμό με στατιστικά κριτήρια [59], [86], [89], εργαλεία ανάλυσης πολυπλοκότητας κυματομορφών - “fractal dimension” [83], [84], καθώς και εναλλακτικές προσεγγίσεις στατιστικής ανάλυσης ανώτερης τάξης - “Higher Order Statistics -HOS” [81], [85].

Μία ιδιαίτερα «δημοφιλής» εργασία που αναπτύχθηκε για την αποθορυβοποίηση βιοϊατρικών σημάτων και εικόνων, είναι η επαναληπτική μέθοδος κυματιδιακής επεξεργασίας των Coifman και Wickerhauser [209]. Πρόκειται για μια διαδικασία διαδοχικής κυματιδιακής κατωφλίωσης, με επιλογή της βέλτιστης κάθε φορά βάσης (best basis selection). Γενικά, η βέλτιστη βάση κυματιδιακής ανάλυσης σχετίζεται τόσο με την επιλογή του κατάλληλου μητρικού κυματιδίου, όσο και με την αντίστοιχη τοπολογία ανάλυσης - σύνθεσης [192]-[193], [208]-[209]. Έτσι, η προσαρμοστική μέθοδος κυματιδιακής αποθορυβοποίησης (Adapted Waveform de-noising) των Coifman και Wickerhauser, περιλαμβάνει τα ακόλουθα βήματα [208]-[210]:

α) το σήμα υφίσταται κυματιδιακή ανάλυση, χρησιμοποιώντας τη βέλτιστη από τις διαθέσιμες βάσεις κυματιδιακής προβολής,

β) οι κυματιδιακοί συντελεστές υφίστανται κατωφλίωση, συνήθως με τη βοήθεια αλγορίθμων αυτόματου ορισμού των σχετικών επιπέδων απόρριψης,

γ) οι εναπομείναντες συντελεστές που αντιστοιχούν στο «καθαρό» σήμα (μη στάσιμο τμήμα) μετασχηματίζονται στο πεδίο του χρόνου και αποθηκεύονται για μελλοντική χρήση,

δ) οι απορριπτέοι συντελεστές (συνιστώσες θορύβου, στάσιμο τμήμα) μετασχηματίζονται στο πεδίο του χρόνου και κατόπιν οδηγούνται στην είσοδο (βήμα “α”) για ένα νέο γύρο επεξεργασίας,

ε) η διαδικασία επαναλαμβάνεται για όσο διάστημα η αποθορυβοποίηση έχει νόημα, δηλαδή για όσο οι «στάσιμες ακολουθίες» μεταξύ δύο διαδοχικών βημάτων διαφέρουν αισθητά (αυτό συνήθως καθορίζεται από μια μικρή ενεργειακή σταθερά σύγκρισης “ε”),

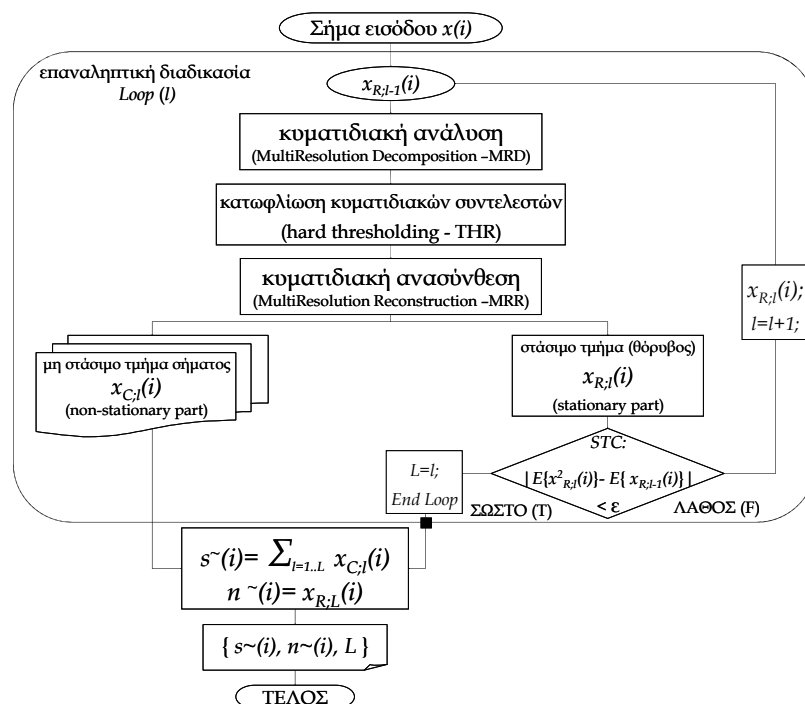
στ) όλες οι μη στάσιμες ακολουθίες, που προκύπτουν σε κάθε βήμα της επαναληπτικής διαδικασίας και αντιπροσωπεύουν χρήσιμη πληροφορία, στο τέλος αθροίζονται και δίνουν το εκτιμώμενο, καθαρό από θόρυβο, σήμα.

Έχοντας σαν βάση την παραπάνω μέθοδο, οι Χατζηλεοντιάδης και Πανάς [133] ανέπτυξαν την μέθοδο διαχωρισμού των στάσιμων και μη στάσιμων συνιστωσών, που αντιπροσωπεύουν «προσθετικό θόρυβο» και «καθαρό σήμα», αντίστοιχα. Ο αλγόριθμος αυτός ονομάστηκε *WTST-NST* (wavelet transform-based stationary-non stationary) και χρησιμοποιήθηκε αρχικά για την αποθορυβοποίηση πνευμονικών ήχων [133]. Η βασική διαφορά του, σε σχέση με την πρωτότυπη μέθοδο των Coifman και Wickerhauser, έγκειται στον τρόπο υπολογισμού των κατωφλίων αποθορυβοποίησης, και κυρίως στο γεγονός ότι δεν εμπλέκεται πλέον η επιλογή της βέλτιστης βάσης κυματιδιακής ανάλυσης (σχ. 3-14). Ειδικότερα, οι τιμές των κατωφλίων $THR_{k,l}$ σε κάθε δέσμη κυματιδιακής ανάλυσης k και για κάθε βήμα l της επαναληπτικής διαδικασίας, δίνονται από τη σχέση:

$$THR_{k,l} = F_{adj} \cdot \sigma_{k,l} \tag{3-28}$$

όπου ο πολλαπλασιαστικός παράγοντας F_{adj} χρησιμοποιείται για την ρύθμιση των κατωφλίων στα επιθυμητά επίπεδα, ενώ $\sigma_{k,l}$ είναι η τυπική απόκλιση των κυματιδιακών συντελεστών για τις αντίστοιχες τιμές k, l . Είναι προφανές ότι η παραπάνω μεθοδολογία μπορεί να εφαρμοστεί με οποιαδήποτε τοπολογία ανάλυσης - σύνθεσης, αλλά και με διάφορες συναρτήσεις μητρικού κυματιδίου. Στην αυθεντική υλοποίηση [133] χρησιμοποιήθηκε το μητρικό κυματίδιο Daubechies 4, ενώ παράλληλα επιλέχθηκε η μέθοδος «σκληρής κατωφλίωσης». Η ενεργειακή σταθερά “ ϵ ” που εμπλέκεται στο κριτήριο τερματισμού της επαναληπτικής διαδικασίας (Stopping Criterion - STC), επιλέχθηκε ίση με $\epsilon=10^{-5}$. Επίσης, αξιολογήθηκε η επίδραση του συντελεστή F_{adj} και αποφασίστηκε η επιλογή της τιμής $F_{adj}=3$ [133]. Όσον αφορά στο εμπλεκόμενο «δένδρο» κυματιδιακής προβολής, αποφασίστηκε η χρήση *DWT* με το βέλτιστο δυνατό αριθμό επιπέδων M , ο οποίος προκύπτει από το δυαδικό λογάριθμο του μήκους N του σήματος εισόδου [194]:

$$M = \log_2 (N) \tag{3-29}$$



Σχήμα 3-14. Απλοποιημένο μπλοκ διάγραμμα λειτουργίας του επαναληπτικού αλγορίθμου αποθορυβοποίησης “WTST-NST”, όπου $x_{R,l}(i)$, $x_{C,l}(i)$ το στάσιμο και μη στάσιμο σήμα, για κάθε δέσμη ανάλυσης k και κάθε βήμα επεξεργασίας l ($l=1,2,\dots,L$).

Ένα από τα πλεονεκτήματα της μεθόδου *WTST-NST* είναι η αυτοματοποιημένη διαδικασία αποθορυβοποίησης (αυτόματη εκτίμηση κατωφλίων), δυνατότητα ιδιαίτερα λειτουργική κατά τη χρησιμοποίησή της από ιατρικό προσωπικό. Στα θετικά στοιχεία συμπεριλαμβάνεται και η εύκολη ρύθμιση του αλγορίθμου, που αφορά στην επιλογή των παραμέτρων (F_{adj} , ϵ), του μητρικού κυματιδίου και του «βάθους» ανάλυσης (M). Στον αντίποδα, ένα μειονέκτημα είναι το μεγάλο υπολογιστικό κόστος που προκύπτει από την επαναληπτική διαδικασία, δηλαδή την πολλαπλή εφαρμογή (L επαναλήψεις) της κυματιδιακής ανάλυσης - σύνθεσης και των κανόνων κατωφλίωσης. Επίσης, η συγκεκριμένη μέθοδος έχει προταθεί κυρίως για επεξεργασία κρουστικών σημάτων [59], [83], [133], με αποτέλεσμα να εμφανίζονται διάφορα προβλήματα κατά την εφαρμογή της σε πιο ομαλές (sustained) κυματομορφές [179], [217]. Εκτός από τους αναπνευστικούς ήχους, η μέθοδος *WTST-NST* χρησιμοποιήθηκε σε εφαρμογές επεξεργασίας και ανάλυσης εντερικών ήχων [59], [71], [86], [88]-[89]. Μάλιστα, προτάθηκαν βελτιώσεις της αρχικής υλοποίησης, σχετικά με το βάθος ανάλυσης M αλλά και την αποφυγή των πολλαπλών εφαρμογών κυματιδιακής ανάλυσης - σύνθεσης, κατά την επαναληπτική διαδικασία [212].

Στο ίδιο πλαίσιο αναπτύχθηκε η μέθοδος *WT-FD* (Wavelet Transform - Fractal Dimension -based) [83]-[84], η οποία προτάθηκε για την επεξεργασία αποθορυβοποίησης των, κρουστικών, αναπνευστικών και εντερικών ήχων (Explosive Lung Sounds -ELS, Explosive Bowel Sounds - EBS). Πρόκειται για την ίδια μεθοδολογία επεξεργασίας που παρουσιάζεται στο σχ. 3-14, με τη διαφορά ότι αλλάζει ριζικά ο τρόπος προσδιορισμού των κατωφλίων. Στην προκειμένη περίπτωση αξιοποιήθηκε ο αλγόριθμος *FDPPA* [82] (Fractal Dimension Peak Peeling Algorithm), ο οποίος έχει αναπτυχθεί για τον εντοπισμό (detection) κρουστικών βιοακουστικών σημάτων (πνευμονικών και εντερικών ήχων), παρουσία θορύβου. Ο αλγόριθμος *FDPPA* χρησιμοποιεί τεχνικές πολυπλοκότητας κυματομορφών "Fractal Dimension" (FD), που εφαρμόζονται σε «κυλιόμενα» παράθυρα επιλογής του σήματος (sliding windows). Έτσι, από τη σύγκριση των εκάστοτε τιμών *FD* με προκαθορισμένες τιμές κατωφλίων, καθορίζεται ποια παράθυρα περιλαμβάνουν συνιστώσες σήματος και ποια θόρυβο. Οι «αποφάσεις» αυτές χρησιμοποιούνται για τη σύσταση των δυαδικών ακολουθιών (binary sequences), βάσει των οποίων γίνεται η κατωφλίωση των κυματιδιακών συντελεστών σε κάθε βήμα l της επαναληπτικής διαδικασίας [83], [84].

Η μέθοδος *WT-FD* έχει αρκετά κοινά χαρακτηριστικά με την *WTST-NST*. Και οι δύο διαθέτουν τη δυνατότητα αυτόματου εντοπισμού κατωφλίων, ενώ αποδίδουν πολύ καλά κατά την αποθορυβοποίηση κρουστικών ηχητικών σημάτων. Όταν τα σήματα δεν έχουν έντονο κρουστικό χαρακτήρα, τότε αρχίζουν να εμφανίζονται διάφορες αδυναμίες [179], [217] όπως είναι η μη ικανοποιητική απομάκρυνση του θορύβου, ή η καταστροφή της δομής των αρχικών κυματομορφών (morphological structure destruction). Επιπρόσθετα, η μέθοδος *WT-FD* έχει μεγαλύτερο υπολογιστικό κόστος από την *WTST-NST* λόγω των ολισθαινόντων παραθύρων «σημείο - προς - σημείο» (point to point sliding windows) και την εφαρμογή των τελεστών *FD* σε καθένα από αυτά [83]. Μια λύση για την αντιμετώπιση του παραπάνω προβλήματος, είναι η εφαρμογή του αλγορίθμου *FDPPA* σε περιορισμένο αριθμό κυματιδιακών ζωνών (στην απλούστερη περίπτωση σε μία ζώνη) [84]. Η προσέγγιση αυτή λειτουργεί ικανοποιητικά εφόσον το σήμα έχει κρουστικό χαρακτήρα, οπότε οι χρήσιμες συνιστώσες εμφανίζονται σε όλες τις κυματιδιακές δέσμες. Στην αντίθετη περίπτωση, η εφαρμογή του αλγορίθμου *FDPPA* και ο προσδιορισμός των κατωφλίων αποθορυβοποίησης θα πρέπει να γίνει χωριστά σε κάθε επίπεδο ανάλυσης, στοιχείο που επιβαρύνει πρόσθετα το υπολογιστικό κόστος. Η ανάγκη ρύθμισης πρόσθετων παραμέτρων που σχετίζονται με τη

λειτουργία του αλγορίθμου *FDPPA* προσδίδει μεγαλύτερη πολυπλοκότητα στη μέθοδο *WT-FD*, σε σχέση με την προγενέστερη *WTST-NST* [179].

Μια εναλλακτική μέθοδος, η οποία αξιοποίησε με διαφορετικό τρόπο την επαναληπτική διαδικασία κατωφλίωσης του σήματος 3-14, είναι η τεχνική *IKD* (Iterative Kurtosis-based Detector) [81], [85]. Η μέθοδος αυτή επιδιώκει και πάλι το διαχωρισμό της ενθόρυβης ακολουθίας σε στάσιμα και μη στάσιμα τμήματα, με στόχο τον χρονικό εντοπισμό των σημάτων *ELS* και *EBS*. Η λειτουργία της έχει πολλές ομοιότητες με το διάγραμμα του σχ. 3-14, με τη διαφορά ότι αντί για κυματιδιακές τεχνικές εφαρμόζονται τεχνικές *HOS*. Έτσι, τμήματα του σήματος επιλέγονται με την εφαρμογή κυλιόμενων παραθύρων, όπου η εκτίμηση των τιμών κύρτωσης (kurtosis) και η σύγκρισή τους με προκαθορισμένες τιμές κατωφλίων, δίνουν τα αποτελέσματα ανίχνευσης σήματος. Η μέθοδος παρουσιάζει παραπλήσια πλεονεκτήματα και μειονεκτήματα με τις τεχνικές *WTST-NST* και *WT-FD*, δηλαδή εφαρμόζεται ικανοποιητικά σε κρουστικά κυρίως σήματα, ενώ παράλληλα έχει λιγότερες υπολογιστικές απαιτήσεις [81], [85].

Συνοψίζοντας, όλες οι προαναφερθείσες τεχνικές αποθρορυβοποίησης έχουν προταθεί για την επεξεργασία κρουστικών βιοακουστικών φαινομένων, πνευμονικών και εντερικών ήχων. Ωστόσο, εκτός από τους κρουστικούς κοιλιακούς ήχους (Explosive Bowel Sounds - *EBS*, ή Intestinal Bursts -*IB*), υπάρχουν και οι ομαλά παραμένουσες κυματομορφές (Regularly Sustained -*RS*) [61], [179], [215]-[218]. Με βάση πειραματικές παρατηρήσεις [179] έχει διαπιστωθεί ότι οι αυτοματοποιημένες τεχνικές *WTST-NST*, *WT-FD*, και *IKD* λειτουργούν ικανοποιητικά μόνο στην περίπτωση των *IB* προτύπων. Στην περίπτωση επεξεργασίας σημάτων *RS* είναι πολύ πιθανό να παρουσιαστούν διάφορα προβλήματα, όπως είναι ο μη ικανοποιητικός περιορισμός του θορύβου, ή η καταστροφή της μορφολογικής δομής των σημάτων. Ως αποτέλεσμα, οι συγκεκριμένες τεχνικές δεν μπορούν να εφαρμοστούν σε μη εποπτευόμενες διαδικασίες, για την επεξεργασία καταγραφών παρατεταμένης χρονικής διάρκειας, όπου εκτός των άλλων επιβαρυντικό είναι και το αυξημένο υπολογιστικό κόστος. Στην τελευταία περίπτωση έχουν διαπιστωθεί πρόσθετα προβλήματα που ανακύπτουν από την εφαρμογή χρονικών παραθύρων επιλογής (frame-based processing) [217], τα οποία αναλύονται διεξοδικά στο επόμενο κεφάλαιο. Καθίσταται, λοιπόν, σαφές ότι υπάρχει ανάγκη μιας πιο αποτελεσματικής και λειτουργικής μεθόδου αποθρορυβοποίησης, η οποία να δίνει εξίσου καλά αποτελέσματα σε όλες τις κατηγορίες εντερικών ήχων, και κυρίως να μπορεί να εφαρμοστεί εύκολα και αξιόπιστα σε καταγραφές παρατεταμένου χρόνου.

3.2.3 Η προτεινόμενη μέθοδος καταστολής θορύβου

Η προτεινόμενη μέθοδος καταστολής θορύβου αναπτύχθηκε σε μια προσπάθεια να αντιμετωπιστούν όλα τα προβλήματα που προαναφέρθηκαν. Βασικοί στόχοι κατά τη διαδικασία σχεδιασμού και υλοποίησης της νέας τεχνικής υπήρξαν, α) η αποτελεσματική καταστολή του θορύβου βάθους *ABN*, β) η ελάχιστη παραμόρφωση των χρήσιμων συνιστωσών σήματος, γ) η διατήρηση της πολυπλοκότητας της μεθόδου και του αντίστοιχου υπολογιστικού κόστους σε χαμηλά επίπεδα, δ) η εύκολη και αξιόπιστη εφαρμογή της μεθόδου σε καταγραφές παρατεταμένου χρόνου [179]. Για την επίτευξη των παραπάνω στόχων αποφασίστηκε ο συνδυασμός των προσεγγίσεων φασματικής αφαίρεσης και κυματιδιακής κατωφλίωσης, δηλαδή η εφαρμογή του παραμετρικού φίλτρου Wiener σε επίπεδο κυματιδιακών συντελεστών.

3.2.4 Επεξεργασία κυματιδιακών συντελεστών με το παραμετρικό φίλτρο Wiener

Η εφαρμογή του φίλτρου Wiener στο πεδίο του μετασχηματισμού κυματιδίων δεν αποτελεί καινούργια μέθοδο [179], [197], [219]. Αν χρησιμοποιήσουμε το συμβολισμό $X_k(w)$ για την ακολουθία των κυματιδιακών συντελεστών με δείκτες w ($w=0,1,..,W_{Xk}-1$) στην k -ζώνη φασματικής ανάλυσης, τότε, με βάση τη σχέση (3-10), παίρνουμε την εξίσωση του φασματικά εξαρτώμενου παραμετρικού φίλτρου Wiener, στο πεδίο του μετασχηματισμού κυματιδίων:

$$H_k = \begin{cases} \left\{ 1 - c \cdot A_{kw} \cdot \left[\frac{\langle N_{FP-k}(w) \rangle}{\langle X_k(w) \rangle} \right]^a \right\}^b, & \text{αν } c \cdot A_{kw} \cdot \left[\frac{\langle N_{FP-k}(w) \rangle}{\langle X_k(w) \rangle} \right]^a \leq 1 \\ 0, & \text{διαφορετικά} \end{cases} \quad (3-30)$$

όπου H_k είναι το κέρδος του φίλτρου, $N_{FP-k}(w)$ οι κυματιδιακοί συντελεστές του διαθέσιμου «προφίλ» θορύβου (noise footprint) και A_{kw} είναι η κυματιδιακή έκφραση του ψυχοακουστικού φίλτρου A_k [179]:

$$\langle N_{FP-k}(w) \rangle = \frac{1}{W_{Nk}} \sqrt{\sum_{w=0}^{W_{Nk}-1} [N_{FP-k}(w)]^2} \quad (3-31)$$

$$\langle X_k(w) \rangle = \frac{1}{W_{Xk}} \sqrt{\sum_{w=0}^{W_{Xk}-1} [X_k(w)]^2} \quad (3-32)$$

$$A_{kw} = \begin{cases} 1 & f_{c-k} \leq 500 \\ \frac{f_{c-k}}{500} & f_{c-k} > 500 \end{cases} \quad (3-33)$$

Η παράμετρος f_{c-k} της εξίσωσης (3-33) αντιστοιχεί στη κεντρική συχνότητα κάθε ζώνης k , η οποία προκύπτει ως ο γεωμετρικός μέσος της μέγιστης και ελάχιστης συχνότητας (geometric mean of the band's frequency limits). Η εξίσωση (3-29) είναι γενική και ισχύει, είτε τα ψυχοακουστικά κριτήρια ενεργοποιούνται, είτε όχι (οπότε $A_{kw}=A_{k0}=1$). Η έξοδος του κυματιδιακού φίλτρου Wiener (εξ. 3-30) παρέχει τους αποθορυβοποιημένους κυματιδιακούς συντελεστές $S_k \sim(w)$, απ' όπου προκύπτει το καθαρό από θόρυβο σήμα $s \sim(i)$, με χρήση του αντίστροφου μετασχηματισμού κυματιδίων:

$$S_k \sim(w) = H_k \cdot X_k(w) \quad (3-34)$$

Η φασματική αναλυτικότητα της κυματιδιακής επεξεργασίας Wiener εξαρτάται από το συνολικό αριθμό των σχηματιζόμενων ζωνών ανάλυσης, δηλαδή από τις τοπολογίες κυματιδιακής ανάλυσης - σύνθεσης. Για την επεξεργασία μη στάσιμων σημάτων υιοθετείται η χρήση χρονικών παραθύρων (Time Windowed Wavelet Wiener Filter, TW-WWF), όπου συχνά εφαρμόζονται τεχνικές επικάλυψης (overlapping). Έτσι, επιτυγχάνονται καλύτερα αποτελέσματα επεξεργασίας, ενώ παράλληλα περιορίζονται πιθανά σφάλματα που σχετίζονται με τις απότομες φασματικές μεταβάσεις λόγω της χρήσης παραθύρων. Αξίζει να σημειωθεί ότι, για την αντιμετώπιση αντίστοιχων προβλημάτων σε εφαρμογές συμπίεσης, αξιοποιούνται διάφορες τεχνικές καταστολής προ-ηχούς (pre-echo cancellation techniques),

βάσει των οποίων γίνεται ρύθμιση των παραθύρων επιλογής, ανάλογα με τα χαρακτηριστικά του σήματος στην εκάστοτε χρονική στιγμή [185], [202], [210], [220]. Τα στοιχεία αυτά, βέβαια, αυξάνουν την πολυπλοκότητα των εφαρμοζόμενων αλγορίθμων, γ' αυτό και δεν υιοθετήθηκαν στη συγκεκριμένη προσέγγιση [179].

3.2.5 Νέες τεχνικές "Wavelet Domain Wiener Filtering" (WDWF)

Τα χαρακτηριστικά του κυματιδιακού φίλτρου Wiener, που περιγράφεται από τις εξισώσεις (3-30) έως (3-34), δεν ανταποκρίνονται πλήρως στις απαιτήσεις των προδιαγραφών που παρουσιάστηκαν στην §3.2.3. Οι κύριες δυσκολίες εντοπίζονται στα ζητήματα της περιορισμένης χρονικής αναλυτικότητας, που εξαρτώνται από τα χαρακτηριστικά των παραθύρων επιλογής. Η μεγάλη «ποικιλία» στη μορφολογία των εντερικών ήχων και των αντίστοιχων χαρακτηριστικών διάρκειας και κρουστικότητας, καθιστά σχεδόν αδύνατη τη χρήση γενικευμένων ρυθμίσεων, δηλαδή «βέλτιστων παραθύρων» που θα λειτουργούν εξίσου ικανοποιητικά σε όλες τις περιπτώσεις [179]. Τα στοιχεία αυτά αυξάνουν την πολυπλοκότητα, ενώ παράλληλα είναι ιδιαίτερα επιβαρυντικά στις αυτοματοποιημένες διαδικασίες επεξεργασίας των καταγραφών παρατεταμένου χρόνου [179], [217]. Στο πλαίσιο αυτό, προτάθηκαν νέες δομές κυματιδιακής επεξεργασίας Wiener «σημείο - προς - σημείο» (point to point wavelet processing), οι οποίες αποτελούν το κύριο αντικείμενο του τρέχοντος κεφαλαίου και στο εξής θα περιγράφονται με το ακρωνύμιο WDWF (Wavelet Domain Wiener Filtering).

3.2.5.1 Επεξεργασία WDWF σημείο προς σημείο, τύπος-I

Στην παράγραφο αυτή αναλύεται η πρώτη από τις δύο προσεγγίσεις κυματιδιακής επεξεργασίας σημείο προς σημείο, η οποία και αναφέρεται ως τύπος-I: "type-I WDWF". Η διευκρίνιση αυτή γίνεται εξαρχής, καθώς στην επόμενη παράγραφο θα παρουσιαστεί μια δεύτερη προσέγγιση, ο τύπος-II: "type-II WDWF" [179]. Αξίζει να σημειωθεί ότι και στις δύο περιπτώσεις αξιοποιήθηκε η πιο κλασσική ρύθμιση του παραμετρικού Wiener φίλτρου, που ισοδυναμεί με την πράξη φασματικής αφαίρεσης ισχύος ($a=2$, $b=1$). Αν και έγιναν δοκιμές για διαφορετικές τιμές των δύο παραπάνω παραμέτρων, δεν κρίθηκε σκόπιμη η υιοθέτησή τους, καθώς δεν διαφοροποιούσαν αισθητά τα αντίστοιχα αποτελέσματα [179]. Έτσι, η προσοχή επικεντρώθηκε σε άλλα σημεία που σχετίζονται με τις ιδιαιτερότητες των τύπων I και II, τα λειτουργικά χαρακτηριστικά της επεξεργασίας σημείο προς σημείο, τις τοπολογίες κυματιδιακής ανάλυσης - ανακατασκευής, την επίδραση της παραμέτρου κατωφλίωσης c και του ψυχοακουστικού μοντέλου A_{kw} . Σε κάθε περίπτωση, η διαφοροποίηση των παραμέτρων a , b δεν είναι απαγορευτική, ενώ αντιθέτως, παρουσιάζει ενδιαφέρον η πιο συστηματική διερεύνηση της χρήσης εναλλακτικών ρυθμίσεων, στο άμεσο μέλλον. Στη θεωρητική ανάλυση που θα ακολουθήσει χρησιμοποιούνται οι γενικοί τύποι, όπου οι παράμετροι a , b δεν έχουν τις τυπικές τους τιμές ($a=2$ και $b=1$), αλλά αποτελούν ανεξάρτητες μεταβλητές του ρυθμιζόμενου παραμετρικού φίλτρου.

Για την ανάπτυξη των τεχνικών WDWF γίνεται χρήση «φίλτρων εκθετικά κινούμενου μέσου όρου» (Exponential Moving Averaging, -ExpMA). Έτσι, αν συμβολίσουμε με $P_{x;a}(k,w)$ την a -δύναμη της ενεργού τιμής κάθε κυματιδιακής ακολουθίας $\langle X(k,w) \rangle^a$, τότε ο υπολογισμός της μπορεί να γίνει με τη βοήθεια της ακόλουθης επαναληπτικής διαδικασίας:

$$P_{x;a}(k,w)|_I = d \cdot |X(k,w)|^a + (1-d) \cdot P_{x;a}(k,w-1) \quad w = 0,1,\dots,W_{xk} - 1 \quad (3-35)$$

όπου k είναι η εκάστοτε δέση κυματιδιακής ανάλυσης, και w ο δείκτης των κυματιδιακών συντελεστών σε κάθε επίπεδο, δηλαδή ο συμβολισμός $X(k,w)$ αντιστοιχεί στον w κυματιδιακό συντελεστή, της k κλίμακας, του ενθόρυβου σήματος $x(i)$. Είναι προφανές ότι η ακολουθία $X(k,w)$ είναι απόλυτα ισοδύναμη με την $X_k(w)$ που χρησιμοποιήθηκε στην παράγραφο 3.2.4. Ο νέος συμβολισμός εισήχθη κυρίως για λειτουργικούς λόγους, που αφορούν στην πιο παραστατική περιγραφή του δισδιάστατου σχήματος χρόνου - κλιμάκωσης, αλλά και τη διάκριση των δύο προσεγγίσεων. Παρόμοια, οι όροι $S^-(k,w)$ και $N_{FP}(k,w)$ αντιστοιχούν στους κυματιδιακούς συντελεστές των σημάτων $s^-(i)$ και $n_{FP}(i)$, αντίστοιχα [179].

Η εξίσωση (3-35) εισάγει τον εκθετικά κινούμενο μέσο όρο εκτίμησης της ισχύος $P_{x;a}(k,w)|_I$, όπου ο δείκτης I υποδηλώνει ότι η συγκεκριμένη έκφραση χρησιμοποιείται για τον τύπο-I, $WDWF$. Η παράμετρος d παίρνει πραγματικές τιμές μέσα στο διάστημα $[0, 1]$ και χρησιμοποιείται για να ρυθμίσει τη «μνήμη», που υπεισέρχεται στη διαδικασία $ExpMA$. Για το λόγο αυτό συχνά αναφέρεται ως παράμετρος μνήμης (memory term). Με τη νέα προσέγγιση εξασφαλίζεται μεγαλύτερη αναλυτικότητα χρόνου, ενώ περιορίζεται η πολυπλοκότητα και οι υπολογιστικές απαιτήσεις. Ακόμη, οι εκτιμώμενες ενεργειακές παράμετροι είναι ευαίσθητες στις παρελθούσες τιμές του σήματος, στοιχείο το οποίο παρουσιάζει αρκετές ομοιότητες με το «ψυχο-ακουστικό φαινόμενο προήγησης» του Haas (Haas precedence effect of psychoacoustics) [188]. Έτσι, παρέχονται νέες δυνατότητες ενσωμάτωσης ψυχοακουστικών κριτηρίων στην κατεύθυνση «αντιληπτικής επεξεργασίας ήχου» (perceptual audio processing) [179].

Από την εξίσωση (3-35) προκύπτει το ενδεχόμενο να εμφανιστούν χρονικές ασυνέχειες κατά την έναρξη των υπολογισμών (δηλαδή στα αρχικά δείγματα $w=0$), καθώς δεν υπάρχουν προηγούμενες τιμές (μνήμη) για να χρησιμοποιηθούν στην επαναληπτική διαδικασία. Για την αντιμετώπιση αυτής της ιδιαιτερότητας και προκειμένου να περιοριστούν οι πιθανότητες αλλοίωσης χρήσιμων συνιστωσών στο σημείο εκκίνησης, η εξίσωση (3-35) μετασχηματίζεται στην ακόλουθη μορφή:

$$P_{x;a}(k,w)|_I = \begin{cases} |X(k,w)|^a, & w = 0 \\ d \cdot |X(k,w)|^a + (1-d) \cdot P_{x;a}(k,w-1), & w = 1, \dots, W_{xk} - 1 \end{cases} \quad (3-36)$$

Με την τελευταία τροποποίηση, η συνάρτηση μεταφοράς $H_{WDWF|I}$ του πρώτου τύπου κυματιδιακής αποθορυβοποίησης $WDWF_I$, γίνεται:

$$H_{WDWF|I}(k,w) = \begin{cases} \left\{ 1 - c \cdot A_{kw} \cdot \left[\frac{P_{nFP;a}(k)}{P_{x;a}(k,w)|_I} \right]^b \right\}, & \text{αν } c \cdot A_{kw} \cdot \left[\frac{P_{nFP;a}(k)}{P_{x;a}(k,w)|_I} \right] \leq 1 \\ 0, & \text{διαφορετικά} \end{cases} \quad (3-37)$$

όπου, το κέρδος του ψυχο-ακουστικού φίλτρου A_{kw} εξακολουθεί να δίνεται από την εξίσωση (3-33). Η αντίστοιχη ενέργεια του θορύβου $P_{nFP;a}(k)$ εκτιμάται από το μέσο όρο (average) των κυματιδιακών συντελεστών $N_{FP}(k,w)$, που προκύπτουν από το διαθέσιμο αποτύπωμα (noise footprint), αφού πρωτίστως υψωθούν στη δύναμη a :

$$P_{nFP;a}(k) = \frac{1}{W_{Nk}} \sum_{w=0}^{W_{Nk}-1} |N_{FP}(k,w)|^a \quad (3-38)$$

Επομένως, η επεξεργασία καταστολής θορύβου με το φασματικά εξαρτώμενο φίλτρο Wiener στο πεδίο του μετασχηματισμού κυματιδίων, περιγράφεται από τη σχέση:

$$S^{\sim}(k, w) = H_{WDWF}(k, w) \cdot X(k, w), \quad w = 0, 1, \dots, W_{sk} - 1 \quad (3-39)$$

Η ενσωμάτωση της διαδικασίας *ExpMA* στο Wiener φίλτρο κυματιδιακής επεξεργασίας παρουσιάζει αρκετά πλεονεκτήματα, σε σχέση τόσο με τις Fourier μεθόδους φασματικής επεξεργασίας, όσο και με τις τεχνικές κυματιδιακής κατωφλίωσης. Έτσι, η μέθοδος *WDWF* διατηρεί τα χαρακτηριστικά της λογαριθμικής-οκταβικής φασματικής ανάλυσης (σε αντίθεση με τις *STFT* υλοποιήσεις Wiener), ενώ παράλληλα παρουσιάζει ικανοποιητική χρονική αναλυτικότητα (σε αντίθεση με την κυματιδιακή επεξεργασία *TW-WWF*, με χρήση παραθύρων). Επιπρόσθετα, έχει καλύτερα χαρακτηριστικά από τις αντίστοιχες υλοποιήσεις με τράπεζες φίλτρων σταθερού ρυθμού δειγματοληψίας (*Constant Rate Filter Banks -CRFB*), οι οποίες επίσης χρησιμοποιούν διαδικασίες *ExpMA* [177], [187]. Το τελευταίο στοιχείο σχετίζεται με την πολύ-διακριτική ανάλυση του μετασχηματισμού *WT*, όπου ο περιορισμός του πλήθους των κυματιδιακών συντελεστών συντελεί στην εξοικονόμηση υπολογιστικής μνήμης και ισχύος. Έτσι, ο αρχικός αριθμός δειγμάτων παραμένει αμετάβλητος, σε αντίθεση με τις *CRFB* Wiener υλοποιήσεις, όπου το σύνολο των δειγμάτων στις εξόδους της τράπεζας φίλτρων, είναι *k*-φορές πολλαπλάσιο του αρχικού (με *k* το συνολικό αριθμό των φίλτρων).

Επιπρόσθετα, το σχήμα πολύ-διακριτικής ανάλυσης παρέχει τη δυνατότητα χρήσης ταυτόσημων ρυθμίσεων του φίλτρου (*identical filter-configuration*) σε όλες τις κυματιδιακές κλίμακες, στοιχείο που συνδέεται κυρίως με τη ρύθμιση της παραμέτρου μνήμης *d*. Στις περιπτώσεις των *CRFB* προσεγγίσεων αυτό δεν είναι αποδοτικό κι έτσι απαιτείται η χρήση διαφορετικής μνήμης εκθετικού μέσου όρου (*d*) για κάθε φασματική περιοχή [177], [187]. Η δυνατότητα ενιαίας μορφοποίησης των παραμέτρων του φίλτρου, έχει ως αποτέλεσμα τη χρήση λιγότερων μεταβλητών εισόδου, συντελώντας στην ευκολότερη εφαρμογή και τον αποδοτικότερο χειρισμό της νέας παραμετρικής προσέγγισης *WDWF* [179]. Συνοψίζοντας, ο νέος αλγόριθμος *WDWF* διαφέρει από τις προηγούμενες παραμετρικές προσεγγίσεις φασματικής αφαίρεσης Wiener [178], [184], στη χρήση του μετασχηματισμού κυματιδίων, και κυρίως στο γεγονός ότι υιοθετεί την διαδικασία του εκθετικά κινούμενου μέσου, αντί των χρονικών παραθύρων επιλογής. Το στοιχείο αυτό εισάγει μια πρόσθετη παράμετρο στις υπάρχουσες *a*, *b*, *c*, που δεν είναι άλλη από τον όρο μνήμης *d*. Η επίδραση των τριών πρώτων μεταβλητών στη συνάρτηση μεταφοράς του φίλτρου είναι σε γενικές γραμμές γνωστή, καθώς αποτέλεσε αντικείμενο έρευνας πολλών εργασιών στο παρελθόν [177], [184], [187] (σχετικά διαγράμματα έχουν συμπεριληφθεί και σχολιαστεί στην §3.1.2.1). Εκείνο που δεν είναι τόσο δεδομένο, είναι η επίδραση της παραμέτρου *d*, στοιχείο που αποτελεί ένα από τα κύρια αντικείμενα της τρέχουσας ενότητας.

Οι τεχνικές *ExpMA* συναντώνται αρκετά συχνά κατά τον υπολογισμό ηχητικών σταθμών (*Sound Levels*) και γενικότερα σε εφαρμογές ακουστικής - επεξεργασίας ήχου [177], [179], [187], [193], [199]. Ένας από τους κύριους λόγους αξιοποίησής τους, εντοπίζεται στο γεγονός ότι τα περισσότερα ακουστικά φαινόμενα έχουν αντίστοιχες περιβάλλουσες ανόδου - καθόδου (*attack-release envelopes*), που μοιάζουν με τις εκθετικές καμπύλες. Σημαντική είναι επίσης η σύνδεσή τους με το ψυχοακουστικό φαινόμενο του Haas [188] που αναφέρθηκε πιο πάνω, καθώς και τα επιθυμητά χαρακτηριστικά εύκολης υλοποίησης και μικρότερου υπολογιστικού κόστους από τις τεχνικές κινούμενου μέσου όρου (*Moving Average -MA*) [179], [199]. Κατά τον υπολογισμό ηχητικών σταθμών, η διαδικασία *ExpMA* εφαρμόζεται οδηγώντας την ακολουθία στιγμιαίας ισχύος του σήματος ($\sim \sigma^2$) σε ένα

φίλτρο άπειρης κρουστικής απόκρισης (Infinite Impulse Response -IIR) 1ης τάξης, που έχει «ανδρομικούς συντελεστές» (reverse coefficients) $\{a_0, a_1\}$ και «πρόσθιους συντελεστές» (forward coefficients) $\{b_1\}$:

$$a_0 = 1, \quad a_1 = -e^{-\frac{1}{f_s \cdot t_w}} \quad (3-40)$$

$$b_1 = 1 - e^{-\frac{1}{f_s \cdot t_w}} = 1 + a_1$$

όπου f_s είναι η συχνότητα δειγματοληψίας σε Hz και t_w η ισοδύναμη διάρκεια ολοκλήρωσης, που προκύπτει από τη διαδικασία, $ExpMA$ σε sec [179], [199]. Συγκρίνοντας τις εξισώσεις (3-36) και (3-41), είναι προφανές ότι ο πρόσθιος συντελεστής b_1 του IIR φίλτρου και η παράμετρος μνήμης d , αποτελούν ένα και το αυτό μέγεθος. Έτσι, οι συντελεστές $\{a_0, a_1\}$ και $\{b_1\}$ του «φίλτρου εκθετικού-μέσου» ισοδυναμούν αντίστοιχα με $\{1, d-1\}$ και $\{d\}$.

Από την εξίσωση (3-40), είναι δυνατό να υπολογιστεί η διάρκεια t_w για διάφορες τιμές της παραμέτρου d . Για να γίνει αντιληπτή η σημασία μιας τέτοιας διερεύνησης, θα πρέπει πρωτίστως να γίνει αναλυθεί η επίδραση της χρονικής διάρκειας t_w στην διαδικασία φιλτραρίσματος. Δεδομένου ότι στα περισσότερα ακουστικά φαινόμενα οι χρήσιμες πληροφοριακές συνιστώσες δεν παρουσιάζονται συνεχώς, η χρήση παραθύρων μικρής διάρκειας επιτρέπει την αποτελεσματικότερη καταστολή του θορύβου φόντου στα χρονικά διαστήματα σιγής ("noise-only" segments, silence/ pause intervals). Έτσι, όσο πιο μικρό είναι το μήκος του παραθύρου, τόσο πιο προσαρμοστικό είναι το φίλτρο Wiener. Ωστόσο, υπάρχει ένα ελάχιστο επιτρεπτό όριο, κάτω από το οποίο είναι δυνατόν να παρουσιαστούν εσφαλμένοι υπολογισμοί των εμπλεκόμενων ενεργειακών παραμέτρων. Αυτό συμβαίνει καθώς για κάθε «παρατηρούμενη» συνιστώσα του σήματος απαιτείται χρονική διάρκεια ολοκλήρωσης t_w που ισοδυναμεί (κατ' ελάχιστο) με έναν πλήρη περιοδικό κύκλο της αντίστοιχης συχνότητας. Για παράδειγμα, ένας καθαρός τόνος συχνότητας 200Hz, απαιτεί σταθερά ολοκλήρωσης $t_w=5msec$ ($=1/200Hz$) ή περισσότερο, ενώ με δεδομένο το ρυθμό δειγματοληψίας ($f_s=8kHz$), η παράμετρος μνήμης διαμορφώνεται στην τιμή $d=0,025$ (εξ. 3-40). Στην περίπτωση, λοιπόν, που ο ρυθμός δειγματοληψίας f_s παραμένει σταθερός, απαιτείται η χρήση διαφορετικών παραμέτρων $d(k)$ για καθεμιά από τις εμπλεκόμενες περιοχές k [177], [179], [187].

Στον πίνακα Π3-1, παρουσιάζονται οι φασματικές ζώνες των κλασικών σχημάτων οκταβικής και τριτο-οκταβικής ανάλυσης και τα αντίστοιχα επιτρεπτά όρια των παραμέτρων $d(k)$, για συχνότητες μέχρι 4kHz ($f_s=8kHz$). Όπου f_L, f_C και f_H είναι αντίστοιχα η χαμηλότερη (Low), κεντρική (Center) και υψηλότερη (High) συχνότητα κάθε περιοχής. Ο δείκτης k χρησιμοποιείται για την αρίθμηση των αντίστοιχων φασματικών ζωνών, ενώ η παράμετρος $tw-min$ αντιπροσωπεύει την ελάχιστη δυνατή χρονική διάρκεια t_w (εκφρασμένη σε msec). Αντίστοιχα διαμορφώνονται και οι μέγιστες επιτρεπτές τιμές της παραμέτρου $d(k)$. Παρατηρούμε ότι προκειμένου να εξασφαλιστεί η βέλτιστη προσαρμοστικότητα του φίλτρου σε κάθε οκταβική (και τριτο-οκταβική) περιοχή k , η παράμετρος $d(k)$ θα πρέπει να μεταβάλλεται αισθητά. Αυτό σημαίνει ότι σε εφαρμογές σταθερού ρυθμού δειγματοληψίας (CRFB) χρειάζονται k διαφορετικές μεταβλητές εισόδου για την παράμετρο μνήμης $d(k)$, στοιχείο που αυξάνει την πολυπλοκότητα και δυσχεραίνει τις διαδικασίες ρύθμισης - εφαρμογής του φίλτρου, ανάλογα και με την τοπολογία φασματικής ανάλυσης [179].

octaves [k]	f _L [Hz]	f _c [Hz]	f _H [Hz]	tw-min [msec]	d(k) [max]	third-octaves [k]	f _L [Hz]	f _c [Hz]	f _H [Hz]	tw-min [msec]	d(k) [max]
1	44,2	62,5	88,4	22,6	0,006	1	55,7	62,5	70,2	18,0	0,007
						2	70,2	78,7	88,4	14,3	0,009
2	88,4	125	176,8	11,3	0,011	3	88,4	99,2	111,4	11,3	0,011
						4	111,4	125,0	140,3	9,0	0,014
						5	140,3	157,5	176,8	7,1	0,017
3	176,8	250	353,6	5,7	0,022	6	176,8	198,4	222,7	5,7	0,022
						7	222,7	250,0	280,6	4,5	0,027
						8	280,6	315,0	353,6	3,6	0,034
4	353,6	500	707,1	2,8	0,043	9	353,6	396,9	445,4	2,8	0,043
						10	445,4	500,0	561,2	2,2	0,054
						11	561,2	630,0	707,1	1,8	0,068
5	707,1	1.000	1.414,2	1,4	0,085	12	707,1	793,7	890,9	1,4	0,085
						13	890,9	1.000,0	1.122,5	1,1	0,105
						14	1.122,5	1.259,9	1.414,2	0,9	0,131
6	1.414,2	2.000	2.828,4	0,7	0,162	15	1.414,2	1.587,4	1.781,8	0,7	0,162
						16	1.781,8	2.000,0	2.244,9	0,6	0,200
						17	2.244,9	2.519,8	2.828,4	0,4	0,245
7	2.828,4	4.000	* 4.000	0,4	0,298	18	2.828,4	3.174,8	3.563,6	0,4	0,298
						19	3.563,6	4.000,0	* 4.000	0,3	0,359

Πίνακας Π3-1. Παρουσίαση των οκταβικών και τρίτο-οκταβικών περιοχών φασματικής ανάλυσης και των αντίστοιχων επιτρεπτών ορίων για τις παραμέτρων μνήμης $d(k)$, που υπεισέρχονται στους υπολογισμούς του εκθετικού μέσου όρου (*με δεδομένη τη σταθερή συχνότητα $f_s=8kHz$, τα ανώτερα συχνοτικά όρια της 7ης οκταβικής περιοχής και 19ης τρίτο-οκταβικής περιοχής διαμορφώνονται στα 4kHz, λόγω του κριτηρίου Nyquist).

Στους πίνακες Π3-2 και Π3-3 παρουσιάζεται η επίδραση της παραμέτρου d για τις τοπολογίες κυματιδιακής ανάλυσης - σύνθεσης “DWT- 6 ζωνών” και “WP- 17 ζωνών” που τελικά υιοθετήθηκαν στις μεθόδους WDWF [179]. Με βάση το ρυθμό δειγμάτων (rate - [Hz]) σε κάθε κόμβο ανάλυσης (WT node) εξετάζεται αν ικανοποιείται η συνθήκη (Cond.: $t_w \geq min$), έτσι ώστε η εκάστοτε «εκθετική διάρκεια» t_w να είναι μεγαλύτερη από το ελάχιστο επιτρεπτό όριο $tw-min$ (καθορίζεται από τη συχνότητα f_L κάθε φασματικής περιοχής k). Παρατηρούμε ότι τα αποτελέσματα της λογικής συνθήκης είναι «αληθή» (TRUE) για όλες τις τιμές της παραμέτρου d , με εξαίρεση τη ρύθμιση $d=0,4$ που οδηγεί σε πολύ μικρά μήκη παραθύρων (κυρίως στο σχήμα DWT). Το στοιχείο αυτό επαληθεύει τη δυνατότητα ενιαίας ρύθμισης του φίλτρου, με όλα τα συνοδευτικά πλεονεκτήματα που σχολιάστηκαν πιο πάνω [179].

WT node	band [k]	rate [Hz]	f _L [Hz]	f _H [Hz]	tw-min [msec]	d=0,05		d=0,1		d=0,2		d=0,4	
						t _w [msec]	Cond.: [t _w ≥min]	t _w [msec]	Cond.: [t _w ≥min]	t _w [msec]	Cond.: [t _w ≥min]	t _w [msec]	Cond.: [t _w ≥min]
1	6	4000	2000	4000	0,5	4,9	TRUE	2,4	TRUE	1,1	TRUE	0,49	FALSE
01	5	2000	1000	2000	1,0	9,7	TRUE	4,7	TRUE	2,2	TRUE	1,0	FALSE
001	4	1000	500	1000	2,0	19,5	TRUE	9,5	TRUE	4,5	TRUE	2,0	FALSE
0001	3	500	250	500	4,0	39,0	TRUE	19,0	TRUE	9,0	TRUE	3,9	FALSE
00001	2	250	125	250	8,0	78,0	TRUE	38,0	TRUE	17,9	TRUE	7,8	FALSE
00000	1	250	0	125	* 16,0	78,0	TRUE	38,0	TRUE	17,9	TRUE	7,8	FALSE

Πίνακας Π3-2. Επίδραση της παραμέτρου d στην χρονική διάρκεια t_w (λόγω της διαδικασίας εκθετικού μέσου όρου), κατά την κυματιδιακή τοπολογία ανάλυσης DWT- 6 ζωνών (*για την αποφυγή διαίρεσης με το 0, η τιμή “tw-min” στη ζώνη $k=1$ υπολογίζεται με βάση την κεντρική συχνότητα (62,5Hz)).

WT node	band [k]	rate [Hz]	f _L [Hz]	f _H [Hz]	tw-min [msec]	d=0,05		d=0,1		d=0,2		d=0,4	
						t _w [msec]	Cond.: [t _w ≥min]	t _w [msec]	Cond.: [t _w ≥min]	t _w [msec]	Cond.: [t _w ≥min]	t _w [msec]	Cond.: [t _w ≥min]
111	17	1000	3500	4000	0,29	19,5	TRUE	9,491	TRUE	4,481	TRUE	1,958	TRUE
110	16	1000	3000	3500	0,33	19,5	TRUE	9,491	TRUE	4,481	TRUE	1,958	TRUE
101	15	1000	2500	3000	0,40	19,5	TRUE	9,491	TRUE	4,481	TRUE	1,958	TRUE
1001	14	500	2250	2500	0,44	38,99	TRUE	18,98	TRUE	8,963	TRUE	3,915	TRUE
1000	13	500	2000	2250	0,50	38,99	TRUE	18,98	TRUE	8,963	TRUE	3,915	TRUE
0111	12	500	1750	2000	0,57	38,99	TRUE	18,98	TRUE	8,963	TRUE	3,915	TRUE
0110	11	500	1500	1750	0,67	38,99	TRUE	18,98	TRUE	8,963	TRUE	3,915	TRUE
0101	10	500	1250	1500	0,80	38,99	TRUE	18,98	TRUE	8,963	TRUE	3,915	TRUE
0100	9	500	1000	1250	1,00	38,99	TRUE	18,98	TRUE	8,963	TRUE	3,915	TRUE
00111	8	250	875	1000	1,14	77,98	TRUE	37,96	TRUE	17,93	TRUE	7,83	TRUE
00110	7	250	750	875	1,33	77,98	TRUE	37,96	TRUE	17,93	TRUE	7,83	TRUE
00101	6	250	625	750	1,60	77,98	TRUE	37,96	TRUE	17,93	TRUE	7,83	TRUE
00100	5	250	500	625	2,00	77,98	TRUE	37,96	TRUE	17,93	TRUE	7,83	TRUE
00011	4	250	375	500	2,67	77,98	TRUE	37,96	TRUE	17,93	TRUE	7,83	TRUE
00010	3	250	250	375	4,00	77,98	TRUE	37,96	TRUE	17,93	TRUE	7,83	TRUE
00001	2	250	125	250	8,00	77,98	TRUE	37,96	TRUE	17,93	TRUE	7,83	FALSE
00000	1	250	0	125	* 16,00	77,98	TRUE	37,96	TRUE	17,93	TRUE	7,83	FALSE

Πίνακας Π3-3. Επίδραση της παραμέτρου d στην χρονική διάρκεια t_w (λόγω της διαδικασίας εκθετικού μέσου όρου), κατά την κυματιδιακή τοπολογία ανάλυσης WP- 17 ζωνών (*για την αποφυγή διαίρεσης με το 0, η τιμή “ $tw-min$ ” στη ζώνη $k=1$ υπολογίζεται με βάση την κεντρική συχνότητα (62,5Hz)).

Ο τύπος $WDWF_l$ ρυθμίστηκε με βάση πειραματικές παρατηρήσεις, αλλά και τους περιορισμούς της παραμέτρου d , όπως περιγράφονται από τη σχέση (3-40) και τον πίνακα Π3-1. Η ενιαία παράμετρος μνήμης επιλέχθηκε ίση με $d=0,2$ ενώ οι υπόλοιπες παράμετροι διαμορφώθηκαν στις τιμές $a=2$, $b=1$ και $c=3$ [179]. Σχετικά παραδείγματα «πιστοποίησης» της αποδοτικότητας αυτών των επιλογών δίνονται στην παράγραφο 3.2.6.1.

3.2.5.2 Επεξεργασία WDWF σημείο προς σημείο, τύπος-II

Σύμφωνα με πολλούς ερευνητές [177]-[178], [184]-[187], [189], [197], οι τεχνικές Wiener αποτελούν τη βέλτιστη λύση επεξεργασίας για την καταστολή του θορύβου βάθους (ελαχιστοποιούν το μέσο τετραγωνικό σφάλμα ανακατασκευής) υπό την προϋπόθεση ότι το επιλεγμένο τμήμα του ενθόρυβου σήματος περιλαμβάνει χρήσιμες πληροφοριακές συνιστώσες (δεν αντιστοιχεί σε περίοδο σιγής). Αυτό, όμως, δεν ισχύει πάντα στα φυσικά ακουστικά φαινόμενα, καθώς κατά μήκος των ηχητικών καταγραφών παρεμβάλλονται διαστήματα εφησυχασμού (pause - silence intervals), τα οποία περιέχουν μόνο θόρυβο [178], [179]. Για το λόγο αυτό, η εξ. (3-1) που περιγράφει το βασικό μοντέλο εμφάνισης του προσθετικού θορύβου, συχνά παίρνει την ακόλουθη τροποποιημένη μορφή [178]:

$$x(i) = bn_s(i) \cdot s(i) + n(i) \quad (3-41)$$

Η δυαδική ακολουθία (binary state sequence) $bn_s(i)$ παίρνει τιμές 0 και 1, έτσι ώστε η κατάσταση $bn_s(i)=1$ αντιστοιχεί σε παρουσία του σήματος $s(i)$, ενώ η κατάσταση $bn_s(i)=0$ σε παρουσία μόνο του θορύβου $n(i)$. Το τροποποιημένο αυτό μοντέλο δείχνει πιο ορθό, ειδικά στην περίπτωση των ενθόρυβων εντερικών ήχων, καθώς τα περισσότερα γαστρεντερικά ακουστικά φαινόμενα προκύπτουν από τυχαίους (χρονικά) συνδυασμούς μεμονωμένων και «αλυσιδωτών» συστολών (solitary and clustered contractions) [61], [86], [179], [217]-[218].

Έχει διαπιστωθεί ότι, κατά την εφαρμογή του φίλτρου Wiener στα διαστήματα σιγής, τα αποτελέσματα αποθορυβοποίησης δεν είναι τόσο ικανοποιητικά [177], [187], [189]. Ένα δυσάρεστο αποτέλεσμα, είναι η παρουσία σφαλμάτων αποθορυβοποίησης (de-noising artifacts), όπως είναι τα «υπολείμματα θορύβου» (noise residuals, “birdy noise”), τα οποία οφείλονται στις μικρές και τυχαίες ενεργειακές διακυμάνσεις του θορύβου [177], [179], [184], [187], [189]-[190]. Είναι φανερό ότι λόγω της απουσίας σήματος, δεν είναι δυνατή η ακουστική απόκρυψη (masking) των υπολειμμάτων θορύβου [188], [202]. Έτσι, οι ήχοι “birdy noise” είναι ιδιαίτερα ενοχλητικοί καθώς γίνονται εύκολα αντιληπτοί (είναι ακουστοί ακόμη και σε πολύ μικρές ηχητικές στάθμες), αλλά και εξαιτίας της εντελώς τυχαίας εμφάνισης και εξαφάνισής τους. Το ανεπιθύμητο αυτό φαινόμενο γίνεται ακόμη πιο έντονο στην περίπτωση των γαστρεντερικών ήχων, όπου μεταξύ των χρήσιμων ακουστικών συνιστωσών παρεμβάλλονται πολλά διαστήματα σιγής [179].

Για την αντιμετώπιση τέτοιου είδους προβλημάτων έχουν προταθεί διάφορες τεχνικές, όπως η περαιτέρω επεξεργασία του αποθορυβοποιημένου σήματος [190] και η εφαρμογή πιο αυστηρών κανόνων κατωφλίωσης, στοιχείο που συχνά οδηγεί και στην αλλοίωση χρήσιμων συνιστωσών. Κλασικά παραδείγματα της δεύτερης περίπτωσης αποτελούν οι τεχνικές σκληρής κατωφλίωσης (hard thresholding) και οι διαδικασίες επαναληπτικής (iterative) κατωφλίωσης [133], [183], [208]-[211], [221]. Για τον ίδιο λόγο, η παράμετρος κατωφλίωσης στην περίπτωση του αλγορίθμου $WDWF_I$ τέθηκε ίση με $c=3$, αντί της κλασσικής φασματικής αφαίρεσης, όπου $c=1$. Μια ποια ενδιαφέρουσα προσέγγιση για την αντιμετώπιση του συγκεκριμένου προβλήματος έγινε στην εργασία των McAulay και Malpass [189]. Οι τελευταίοι ανέπτυξαν έναν αλγόριθμο «απαλής απόφασης» (soft-decision algorithm), ο οποίος χρησιμοποιείται πριν την διαδικασία καταστολής θορύβου, παρέχοντας πληροφορίες για το αν το εκάστοτε χρονικό παράθυρο περιλαμβάνει πληροφοριακό σήμα.

Ειδικότερα, οι McAulay και Malpass διαπίστωσαν ότι το φίλτρο φασματικής αφαίρεσης δεν ήταν σε θέση να εκτιμήσει σωστά την παρουσία «μόνο θορύβου», έτσι ώστε να συμπίεσει εντελώς τις ενθόρυβες συνιστώσες, στις αντίστοιχες χρονικές στιγμές πάυσης-σιγής. Για το σκοπό αυτό πρότειναν μία ιεραρχική δομή ανάλυσης δύο καταστάσεων, όπου στο πρώτο επίπεδο επεξεργασίας εκτιμάται εάν το χρονικό παράθυρο επεξεργασίας περιλαμβάνει μόνο θόρυβο, ή σήμα και θόρυβο μαζί. Αν υπάρχει μόνο θόρυβος, οι αντίστοιχες συνιστώσες απορρίπτονται χωρίς περαιτέρω επεξεργασία. Στην περίπτωση που «αποφασίζεται παρουσία σήματος», τότε ενεργοποιείται το φίλτρο φασματικής αφαίρεσης. Ο αλγόριθμος απαλής απόφασης χρησιμοποιεί απριόρι λόγους σήματος προς θόρυβο σε όλες τις φασματικές συνιστώσες, για να εκτιμήσει τις πιθανότητες των τελευταίων να περιλαμβάνουν (ή όχι) χρήσιμο πληροφοριακό σήμα. Με βάση, λοιπόν, τις εξαρτώμενες από τις σηματοθορυβικές σχέσεις πιθανότητες, ο εκτιμητής μέγιστης πιθανοφάνειας (maximum likelihood estimator), παρέχει τη δυαδική έξοδο απόφασης: η κατάσταση $H_{bn}=1$ σημαίνει ότι το σήμα είναι ενεργό, ενώ η κατάσταση $H_{bn}=0$ υπονοεί περίοδο σιγής [179], [189].

Η παραπάνω προσέγγιση αποτέλεσε το κίνητρο για την υλοποίηση μιας εναλλακτικής μεθόδου $WDWF$, η οποία αποτελεί τον «τύπο-II κυματιδιακής αποθορυβοποίησης Wiener σημείο προς σημείο» - $WDWF_{II}$ [179]. Παρατηρώντας τη σχέση 3-37, η συνάρτηση μεταφοράς $H_{WDWF_{II}}$ είναι δυνατό να παραγοντοποιηθεί σε δύο όρους: στον όρο του φίλτρου «δυαδικής απόφασης» (binary decision filter) $H_{bn|I}$ και στον όρο του φίλτρου αποθορυβοποίησης (de-noising filter) $H_{dn|I}$. Η τροποποίηση αυτή γίνεται για λόγους ομοιομορφίας με το μοντέλο παρουσίας προσθετικού θορύβου της εξ. (3-41), αλλά και με την ιεραρχική προσέγγιση δύο επιπέδων των McAulay και Malpass:

$$\begin{aligned}
H_{WDWF|I}(k, w) &= H_{bn|I}(k, w) \cdot H_{dn|I}(k, w) \\
H_{bn|I}(k, w) &= \left\{ c \cdot A_{kw} \cdot P_{nFP,a}(k) \leq P_{x,a}(k, w) \Big|_I \right\} \\
H_{dn|I}(k, w) &= \left\{ 1 - c \cdot A_{kw} \cdot \left[\frac{P_{nFP,a}(k)}{P_{x,a}(k, w)} \Big|_I \right] \right\}^b
\end{aligned} \tag{3-42}$$

Οι δείκτες “I” στους παραπάνω όρους χρησιμοποιούνται για να δηλώσουν τον αντίστοιχο τύπο (I ή II) της κυματιδιακής επεξεργασίας *WDWF*, καθώς και των ενδιάμεσων φίλτρων H_{bn} και H_{dn} . Η ενσωμάτωση του ψυχοακουστικού όρου A_{kw} παρέχει και πάλι τη δυνατότητα επιλογής ή όχι των κριτηρίων απόκρυψης, οπότε το φίλτρο A_{kw} ενεργοποιείται (υπολογίζεται βάσει της εξ. 3-33) ή απενεργοποιείται ($A_{kw}=A_{k0}=1$), αντίστοιχα [179].

Είναι προφανές ότι το φίλτρο απόφασης $H_{bn|I}(k, w)$ είναι αρκετά ευαίσθητο στις τυχαίες μεταβολές του θορύβου, ειδικότερα για τις ρυθμίσεις $c=1$ και $A_{kw}=A_{k0}=1$. Αυτός είναι και ο λόγος που η φασματική αφαίρεση δεν αποδίδει το ίδιο ικανοποιητικά στις περιόδους σιγής, καθώς πολύ μικρές διακυμάνσεις μπορεί να θεωρηθούν ως σήμα. Η πιο απλή λύση στο παραπάνω πρόβλημα είναι να αυξηθεί η τιμή της παραμέτρου κατωφλίωσης c (όπως συνέβη και στην περίπτωση της μεθόδου *WDWF_I*, όπου $c=3$). Μια από τις αρνητικές επιπτώσεις της προηγούμενης διευθέτησης είναι η πιθανή αλλοίωση χρήσιμων συνιστωσών (όταν το σήμα είναι παρών), καθώς η καμπύλη απόρριψης του φίλτρου γίνεται πιο απότομη [177], [184], [187]. Επίσης, είναι δυνατόν να εμφανιστούν σφάλματα μετα-ηχούς (post-echo artifacts), γεγονός που συνδέεται με τη καθυστέρηση πλήρους απόσβεσης του θορύβου βάθους, αμέσως μετά το «πέρασ» του σήματος (οφείλεται στη «μνήμη» της διαδικασίας «εκθετικού μέσου όρου»). Παρά το γεγονός ότι τα φαινόμενα αυτά δεν γίνονται εύκολα ακουστά - αντιληπτά (στις περισσότερες περιπτώσεις ανιχνεύονται κυρίως γραφικά), ωστόσο η παρουσία τους μπορεί να αλλοιώσει τη μορφολογική δομή των αντίστοιχων κυματομορφών, με επιπτώσεις στα αποτελέσματα αυτοματοποιημένης ανάλυσης [179], [217].

Μια πιθανή λύση για την αντιμετώπιση των παραπάνω δυσκολιών είναι η χρήση μεταβλητών παραμέτρων κατωφλίωσης $c(k, w)$, ή μεταβλητών παραμέτρων «εκθετικής μνήμης» $d(k, w)$, οι οποίες θα μπορούσαν να προσαρμόζονται ανάλογα με τις καταστάσεις εξόδου του φίλτρου $H_{bn|I}(k, w)$. Βέβαια, η τροποποίηση αυτή θα είχε ως αποτελέσματα την αυξημένη πολυπλοκότητα της μεθόδου, γι’ αυτό και εγκαταλείφθηκε. Στη θέση της, προτάθηκε μια εναλλακτική προσέγγιση εκτίμησης της ισχύος $P_{x,a}(k, w)$ στο φίλτρο απόφασης $H_{bn}(k, w)$, η οποία είναι περισσότερο προσαρμοστική στις μεταβολές του «καθαρού» σήματος. Έτσι, αποφασίστηκε η αξιοποίηση των παρελθόντων «καθαρών» συντελεστών και η διαμόρφωση μιας νέας παραμέτρου ισχύος (“*type II a-powered estimation*”) $P_{x,a}(k, w) |_{II}$:

$$P_{x,a}(k, w) |_{II} = \begin{cases} |X(k, w)|^a, & w = 0 \\ d \cdot |X(k, w)|^a + (1-d) \cdot P_{s',a}(k, w-1), & w = 1, \dots, W_{Xk} - 1 \end{cases} \tag{3-43}$$

Η παράμετρος ισχύος $P_{x,a}(k, w) |_{II}$ που εισάγει η εξ. (3-43) είναι σχετική με τον τύπο II επεξεργασίας *WDWF* (“*type II a-powered estimation*”), ενώ η d_{ps} (σχ. 3-44) χρησιμοποιείται κατά αντιστοιχία με την γνωστή παράμετρο μνήμης d . Στην προκειμένη περίπτωση μας ενδιαφέρει η εξασφάλιση πολύ μικρής μνήμης για τον υπολογισμό της ισχύος του καθαρού

σήματος ($d_{PS} \rightarrow 1$). Έτσι, επιλέχθηκε η οριακή τιμή $d_{PS}=1$, η οποία εκτός από πολύ γρήγορη προσαρμογή, επιτρέπει την αποφυγή εισαγωγής μιας νέας παραμέτρου (της d_{PS}) [179]:

$$P_{s^-,a}(k,w) = d_{PS} \cdot |S^-(k,w)|^a + (1-d_{PS}) \cdot P_{s^-,a}(k,w-1), \quad w=1, \dots, W_{Xk}-1, \quad d_{PS} \rightarrow 1$$

ή, τελικά,

(3-44)

$$P_{s^-,a}(k,w) = |S^-(k,w)|^a, \quad (d_{PS} \cong 1)$$

Συγκρίνοντας τις εξισώσεις (3-36) και (3-44), και λαμβάνοντας υπόψη το γεγονός ότι ο προσθετικός θόρυβος ABN είναι ασυσχέτιστος με το πληροφοριακό σήμα, μπορούμε να επισημάνουμε τα εξής: α) και οι δύο εκτιμήσεις ισχύος $P_{x;a}(k,w)|_I$ και $P_{x;a}(k,w)|_{II}$ περιλαμβάνουν το τμήμα ενέργειας του σήματος (signal-part) και αυτό του θορύβου (noise-part), β) η διαδικασία εκθετικού μέσου που εισάγει ο τύπος II είναι πιο προσαρμοσμένη στην κυματομορφή του εκτιμώμενου «καθαρού» σήματος, και άρα η ισοδύναμη διάρκεια των «εκθετικών παραθύρων» είναι μικρότερη (δηλαδή η επίδραση των παρελθόντων ενθόρυβων δειγμάτων είναι πιο μικρή), γ) οι εκτιμώμενες τιμές ισχύος ικανοποιούν πάντα την ανισότητα: $P_{x;a}(k,w)|_{II} < P_{x;a}(k,w)|_I$. Αν τώρα χρησιμοποιήσουμε τον όρο $P_{x;a}(k,w)|_{II}$ αντί του $P_{x;a}(k,w)|_I$ στην εξ. (3-42β), τότε προκύπτει ο τύπος- II για το φίλτρο ανίχνευσης $H_{bn|II}(k,w)$:

$$H_{bn|II}(k,w) = \left\{ c \cdot A_{kw} \cdot P_{nFP;a}(k) \leq P_{x;a}(k,w)|_{II} \right\} \quad (3-45)$$

Η προηγούμενη σχέση παρουσιάζει αρκετές λειτουργικές ομοιότητες με τον αλγόριθμο απαλής απόφασης των McAulay και Malpass. Το φίλτρο εντοπισμού $H_{bn|II}(k,w)$ παρέχει στην έξοδο του την πιθανότητα για τις επερχόμενες χρονικές στιγμές να περιλαμβάνουν σήμα ή θόρυβο, λαμβάνοντας υπόψη τις προηγούμενες καταστάσεις (παρουσίας ή όχι σήματος), σε συνδυασμό με τα ενεργειακά επίπεδα του θορύβου. Υπό την έννοια αυτή, είναι προφανές ότι οι πιθανότητες να υπάρχουν χρήσιμες πληροφοριακές συνιστώσες αυξάνονται όταν η προηγούμενη κατάσταση στην έξοδο του φίλτρου είναι ενεργή (παρουσία σήματος), απ' ότι στην αντίθετη περίπτωση. Με άλλα λόγια, το φίλτρο παρέχει τη δεσμευμένη πιθανότητα της λογικής κατάστασης εντοπισμού σήματος, υπό τη συνθήκη της ακριβώς προηγούμενης κατάστασης. Αυτό συμβαίνει καθώς, από τη στιγμή που θα παρουσιαστεί κάποιο σήμα, τότε θα πρέπει να περάσει ένα ελάχιστο χρονικό διάστημα που σχετίζεται με τους χρόνους ανόδου - καθόδου (attack - release cycle) της αντίστοιχης καμπύλης, προτού εμφανιστεί και πάλι διάστημα σιγής (απουσίας σήματος).

Η χρήση της παραμέτρου d στη σχέση (3-45) επιτρέπει το συνδυασμό παρελθόντων αποθορυβοποιημένων συντελεστών με τα τρέχοντα ενθόρυβα δείγματα, έτσι ώστε να αποφεύγονται οι εσφαλμένες αποφάσεις κατά την έναρξη και τον τερματισμό των ακουστικών φαινομένων, δηλαδή στα όρια των «ενεργών τμημάτων». Η παράμετρος μνήμης d χρησιμοποιείται για να ρυθμίσει τις αντίστοιχες κατανομές πυκνότητας πιθανότητας (probability density functions -pdf) του φίλτρου απόφασης, ενώ η φυσική της ερμηνεία είναι εντελώς διαφορετική από ότι στην περίπτωση του τύπου I (δεν είναι πλέον η κλασική παράμετρος «εκθετικής μνήμης») [179]. Στην προκειμένη περίπτωση (τύπος II) λειτουργεί ως ένας ρυθμιστικός παράγοντας ισορροπίας που υπεισέρχεται στον υπολογισμό της παραμέτρου ισχύος $P_{x;a}(k,w)|_{II}$, βάσει της οποίας προκύπτει η απόφαση παρουσίας ή απουσίας σήματος [179].

Λαμβάνοντας υπόψη τα παραπάνω, είναι εύκολο να σχηματιστεί ένα υβριδικό σύστημα το οποίο να περιλαμβάνει τα φίλτρα $H_{bn|II}(k,w)$ και $H_{dn|I}(k,w)$, προσαρμόζοντας την πρόταση των McAulay και Malpass στο Wiener φίλτρο κυματιδιακής αποθορυβοποίησης ($WDWF_{I\&II}$). Σύμφωνα δε με την προηγούμενη ανάλυση, ο τύπος II του φίλτρου απόφασης $H_{bn|II}(k,w)$ είναι πιο αποτελεσματικός στον εντοπισμό και την απόρριψη των τμημάτων σιγής, απ' ό,τι ο τύπος I . Επιπρόσθετα, οι συνιστώσες χαμηλού ενεργειακού επιπέδου που συνήθως αποκρύπτονται ακουστικά από το θόρυβο, είναι πολύ πιθανό να απορριφθούν από το φίλτρο $H_{bn|II}(k,w)$, ενσωματώνοντας πρόσθετα «αντιληπτικά» (perceptual) κριτήρια. Η συμπεριφορά αυτή βρίσκεται στο ίδιο μήκος κύματος με τη λειτουργία του όρου A_{kw} , καθώς και οι δύο βασίζονται στο ίδιο ψυχο-ακουστικό φαινόμενο: «οι αδύναμες συνιστώσες σήματος που συναντώνται μέσα σε ενθόρυβες ακολουθίες, συνήθως αποκρύπτονται ακουστικά (maskee) από το θόρυβο (masker)». Έτσι, είναι προτιμητέα η πλήρης απόρριψη των μη ακουστών συνιστωσών, από την διατήρηση κάποιων πολύ ασθενικών σημάτων μαζί με ανεπιθύμητα υπολείμματα θορύβου [177], [179], [181], [186]-[188].

Το φίλτρο αποθορυβοποίησης $H_{dn|I}(k,w)$ της υβριδικής μεθόδου $WDWF_{I\&II}$ παραμένει ως έχει, όπως στην περίπτωση του τύπου I ($WDWF_I$). Μάλιστα, η τροποποίηση του φίλτρου απόφασης επιτρέπει τη ρύθμιση της παραμέτρου κατωφλίωσης σε χαμηλότερα επίπεδα ($c=1$), όπως συμβαίνει στην κλασική φασματική αφαίρεση. Δεδομένου ότι υπάρχουν συνιστώσες σήματος την τρέχουσα χρονική στιγμή, τα τυχόν υπολείμματα θορύβου (low-level noise residuals) που ενδέχεται να προκύψουν από την ηπιότερη κατωφλίωση, συμπίπτουν χρονικά με τις χρήσιμες συνιστώσες, με αποτέλεσμα να αποκρύπτονται ακουστικά από τις τελευταίες. Έτσι, η παρουσία τους γίνεται πολύ δύσκολα αντιληπτή, τόσο ακουστικά όσο και οπτικά, επηρεάζοντας ελάχιστα τα αποτελέσματα ανάλυσης [179]. Η κύρια διαφορά με την αντιληπτική προσέγγιση A_{kw} είναι ότι οι ρόλοι των συνιστωσών σήματος και θορύβου αντιστρέφονται, καθώς ο θόρυβος (maskee) αποκρύπτεται πλέον από το πληροφοριακό σήμα (masker). Η εφαρμογή ηπιότερης κατωφλίωσης παρέχει πρόσθετα πλεονεκτήματα, καθώς είναι λιγότερο πιθανό να οδηγήσει σε παραμορφώσεις του πληροφοριακού σήματος. Η νέα υβριδική μέθοδος που συνδυάζει τα φίλτρα $H_{bn|II}(k,w)$ και $H_{dn|I}(k,w)$ αποτελεί τον τύπο $I\&II$ κυματιδιακής επεξεργασίας Wiener ($WDWF_{I\&II}$) και περιγράφεται από τη σχέση:

$$H_{WDFW|I\&II}(k,w) = H_{bn|II}(k,w) \cdot H_{dn|I}(k,w) \quad (3-46)$$

όπου τα φίλτρα $H_{bn|II}(k,w)$ και $H_{dn|I}(k,w)$ δίνονται από τις εξισώσεις (3-46) και (3-43), αντίστοιχα. Μια διαφορά ανάμεσα στους τύπους I και $I\&II$ έγκειται στο γεγονός ότι οι παράμετροι a , b , c , d μπορούν πλέον να διαμορφωθούν - ρυθμιστούν αυτόνομα για κάθε φίλτρο (a_I , b_I , c_I , d_I , και a_{II} , b_{II} , c_{II} , d_{II}). Ωστόσο, είναι δυνατή η χρήση κοινών παραμέτρων έτσι ώστε να αποφευχθεί η εισαγωγή πρόσθετων μεταβλητών εισόδου. Με βάση εμπειρικές παρατηρήσεις και λαμβάνοντας υπόψη τους περιορισμούς (για την παράμετρο d) που συζητήθηκαν στην §3.2.5.1, η βέλτιστη ενιαία ρύθμιση του φίλτρου $WDWF_{I\&II}$ οδήγησε στην επιλογή των ακόλουθων παραμέτρων: $a=2$, $b=1$, $c=1$ and $d=0,1$ [179]. Το υβριδικό μοντέλο $WDWF_{I\&II}$ παρέχει ικανοποιητικά αποτελέσματα καταστολής θορύβου και στις δύο καταστάσεις απόφασης εντοπισμού (είτε υπάρχει σήμα, είτε όχι). Ωστόσο, θα μπορούσε κανείς να θεωρήσει ως μειονέκτημα το γεγονός ότι εισάγονται περισσότεροι παράμετροι που χρίζουν ρύθμισης (στη γενική περίπτωση), ενώ παράλληλα απαιτούνται και περισσότερες πράξεις για τον υπολογισμό των δύο παραμέτρων ισχύος. Έτσι, αποφασίστηκε η χρήση της $P_{x,a}(k,w)|_{II}$, τόσο για το φίλτρο απόφασης $H_{bn-II}(k,w)$, όσο και για το φίλτρο καταστολής θορύβου $H_{dn-II}(k,w)$, με αποτέλεσμα τη διαμόρφωση του τύπου II ($WDWF_{II}$) [179]:

$$H_{WDWF|II}(k, w) = H_{bn|II}(k, w) \cdot H_{dn|II}(k, w) \quad (3-47)$$

$$H_{dn|II}(k, w) = \left\{ 1 - c \cdot A_{kw} \cdot \left[\frac{P_{nFP;a}(k)}{P_{x;a}(k, w)|_{II}} \right] \right\}^b$$

Με βάση τις σχέσεις (3-46) και (3-47), ο αλγόριθμος $WDWF_{II}$ διαφέρει από το αντίστοιχο υβριδικό σύστημα $WDWF_{I\&II}$ μόνο στο φίλτρο αποθορυβοποίησης $H_{dn}(k, w)$, το οποίο ενεργοποιείται κατά την «απόφαση παρουσίας σήματος». Από τις εξ. (3-36), (3-42), (3-43) και (3-47), είναι προφανές ότι, για τις ίδιες τιμές των παραμέτρων a , b , c και d , ο τύπος- II του φίλτρου $H_{dn|II}(k, w)$ εισάγει πιο «αυστηρούς» κανόνες κατωφλίωσης (πιο απότομες καμπύλες) σε σχέση με τον τύπο I , δεδομένου ότι $P_{x;a}(k, w)|_{II} < P_{x;a}(k, w)|_I$. Ωστόσο, στην κατάσταση παρουσίας σήματος, οι χρήσιμες συνιστώσες έχουν μεγαλύτερο ενεργειακό περιεχόμενο από τις αντίστοιχες του θορύβου, κι έτσι καθορίζουν σε μεγαλύτερο βαθμό την εκτιμώμενη ισχύ $P_{x;a}(k, w)$. Δηλαδή, η εκτίμηση ισχύος των ενθόρυβων συντελεστών με βάση τον τύπο II , πλησιάζει την αντίστοιχη εκτίμηση του τύπου I , όταν η κατάσταση παρουσίας σήματος είναι ενεργή. Αν συγκρίνουμε την τελικά διαμορφωμένη μέθοδο $WDWF_{II}$ (με $c=1$) με την αντίστοιχη $WDWF_I$ ($c=3$), η $WDWF_{II}$ παρέχει ηπιότερη επεξεργασία κατωφλίωσης όταν υπάρχουν συνιστώσες σήματος, και αντίστροφα αυστηρότερους κανόνες απόρριψης όταν το σήμα βρίσκεται σε επίπεδα πολύ χαμηλότερα από το θόρυβο (buried below noise levels) ή απουσιάζει εντελώς [179]. Σε κάθε περίπτωση, είναι δυνατή η διαμόρφωση μιας γενικής έκφρασης που να περιγράφει την κυματιδιακή επεξεργασία $WDWF$ (εξ. 3-48), η οποία ισχύει και για τους δύο τύπου (I και II), με διαφορετικές κάθε φορά εκτιμήσεις της παραμέτρου ισχύος $P_{x;a}(k, w)$. Ο τύπος $I\&II$ αποτέλεσε μια ενδιάμεση κατάσταση στην προσπάθεια ανάπτυξης της μεθόδου $WDWF_{II}$.

$$H_{WDWF}(k, w) = \begin{cases} \left\{ 1 - c \cdot A_{kw} \cdot \left[\frac{P_{nFP;a}(k)}{P_{x;a}(k, w)} \right] \right\}^b, & \text{αν } c \cdot A_{kw} \cdot \left[\frac{P_{nFP;a}(k)}{P_{x;a}(k, w)} \right] \leq 1 \\ 0, & \text{διαφορετικά} \end{cases} \quad (3-48)$$

Η ανάπτυξη της μεθόδου $WDWF_{II}$, αν και βασίστηκε κατά κύριο λόγο σε εμπειρικές παρατηρήσεις και παραδοχές, διαθέτει κάποια μοναδικά χαρακτηριστικά [179]:

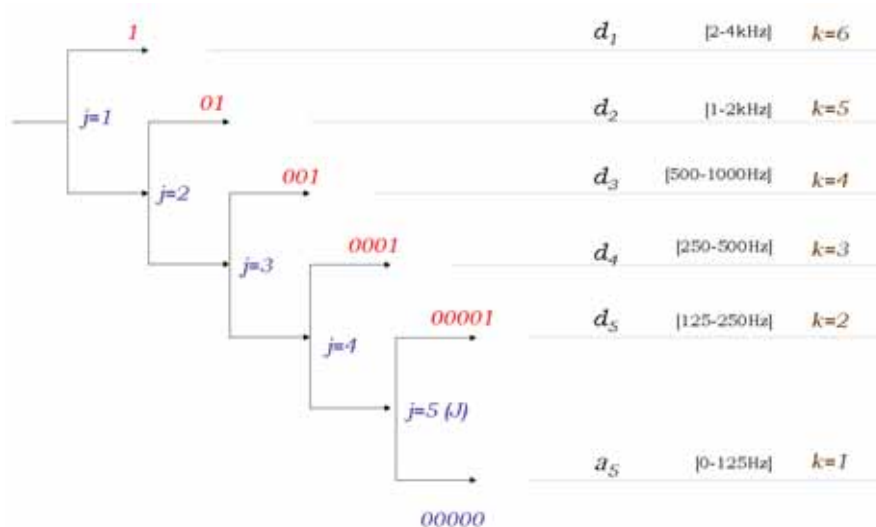
- α) εξακολουθεί να είναι ιδιαίτερα απλή και βολική, τόσο λόγω της περιορισμένης πολυπλοκότητας υλοποίησης, όσο και λόγω του σχετικά μειωμένου υπολογιστικού κόστους,
- β) ενσωματώνει την προσέγγιση του φίλτρου απαλής απόφασης που προτάθηκε από τους McAulay και Malpass [189], για την αποδοτικότερη αντιμετώπιση των περιόδων σιγής,
- γ) ενσωματώνει χαρακτηριστικά αντιληπτικής επεξεργασίας (perceptual), εξαλείφοντας τις αδύναμες συνιστώσες σήματος που συνήθως αποκρύπτονται από το θόρυβο,
- δ) προσαρμόζει συνεχώς (σημείο προς σημείο) τη συνάρτηση μεταφοράς του φίλτρου κι έτσι δεν χρειάζεται περαιτέρω επεξεργασία για τον περιορισμό ενδεχομένων σφαλμάτων και υπολειμμάτων αποθορυβοποίησης [190], ή για την αποφυγή απότομων μεταβάσεων [219].

Η συνάρτηση μεταφοράς του φίλτρου $WDWF_{II}$ αποτελεί μια αυτοπαλινδρομική διαδικασία πρώτης τάξης (first order Autoregressive model - AR) ως προς τη μεταβλητή

στιγματίας a -ιοχύος $|X(k,w)|^a$, καθώς η έξοδος του φίλτρου προκόπεται από την τρέχουσα τιμή εισόδου και την αμέσως προηγούμενη τιμή εξόδου. Ο εμπλεκόμενος βρόχος ανάδρασης οδηγεί σε περαιτέρω «συμπίεση» (shrinkage) των κυματιδιακών συντελεστών, κι έτσι δεν εμπεριέχονται κίνδυνοι οδήγησης σε αστάθεια, συμπεριφορά που συναντάται συχνά στις AR διαδικασίες. Η ρύθμιση των παραμέτρων του φίλτρου $WDWF_{II}$ βασίστηκε σε εμπειρικές παρατηρήσεις αλλά και στις θεωρητικές παραδοχές που σχετίζονται με τη συμπεριφορά του υβριδικού τύπου $WDWF_{I&II}$. Τα σχετικά αποτελέσματα οδήγησαν στην επιλογή των τιμών $a=2$, $b=1$, $c=1$ και $d=0,1$ οι οποίες είναι οι ίδιες με τις ρυθμίσεις της προσέγγισης $WDWF_{I&II}$ [179]. Η επαλήθευση της ορθότητας όλων των παραπάνω επιλογών και γενικότερα η αξιολόγηση των δύο τύπων επεξεργασίας $WDWF$, αναλύονται στην §3.2.6 του παρόντος.

3.2.5.3 Επεξεργασία $WDWF$ μέσω διακριτού μετασχηματισμού κυματιδίων 6 ζωνών

Το ζήτημα επιλογής της «βέλτιστης» τοπολογίας ανάλυσης - σύνθεσης στην ανάλυση DWT και WP είναι συνηθισμένο στις κυματιδιακές τεχνικές. Για τη διευθέτησή του έχουν δοκιμαστεί διάφορες συναρτήσεις «κόστους» -“cost functions” (εκφράσεις εντροπίας, αντιληπτικά κριτήρια, κ.ά.), οι οποίες αποτιμούν τη συνεισφορά των κόμβων ανάλυσης στο συνολικό «κυματιδιακό δένδρο» και αποφαινόνται για τη βέλτιστη δομή [208]-[210]. Μια άλλη προσέγγιση αφορά στη χρήση του βέλτιστου δυνατού αριθμού επιπέδων ανάλυσης DWT , όπως προκόπεται από τη εξ. (3-29) [194]. Η διευθέτηση αυτή, η οποία εξαρτά τα επίπεδα ανάλυσης από το μήκος N του σήματος εισόδου, έχει αξιοποιηθεί στις μεθόδους $WTST-NST$ και $WT-FD$ για την επεξεργασία πνευμονικών και εντερικών ήχων [59], [83], [84], [133]. Ωστόσο, η υιοθέτηση της πιο πάνω λογικής οδηγεί μάλλον σε υπερβολικές αναλύσεις, καθώς για μια ακολουθία μήκους 2.048 δειγμάτων ορίζονται 11 επίπεδα κυματιδιακής ανάλυσης. Δεδομένης της δειγματοληψίας ($f_s=8kHz$), πολλές από τις σχηματιζόμενες φασματικές ζώνες βρίσκονται σε συχνότητες κάτω των 100 Hz, όπου η γαστρεντερική ηχητική δραστηριότητα είναι σχεδόν ανύπαρκτη [179]. Τα παραπάνω στοιχεία επιβαρύνουν υπολογιστικά τη διαδικασία (ιδιαίτερα για σήματα μεγαλύτερης διάρκειας) χωρίς σημαντικό όφελος [179], [212]. Έτσι, στη θέση της προτιμήθηκε η σταθερή τοπολογία ανάλυσης του σχ. 3-15, που αποτελεί έναν κλασσικό μετασχηματισμό DWT 5 επιπέδων (6 ζώνες ανάλυσης).

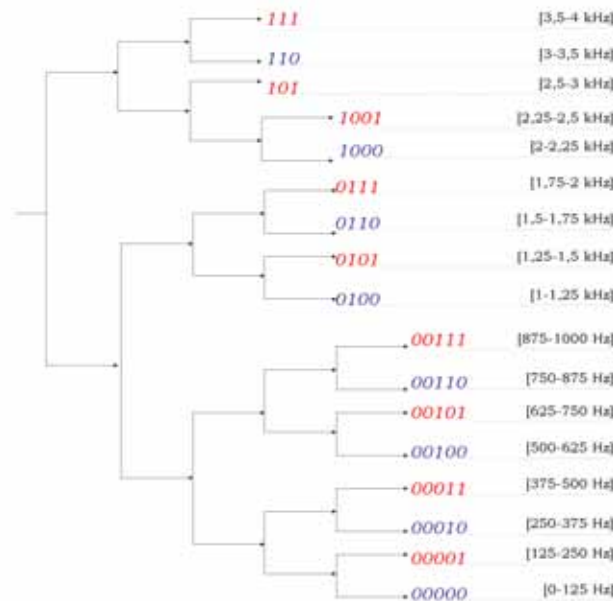


Σχήμα 3-15. Η τοπολογία DWT -6 ζωνών που χρησιμοποιήθηκε στις μεθόδους $WDWF$.

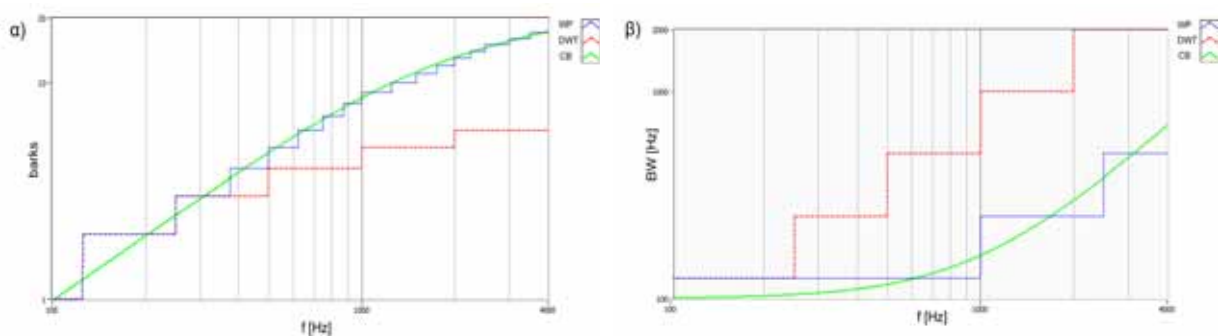
Η τοπολογία *DWT-6* είναι ισοδύναμη με την κλασσική οκταβική ανάλυση στην περιοχή [0-4kHz], και αποδείχθηκε ότι παρέχει την αναγκαία φασματική διακριτικότητα, για την επεξεργασία αποθορυβοποίησης εντερικών ήχων. Κατά τη διαδικασία ρύθμισης, προέκυψαν οι τιμές $a=2$, $b=1$, $c=3$ και $d=0,2$ για τον τύπο I ($WDWF_{I-6}$) και $a=2$, $b=1$, $c=1$ και $d=0,1$ για τον τύπο II ($WDWF_{II-6}$), ενώ επιλέχθηκε το μητρικό κοματιδίδια *Daubechies 6*. Η ενσωμάτωση του ψυχοακουστικού φίλτρου A_{kw} αποδείχθηκε ιδιαίτερα χρήσιμη, παρέχοντας πολύ καλά ακουστικά αποτελέσματα (audible results), με μεγάλη ανοχή σε σφάλματα και υπολείμματα θορύβου [179]. Η επαλήθευση της ορθότητας των προηγούμενων ρυθμίσεων δίνονται στην επόμενη ενότητα αξιολόγησης των μεθόδων *WDWF*.

3.2.5.4 Επεξεργασία *WDWF* μέσω πακέτων κυματιδιακής ανάλυσης 17 ζωνών

Αν και οι τεχνικές $WDWF_{I-6}$, $WDWF_{II-6}$ αποδείχτηκαν ιδιαίτερα αποδοτικές, ωστόσο, δοκιμάστηκε μια εναλλακτική τοπολογία κυματιδιακών πακέτων (WP) 17 ζωνών (σχ. 3-16), για δύο λόγους: α) οι τεχνικές φασματικής αφαίρεσης λειτουργούν καλύτερα όσο μεγαλώνει ο αριθμός των ζωνών ανάλυσης [177], [187], β) η επιλογή μιας τοπολογίας που να προσομοιώνει καλύτερα τα χαρακτηριστικά της ανθρώπινης ακοής θα παρείχε δυνατότητες καλύτερης ακροαστικής ανάλυσης [179]. Σε σχέση με την τοπολογία *DWT-6*, η νέα δομή *WP-17* προσεγγίζει καλύτερα τα χαρακτηριστικά της ανάλυσης κρισίμων ζωνών (σχ. 3-17).



Σχήμα 3-16. Η τοπολογία *WP-17* ζωνών που χρησιμοποιήθηκε στις μεθόδους *WDWF*.



Σχήμα 3-17. Σύγκριση της ανάλυσης κρισίμων ζωνών (CB) με τα σχήματα «6-ζωνών *DWT*» και «17-ζωνών *WP*», των μεθόδων *WDWF*: α. αριθμός διαμορφωμένων ζωνών (barks), β. αντίστοιχο φασματικό εύρος (BW).

Οι δύο νέες τεχνικές αποθορυβοποίησης $WDWF_{I-17}$ και $WDWF_{II-17}$ παρέχουν καλύτερα ακουστά αποτελέσματα, ιδιαίτερα στις περιπτώσεις που τα χρήσιμα σήματα έχουν στενό εύρος ζώνης (narrow-banded). Στον αντίποδα, το νέο «δένδρο» κυματιδιακής ανάλυσης έχει περισσότερες υπολογιστικές απαιτήσεις. Έτσι, η τοπολογία 17 ζωνών προτείνεται κυρίως για επεξεργασία απομονωμένων βιοακουστικών συμβάντων (short-term analysis) και όχι τόσο για καταγραφές παρατεταμένης διάρκειας. Από τις αντίστοιχες διαδικασίες ρύθμισης των παραμέτρων, προέκυψαν οι τιμές $a=2, b=1, c=3, d=0,2$ για τον τύπο I και $a=2, b=1, c=1, d=0,1$ για τον τύπο II, δηλαδή όπως και στην περίπτωση της τοπολογίας «5-ζωνών DWT». Τα ίδια συμπεράσματα προέκυψαν σχετικά με την επιλογή του μητρικού κυματιδίου (*Daubechies 6*) και τη συμπεριφορά του ψυχοακουστικού φίλτρου A_{kw} [179].

3.2.5.5 Επεξεργασία WDWF μέσω αυτόματου εντοπισμού κατωφλίων

Οι τεχνικές αυτόματου εντοπισμού κατωφλίων δεν αποτέλεσαν βασικό αντικείμενο έρευνας των τεχνικών $WDWF$. Εξάλλου, η εισαγωγή του τύπου II διευθέτησε σε ικανοποιητικό βαθμό τα τυχόν προβλήματα που ανακύπτουν λόγω της απόκλισης από τα χαρακτηριστικά «πλήρους» στασιμότητας του θορύβου. Ωστόσο, για λόγους πληρότητας του όλου ερευνητικού εγχειρήματος, καθώς για την περαιτέρω διευκόλυνση του ιατρικού προσωπικού κατά την χρήση των φίλτρων $WDWF$ σε ακροαστικές αναλύσεις μεμονωμένων συμβάντων (short-term), δοκιμάστηκαν και κάποιες προσεγγίσεις αυτοματοποιημένης κατωφλίωσης. Στο πλαίσιο αυτό, υιοθετήθηκαν προσεγγίσεις «εμπειρικής επεξεργασίας Wiener» [222], όπου αξιοποιούνται οι τεχνικές αυτόματης κυματιδιακής κατωφλίωσης, σε συνδυασμό με το φίλτρο Wiener. Πρόκειται για μια ιεραρχική επεξεργασία όπου, σε πρώτη φάση χρησιμοποιείται μία από τις υπάρχουσες τεχνικές αυτόματου εντοπισμού κατωφλίων $th(k)$, και στη συνέχεια ακολουθεί η επεξεργασία $WDWF$, θέτοντας τις τιμές των κατωφλίων στα αντίστοιχα επίπεδα θορύβου " $N_{FP}(k) ← th(k)$ ".

Στο πλαίσιο αυτό, χρησιμοποιήθηκαν οι τεχνικές αυτόματης κατωφλίωσης $RIGSURE$ [192]-[193], αλλά και η στατιστική εκτίμηση της μεθόδου $WTST-NST$ (εξ. 3-28). Όσον αφορά την τελευταία, ο ρυθμιστικός παράγοντας επιλέχθηκε ίσος με $F_{adj}=3$, ενώ η παράμετρος $\sigma_{k,l}$ προκύπτει πλέον από τη διασπορά των ενθόρυβων κυματιδιακών συντελεστών ($\sigma_{k,l}=\sigma_{k,0}=\sigma_k$). Και οι δύο προηγούμενες προσεγγίσεις δίνουν ικανοποιητικά αποτελέσματα στην πλειονότητα των περιπτώσεων, αρκεί να μην υπάρχουν ασυνέχειες (interruptions) των σημάτων. Αυτό σημαίνει ότι, κατά την έναρξη και λήξη των απομονωμένων συμβάντων θα πρέπει να υπάρχουν περίοδοι σιγής, δηλαδή διαστήματα με θόρυβο ABN (noise overheads). Διαφορετικά, οι τεχνικές αυτόματης εκτίμησης των κατωφλίων ενδέχεται να οδηγήσουν σε σημαντικά λάθη (ειδικά για την περίπτωση της εξ. 3-28). Σε μια προσπάθεια διευθέτησης αυτού του προβλήματος, δοκιμάστηκαν στατιστικές τεχνικές ενεργειακής ταξινόμησης τύπου "penalize low". Δεδομένου ότι το πληροφοριακό σήμα καταλαμβάνει ένα περιορισμένο τμήμα του συνολικού χρόνου παρατήρησης, μπορούμε να επιλέξουμε τους $W_{min}(k)$ κυματιδιακούς συντελεστές με το μικρότερο πλάτος (σε κάθε ζώνη k), οι οποίοι λογικά θα αντιστοιχούν σε θόρυβο. Από τις μέσες ενεργειακές τιμές αυτών των συντελεστών καθορίζονται τα επίπεδα κατωφλίωσης (κατά αντιστοιχία με την εξ. 3-31), και στη συνέχεια εφαρμόζονται οι σχέσεις κυματιδιακής κατωφλίωσης $WDWF$. Η ρύθμιση των παραμέτρων $W_{min}(k)$ σε τιμές μεταξύ 20% και 50% του συνολικού μήκους κάθε κυματιδιακής ακολουθίας, αποδείχθηκε ότι λειτουργεί αρκετά αποτελεσματικά.

Η πιο πάνω προσέγγιση ταξινόμησης (ordering) των κυματιδιακών συντελεστών, αποτέλεσε κίνητρο για την ανάπτυξη μιας ακόμη εναλλακτικής μεθόδου. Η τελευταία

ανέκυψε στην προσπάθεια να αποφευχθούν λανθασμένες ενεργειακές εκτιμήσεις, στην περίπτωση που η επιλογή των «χαμηλότερων» συντελεστών δεν αποδίδει την πλήρη διακόμανση του θορύβου. Έτσι, αποφασίστηκε η αξιοποίηση κινούμενων παραθύρων επιλογής, έτσι ώστε οι $W_{min}(k)$ συνεχόμενοι συντελεστές με τη μικρότερη ενέργεια, να καταχωρούνται ως τμήμα θορύβου-σιγής, βάσει των οποίων γίνεται στη συνέχεια η εκτίμηση των κατωφλίων. Μάλιστα, η συγκεκριμένη προσέγγιση μπορεί να εφαρμοστεί απευθείας στο πεδίο του χρόνου, έτσι ώστε να εντοπιστεί το αποτύπωμα θορύβου $n_{FP}(i)$ και στη συνέχεια να ακολουθηθεί η γνωστή διαδικασία επεξεργασίας WDWF.

Σε επέκταση των παραπάνω, εφαρμόστηκαν και πιο «δυναμικές» τεχνικές εντοπισμού των περιοχών θορύβου, όπου εκτός από τις ενεργειακές συγκρίσεις αξιοποιήθηκαν πρόσθετοι παράμετροι κρουστικότητας, «πολυπλοκότητας κυματομορφών» και δυναμικής περιοχής των επιλεγμένων τμημάτων. Βέβαια, οι τεχνικές αυτές είναι συχνά επιβαρυντικές σε ζητήματα υλοποίησης και υπολογιστικού κόστους, και γι' αυτό δεν έτυχαν πλήρους διερεύνησης - αξιολόγησης. Σε κάθε περίπτωση, η περαιτέρω εξέλιξή τους αποτελεί αντικείμενο διερεύνησης, καθώς αναμένεται να παράσχουν νέες δυνατότητες επεξεργασίας.

3.2.6 Αξιολόγηση των τεχνικών Wavelet Domain Wiener Filtering (WDWF)

Για τις ανάγκες ρύθμισης και αξιολόγηση των μεθόδων WDWF αξιοποιήθηκαν κυρίως ενθόρυβοι κουλιακοί ήχοι, που επιλέχθηκαν από τις διαθέσιμες καταγραφές παρατεταμένου χρόνου. Επιπρόσθετα, χρησιμοποιήθηκαν και ντετερμινιστικά σήματα δοκιμής (test signals), όπως είναι οι κλασσικοί τόνοι σταθερής και μεταβαλλόμενης συχνότητας (sweeps, chirp signals). Ωστόσο, τα σήματα αυτά έχουν τεράστιες μορφολογικές διαφορές με τις κυματομορφές των γαστρεντερικών ήχων, και γι' αυτό δεν μπορούν να παράσχουν αντιπροσωπευτικά αποτελέσματα αξιολόγησης. Έτσι, στη θέση τους χρησιμοποιήθηκαν τόνοι διαμορφωμένοι με γκαουσιανούς φακέλους, ή αθροίσματα των παραπάνω σημάτων ("sums of weighted linear adaptive modulated Gaussian functions"), που αποτέλεσαν τους «συνθετικούς κουλιακούς ήχους» (Synthetic AS - s_{SAS}):

$$s_{SAS}[i] = \sum_{g=0}^{G-1} C_g \cdot tn_g[i], \quad (3-49)$$

$$tn_g[i] = (e_g \pi)^{-0.25} \exp \left\{ -\frac{[i - i_g]^2}{2e_g} + j(2\pi\theta_g[i - i_g]) \right\}$$

Οι παραπάνω εξισώσεις περιγράφουν την «Προσαρμοστική Αναπαράσταση» (Adaptive Representation), που αποτελεί μια από τις πιο γνωστές τεχνικές συνδυασμένης ανάλυσης χρόνου συχνότητας (Joint Time Frequency Analysis - JTFA), καθώς παρουσιάζει πολύ καλά χαρακτηριστικά διακριτικότητας [193], [196], [223]. Η αντίστροφη διαδικασία, που αφορά στην «προβολή» ενός φυσικού σήματος στις συνιστώσες της μορφής $tn_g(i)$ (δηλαδή στην εύρεση των G , e_g , i_g , και θ_g), είναι γνωστή ως «Προσαρμοστικός Μετασχηματισμός» - "Adaptive Transform". Οι καμπύλες s_{SAS} αποτελούν καλύτερες προσεγγίσεις των κουλιακών ήχων σε σχέση με τους κλασσικούς τόνους ή τις σαρώσεις τόνων (ιδιαίτερα για τα κρουστικά συμβάντα), και γι' αυτό χρησιμοποιήθηκαν κατά τη φάση ρύθμισης των φίλτρων WDWF, και στην αξιολόγηση της επίδρασης των a , b , c , d . Στο πλαίσιο αυτό, ιδιαίτερα βολική υπήρξε η δυνατότητα καθορισμού διαφόρων χαρακτηριστικών των σημάτων s_{SAS} (πλήθος και

ενεργειακό περιεχόμενο των φασματικών συνιστωσών, χρονική στιγμή εμφάνισης, διάρκεια και κρουστικότητα των επιμέρους συμβάντων), από τη ρύθμιση των παραμέτρων G , θ_g , i_g , e_g . Αν και τα στοιχεία νομοτέλειας των ντετερμινιστικών σημάτων $SSAS$ λειτουργούν επιβαρυντικά ως προς τα αποτελέσματα της επεξεργασίας $WDWF$, η ενσωμάτωσή τους αποφασίστηκε στη φιλοσοφία του χειρίστου σεναρίου (worst case scenario) [179].

Εκτός από την παραπάνω διαδικασία πιστοποίησης, η ποιοτική και ποσοτική αξιολόγηση των τεχνικών $WDWF$ βασίστηκε στα αποτελέσματα επεξεργασίας φυσικών κοιλιακών ήχων, με στόχο μια πιο αντικειμενική - αξιόπιστη ανάλυση. Κατά τη διαδικασία ποσοτικής αξιολόγησης, επιλέχθηκαν ενθόρυβοι κοιλιακοί ήχοι υψηλής σηματοθορυβικής σχέσης (>10-15dB) και διάφορων μορφολογικών χαρακτηριστικών, οι οποίοι στη συνέχεια αποθορυβοποιήθηκαν ώστε να αποτελέσουν τα «καθαρά» (De-noised), σήματα δοκιμής $SDAS$. Η καταστολή του θορύβου έγινε με κλασσικές τεχνικές φασματικής αφαίρεσης (όχι $WDWF$), οι οποίες είναι ιδιαίτερα αποδοτικές στις παραπάνω συνθήκες SNR [179].

3.2.6.1 Έλεγχος της αποδοτικότητας των επιλεγμένων παραμέτρων επεξεργασίας

Η χρήση των φυσικών και ντετερμινιστικών σημάτων δοκιμής που περιγράφηκαν πιο πάνω, έδωσε τη δυνατότητα διερεύνησης των αποτελεσμάτων αποθορυβοποίησης, ανάλογα με τα χαρακτηριστικά των πληροφοριακών σημάτων και του προσθετικού θορύβου. Η διαδικασία αυτή εφαρμόστηκε για διάφορες ρυθμίσεις των παραμέτρων c , d (για τις άλλες δύο επιλέχθηκαν οι κλασσικές τιμές $a=2$, $b=1$) και λειτούργησε ως βρόχος ανατροφοδότησης για τη διαμόρφωση και τελική ρύθμιση των τεσσάρων φίλτρων $WDWF$. Στα σχ. 3-18 και 3-19 παρουσιάζονται τα αποτελέσματα της επεξεργασίας $WDWF$, κατά την εφαρμογή της σε σήματα δοκιμής $SSAS$ «μολυσμένων» με λευκό γκαουσιανό θόρυβο (GN). Το πρώτο σχήμα αφορά την ανάλυση στο πεδίο του χρόνου (Time Domain -TD), ενώ στο δεύτερο αποτυπώνονται σπεκτρογραφικοί πίνακες $X_w(t,k)$ χρόνου - συχνότητας που προέκυψαν με το μετασχηματισμό $STFT$. Εκτός από το θόρυβο GN , δοκιμάστηκαν και άλλα προφίλ, όπως είναι ο ροζ θόρυβος (Pink Noise -PN) και ο ομοιόμορφος λευκός θόρυβος (Uniform White Noise -UWN), όπου διαπιστώθηκαν παραπλήσια αποτελέσματα [179].

Τα δύο σχήματα παρουσιάζουν τα αποτελέσματα αποθορυβοποίησης για όλους τους τύπους $WDWF$ (I , II , $I\&II$), με και χωρίς την ενεργοποίηση του ψυχοακουστικού φίλτρου A_{kw} (αριστερή και δεξιά πλευρά των σχημάτων αντίστοιχα). Τα σήματα δοκιμής περιλαμβάνουν έξι συνιστώσες $tn_g(i)$, $g=1,..6$ ($G=6$) σε διαφορετικές χρονικές στιγμές και με κλιμακούμενο φασματικό περιεχόμενο. Επίσης, χρησιμοποιήθηκαν τρεις διαφορετικές ρυθμίσεις διάρκειας - κρουστικότητας (με έλεγχο της παραμέτρου e_g). Παρατηρούμε ότι η προσέγγιση της φασματικής αφαίρεσης ισχύος $WDWF$ με $c=1$ και $d=0,1$ (διαγράμματα b1-c1, b2-c2) αδυνατεί να απομακρύνει το θόρυβο σε ικανοποιητικά επίπεδα αφήνοντας έντονα υπολείμματα θορύβου. Αντίθετα, οι ρυθμίσεις $c=3$, $d=0,2$ ($WDWF_I$) και $c=1$, $d=0,1$ ($WDWF_{I\&II}$, $WDWF_{II}$) λειτουργούν ικανοποιητικά σε όλες τις τοπολογίες ανάλυσης ($DWT-6$, $WP-17$). Παράλληλα, είναι ορατά τα φαινόμενα μετά-ήχους, τα οποία εμφανίζονται στον τύπο I (διαγράμματα d1-d2, e1-e2), ενώ εξαλείφονται πλήρως στους τύπους II και $I\&II$ (διαγράμματα f1-j1, f2-j2). Σε γενικές γραμμές, οι τύποι II και $I\&II$ παρουσιάζουν σχεδόν ταυτόσημη συμπεριφορά, γεγονός που αποδεικνύει και πειραματικά τα πλεονεκτήματα του τύπου II (δεδομένης της απλούστερης υλοποίησης και του μικρότερου υπολογιστικού κόστους). Συγκρίνοντας τα αποτελέσματα των τοπολογιών $DWT-6$ και $WP-17$, παρατηρούμε ότι η πρώτη λειτουργεί καλύτερα στις περιόδους σιγής, απορρίπτοντας πλήρως τους συντελεστές θορύβου, ενώ η δεύτερη είναι πιο «ευαίσθητη» με τις χρήσιμες πληροφοριακές συνιστώσες, κατά την

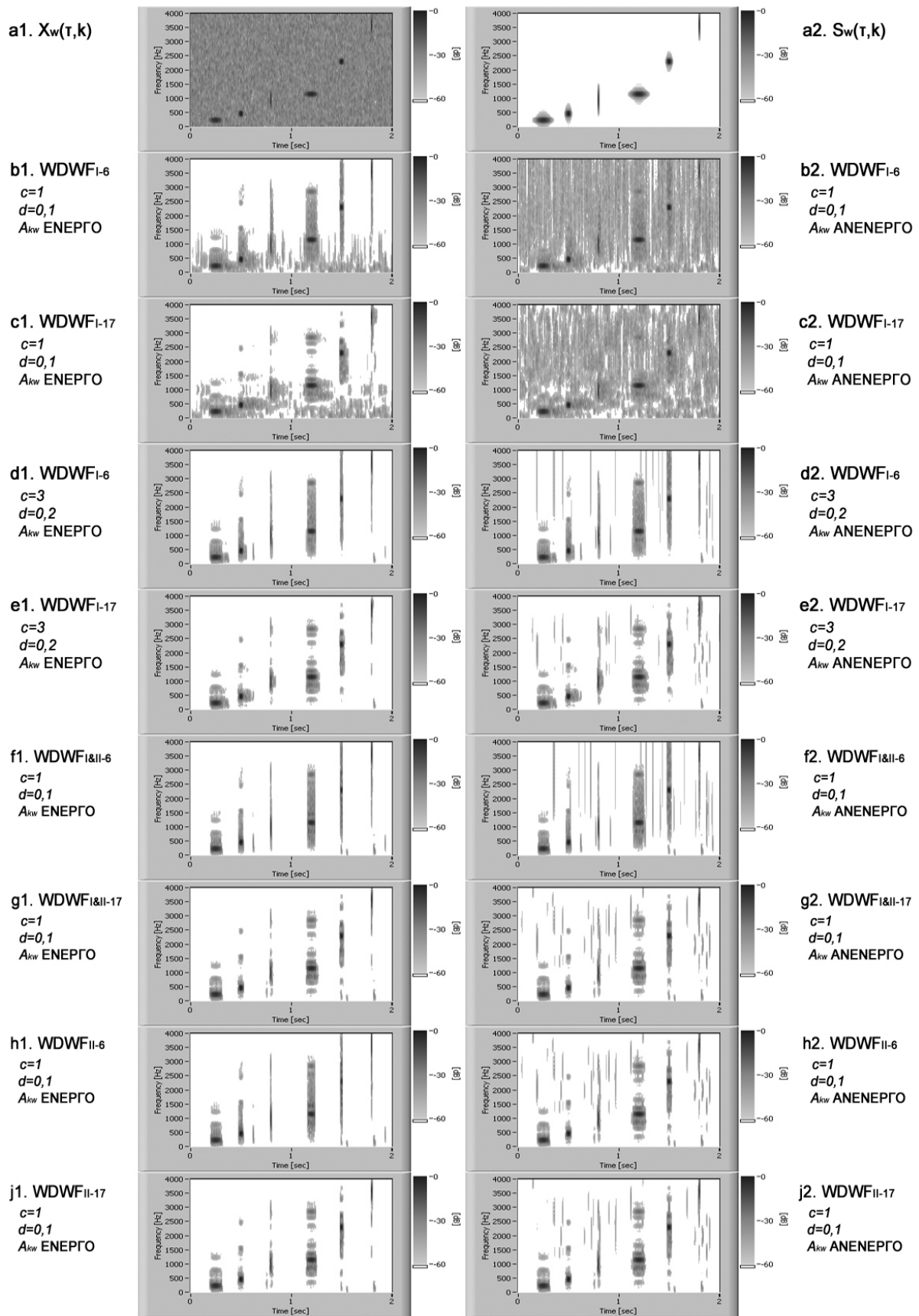
επεξεργασία αποθορυβοποίησης (σχ. 3-19). Τέλος, η ενσωμάτωση του όρου $A_{κω}$, παρέχει μεγάλη ανοχή στα υπολείμματα θορύβου τύπου “birdy-noise”, «καθαρίζοντας» πλήρως τις περιόδους σιγής. Έτσι, οι συνιστώσες χαμηλής ισχύος που βρίσκονται στα όρια, ή κάτω από τα επίπεδα του θορύβου, υφίστανται έντονη συμπίεση, στοιχείο που αποτελεί περισσότερο πλεονέκτημα παρά μειονέκτημα, αν εξεταστεί βάσει αντιληπτικών κριτηρίων [179].

Εκτός από την ποιοτική αποτίμηση της συμπεριφοράς των παραμέτρων (a , b , c , d), αποφασίστηκε και η εφαρμογή μιας διαδικασίας ποσοτικής αξιολόγησης. Δεδομένου ότι η συμπεριφορά των a , b , c είναι γνωστή [177], [184], [187], η διαδικασία επικεντρώθηκε στη διερεύνηση της παραμέτρου μνήμης d , η οποία παίζει ιδιαίτερα σημαντικό ρόλο στη συμπεριφορά των φίλτρων $WDWF$. Έτσι, αποφασίστηκε η χρήση σημάτων δοκιμής $tn_g(i)$ (εξ. 3-49β), τα οποία μολύνθηκαν με διάφορα προφίλ θορύβου (GWN , UWN και PN) διαμορφώνοντας μια «τοπική» σηματοθορυβική σχέση της τάξης των 6dB (Localized Signal to Noise Ratio, $LSNR$ [81]): $LSNR=6dB$ [179]. Η παράμετρος θ_g ρυθμίστηκε σε ένα «πεδίο» 16 τιμών (όμοιο με το τρίτο-οκταβικό πλέγμα), σχηματίζοντας τις ακόλουθες συχνότητες: $\{100/125/160/200/250/315/400/500/630/810/1000/1250/1600/2000/2500/3150 \text{ Hz}\}$ · η e_g ρυθμίστηκε ανάλογα σε ένα πλέγμα 8 τιμών $\{0,002/0,005/0,01/0,02/0,05/0,1/0,15/0,2\}$, ορίζοντας ένα αντίστοιχο σχήμα «λογαριθμικών εκφράσεων κρουστικότητας» (Logarithmic Crest Factor - LCF, εξ. 3-50): $\{10,6/11,5/12,8/15,2/18,6/21,3/24,1/27,8/30,4 \text{ dB}\}$ [179]. Έτσι, προέκυψε ένα διδιάστατο πλέγμα συχνότητας - κρουστικότητας, που περιλαμβάνει 16×8 κόμβους, δηλαδή 128 σήματα δοκιμής. Τα τελευταία μολύνθηκαν με τυχαίες επαναλήψεις διαφορετικών προφίλ θορύβου ($10 \times GWN$, $10 \times UWN$, $10 \times PN$), παρέχοντας 30 διαφορετικά «ενθόρυβα δείγματα» σε κάθε κόμβο (συνολικά 3.840 περιπτώσεις). Στα δείγματα αυτά εφαρμόστηκαν οι 4 τεχνικές αποθορυβοποίησης $WDWF$, ενώ τα αποτελέσματα αξιολόγησης προέκυψαν βάσει του δείκτη συσχέτισης CCI (Cross Correlation Index, εξ. 3-51): [81], [179]

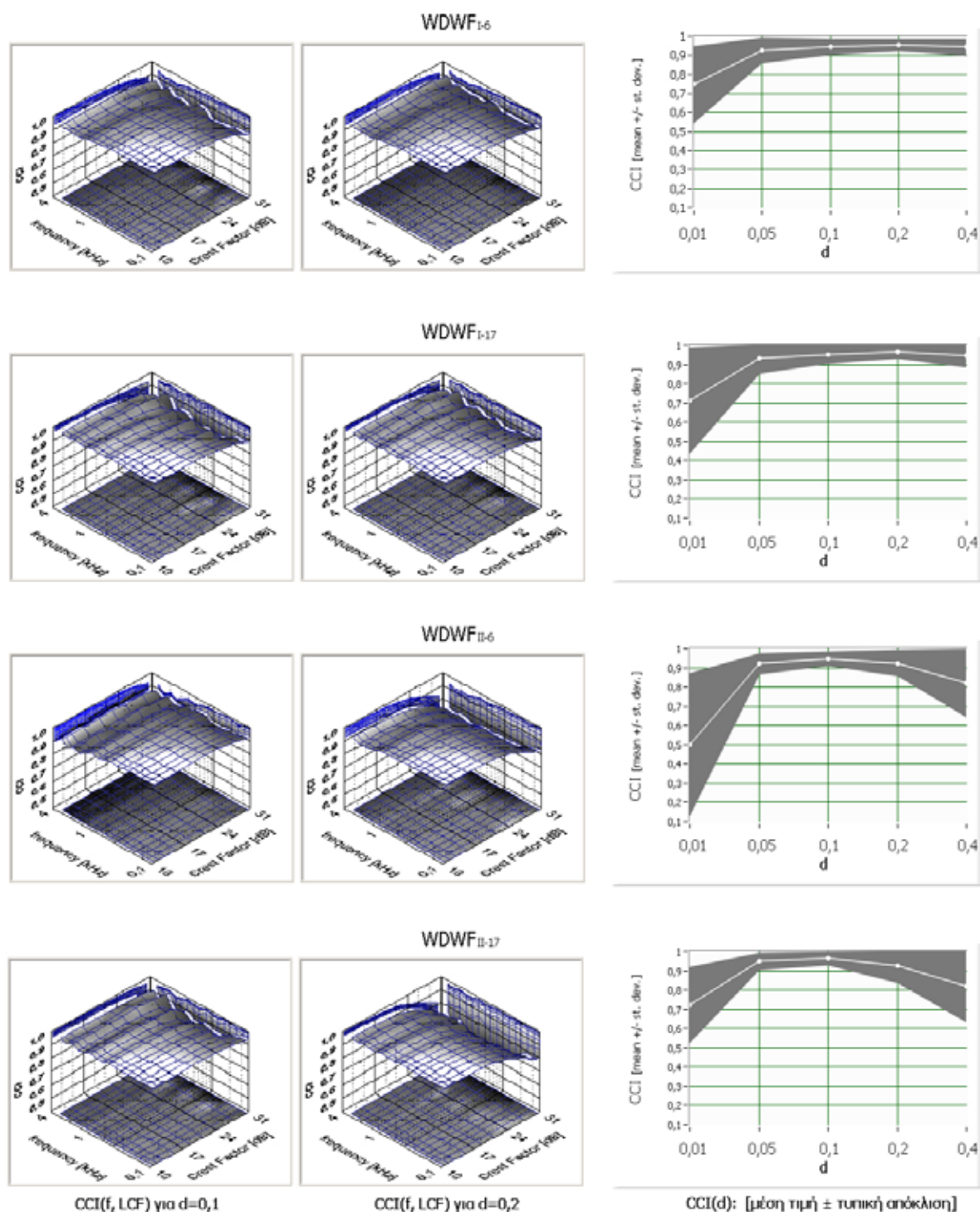
$$LCF(s) = 20 \cdot \log_{10} \frac{rms(s)}{\max(|s|)} \quad [dB], \quad \begin{array}{l} rms: \text{τελεστής ενεργού τιμής} \\ max: \text{τελεστής μέγιστης τιμής} \end{array} \quad (3-50)$$

$$CCI = \frac{\sum_i [s(i) - \text{mean}\{s(i)\}] \cdot [s^-(i) - \text{mean}\{s^-(i)\}]}{\sqrt{\sum_i [s(i) - \text{mean}\{s(i)\}]^2} \cdot \sqrt{\sum_i [s^-(i) - \text{mean}\{s^-(i)\}]^2}}, \quad \text{mean: τελεστής μέσης τιμής} \quad (3-51)$$

Τα αποτελέσματα ποσοτικής αξιολόγησης από τη δοκιμή πέντε διαφορετικών τιμών της παραμέτρου d $\{0,01 / 0,05 / 0,1 / 0,2 / 0,4\}$, παρουσιάζονται στο σχ. 3-20. Στις δύο πρώτες στήλες απεικονίζεται η επιφάνεια “ $CCI(f, LCF)$ ” για $d=0,1$ και $d=0,2$. Στην 3η στήλη απεικονίζονται οι διακυμάνσεις του δείκτη CCI που αντιστοιχούν σε όλο του εύρος του πλέγματος συχνότητας - κρουστικότητας (μέση τιμή \pm τυπική απόκλιση, “ $mean \pm st. dev$ ”). Από τα διαγράμματα προκύπτει ότι οι ρυθμίσεις $d=0,2$ και $d=0,1$ αποτελούν τις βέλτιστες επιλογές για τους τύπους I και II αντίστοιχα, καθώς έχουν τον καλύτερο μέσο δείκτη συσχέτισης με τη μικρότερη διασπορά. Αξίζει να σημειωθεί ότι στόχος της παρούσης διαδικασίας ήταν να διερευνηθεί η επίδραση της παραμέτρου d , και όχι τα αποτελέσματα αποθορυβοποίησης αυτά κάθε αυτά (σχετικά αποτελέσματα δίνονται στην «ποσοτική αξιολόγηση», §3.2.6.3). Παράλληλα, διενεργήθηκαν προσομοιώσεις με περισσότερα δείγματα θορύβου (300 για κάθε σήμα δοκιμής: $100 \times GWN$, $100 \times UWN$, $100 \times PN$) και διαφορετικούς λόγους $LSNR$, χωρίς ουσιαστικές διαφοροποιήσεις των αποτελεσμάτων. Ωστόσο, αποφεύχθηκαν ιδιαίτερα χαμηλές τιμές $LSNR$, καθώς ο τύπος II τείνει να εισάγει



Σχήμα 3-19. Αποτελέσματα αποθρονοποίησης (σπεκτρογράμματα), κατά την εφαρμογή των φίλτρων WDF σε σήματα δοκιμής (τόνοι διαμορφωμένοι με γκαουσιανές περιβάλλουσες), μολυσμένα με λευκό θόρυβο (GWN) - οι καμπύλες αντιστοιχούν σε διάφορες ρυθμίσεις των παραμέτρων c, d ($a=2, b=1$), και για διαφορετικές καταστάσεις του φίλτρου A_{kw} (αριστερά: ενεργοποιημένο / δεξιά: απενεργοποιημένο).



Σχήμα 3-20. Επίδραση της παραμέτρου d στα φίλτρα WDWF (ποσοτική εκτίμηση βάσει του δείκτη CCI): οι 2 πρώτες στήλες αντιστοιχούν στις επιφάνειες “CCI(f,LCF)” για $d=0,1$ και $d=0,2$. στην 3^η δίνονται οι διακυμάνσεις του δείκτη CCI για διάφορες τιμές του d , σε όλο το εύρος του πλέγματος «συχνότητας - κρουστικότητας».

3.2.6.2 Ποιοτική αξιολόγηση των προτεινόμενων τεχνικών αποθροβοποίησης

Κατά τη διαδικασία ποιοτικής αξιολόγησης, ενθόρυβοι γαστρεντερικοί ήχοι διαφόρων μορφολογικών χαρακτηριστικών και διαφόρων σηματοθροβικών λόγων, επιλέχθηκαν και αποθροβοποιήθηκαν με τις τεχνικές WDWF. Σχετικά με τη χρήση των σηματοθροβικών σχέσεων, είναι προφανές ότι οι κλασικές εκφράσεις SNR δεν ενδείκνυνται στην προκειμένη

περίπτωση, καθώς οι κοιλιακοί ήχοι περιλαμβάνουν πολλά διαστήματα παύσης-σιγής. Έτσι, στην περίπτωση κρουστικών βιοακουστικών σημάτων χρησιμοποιήθηκε η παράμετρος $LSNR$, που εκφράζει το λόγο σήματος προς θόρυβο SNR , στην «τοπική περιοχή» παρουσίας σήματος [81], [83], [133]. Η έκφραση $LSNR$ προσφέρεται όταν το σήμα είναι «χρονικά εντοπισμένο» (time-localized) όπως συμβαίνει με τους κρουστικούς ήχους EBS (ή IB), ELS αλλά και τους συνθετικούς τόνους $tn_g(i)$. Στην περίπτωση σημάτων RS ή συνδυασμών RS και IB σε τυχαίες χρονικές θέσεις, η χρήση της έκφρασης $LSNR$ είναι προβληματική. Στο πλαίσιο αυτό, εισήχθη η έννοια του «αποτελεσματικού λόγου σήματος προς θόρυβο» (Effective Signal to Noise Ratio - $ESNR$), ο οποίος ορίζεται ως ακολούθως [179]:

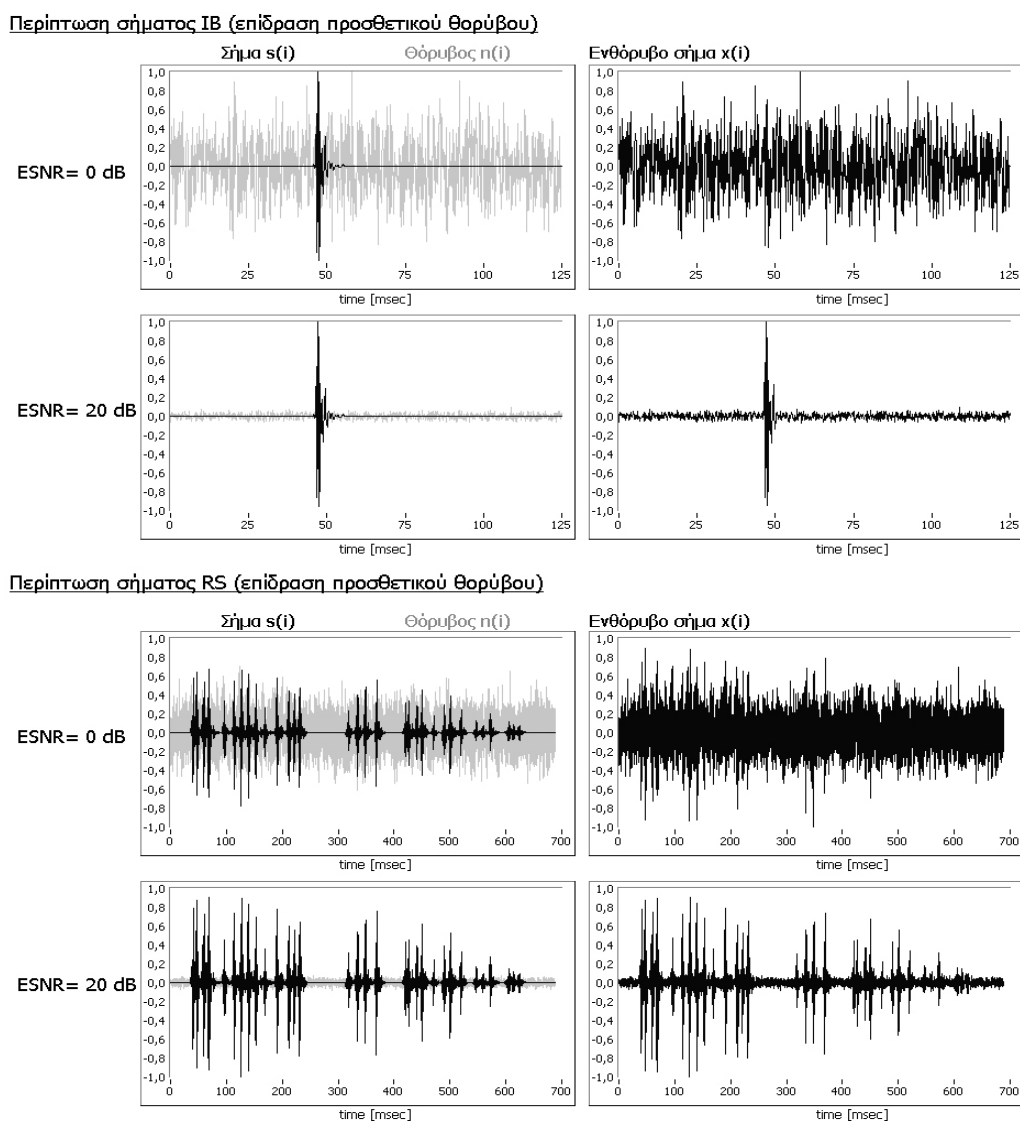
$$ESNR(th) = 10 \cdot \log_{10} \left(\frac{\sum_{i, |s(i)| > th_{ESNR}} s(i)^2}{\sum_{i, |s(i)| > th_{ESNR}} n(i)^2} \right) \cong 10 \cdot \log_{10} \left(\frac{\sum_{i, |s(i)| > th_{ESNR}} s(i)^2}{\sum_i n(i)^2} \right) \quad (3-52)$$

όπου s είναι το αρχικό «καθαρό» σήμα (ή η εκτίμηση s^-), και n είναι ο προσθετικός θόρυβος μόλυνσης (ή το εξαγόμενο σήμα n^-). Η χρήση του λόγου $ESNR$ είναι πιο αντικειμενική, καθώς η ισχύς του σήματος εκτιμάται μόνο από τις σημαντικές «συνιστώσες», δηλαδή στα χρονικά διαστήματα όπου το σήμα είναι ενεργό. Η εφαρμογή της εξ. (3-52) περιλαμβάνει τα ακόλουθα βήματα [179]: α) αρχικά εντοπίζεται ένα διάστημα τοπικού μεγίστου, όπου υπολογίζεται η στάθμη ισχύος του σήματος, β) η τιμή κατωφλίου th_{ESNR} προσδιορίζεται έτσι ώστε, οι περιοχές σήματος που υπολείπονται κατά 20 dB του «τοπικού μεγίστου», δε «προσμετρώνται» ως σήμα, αλλά ως περίοδος σιγής. Για τον υπολογισμό των σταθμών ισχύος χρησιμοποιείται η διαδικασία εκθετικού μέσου (εξ. 3-40) με συντελεστές $\{1, -0.92\}$ και $\{0.08\}$ [179]. Στο σχ. 3-21 παρουσιάζονται οι οριακές καταστάσεις ($ESNR=0dB$ και $ESNR=20dB$), για δύο αντιπροσωπευτικά πρότυπα IB και RS . Τα σήματα στο συγκεκριμένο γράφημα (όπως και σε όλα τα παραδείγματα που θα ακολουθήσουν), κανονικοποιούνται ως εξής: α) το ενθόρυβο σήμα $x(i)$ ρυθμίζεται έτσι ώστε η μέγιστη τιμή πλάτους του να αντιστοιχεί στη μονάδα, β) τα υπόλοιπα σήματα $s(i)$, $s^-(i)$, $n(i)$, και $n^-(i)$ ρυθμίζονται ανάλογα, αφού πολλαπλασιαστούν με τον ίδιο παράγοντα κλιμάκωσης. Στην περίπτωση που η ανάλυση περιλαμβάνει και τα φάσματα των παραπάνω σημάτων $X(k)$, $S(k)$, $S^-(k)$, $N(k)$, $N^-(k)$, τότε, η μέγιστη στάθμη της καμπύλης $|X(k)|$ διαμορφώνεται στα 0 dB, ενώ ανάλογα προσαρμόζονται και τα υπόλοιπα φάσματα [179].

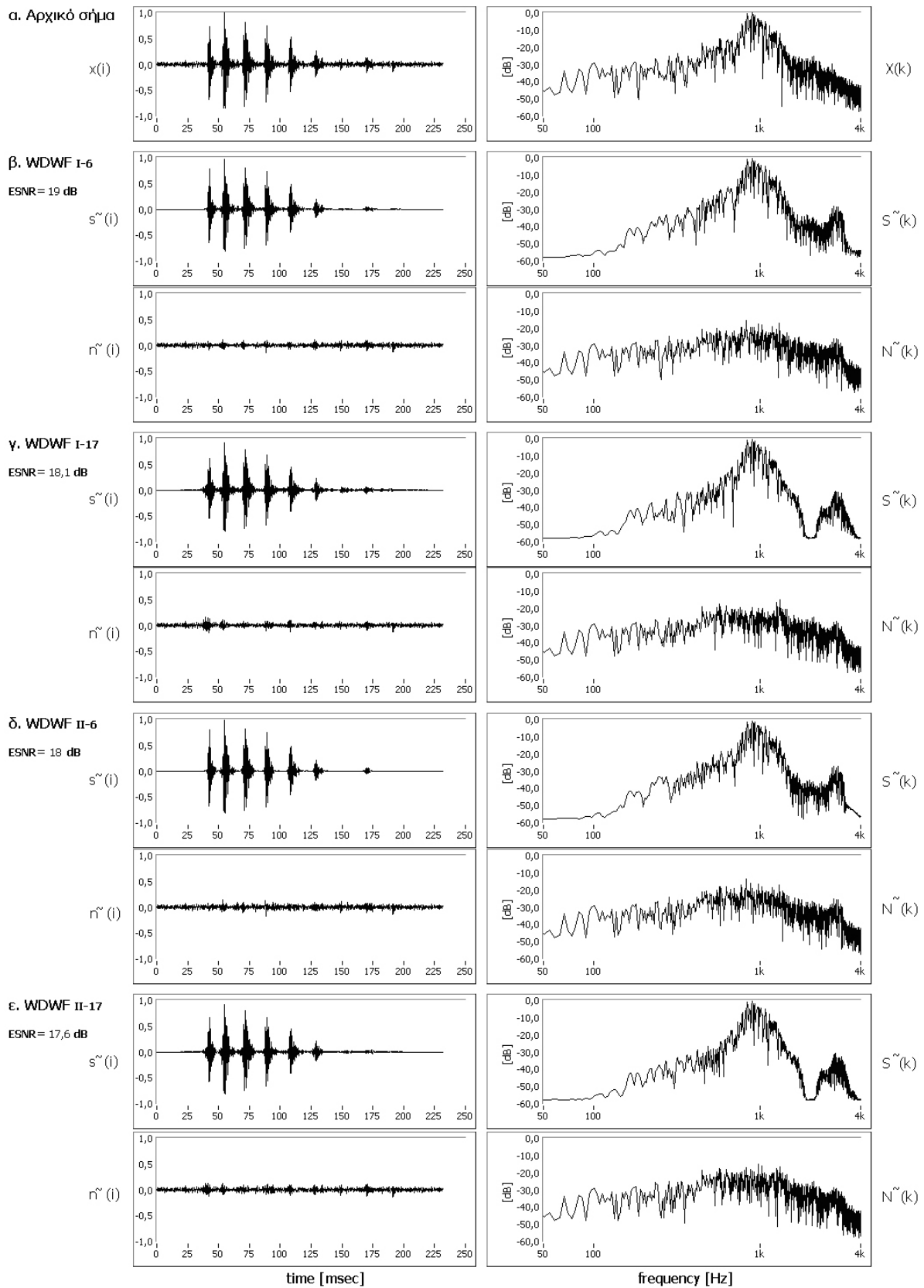
Στα σχ. 3-22 και 3-23 απεικονίζονται τα αποτελέσματα αποθορυβοποίησης δύο διαφορετικών ενθόρυβων σημάτων, στα οποία έχουν εφαρμοστεί οι 4 τεχνικές $WDWF$. Διαπιστώνουμε ότι όλες οι προσεγγίσεις δίνουν πολύ καλά αποτελέσματα, τόσο σε υψηλές ($ESNR > 0$) όσο και σε χαμηλές ($ESNR < 0$) σηματοθορυβικές σχέσεις. Επίσης επαληθεύεται ότι η τοπολογία $WP-17$ εφαρμόζει πιο «προσεκτική» καταστολή θορύβου, ενώ η $DWT-6$ είναι πιο «αυστηρή» στην απόρριψη των περιόδων σιγής. Τα αποτελέσματα του τύπου II δείχνουν καλύτερα, καθώς συμπιέζουν ισχυρά τις περιόδους σιγής, ενώ διατηρούν τα αρχικά πλάτη κατά την παρουσία σήματος. Στα σχ. 3-24, 3-25 παρατίθενται τα αντίστοιχα αποτελέσματα στην περίπτωση κακής σηματοθορυβικής σχέσης, με και χωρίς το φίλτρο A_{kw} . Παρατηρούμε ότι η ενεργοποίηση των ψυχοακουστικών κριτηρίων έχει ως αποτέλεσμα την απόρριψη των συνιστωσών που βρίσκονται σε χαμηλά επίπεδα. Από την ακουστική ανάλυση των αντίστοιχων σημάτων, προκύπτει ότι η χρήση του A_{kw} είναι μάλλον προτιμητέα, καθώς παρέχει καλύτερα (πιο φυσικά) ακουστά αποτελέσματα, περιορίζοντας την παρουσία των ανεπιθύμητων φαινομένων “birdy noise”. Στο παράδειγμα του σχ. 3-26, όπου μεταξύ των

συνιστώσων θορύβου περιλαμβάνεται βόμβος, τα φίλτρα *WDWF* εξακολουθούν να λειτουργούν ικανοποιητικά, αρκεί βέβαια να παρέχεται το σωστό προφίλ θορύβου [179].

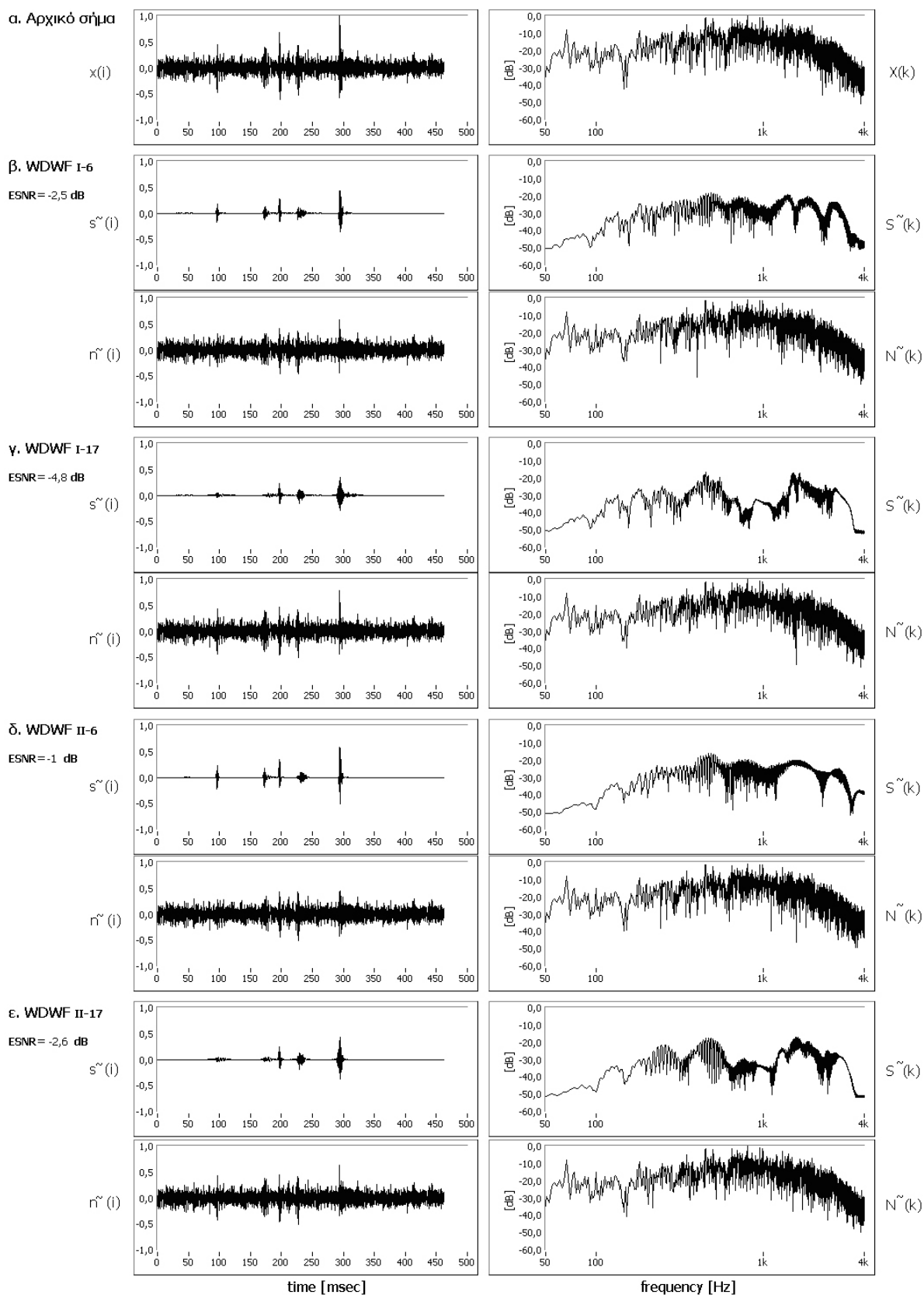
Στο σχ. 3-27 παρουσιάζονται τα αποτελέσματα σύγκρισης των αλγορίθμων *WDWF* με τις κυματιδιακές μεθόδους *RIGSURE*, *SQWLOG* και τις αντίστοιχες προσεγγίσεις απαλής και σκληρής κατωφλίωσης (soft/hard thresholding) [192]-[193]. Από τους 4 συνδυασμούς, οι τεχνικές *RIGSURE*_(soft) και *SQWLOG*_(hard) δίνουν τα καλύτερα αποτελέσματα, όπου η πρώτη αλλοιώνει κατ' ελάχιστον τις χρήσιμες συνιστώσες (minimal signal distortion), ενώ η δεύτερη εφαρμόζει αυστηρή «καταστολή» (rough de-noising), ιδιαίτερα στις περιόδους σιγής [179]. Στα σχ. 3-28, 3-29 δίνονται τα αντίστοιχα αποτελέσματα σύγκρισης με τις μεθόδους *WTST-NST*, *WT-FD* και *IKD*, όπου επαληθεύονται και πειραματικά οι παρατηρήσεις της §3.2.2. Δηλαδή, παρέχονται ικανοποιητικά αποτελέσματα στην περίπτωση σημάτων *IB* (σχ. 3-28), σε αντίθεση με τα πρότυπα *RS* (σχ. 3-29), όπου, παρά το μεγάλο αριθμό επαναλήψεων *L*, παρατηρείται μη ικανοποιητική καταστολή θορύβου ή/και προβλήματα έντονης παραμόρφωσης. Χαρακτηριστικά παραδείγματα αποτελούν οι ρυθμίσεις *WTST-NST* με $F_{adj}=3$ και *WTST-NST* με $F_{adj}=4$, (για τις δύο περιπτώσεις αντίστοιχα). Γενικά, ο αλγόριθμος *WTST-NST* είναι πιο «σταθερός» των αντίστοιχων *WT-FD* και *IKD* [179].



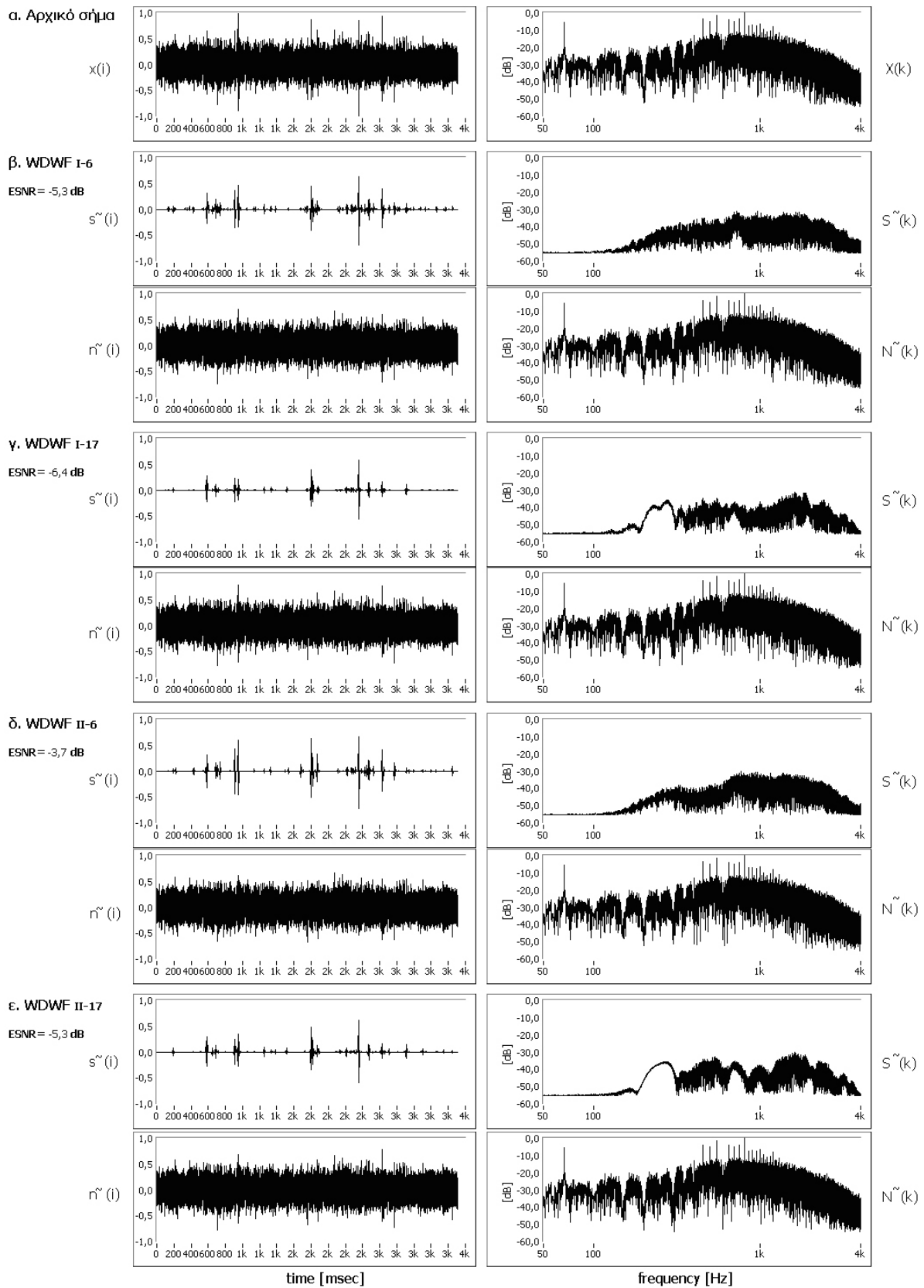
Σχήμα 3-21. Χαρακτηριστικές περιπτώσεις εμφάνισης προσθετικού θορύβου σε σήματα *RS*, *IB* με σηματο-θορυβικές σχέσεις $ESNR=0dB$ και $ESNR=20dB$.



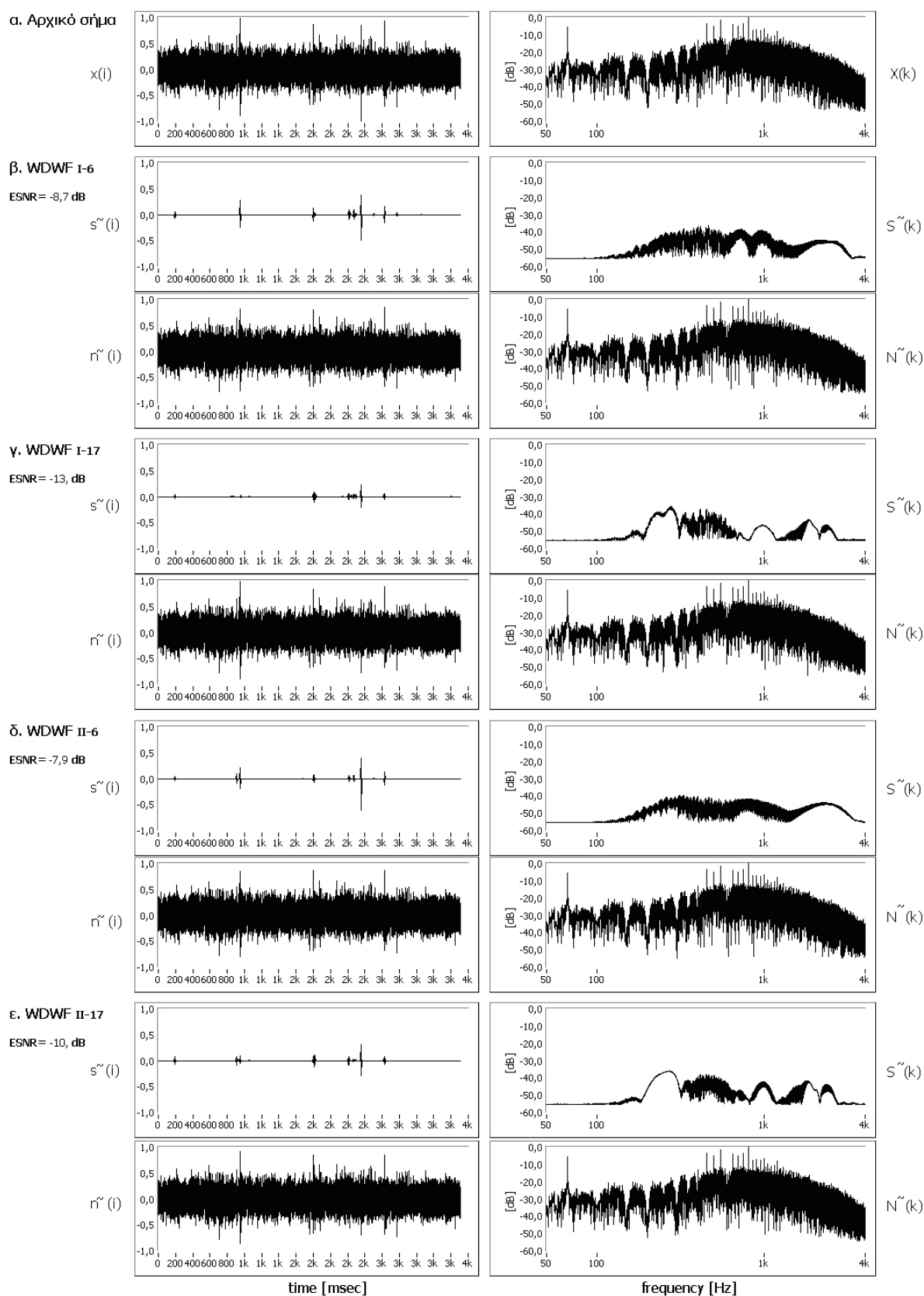
Σχήμα 3-22. Αποτελέσματα αποθορυβοποίησης στο πεδίο του χρόνου (αριστερά) και της συχνότητας (δεξιά) κατά την εφαρμογή των φίλτρων WDWF σε σήμα σχετικά υψηλής σηματοθορυβικής σχέσης ESNR.



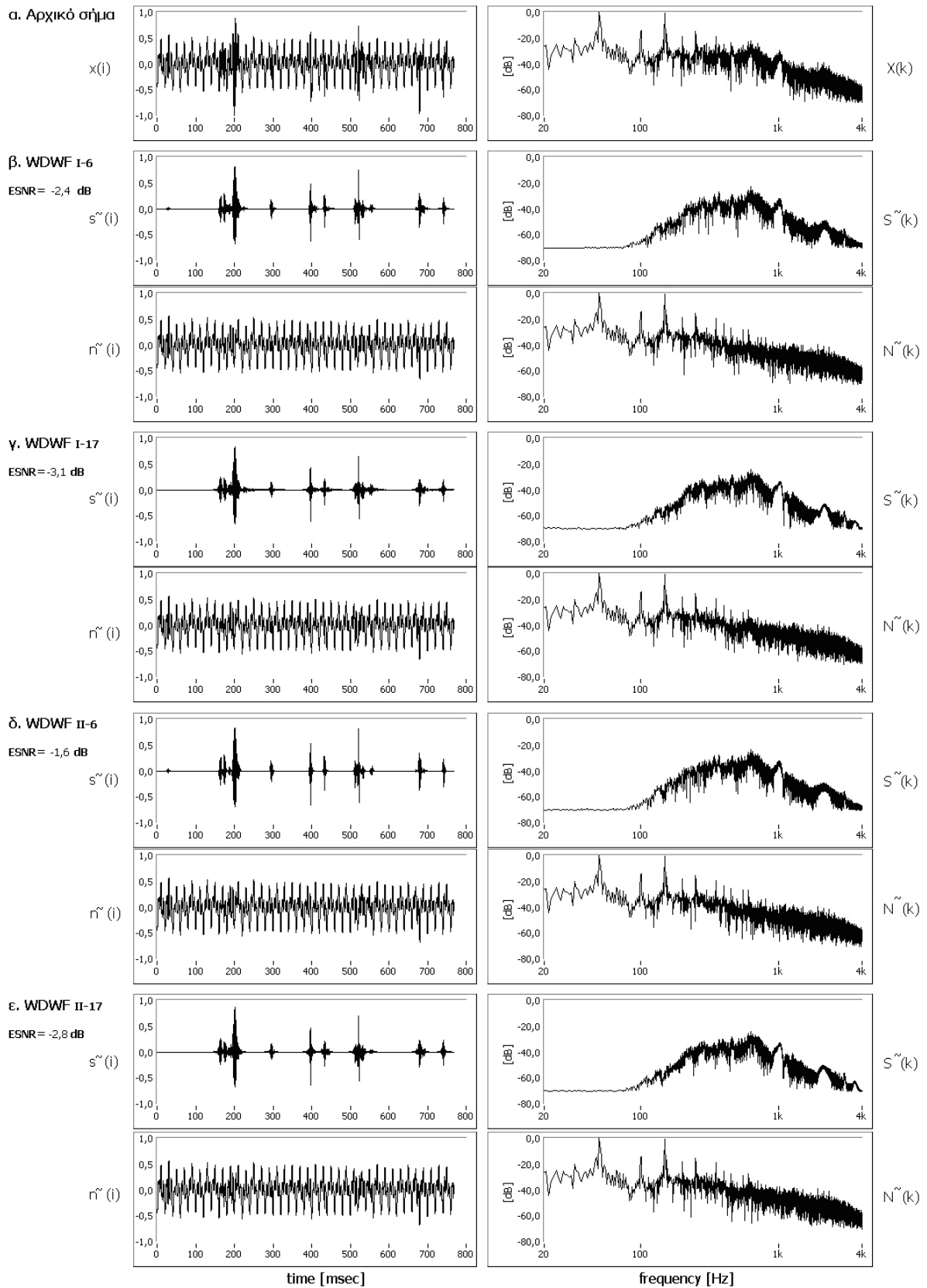
Σχήμα 3-23. Αποτελέσματα αποθρομβοποίησης στο πεδίο του χρόνου (αριστερά) και της συχνότητας (δεξιά) κατά την εφαρμογή των φίλτρων WDWF σε σήμα σχετικά χαμηλής σηματοθρομβικής σχέσης ESNR.



Σχήμα 3-24. Αποτελέσματα αποθρομβοποίησης στο πεδίο του χρόνου (αριστερά) και της συχνότητας (δεξιά) κατά την εφαρμογή των φίλτρων WDF σε σήμα χαμηλής σηματοθρομβικής σχέσης ESNR, χωρίς τη χρήση του ψυχοακουστικού φίλτρου A_{kw} .

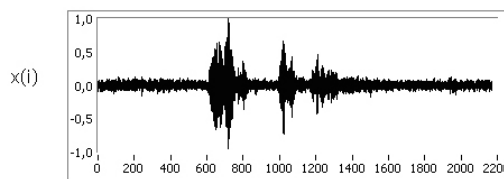


Σχήμα 3-25. Αποτελέσματα αποθρομβοποίησης στο πεδίο του χρόνου (αριστερά) και της συχνότητας (δεξιά) κατά την εφαρμογή των φίλτρων WDF στο ίδιο με το σχ.3-24 σήμα χαμηλής σηματοθρομβικής σχέσης ESNR, με χρήση του ψυχοακουστικού φίλτρου $A_{k\omega}$ αυτή τη φορά.

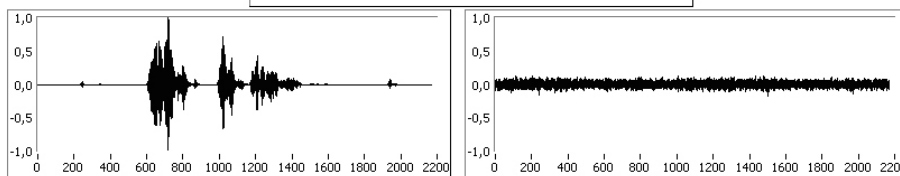


Σχήμα 3-26. Αποτελέσματα αποθορυβοποίησης στο πεδίο του χρόνου (αριστερά) και της συχνότητας (δεξιά) κατά την εφαρμογή των φίλτρων WDF σε σήμα χαμηλής σηματοθορυβικής σχέσης ESNR, όπου στο «μολυσματικό» θόρυβο συμπεριλαμβάνεται και βόμβος.

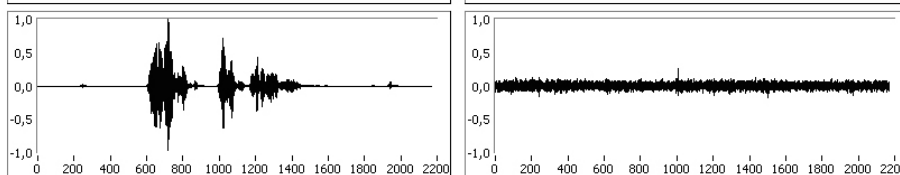
α. Αρχικό σήμα



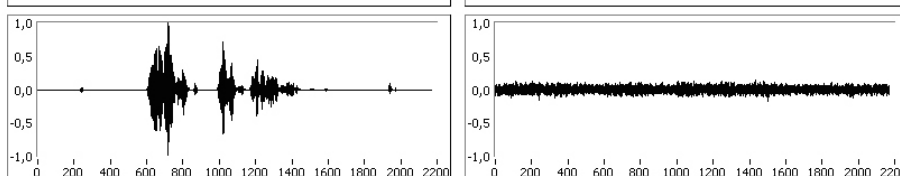
β. WDFW I-6
ESNR = 17,5 dB



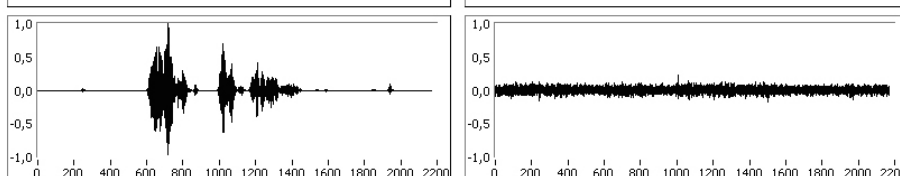
γ. WDFW I-17
ESNR = 17,5 dB



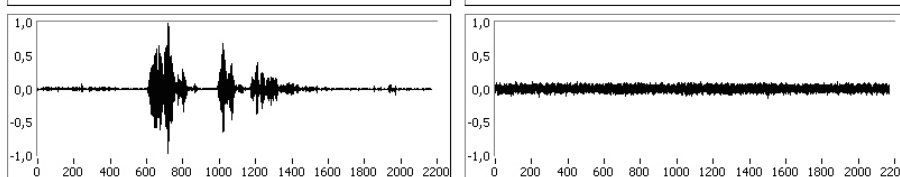
δ. WDFW II-6
ESNR = 17,2 dB



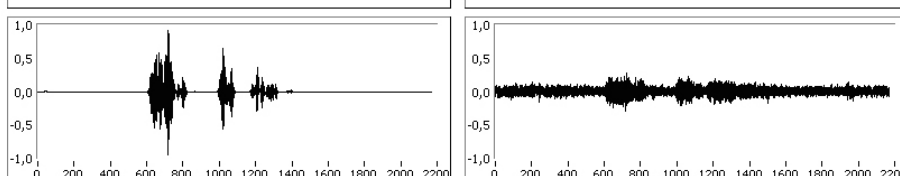
ε. WDFW II-17
ESNR = 17,2 dB



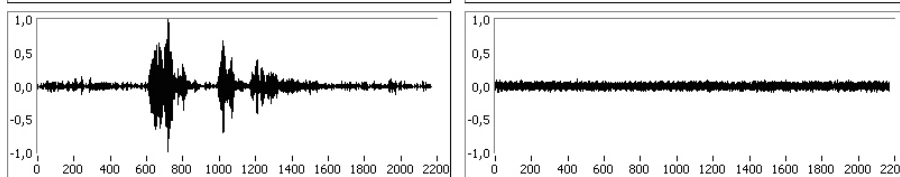
στ. RIGSURE (soft)
ESNR = 17,8 dB



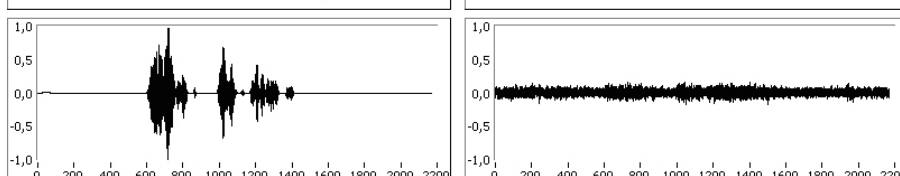
ζ. SQTWOLOG (soft)
ESNR = 13,9 dB



η. RIGSURE (hard)
ESNR = 19,2 dB



θ. SQTWOLOG (hard)
ESNR = 16,4 dB

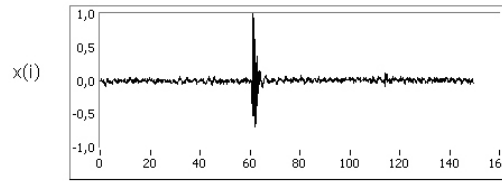


$\tilde{s}(i)$ denoised signal vs time [msec]

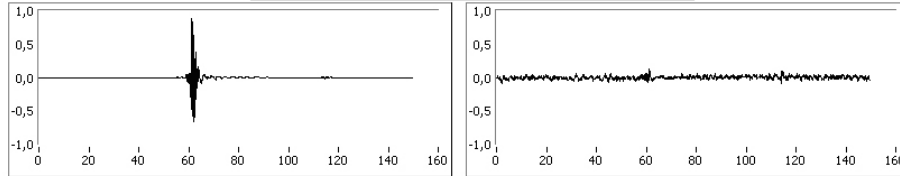
$\tilde{n}(i)$ extracted noise vs time [msec]

Σχήμα 3-27. Σύγκριση των αποτελεσμάτων αποθρομβοποίησης των τεχνικών WDFW (β-ε) με τις κλασικές τεχνικές κυματιδιακής κατωφλίωσης (στ-θ) RIGSURE και SQTWOLOG (ανάλυση στο πεδίο χρόνου: στη αριστερή πλευρά απεικονίζεται το αποθρομβοποιημένο σήμα \tilde{s} και στη δεξιά ο εξαγόμενος θόρυβος \tilde{n}).

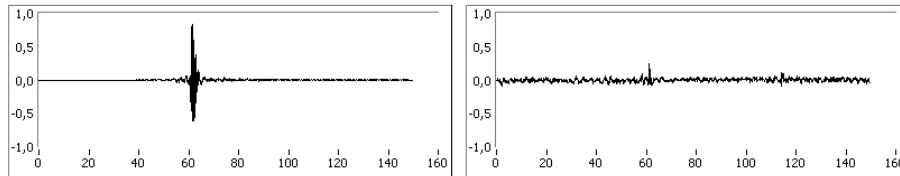
α. Αρχικό σήμα



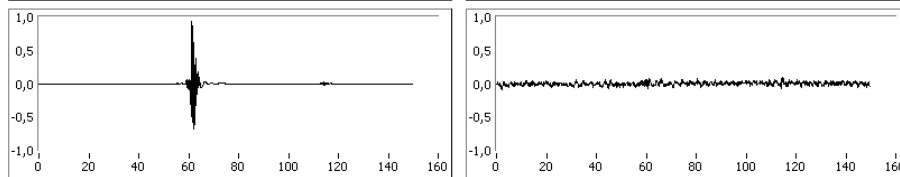
β. WDF I-6
ESNR = 19,3 dB



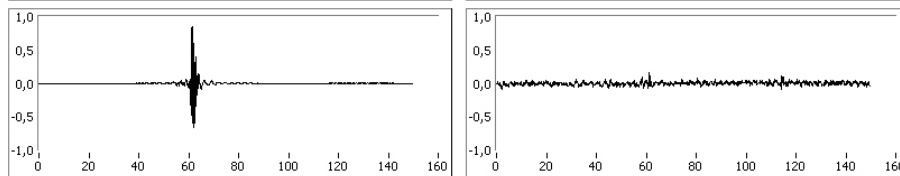
γ. WDF I-17
ESNR = 18,7 dB



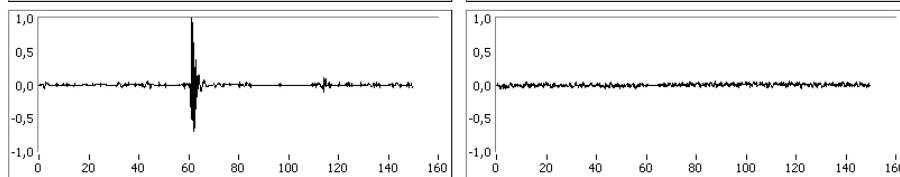
δ. WDF II-6
ESNR = 19,5 dB



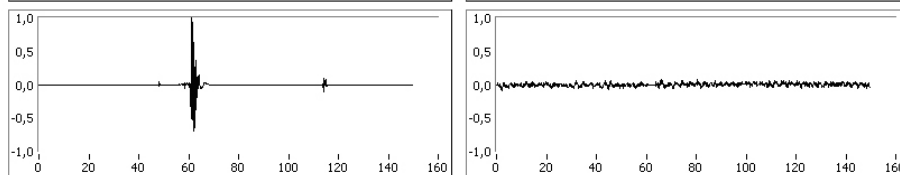
ε. WDF II-17
ESNR = 19,2 dB



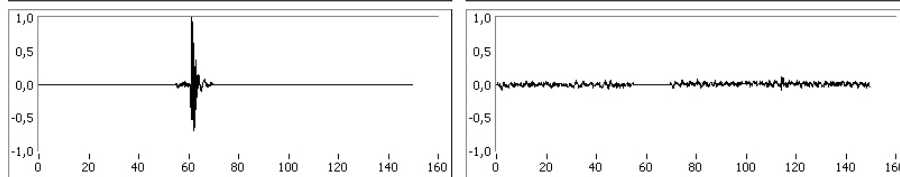
στ. WTST-NST
"DB-4", $F_{adj}=3$, $L=7$
ESNR = 21,3 dB



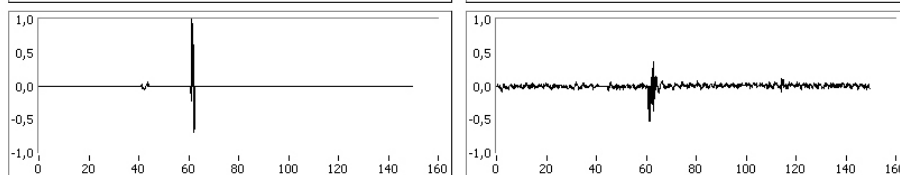
ζ. WTST-NST
"DB-4", $F_{adj}=4$, $L=5$
ESNR = 20,2 dB



η. WT-FD
"DB-4", $L=2$
ESNR = 20,3 dB



θ. IKD
 $L=3$
ESNR = 12,6 dB

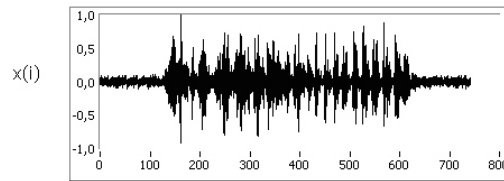


$\tilde{s}(i)$ denoised signal vs time [msec]

$\tilde{n}(i)$ extracted noise vs time [msec]

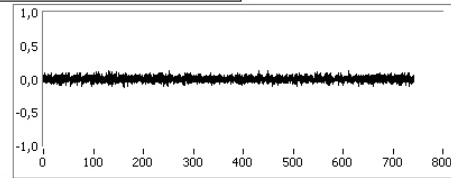
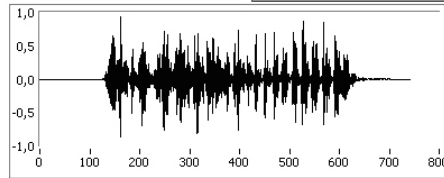
Σχήμα 3-28. Σύγκριση των αποτελεμάτων αποθρορυβοποίησης των τεχνικών WDF (β-ε) με τις τεχνικές WTST-NST ($F_{adj}=3, 4$), WT-FD και IKD (στ-θ) στην περίπτωση σήματος IB (ανάλυση στο πεδίο χρόνου: στην αριστερή πλευρά απεικονίζεται το αποθρορυβοποιημένο σήμα \tilde{s} και στη δεξιά ο εξαγόμενος θόρυβος \tilde{n}).

α. Αρχικό σήμα



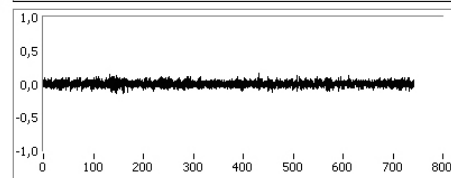
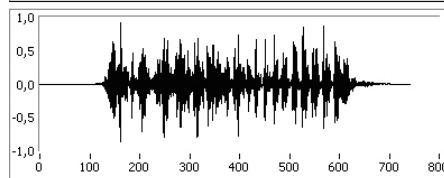
β. WDWF I-6

ESNR = 17,1 dB



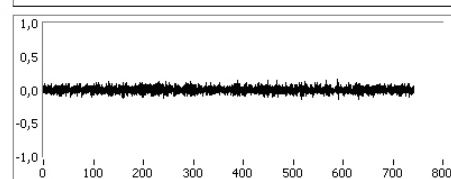
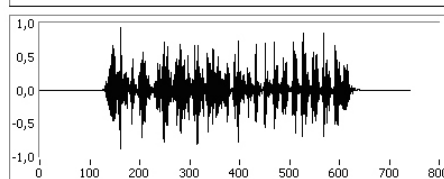
γ. WDWF I-17

ESNR = 16,6 dB



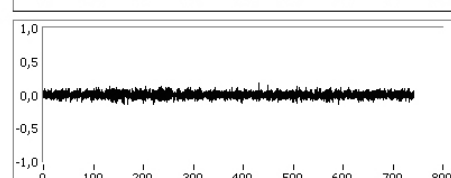
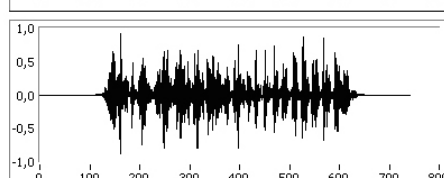
δ. WDWF II-6

ESNR = 16 dB



ε. WDWF II-17

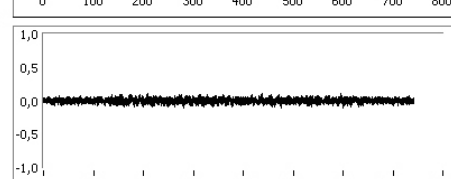
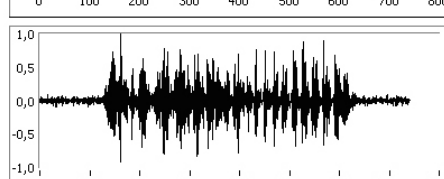
ESNR = 16 dB



στ. WTST-NST

"DB-4", $F_{adj}=3, L=24$

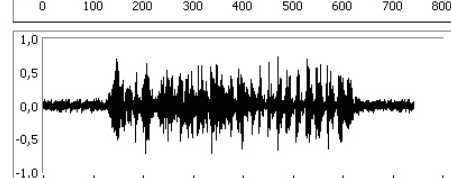
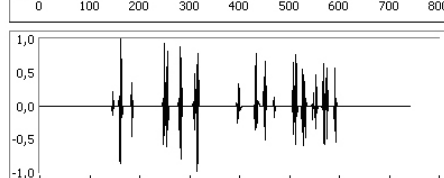
ESNR = 18 dB



ζ. WTST-NST

"DB-4", $F_{adj}=4, L=8$

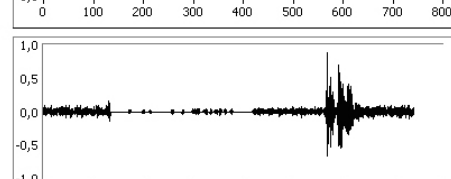
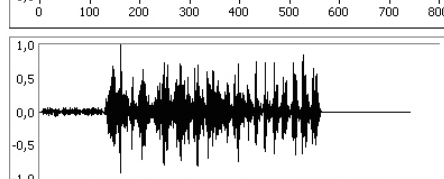
ESNR = 2,7 dB



η. WT-FD

"DB-4", $L=18$

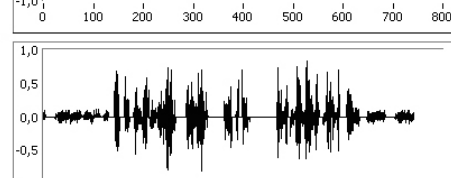
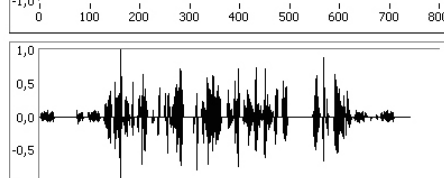
ESNR = 11,2 dB



θ. IKD

$L=2$

ESNR = 4,1 dB



$\tilde{s}(i)$ denoised signal vs time [msec]

$\tilde{n}(i)$ extracted noise vs time [msec]

Σχήμα 3-29. Σύγκριση των αποτελεσμάτων αποθρονοποίησης των τεχνικών WDWF (β-ε) με τις τεχνικές WTST-NST ($F_{adj}=3, 4$), WT-FD και IKD (στ-θ) στην περίπτωση σήματος RS (ανάλυση στο πεδίο χρόνου: στην αριστερή πλευρά απεικονίζεται το αποθρονοποιημένο σήμα \tilde{s} και στη δεξιά ο εξαγόμενος θόρυβος \tilde{n}).

3.2.6.3 Ποσοτική αξιολόγηση των προτεινόμενων τεχνικών αποθορυβοποίησης

Εκτός από τη διαδικασία ποιοτικής αξιολόγησης, αξιοποιήθηκαν πρόσθετοι δείκτες για την αποτίμηση των αποτελεσμάτων αποθορυβοποίησης με ποσοτικά κριτήρια. Στο πλαίσιο αυτό χρησιμοποιήθηκαν 100 αντιπροσωπευτικά πρότυπα *IB* και *RS* (200 συνολικά δείγματα), τα οποία αποτέλεσαν τα «καθαρά» σήματα δοκιμής που περιγράφηκαν στην §3.2.6 ($s \leftarrow s_{DAS}$). Οι ήχοι αυτοί μολύνθηκαν με προσθετικό θόρυβο *AWN* μηδενικής μέσης τιμής και μοναδιαίας διασποράς ($\mu_N=0, \sigma_N^2=1$), διαφορετικής κάθε φορά στάθμης. Πρόκειται για παγιωμένη διαδικασία αξιολόγησης των μεθόδων καταστολής θορύβου [59], [84], [133], στην οποία χρησιμοποιούνται διάφοροι ποσοτικοί δείκτες, όπως είναι ο δείκτης συσχέτισης *CCI*. Επίσης δοκιμάστηκαν πρόσθετα προφίλ θορύβου (*UWN, PN*), χωρίς ωστόσο να παρατηρηθεί διαφοροποίηση των αποτελεσμάτων [179]. Έτσι, σχηματίστηκαν 200 διαφορετικά ενθόρυβα δείγματα $x(i)$, για σηματοθρυβικές σχέσεις (*ESNR*) από 0dB έως 20dB, με βήμα μεταβολής $\Delta_{ESNR}=0,5dB$ (συνολικά προέκυψαν $200 \times 40 = 8.000$ δείγματα θορύβου $n(i)$ και άλλα τόσα ενθόρυβα σήματα $x(i)$, σε όλα τα επίπεδα *ESNR*). Εκτός από την παράμετρο *CCI*, συμπεριλήφθηκαν δύο ακόμη ποσοτικοί δείκτες (ένας στο πεδίο του χρόνου και ένας στο πεδίο της συχνότητας): ο «λόγος σήματος προς απόκλιση» (*Signal to Deviation Ratio - SDR*, εξ. 3-53), και η «μέτρηση φασματικής παραμόρφωσης» (*Spectral Distortion Measure - SDM*, εξ. 3-54), που έχουν επίσης αξιοποιηθεί σε αντίστοιχες διαδικασίες [179], [224]-[225]:

$$SDR = 10 \cdot \log \left[\frac{\sum_{i=0}^{N-1} [s(i)]^2}{\sum_{i=0}^{N-1} [s^{\sim}(i) - s(i)]^2} \right] [dB] \quad (3-53)$$

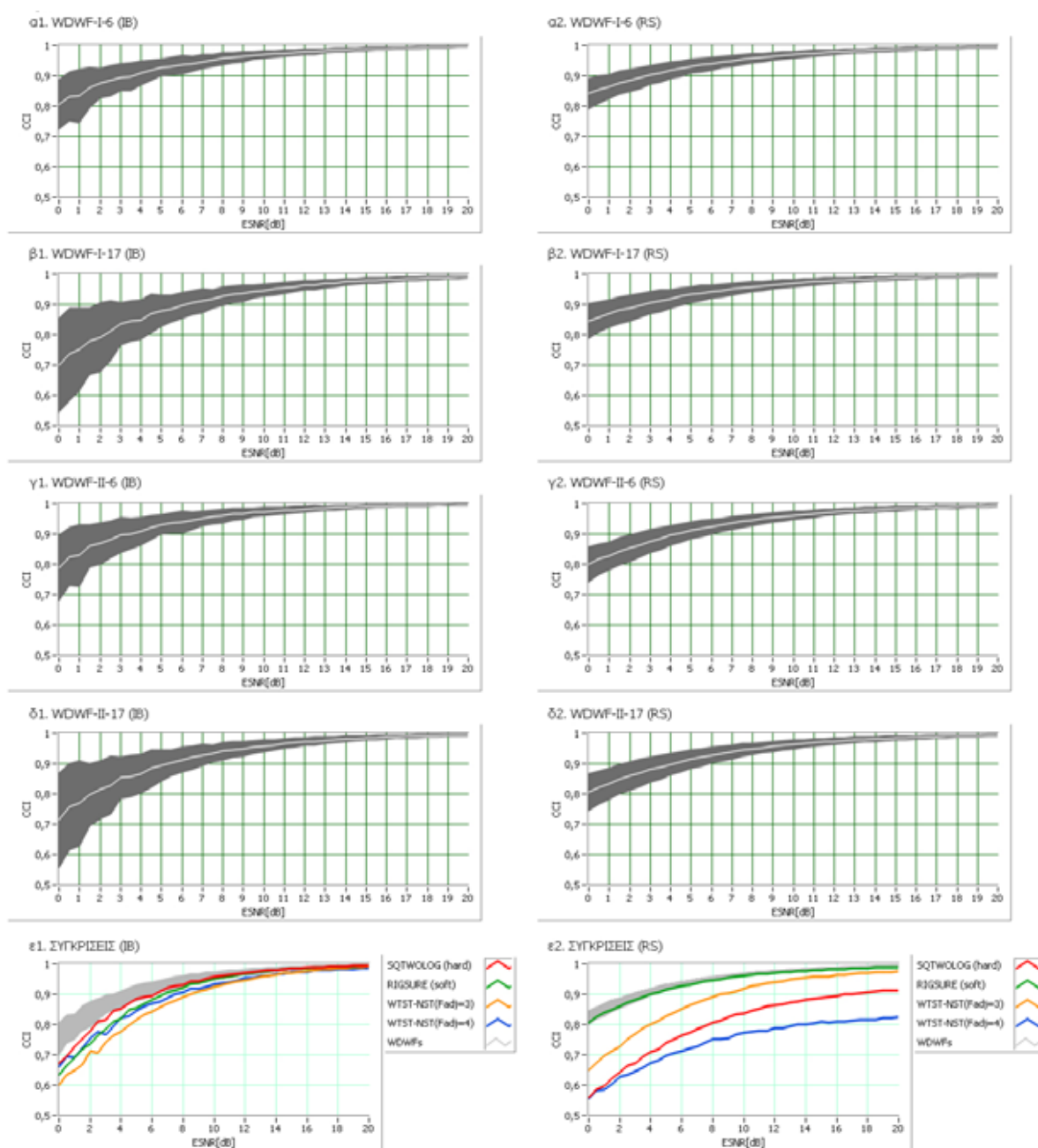
$$SDM(s, s^{\sim}) = \frac{1}{P} \cdot \frac{1}{256} \sum_{i=1}^P \sum_{k=0}^{256} 20 \cdot \left| \log(|S_{P,n}(k)|) - \log(|S_{P,n}^{\sim}(k)|) \right| \quad (3-54)$$

όπου $s(i)$ και $s^{\sim}(i)$ είναι και πάλι το καθαρό και το αποθορυβοποιημένο σήμα (αντίστοιχα), ενώ οι όροι $S_{P,n}(k)$ και $S_{P,n}^{\sim}(k)$ προκύπτουν από την εφαρμογή του μετασχηματισμού STFT (p -χρονικό παράθυρο, k -φασματική συνιστώσα) στα σήματα $s_n(i)$ και $s_n^{\sim}(i)$. Τα τελευταία αποτελούν κανονικοποιημένες εκφράσεις των $s(i)$, $s^{\sim}(i)$ (εξ. 3-55), όπου το διάνυσμα λευκού θορύβου $n_0(i)$ έχει πολύ μικρή ενέργεια (-30 dB) και η χρήση του συνίσταται στην αποφυγή του υπολογισμού "log(0)" [179], [225]:

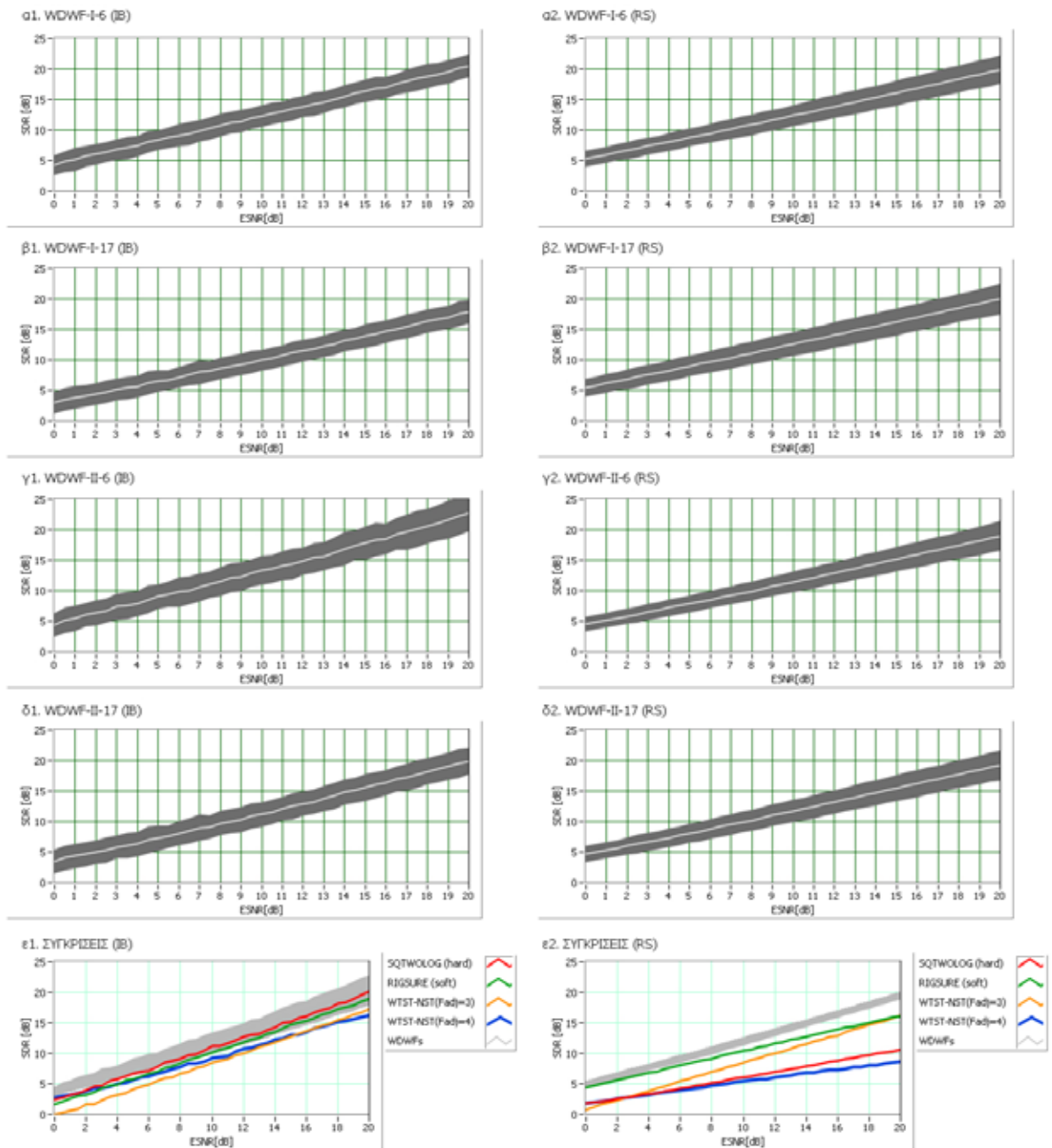
$$s_n(i) = \frac{s(i)}{\|s\|} + n_0(i), \quad s_n^{\sim}(i) = \frac{s^{\sim}(i)}{\|s^{\sim}\|} + n_0(i) \quad (3-55)$$

Η διαδικασία ποσοτικής αξιολόγησης πραγματοποιήθηκε με βάση και τους τρεις παραπάνω δείκτες, ξεχωριστά για τα πρότυπα *IB* και *RS*. Εκτός από τις τεχνικές *WDWF*, συμπεριλήφθηκαν (για λόγους σύγκρισης) οι τεχνικές *RIGSURE*_(soft), *SQTWOLOG*_(hard), *WTST-NST* ($F_{adj}=3$) και *WTST-NST* ($F_{adj}=4$), οι οποίες έδωσαν τα καλύτερα αποτελέσματα (από τις «μεθόδους αναφοράς») κατά τη φάση ποιοτικής αξιολόγησης (§ 3.2.6.2). Η χωριστή αξιολόγηση των προτύπων *IB* και *RS*, αποφασίστηκε καθώς οι εμπλεκόμενες τεχνικές σύγκρισης λειτουργούν αποτελεσματικά, κυρίως στην περίπτωση κρουστικών ήχων [59], [84], [133], [179]. Παράλληλα, δόθηκε η δυνατότητα διερεύνησης της επίδοσης των τεχνικών *WDWF* σε σήματα διαφόρων μορφολογικών χαρακτηριστικών. Τα αποτελέσματα ποσοτικής

αξιολόγησης με βάση τους τρεις δείκτες *CCI*, *SDR* και *SDM* δίνονται στα σχήματα 3-30, 3-31 και 3-32, αντίστοιχα. Στα διαγράμματα αυτά παρουσιάζονται οι διακυμάνσεις των δεικτών (μέση τιμή ± τυπική απόκλιση) για κάθε μια από τις τεχνικές *WDWF* ($\alpha 1$ - $\delta 1$, $\alpha 2$ - $\delta 2$), καθώς επίσης και συγκρίσεις με τις μέσες τιμές των μεθόδων αναφοράς ($\epsilon 1$, $\epsilon 2$). Παρατηρούμε ότι ο τύπος *WDWF-II* παρουσιάζει λιγότερο καλή απόδοση από τον αντίστοιχο τύπο *I*, στις χαμηλές σηματοθορυβικές σχέσεις ($ESNR < 5dB$). Αυτό συμβαίνει, καθώς τα φίλτρα *WDWF-II* λειτουργούν περισσότερο με βάση αντιληπτικά κριτήρια στις παραπάνω συνθήκες, τα οποία δεν είναι δυνατόν να αποδοθούν με τους ποσοτικούς (αντικειμενικούς) δείκτες. Για τον ίδιο λόγο, το ψυχοακουστικό φίλτρο A_{kw} δεν συμπεριλήφθηκε στις διαδικασίες ποσοτικής αξιολόγησης [179].



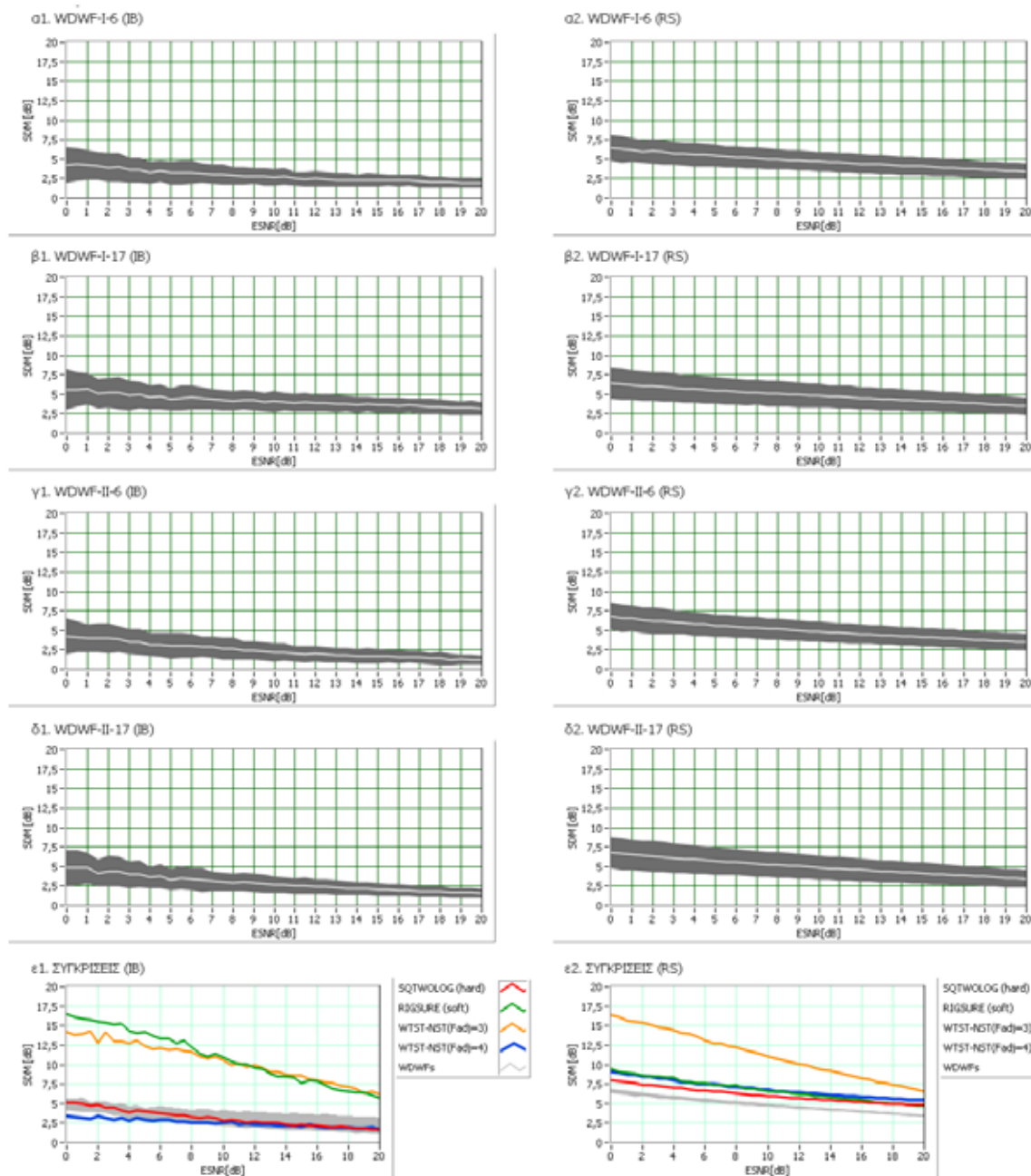
Σχήμα 3-30. Ποσοτική αξιολόγηση των αποτελεσμάτων αποθορυβοποίησης σε διάφορες σηματοθορυβικές σχέσεις ($ESNR=0:0,5:20$ dB) με βάση το δείκτη *CCI*: παρέχονται τα αποτελέσματα (μέση τιμή ± τυπική απόκλιση) των τεσσάρων τεχνικών *WDWF* (α - δ), καθώς και συγκρίσεις με μεθόδους αναφοράς (ϵ), χωριστά για τα πρότυπα *IB* (αριστερά, $\alpha 1 - \epsilon 1$) και τα αντίστοιχα *RS* (δεξιά, $\alpha 2 - \epsilon 2$).



Σχήμα 3-31. Ποσοτική αξιολόγηση των αποτελεσμάτων αποθορυβοποίησης σε διάφορες σηματοθορυβικές σχέσεις ($ESNR=0;0,5;20$ dB) με βάση το δείκτη SDR: παρέχονται τα αποτελέσματα (μέση τιμή \pm τυπική απόκλιση) των τεσσάρων τεχνικών WDF (α-δ), καθώς και συγκρίσεις με μεθόδους αναφοράς (ε), χωριστά για τα πρότυπα IB (αριστερά, α1 - ε1) και τα αντίστοιχα RS (δεξιά, α2- ε2).

Είναι φανερό ότι όλες οι τεχνικές WDF παρέχουν πολύ καλά αποτελέσματα, και στις δύο περιπτώσεις προτύπων (IB, RS) και σε όλο το εύρος των σηματοθορυβικών λόγων ESNR. Μάλιστα, για τιμές ESNR της τάξης των 10dB ($CCI \approx 0,95$) η WDF επεξεργασία πλησιάζει τα χαρακτηριστικά «τέλειας αποθορυβοποίησης» (perfect de-noising) [179]. Από τις εκτιμήσεις του δείκτη CCI (σχ. 3-30), διαπιστώνουμε ότι η τοπολογία DWT-6 είναι αποδοτικότερη από την αντίστοιχη WP-17 στην περίπτωση των προτύπων IB και ιδιαίτερα για χαμηλές τιμές του λόγου ESNR. Στον αντίποδα, οι υλοποιήσεις WDF μέσω κυματιδιακών πακέτων (WP-17), αποδίδουν λίγο καλύτερα στην περίπτωση των προτύπων RS (σχ. 3-31). Από την ανάλυση του σχ. 3-32, προκύπτει ότι ο τύπος II προκαλεί μικρότερη φασματική παραμόρφωση (SDR),

στοιχείο που συνδέεται με την ηπιότερη συμπίεση των συντελεστών στα διαστήματα παρουσίας του σήματος (σχετική ανάλυση έγινε στην §3.2.5.2). Όσον αφορά τη σύγκριση με τις μεθόδους αναφοράς, παρατηρούμε ότι οι $SQWLOG_{(hard)}$ και $WTST-NST (F_{adj}=4)$ είναι καλές επιλογές για την περίπτωση των προτύπων IB , ενώ οι $RIGSURE_{(soft)}$ και $WTST-NST (F_{adj}=3)$ λειτουργούν καλύτερα με πρότυπα RS (σχ. 3-30). Σε κάθε περίπτωση, η επίδοσή τους υπολείπεται κατά πολύ της αντίστοιχης των $WDWF$ μεθόδων, ενώ οι διαφορές μεγαλώνουν κατά την επεξεργασία RS σημάτων. Μπορούμε, λοιπόν, να ισχυριστούμε ότι οι τεχνικές $WDWF$ ανταποκρίνονται πλήρως στις προδιαγραφές του αρχικού σχεδιασμού, παρέχοντας πολύ καλά αποτελέσματα καταστολής θορύβου (robust de-noising), χωρίς αλλοίωση των χρήσιμων συνιστωσών (minimal signal distortion) [179].



Σχήμα 3-32. Ποσοτική αξιολόγηση των αποτελεσμάτων αποθορυβοποίησης σε διάφορες σηματοθρομβικές σχέσεις ($ESNR=0:0,5:20$ dB) με βάση το δείκτη SDM: παρέχονται τα αποτελέσματα (μέση τιμή \pm τοπική απόκλιση) των τεσσάρων τεχνικών $WDWF$ (α-δ), καθώς και συγκρίσεις με μεθόδους αναφοράς (ε), χωριστά για τα πρότυπα IB (αριστερά, α1 - ε1) και τα αντίστοιχα RS (δεξιά, α2- ε2).

3.3 Επισκόπηση 3ου κεφαλαίου

Στο κεφάλαιο αυτό αναλύθηκαν τα θέματα που σχετίζονται με την παρουσία του προσθετικού θορύβου βάρους *ABN* και τα αντίστοιχα προβλήματα που ανακύπτουν κατά την επεξεργασία και ανάλυση των κοιλιακών ήχων. Στο πλαίσιο αυτό παρουσιάστηκαν οι τεχνικές καταστολής θορύβου, που έχουν χρησιμοποιηθεί στο συγκεκριμένο πεδίο βιοϊατρικής έρευνας, ενώ αναπτύχθηκε μια νέα μεθοδολογία επεξεργασίας, που αποτέλεσε ένα από τα κύρια αντικείμενα της διατριβής. Η νέα προσέγγιση αφορά στο συνδυασμό της φασματικής επεξεργασίας Wiener με την κυματιδιακή ανάλυση, ενώ κατά την ποιοτική και ποσοτική αξιολόγηση των αλγορίθμων που υλοποιήθηκαν, διαπιστώθηκε πολύ καλή απόδοση στις διάφορες κατηγορίες βιοακουστικών σημάτων.

3.3.1 Αποτίμηση προτεινόμενων παρεμβάσεων

Τα παραπάνω συμπεράσματα είναι κεφαλαιώδους σημασίας, καθώς οι τεχνικές καταστολής θορύβου *WDWF* έχουν χρησιμοποιηθεί σε διάφορες εφαρμογές στο πλαίσιο της παρούσης διατριβής, όπως είναι οι αλγόριθμοι αυτόματης σάρωσης για τον εντοπισμό και την τοπογραφική ερμηνεία των διαφόρων συμβάντων, η εξαγωγή ενεργειακών παραμέτρων σύντομης χρόνου-συχνότητας (απαλλαγμένων από θόρυβο), η συνεισφορά στην αναγνώριση προτύπων κοιλιακών ήχων και στις προσεγγίσεις ακουστικής χαρτογράφησης (σχετικές λεπτομέρειες δίνονται στα κεφάλαια 4, 5 και 6). Ειδικότερα, αναπτύχθηκαν τέσσερις εναλλακτικές μέθοδοι κυματιδιακής επεξεργασίας, που παρουσιάζουν πολύ καλή απόδοση, δυνατότητα εύκολης ρύθμισης, περιορισμένη πολυπλοκότητα και μικρό υπολογιστικό κόστος. Στα θετικά των νέων τεχνικών συμπεριλαμβάνεται η δυνατότητα εφαρμογής τους σε σήματα οσοδήποτε μεγάλου μήκους (χωρίς επιπτώσεις στα αποτελέσματα επεξεργασίας), ενώ πολύ εύκολη είναι και η αξιοποίησή τους σε καταγραφές παρατεταμένου χρόνου. Όπως προέκυψε από τη σύγκριση με τις προϋπάρχουσες μεθόδους, οι νέες τεχνικές παρουσιάζουν πολύ καλύτερες επιδόσεις, ενώ είναι η πρώτη φορά που παρατηρείται η εξίσου επιτυχής αποθορυβοποίηση μη κρουστικών κοιλιακών ήχων. Ιδιαίτερο ενδιαφέρον παρουσιάζουν και οι προσεγγίσεις αυτόματου εντοπισμού των κυματιδιακών κατωφλίων, οι οποίες, αν και δεν αποτέλεσαν βασικό αντικείμενο της διατριβής, αποδείχθηκαν ιδιαίτερα βολικές, ως προς την απλότητα υλοποίησης αλλά και την αποτελεσματικότητά τους.

3.3.2 Στοιχεία πρωτοτυπίας & ερευνητική συνεισφορά

Στα στοιχεία πρωτοτυπίας συμπεριλαμβάνεται, καταρχάς, η φιλοσοφία επεξεργασίας «σημείο προς σημείο», που προέκυψε από το συνδυασμό του κυματιδιακού φίλτρου Wiener και της αντίστοιχης διαδικασίας εκθετικού μέσου, ενώ ο 2^{ος} τύπος επεξεργασίας (*WDWF-II*), αποτελεί καινοτομία που αναπτύχθηκε εξολοκλήρου στο πλαίσιο της παρούσης διατριβής. Η νέα αυτή προσέγγιση συνδυάζει πολύ καλά χαρακτηριστικά λειτουργίας όπως είναι η αποτελεσματική καταστολή του θορύβου, τόσο σε περιόδους σιγής, όσο και σε περιόδους παρουσίας σήματος, η δυνατότητα ηπιότερης «συμπίεσης» κατά την εμφάνιση χρήσιμων συνιστωσών και η «ελάχιστη αλλοίωση» των τελευταίων, η μεγαλύτερη ανοχή σε σφάλματα αποθορυβοποίησης και υπολείμματα θορύβου, η διατήρηση των στοιχείων περιορισμένης πολυπλοκότητας και η ενσωμάτωση αντιληπτικών κριτηρίων. Σε ότι αφορά στις προσεγγίσεις αντιληπτικής επεξεργασίας, καινοτόμα είναι και η χρήση του ψυχοακουστικού φίλτρου A_{kw} , που σχετίζεται με τα χαρακτηριστικά του θορύβου ομοιόμορφης απόκρυψης και εφαρμόζεται για πρώτη φορά στην επεξεργασία βιοακουστικών σημάτων. Η

διαμόρφωση του «φασματικά εξαρτώμενου κυματιδιακού φίλτρου Wiener» αποδείχθηκε ιδιαίτερα χρήσιμη, τόσο κατά την ακροαστική ανάλυση των γαστρεντερικών ακουστικών σημάτων, όσο και κατά την εφαρμογή τεχνικών αυτόματης επεξεργασίας. Σχετικά με τις διαδικασίες ποιοτικής και ποσοτικής αξιολόγησης, είναι επίσης η πρώτη φορά που κάτι τέτοιο διεξάγεται χωριστά για τα πρότυπα *IB* και *RS*, απ' όπου προέκυψε ότι οι τεχνικές *WDWF* μπορούν να εφαρμοστούν εξίσου ικανοποιητικά σε όλες τις κατηγορίες σημάτων, ξεπερνώντας τα προβλήματα των προϋπαρχόντων μεθόδων.

3.3.3 Προτεινόμενες μελλοντικές εξελίξεις

Οι προτεινόμενες μελλοντικές εξελίξεις των τεχνικών *WDWF* μπορούν να εντοπιστούν σε δύο σημεία. Το ένα αφορά στην αξιοποίησή τους σε άλλες εφαρμογές επεξεργασίας ήχου, όπως είναι η ανάλυση σημάτων ομιλίας - φωνής, η αποκατάσταση ενθόρυβων μουσικών ηχογραφήσεων, η διεξαγωγή ακουστικών μετρήσεων σε περιβάλλον θορύβου, η επεξεργασία δονήσεων δομικών συστημάτων. Ήδη έχουν πραγματοποιηθεί κάποιες πρώτες πειραματικές εφαρμογές σε σήματα των παραπάνω κατηγοριών, όπου τα αποτελέσματα είναι ιδιαίτερα ενθαρρυντικά. Παράλληλα, ενδιαφέρον παρουσιάζει η χρήση των νέων τεχνικών σε άλλες κατηγορίες βιοακουστικών φαινομένων, όπως είναι οι καρδιακοί, οι πνευμονικοί και οι μυοσκελετικοί ήχοι. Το δεύτερο σημείο μελλοντικής επέκτασης αφορά στην εξέλιξη των αλγορίθμων αυτόματου εντοπισμού κατώφλιων, με την αξιοποίηση στοιχείων των τεχνικών *WTST-NST*, *WT-FD*, *IKD*, αλλά και τη χρήση νέων εργαλείων ανάλυσης. Στο πλαίσιο αυτό, έχει προταθεί η ενσωμάτωση διαδικασιών εκθετικού μέσου, με στόχο τη διαρκή ανανέωση των ενεργειακών-φασματικών χαρακτηριστικών του θορύβου και την επίτευξη μεγαλύτερης προσαρμοστικότητας κατά την αποθορυβοποίηση. Το παραπάνω εγχείρημα βρίσκεται ήδη σε φάση σχεδιασμού, ενώ με ενδιαφέρον αναμένονται τα πρώτα αποτελέσματα αυτών των υλοποιήσεων. Υπό εξέταση είναι και μια εναλλακτική προσέγγιση καταστολής θορύβου, αξιοποιώντας πρόσθετα σήματα πολυκαναλικής λήψης και αντίστοιχους αλγόριθμους επεξεργασίας (*multi-channel noise reduction*). Σε κάθε περίπτωση, οι τεχνικές αυτόματης αποθορυβοποίησης χρήζουν πλήρους διερεύνησης των επιδόσεών τους, με αξιοποίηση των κατάλληλων ποιοτικών και ποσοτικών δεικτών.

ΚΕΦΑΛΑΙΟ 4^ο

4. Επεξεργασία παρατεταμένου χρόνου

Στο κεφάλαιο αυτό αναλύονται τα ζητήματα επεξεργασίας και ανάλυσης των καταγραφών παρατεταμένου χρόνου. Παρουσιάζονται τα γενικά προβλήματα διαχείρισης, που γεννιούνται από το μεγάλο όγκο των πληροφοριών, καθώς και οι δυσκολίες που σχετίζονται με την εμφάνιση ανεπιθύμητων παρεμβολών κατά μήκος των χρήσιμων πειραματικών δεδομένων. Κατόπιν, παρατίθενται οι βασικές μεθοδολογίες που υιοθετούνται συνήθως κατά την επεξεργασία ηχογραφήσεων μεγάλης χρονικής διάρκειας, σε συνδυασμό με την αξιοποίηση αλγορίθμων ανίχνευσης και εντοπισμού σήματος. Η ανάλυση εστιάζει στις ιδιαιτερότητες των μεθόδων ακουστικής μελέτης της γαστρεντερικής δραστηριότητας και τα χαρακτηριστικά των μεθόδων που έχουν χρησιμοποιηθεί έως τώρα. Στο πλαίσιο αυτό, παρουσιάζονται οι αλγόριθμοι που αναπτύχθηκαν για τις ανάγκες της παρούσης διατριβής, καθώς επίσης και τα πειραματικά αποτελέσματα αξιολόγησης της συμπεριφοράς τους.

4.1 Επεξεργασία ηχητικών καταγραφών παρατεταμένου χρόνου

Η επεξεργασία ηχητικών σημάτων παρατεταμένης χρονικής διάρκειας, συναντάται σε διάφορες εφαρμογές της «μηχανικής ήχου» (sound engineering), όπως είναι η διεξαγωγή και επεξεργασία πολυκαναλικών μουσικών ηχογραφήσεων, η ανάλυση των διαφόρων μορφών ακουστικού θορύβου και η αξιολόγηση της επίδρασής τους στον άνθρωπο, η αξιοποίηση ηλεκτρακουστικών συστημάτων σε συνεδριακές μονάδες καταγραφής διαλέξεων και σε συστήματα ασφαλείας, η εφαρμογή πρωτοκόλλων ακουστικών μετρήσεων και μεθόδων «ακουστικής» διάγνωσης [94]-[95], [161]-[162], [217], [226]. Όσον αφορά την τελευταία περίπτωση, εδώ συμπεριλαμβάνονται οι τεχνικές ακουστικού ελέγχου των δομικών και μηχανικών συστημάτων (structural - mechanical acoustics), και φυσικά η κατηγορία των βιοακουστικών μεθόδων, όπου εμπίπτει και η παρούσα εργασία.

Όπως έχει ήδη επισημανθεί, το σημαντικότερο πρόβλημα αυτών των εφαρμογών εντοπίζεται στη διαχείριση και ανάλυση ενός μεγάλου όγκου δεδομένων, καθώς και στις δυσκολίες εφαρμογής των απαραίτητων τεχνικών επεξεργασίας. Για την αντιμετώπισή τους υιοθετούνται τεχνικές ολοκλήρωσης και αντίστοιχα σήματα σύνοψης (summarization curves), ενώ η επεξεργασία των δεδομένων γίνεται με χρήση χρονικών παραθύρων επιλογής. Στο πλαίσιο αυτό αναπτύχθηκε και το πρωτόκολλο MPEG-7 [164], [166], [227] που περιλαμβάνει ένα σύνολο από κανόνες και σχήματα περιγραφής για την πλοήγηση και διαχείριση οπτικοακουστικών πληροφοριών, όπως είναι οι βιοακουστικές καταγραφές με τα συνοδευτικά πολυμεσικά στοιχεία περιγραφής, ελέγχου και οπτικής παρακολούθησης της πειραματικής διαδικασίας [161]-[162], [217], [226].

Στις περισσότερες περιπτώσεις, εκτός από τις χρήσιμες πληροφοριακές συνιστώσες εντοπίζονται διάφορα σήματα παρεμβολών, καθώς και μεγάλα διαστήματα σιγής. Έτσι, εφαρμόζονται τεχνικές σάρωσης - ανίχνευσης για τον εντοπισμό των περιόδων «ακουστικής δραστηριότητας», καθώς και έμπειρα συστήματα αναγνώρισης προτύπων, για την απομάκρυνση των ανεπιθύμητων παρεμβολών [161]-[162], [215]-[218]. Τα ακουστικά συμβάντα (events) που προκύπτουν από την παραπάνω διαδικασία και δεν αντιστοιχούν σε

σήματα θορύβου, αποτελούν τις «δομικές μονάδες» περιγραφής του συνολικού φαινομένου, στα οποία μπορούν να εφαρμοστούν πιο εξελιγμένοι μέθοδοι επεξεργασίας και ανάλυσης. Στο πλαίσιο αυτό, ένα γενικό μοντέλο διαχείρισης ηχητικών καταγραφών παρατεταμένου χρόνου περιλαμβάνει τα ακόλουθα στάδια: α) εντοπισμό των σημαντικών ακουστικών συμβάντων κατά μήκος του χρόνου παρακολούθησης, β) απομάκρυνση των ανεπιθύμητων παρεμβολών, γ) εφαρμογή οντολογιών ανάλυσης και σχημάτων περιγραφής για τα απομονωμένα ακουστικά στιγμιότυπα σημαντικής δραστηριότητας (significant events), δ) εξαγωγή παραμέτρων σύνοψης για την περιγραφή του συνολικού φαινομένου [217], [226].

4.1.1 Τεχνικές ανίχνευσης και εντοπισμού σήματος

Υπάρχουν διάφορες προσεγγίσεις για τον εντοπισμό σημάτων μέσα σε ενθόρυβες ηχητικές καταγραφές. Χαρακτηριστικές εφαρμογές αποτελούν οι τεχνικές διαχείρισης πολυμεσικού υλικού μέσω ακουστικών εργαλείων, οι αλγόριθμοι εντοπισμού των έμφωνων και άφωνων τμημάτων σε σήματα ομιλίας, οι μέθοδοι αναγνώρισης ομιλίας και βιοακουστικών προτύπων, τα εργαλεία ανάλυσης βιοακουστικών καταγραφών. Στο πλαίσιο αυτό χρησιμοποιούνται συνδυασμοί παραμέτρων (features) χρόνου - συχνότητας, όπως είναι η ενεργειακή σύγκριση ηχητικών σταθμών, η αξιοποίηση κυματομορφών περιβάλλουσας (envelope curves) και ο υπολογισμός των χρόνων ανόδου - καθόδου (attack-release time), οι τεχνικές εντοπισμού και ανάλυσης «κορυφών» (peak detection - analysis), η εύρεση του ρυθμού μετάβασης από το «μηδέν» (Zero Crossing Rate -ZCR) και οι εκφράσεις των διακυμάνσεων πλάτους (Average Magnitude Function -AMF, Average Magnitude Difference Function -AMDF) [161]-[162], [215]-[218], [228]-[234]. Σε εφαρμογές «σύντομου χρόνου» (short-term processing), όπου το υπολογιστικό κόστος δεν αποτελεί τόσο κρίσιμη παράμετρο, εφαρμόζονται πιο απαιτητικοί αλγόριθμοι, που συνδυάζουν τεχνικές στατιστικής ανάλυσης ανώτερης τάξης (HOS) [85], [97], [235], μεθόδους ιδιοδιανυσμάτων και ιδιαζουσών τιμών (Singular Value Decomposition, - SVD) [236]-[237], εργαλεία πολυπλοκότητας κυματομορφών (FD) [82]-[84], κ.ά.

Από τις παραπάνω μεθόδους, οι «ενεργειακές τεχνικές» βασίζονται στη σύγκριση ενεργειακών παραμέτρων (ηχητική στάθμη, ενέργεια, ισχύς) των επιλεγμένων σημάτων και των αντίστοιχων περιόδων σιγής. Πρόκειται για τις πιο απλές στη σύλληψη προσεγγίσεις, όπου οι «λογικές εκφράσεις εντοπισμού» και τα αντίστοιχα επίπεδα σύγκρισης καθορίζονται εμπειρικά ή με βάση αντιληπτικά κριτήρια. Όπως θα δούμε στη συνέχεια, τέτοιου είδους προσεγγίσεις ακολουθήθηκαν, ως ένα βαθμό, και στο πλαίσιο της παρούσης διατριβής. Παράλληλα αξιοποιήθηκαν εργαλεία ανάλυσης “Fractal Dimension”, τα οποία αξιολογούν τη «γεωμετρική πολυπλοκότητα» των επιλεγμένων κυματομορφών, ανεξάρτητα από τα ενεργειακά τους επίπεδα ή τα αντίστοιχα χαρακτηριστικά των περιόδων σιγής. Τα τελευταία χρόνια, οι τεχνικές ανάλυσης FD χρησιμοποιούνται ολοένα και περισσότερο σε εφαρμογές επεξεργασίας βιοακουστικών σημάτων, συμπεριλαμβανομένων και των γαστρεντερικών ήχων [82]-[84], [217].

4.1.2 Ανάλυση πολυπλοκότητας κυματομορφών “Fractal Dimension”

Ο όρος “fractal dimension” (FD) αναφέρεται στην «κλασματική έκφραση» της διάστασης γεωμετρικών αντικειμένων (*non-integer or fractional dimension of geometric objects*) [217], [238]. Δηλαδή, οι τελεστές FD εκτιμούν την «πολυπλοκότητα» των διαφόρων σημάτων, θεωρώντας τις κυματομορφές τους στο πεδίο του χρόνου, ως γεωμετρικά σχήματα. Διάφορες μέθοδοι έχουν προταθεί για τον υπολογισμό της πολυπλοκότητας FD, είτε απευθείας στο

πεδίο του χρόνου, είτε στο πεδίο της συχνότητας [238]-[239]. Οι παράμετροι ανάλυσης FD είναι αρκετά συνηθισμένες στην επεξεργασία βιοϊατρικών σημάτων, όπως είναι η ταξινόμηση καταγραφών ηλεκτροεγκεφαλογραφίας και ηλεκτροκαρδιογραφίας (EEG, EKG) [239]-[241], η διάγνωση φωνητικών διαταραχών και γενικότερα η ανάλυση βιοακουστικών φαινομένων [242]-[246], ο εντοπισμός των περιόδων «παρουσίας σήματος» και η εύρεση των κατωφλίων αποθρομβοποίησης [82]-[84]. Γενικά, οι προσεγγίσεις FD είναι ιδιαίτερα αποδοτικές στην ανίχνευση απότομων μεταβάσεων (transient detection), με το πρόσθετο πλεονέκτημα του μικρού υπολογιστικού κόστους [239]. Αυτοί είναι και οι βασικοί λόγοι αξιοποίησή τους στο πλαίσιο της παρούσης διατριβής [217].

Δύο από τις πιο κλασσικές προσεγγίσεις υπολογισμού της πολυπλοκότητας FD , που εφαρμόζονται απευθείας στο πεδίο του χρόνου, είναι οι μέθοδοι του Katz (FD_K) [240] και του Petrosian (FD_P) [241]. Ο τελεστής FD_K (εξ. 4-1) έχει το πλεονέκτημα ότι εκτιμά με μεγαλύτερη ακρίβεια την περιβάλλουσα του σήματος (κατά την εφαρμογή του με χρήση κυλιόμενων παραθύρων - "sliding windows") [217], ενώ είναι και πιο αποδοτικός στον εντοπισμό των διαστημάτων σιγής. Από τις διάφορες τεχνικές FD_P , θα εστιάσουμε στο 3ο τύπο $FD_{P;c}$ (εξ. 4-2, type c, Petrosian), ο οποίος παρέχει παραπλήσια αποτελέσματα ανίχνευσης σήματος, με το πλεονέκτημα του πολύ μικρότερου υπολογιστικού κόστους (σχετικά πειράματα έδειξαν ότι ο υπολογισμός του τελεστή $FD_{P;c}$ γίνεται 10 φορές πιο γρήγορα από τον αντίστοιχο FD_K) [217], [239]. Έτσι, ο τύπος $FD_{P;c}$ είναι καταλληλότερος για την επεξεργασία καταγραφών παρατεταμένου χρόνου.

$$FD_K(y) = \frac{\log_{10}(L_{tot})}{\log_{10}(d_m)} \quad (4-1)$$

$$d_m = \max(\text{distance}(1, i))$$

$$FD_{P;c}(y) = \frac{\log_{10}(N)}{\log_{10}(N) + \log_{10}\left(\frac{N}{N + 0,4 \cdot N_{\Delta}}\right)} \quad (4-2)$$

Όπου y είναι το ψηφιοποιημένο σήμα εισόδου, L_{tot} το «συνολικό μήκος» (total length) της κυματομορφής y , το οποίο αντιστοιχεί στο άθροισμα των αποστάσεων μεταξύ διαδοχικών δειγμάτων, και d_m η «διάμετρος», που υπολογίζεται ως η απόσταση ανάμεσα στο πρώτο δείγμα της ακολουθίας $y(i)$ και το δείγμα που παρέχει τη μεγαλύτερη απόσταση. Ο τελεστής "distance" αντιστοιχεί στην Ευκλείδεια απόσταση ανάμεσα σε δύο σημεία (A, B) επάνω στην καμπύλη y , με καρτεσιανές συντεταγμένες $[i_A, y(i_A)]$ και $[i_B, y(i_B)]$. Στην εξ. 4-2, N είναι το μήκος της ακολουθίας εισόδου και N_{Δ} είναι το πλήθος αλλαγών πρόσημου της δυαδικής ακολουθίας $[y(i+1)-y(i)]$, που προκύπτει από τη διαφορά των διαδοχικών σημείων της καμπύλης y [217], [239]-[241]. Εκτός από τις τεχνικές άμεσου υπολογισμού της παραμέτρου FD στο πεδίο του χρόνου, υπάρχουν και οι μέθοδοι «φασματικής ισχύος» (Power Spectrum Methods, -PSM), οι οποίες εκτιμούν την πολυπλοκότητα FD από την κλίση της λογαριθμικής φασματικής καμπύλης [238], [242]-[245]. Έτσι, αν P_i είναι η ακολουθία φασματικής ισχύος, που υπολογίζεται με εφαρμογή του $STFT$, τότε η FD_{PSM} [245]:

$$FD_{PSM} = \frac{5 - \beta_p}{2}, \quad \text{όπου} \quad P_i = c_p \cdot \|k_i\|^{-\beta_p} \quad (4-3)$$

Δηλαδή, οι συνιστώσες φασματικής ισχύος P_i μοντελοποιούνται συναρτήσει των θετικών σταθερών c_p και β_p (υπολογίζονται γραφικά ή μέσω αριθμητικών μεθόδων ελαχίστων τετραγώνων), έτσι ώστε η εξ. 4-3β να αποτελεί γραμμική προσέγγιση της φασματικής πυκνότητας ισχύος P_i , κατά τη λογαριθμική της απεικόνιση (log-log plot) ως προς τη συχνότητα [217], [244]-[245]. Η εκτίμηση της πολυπλοκότητας FD_{PSM} μπορεί να γίνει και με οποιαδήποτε άλλη μέθοδο φασματικής ανάλυσης (εκτός από τον αλγόριθμο FFT), όπως είναι ο μετασχηματισμός κυματιδίων [242]. Μια διαφορετική προσέγγιση ανάλυσης, είναι ο απευθείας υπολογισμός της πολυπλοκότητας FD (βάσει των εξ. 4-1 και 4-2) για καθεμιά από τις ακολουθίες κυματιδιακών συντελεστών, που προκύπτουν με εφαρμογή του διακριτού μετασχηματισμού DWT ή των κυματιδιακών πακέτων WP [83]-[84], [217].

4.2 Ανάλυση γαστρεντερικών ακουστικών καταγραφών παρατεταμένου χρόνου

Όπως έχει ήδη αναφερθεί, υπάρχουν δύο βασικές στρατηγικές που εφαρμόζονται στη μελέτη του «μηχανισμού γαστρεντερικής λειτουργίας», μέσω βιοακουστικών καταγραφών. Η πρώτη αφορά στην ανάλυση σύντομου χρόνου, όπου διαστήματα μικρής διάρκειας και καθορισμένου χρονισμού λαμβάνονται ως αντιπροσωπευτικά δείγματα του συνολικού φαινομένου [42], [74]-[75]. Η δεύτερη, αφορά στην παρακολούθηση των γαστρεντερικών βιοακουστικών φαινομένων για παρατεταμένο χρονικό διάστημα, αξιοποιώντας καμπύλες συνοπτικής απεικόνισης και εργαλεία στατιστικής ανάλυσης [24]-[25], [64], [69], [72]-[73]. Μια τρίτη προσέγγιση, η οποία επιχειρεί να αξιοποιήσει πλεονεκτήματα και των δύο προηγούμενων στρατηγικών, αφορά στον εντοπισμό και την περαιτέρω ανάλυση των σημαντικών ακουστικών συμβάντων, που αποτελούν τα «δομικά στοιχεία» γαστρεντερικής δραστηριότητας [161]-[162], [215]-[218]. Η υβριδική αυτή προσέγγιση επιδιώκει τη χρήση νέων-εξελιγμένων σχημάτων περιγραφής περιεχομένου (συμπεριλαμβανομένων των παραμέτρων σύνοψης - περίληψης) και αποτέλεσε τη βάση σχεδιασμού και ανάπτυξης του προτεινόμενου πρωτοκόλλου βιοϊατρικής παρακολούθησης [215]-[218], [226].

4.2.1 Υπάρχουσα κατάσταση και ιδιαιτερότητες του προβλήματος

Όσον αφορά τις υπάρχουσες τεχνικές ανίχνευσης και εντοπισμού των κοιλιακών βιοακουστικών φαινομένων, σχετική είναι η μέθοδος $FDPPA$ των Χατζηλεοντιάδη και Ρέκανου [82], καθώς και η εργασία των Ranta κ.ά. [86]. Η πρώτη χρησιμοποιεί ολισθαίνοντα παράθυρα «σημείο προς σημείο» σε συνδυασμό με τελεστές FD_K , και όπως έχει ήδη σχολιαστεί στο προηγούμενο κεφάλαιο, έχει προταθεί για τον εντοπισμό κρουστικών βιοακουστικών σημάτων. Πειραματικές δοκιμές έχουν δείξει [179] ότι η μέθοδος δεν αποδίδει εξίσου ικανοποιητικά στην περίπτωση ομαλών κυματομορφών, όπως είναι τα πρότυπα RS , ενώ σε κάθε περίπτωση το υπολογιστικό κόστος είναι απαγορευτικό για χρήση σε καταγραφές πολύ μεγάλης διάρκειας [217]. Η δεύτερη προσέγγιση, όπως προκύπτει και από τον τίτλο της αντίστοιχης εργασίας [86] (*Wavelet-based bowel sounds denoising, segmentation and characterization*), περιλαμβάνει την αποθρομβοποίηση, το διαχωρισμό και την ερμηνεία επιλεγμένων τμημάτων των βιοακουστικών καταγραφών, χρησιμοποιώντας μια τροποποιημένη εκδοχή του αλγορίθμου $WTST-NST$. Έτσι, παρουσιάζει αρκετές δυσκολίες κατά την εφαρμογή της στην (μη εποπτευόμενη) επεξεργασία καταγραφών μεγάλης διάρκειας, όπως είναι η μη ικανοποιητική αποθρομβοποίηση των κυματομορφών RS και το μεγάλο υπολογιστικό κόστος [179], [217]. Επιπρόσθετα, και οι δύο παραπάνω τεχνικές τείνουν να εντοπίζουν έναν υπερβολικό αριθμό συμβάντων, τα οποία ενδέχεται να

απέχουν ελάχιστα μεταξύ τους, αποτέλεσμα όχι και τόσο λειτουργικό για την ανάλυση και διαχείριση των καταγραφών παρατεταμένου χρόνου [217].

Στο πλαίσιο αυτό, αναζητήθηκε μια εναλλακτική προσέγγιση για τον εντοπισμό των σημαντικών βιοακουστικών συμβάντων, δίνοντας ιδιαίτερη βαρύτητα στο ζήτημα των υπολογιστικών απαιτήσεων, αλλά και την αδυναμία εποπτείας της ανάλυσης. Συνοπτικά, η προτεινόμενη διαδικασία επεξεργασίας παρατεταμένου χρόνου περιλαμβάνει τα ακόλουθα: α) εντοπισμό των περιόδων σημαντικής βιοακουστικής δραστηριότητας, β) απομόνωση των σημαντικών συμβάντων από τις περιόδους σιγής και τους ανεπιθύμητους θορύβους παρεμβολών, γ) τοπογραφική ερμηνεία των «καταχωρημένων» συμβάντων (registered events) που εντοπίστηκαν στο “β” και αξιοποίηση πρόσθετων σχημάτων περιγραφής τους, δ) εξαγωγή παραμέτρων σύνοψης - περίληψης του συνολικού φαινομένου, ε) συνδυασμό όλων των παραπάνω παρατηρήσεων σε μια ευέλικτη οντολογία πλοήγησης, διαχείρισης περιεχομένου και ανάλυσης [217].

4.2.2 Τεχνικές σάρωσης με χρήση παραθύρων μικρής διάρκειας

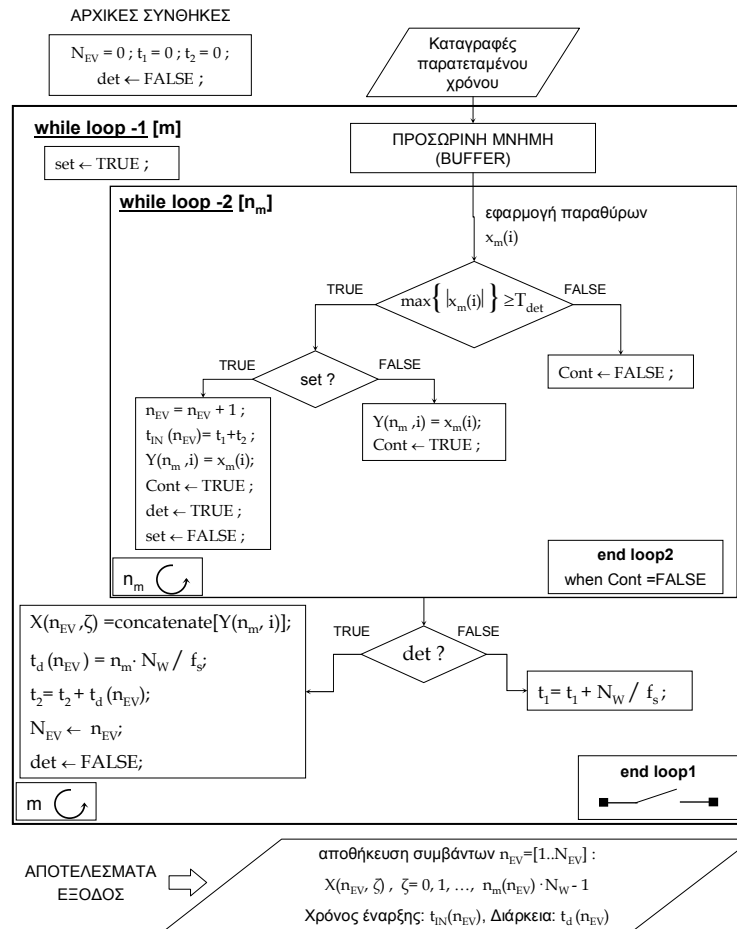
Στην περίπτωση καταγραφών παρατεταμένης διάρκειας (≥ 2 ώρες), είναι προφανές ότι η άμεση επεξεργασία του συνολικού σήματος είναι αδιανόητη, καθώς δεν υπάρχει ρεαλιστικό υπολογιστικό σύστημα που να μπορεί να διαχειριστεί απευθείας ένα τόσο μεγάλο όγκο δεδομένων. Μάλιστα, ακόμη και ο συμβολισμός $x(i)$ είναι προβληματικός, καθώς, για μια τυπική διάρκεια 2 ωρών και με δεδομένη τη συχνότητα δειγματοληψίας ($f_s=8kHz$), ο δείκτης i μεταβάλλεται μέχρι και την τιμή των 58.000.000 δειγμάτων, περίπου. Στις περιπτώσεις αυτές προτιμάται η επεξεργασία των σημάτων με χρήση παραθύρων - πλαισίων επιλογής (windowing techniques). Αν χρησιμοποιήσουμε τους συμβολισμούς m και N_w για το δείκτη και το μήκος (σε δείγματα) των ορθογωνικών (μη επικαλυπτόμενων) παραθύρων επιλογής, τότε τα τμήματα της ενθόρυβης καταγραφής $x_m(i)$ σε κάθε χρονική στιγμή m (m -instant, m -windowed signal), θα έχουν την ακόλουθη μορφή:

$$x_m(i) = x(m \cdot N_w + i), \quad 0 \leq i \leq N_w - 1, \quad 0 \leq m < \left\lceil \frac{\text{Length}(x)}{N_w} \right\rceil \quad (4-4)$$

όπου ο τελεστής “Length(x)” υπολογίζει το συνολικό μήκος (αριθμό δειγμάτων) της καταγραφής x , ενώ ο τελεστής “ $\lceil \rceil$ ” δίνει την αμέσως μεγαλύτερη ακέραια τιμή του ορίσματος (ceiling operator). Στο πλαίσιο αυτό, είναι δυνατό να εφαρμοστούν εκφράσεις ενεργειακής σύγκρισης σε κάθε παράθυρο σήματος $x_m(i)$ με προκαθορισμένες τιμές που αντιστοιχούν στα ενεργειακά χαρακτηριστικά των διαστημάτων σιγής. Έτσι προέκυψε η διαδικασία σάρωσης με χρήση παραθύρων μικρής διάρκειας (Short Block Scanning - SBS), ένα απλοποιημένο διάγραμμα της οποίας δίνεται στο σχ. 4-1 [161]-[162].

Ειδικότερα, τα επιλεγμένα πλαίσια σήματος $x_m(i)$ ελέγχονται για το αν ικανοποιούν τις συνθήκες εντοπισμού δραστηριότητας (det) μέσω της έκφρασης “ $WS_{OP}\{x_m(i)\} \geq T_{det}$ ”, όπου WS_{OP} είναι ένας «ενεργειακός τελεστής» (OPERAND: ενεργός τιμή, ενέργεια, ισχύς, τιμή μεγίστου, κλπ) και T_{det} είναι η προκαθορισμένη τιμή «κατωφλίου εντοπισμού». Η εξωτερική επαναληπτική διαδικασία του σχ. 4-1 (while loop 1) επαναλαμβάνεται μέχρις ότου η σάρωση εφαρμοστεί σε ολόκληρο το μήκος των καταγραφών, ή στα διαστήματα που ορίζονται από το χρήστη. Η ανίχνευση ενός νέου συμβάντος επιτυγχάνεται με τη βοήθεια της βοηθητικής μεταβλητής set , που λειτουργεί ως σημαία (flag). Μόλις εντοπιστεί ένα νέο συμβάν εκτελείται η εσωτερική επαναληπτική διαδικασία (while loop 2), η οποία συνεχίζεται για όσο η λογική συνθήκη $Cont$ παραμένει αληθής. Έτσι, όλα τα διαδοχικά

πλαίσια $n_m(n_{EV})$ που ικανοποιούν τις συνθήκες εντοπισμού (ενεργά πλαίσια) ενσωματώνονται (concatenated) σε ένα ενιαίο συμβάν n_{EV} , το οποίο έχει χρονική έναρξη $t_{IN}(n_{ev})$ και διάρκεια $t_d(n_{ev})$. Για τον υπολογισμό των τελευταίων αξιοποιούνται οι βοηθητικές μεταβλητές t_1, t_2 . Ως αποτέλεσμα, προκύπτουν N_{EV} συνολικά συμβάντα, των οποίων τα ηχητικά δεδομένα $X(n_{EV}, \zeta)$ καταχωρούνται σε δυαδικά αρχεία τύπου «εγγραφών» (binary data records), μαζί με κάποια συνοδευτικά στοιχεία περιγραφής της πειραματικής διαδικασίας. Ο νέος τύπος αρχείων (*.sbs) προέκυψε από το όνομα της τεχνικής *Short Block Scanning*, αλλά και από το ακρωνύμιο *Streaming Bowel Sounds*, που υποδηλώνει ότι τα χρήσιμα δεδομένα μεταφέρονται σε νέα αρχεία [161]-[162], [215].



Σχήμα 4-1. Μπλοκ διάγραμμα λειτουργίας της διαδικασίας σάρωσης με χρήση παραθύρων μικρής διάρκειας (Short Block Scanning - SBS).

Κατά την αρχική εφαρμογή της παραπάνω διαδικασίας χρησιμοποιήθηκε ο τελεστής μεγίστου ($WS_{OP}\{x_m(i)\} \leftarrow \max |x_m(i)|$), επειδή είναι πιο ευαίσθητος στην ανίχνευση μικρών μεταβολών [161]. Ωστόσο, αυτό αποδείχθηκε περισσότερο μειονέκτημα παρά πλεονέκτημα, καθώς όλες οι μικρές διακυμάνσεις, που αντιστοιχούν είτε σε τυχαίο θόρυβο, είτε σε πολύ ασθενικές συνιστώσες σήματος («θαμμένες» στο θόρυβο), καταχωρούνταν ως σημαντικές δραστηριότητες. Για το λόγο αυτό, αποφασίστηκε η χρήση του τελεστή ενεργού τιμής ($WS_{OP}\{x_m(i)\} \leftarrow rms[x_m(i)]$), καθώς και μεγαλύτερων κατωφλίων εντοπισμού T_{det} . Ο τελεστής *rms* αποδείχθηκε πιο σταθερός, παρουσιάζοντας ανοχή στις τυχαίες διακυμάνσεις του θορύβου. Στον αντίποδα, η αύξηση των κατωφλίων T_{det} δημιούργησε ένα νέο πρόβλημα. Έτσι, προκειμένου να μην ανιχνεύονται τα πλαίσια μικρού ενεργειακού περιεχομένου, υπάρχουν περιπτώσεις όπου η «κατάληξη» των ανιχνευμένων συμβάντων δεν είναι ομαλή,

καθώς τα τελευταία στη σειρά πλαίσια δεν ικανοποιούν τις συνθήκες εντοπισμού. Για τη διευθέτηση αυτού του ζητήματος αποφασίστηκε η χρήση δύο κατωφλίων σύγκρισης, όπου το πρώτο (T_{ON}) χρησιμοποιείται κατά τον εντοπισμό νέων συμβάντων και το δεύτερο ($T_{OFF} < T_{ON}$) για τον «τερματισμό» των συμβάντων που έχουν ήδη «ξεκινήσει»:

$$TDR(x; m; T_{ON}; T_{OFF}) = \begin{cases} c_{op} \{WS_{OP}[x_m(i)]; T_{ON}\}, & \text{if } TDR(x; m-1; T_{ON}; T_{OFF})=0 \\ c_{op} \{WS_{OP}[x_m(i)]; T_{OFF}\}, & \text{if } TDR(x; m-1; T_{ON}; T_{OFF})=1 \end{cases} \quad (4-5)$$

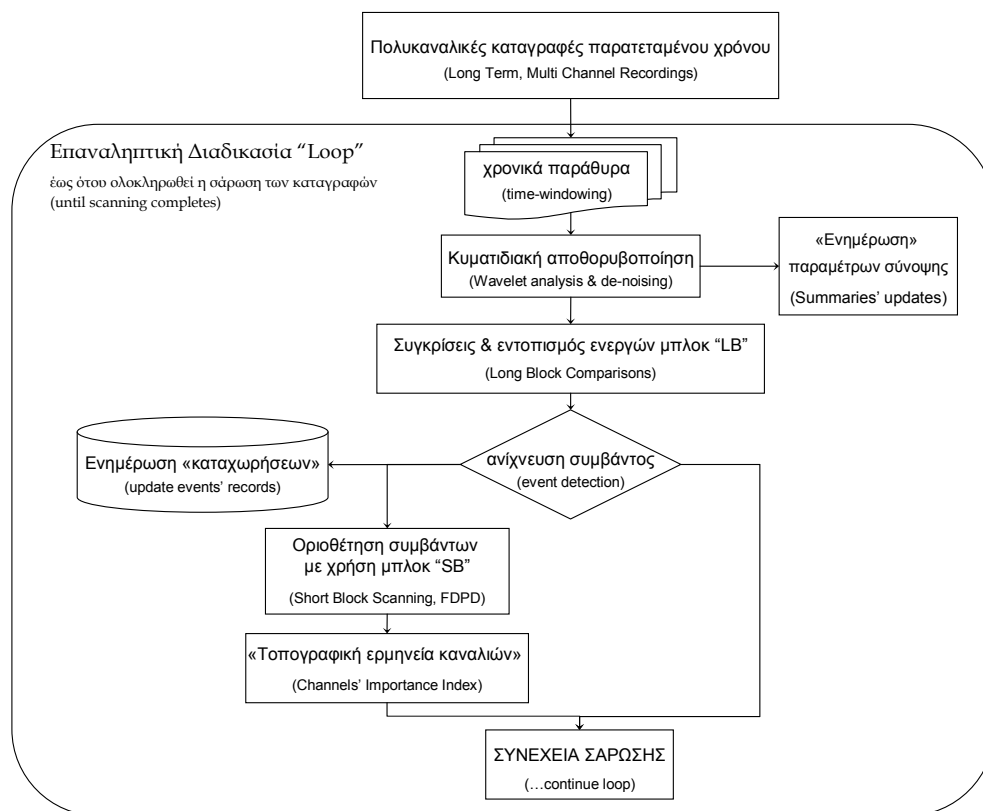
όπου η συνάρτηση $TDR()$ εκφράζει τον κανόνα εντοπισμού (*threshold decision rule*), και ο τελεστής $c_{op}\{\}$ τη λογική πράξη σύγκρισης (*compare operand*). Είναι προφανές ότι η εξ. (4-5) είναι γενική και εξακολουθεί να ισχύει στην περίπτωση χρήσης ενός μόνο κατωφλίου ($T_{ON}=T_{OFF}=T_{det}$), όπως παρουσιάζεται στο σχ. 4-1 [215], [218]. Για την παραπάνω διαδικασία σάρωσης δοκιμάστηκαν διάφορα μήκη παραθύρων, με επικρατέστερα τα $N_W=2.048$ και $N_W=4.096$ (δείγματα). Είναι φανερό, ότι όσο πιο μικρό είναι το μήκος του παραθύρου, τόσο μεγαλύτερη είναι η ακρίβεια (διακριτική ικανότητα) του εντοπισμού, με το μειονέκτημα ότι η διαδικασία σάρωσης διαρκεί περισσότερο (έχει περισσότερες υπολογιστικές απαιτήσεις). Παράλληλα, όσο μεγαλώνει το μήκος του παραθύρου ανάλυσης, η χρήση του τελεστή *rms* καθίσταται προβληματική, καθώς αδυνατεί να εντοπίσει τα κρουστικά φαινόμενα πολύ μικρής διάρκειας. Έτσι, επιλέχθηκε ως βέλτιστη η τιμή $N_W=2.048$ δείγματα [215]-[218].

Η διαδικασία *SBS*, όπως περιγράφηκε ως τώρα, χρησιμοποιήθηκε επιτυχώς για τον εντοπισμό και διαχωρισμό (*segmentation*) των επιμέρους ηχητικών τμημάτων (κατά μήκος των καταγραφών παρατεταμένου χρόνου), και την αξιοποίησή τους στις εργασίες ταξινόμησης και αναγνώρισης προτύπων [215]-[216], [218]. Ωστόσο, υπάρχουν κάποιες αδυναμίες που καθιστούν τη διαδικασία ανάλυσης και διαχείρισης περιεχομένου, προβληματική, όπως είναι ο εντοπισμός ενός υπερβολικά μεγάλου αριθμού συμβάντων (πολλά από τα οποία απέχουν ελάχιστα μεταξύ τους) και το υψηλό υπολογιστικό κόστος. Εκτός από τα χαρακτηριστικά αυτά, που προφανώς δεν συμφωνούν με τις προδιαγραφές των § 4.2 και 4.2.1, ένα άλλο αρνητικό στοιχείο αφορά στην παρουσία θορύβου, που αφενός δυσχεραίνει τη διαδικασία ανίχνευσης, αφετέρου αλλοιώνει τα χαρακτηριστικά των παραμέτρων σύνοψης (καμπύλες «ενέργειας-χρόνου» και «ενέργειας-χρόνου-συχνότητας»). Έτσι, αποφασίστηκε η αξιοποίηση των τεχνικών αποθρομβοποίησης *WDWF* και η ανάπτυξη μιας πιο ευέλικτης μεθόδου, που να μπορεί να εφαρμοστεί αποτελεσματικότερα στη διαδικασία σάρωσης και εντοπισμού των σημαντικών βιοακουστικών συμβάντων [217].

4.2.3 Μέθοδος πολύ-διακριτικής σάρωσης

Η νέα μέθοδος ανάλυσης παρατεταμένου χρόνου σχεδιάστηκε έτσι ώστε να παρουσιάζει μεγαλύτερη προσαρμοστικότητα, τόσο στη διακριτική ικανότητα ανίχνευσης κατά τη σάρωση, όσο και στη ρύθμιση των κατωφλίων εντοπισμού. Επίσης, αποφασίστηκε η ενσωμάτωση της επεξεργασίας αποθρομβοποίησης *WDWF*, για τους λόγους που εξηγήθηκαν στην προηγούμενη παράγραφο. Έτσι, η νέα διαδικασία πολυδιακριτικής σάρωσης (**M**ulti-**R**esolution **S**canning, **MRS**) εφαρμόστηκε σε επίπεδο κυματιδιακών συντελεστών, όπου εκτός από την καταστολή θορύβου, διευθετήθηκαν όλα τα ζητήματα εντοπισμού - διαχωρισμού των σημαντικών βιοακουστικών συμβάντων, και εξαγωγής παραμέτρων σύνοψης-ολοκλήρωσης του συνολικού φαινομένου παρατεταμένου χρόνου (**L**ong-**T**erm **W**avelet-**D**etection, **S**egmentation and **S**ummarization, **LT-WDSS**) [217]. Η διαδικασία **MRS**

συνδυάζει τη χρήση δύο παραθύρων ανάλυσης (σχ. 4-2), όπου το ένα έχει σχετικά μεγάλο μήκος (Long Block -LB) και εφαρμόζεται αρχικά, για την ανίχνευση των συμβάντων σημαντικής ακουστικής δραστηριότητας (long-block significant event detection), ενώ το δεύτερο έχει μικρότερο μήκος (Short Block -SB) και ενεργοποιείται για τον ακριβέστερο καθορισμό των ορίων κάθε δραστηριότητας (short block refining) [217].



Σχήμα 4-2. Συνοπτικό μπλοκ διάγραμμα λειτουργίας της διαδικασίας πολυδιακριτικής σάρωσης (Multi-Resolution Scanning, MRS).

Με βάση το παραπάνω διάγραμμα, η απόφαση εντοπισμού σήματος προκύπτει από τη σύγκριση των συντελεστών κυματιδίων (σύγκριση πολλαπλών φασματικών ζωνών, multi-band comparisons), ενώ οι καμπύλες σύνοψης «ενέργειας-χρόνου» και «ενέργειας-χρόνου-συχνότητα», ανανεώνονται συνεχώς, σε κάθε μπλοκ *LB*. Τα *SB*-μπλοκ αξιοποιούνται με στόχο την εφαρμογή του *FD*-τελεστή ανίχνευσης σιγής (Fractal Dimension Pause Detector - FDPD) και την ακριβέστερη οριοθέτηση των συμβάντων. Από τη σύγκριση των συντελεστών κυματιδίων όλων των «ηχητικών καναλιών», προκύπτει μια τοπογραφική ερμηνεία της σημαντικότητάς τους, που σχετίζεται με την εγγύτητα του αντίστοιχου αισθητήρα λήψης στην εκάστοτε θέση βιοακουστικής διέγερσης*. Μια πρόσθετη καινοτομία της ιεραρχικής σάρωσης *MRS* αφορά στον τρόπο καταχώρησης των ανιχνευμένων συμβάντων. Στην προκειμένη περίπτωση, αποθηκεύονται μόνο τα αποτελέσματα εντοπισμού (αντί των ηχητικών δεδομένων), δηλαδή οι χρονικές «διευθύνσεις» των συμβάντων, μαζί με τα στοιχεία περιγραφής και συνοπτικής παρουσίασης της πειραματικής διαδικασίας. Οι διευθύνσεις καταχωρούνται με τη μορφή «δεσμών» (links) σε αρχεία αναφοράς (indexing), που παραπέμπουν στις αρχικές καταγραφές. Κίνητρο για τη νέα δομή αρχειοθέτησης

* Αν και γενικά είναι δυνατή η χρήση οποιουδήποτε αριθμού καναλιών, στην προκειμένη περίπτωση αξιοποιήθηκαν οι 4 περιφερειακοί αισθητήρες της διάταξης "*A-soundfield 4+3*", για λόγους που αναλύθηκαν στο 2^ο κεφάλαιο.

αποτελέσει η φιλοσοφία του πρωτοκόλλου MPEG-7, σε συνδυασμό με τις αυξημένες αποθηκευτικές δυνατότητες των σύγχρονων μέσων τυχαίας προσπέλασης. Ο νέος τύπος αρχείων (*.mrs) πήρε το όνομά του από τη διαδικασία σάρωσης MRS, καθώς και από το ακρωνύμιο της έκφρασης “*Multi-channel Recordings Source-indexing*”, που αναφέρεται στην αρχειοθέτηση μέσω «δεικτοδότησης». Η χρήση τους παρέχει μεγαλύτερη ευελιξία διαχείρισης των πειραματικών αποτελεσμάτων, ευκολότερη πρόσβαση - πλοήγηση του περιεχομένου και δυνατότητα εφαρμογής πρόσθετων στοιχείων επεξεργασίας και «τροποποιήσεων» [217] (περισσότερες λεπτομέρειες δίνονται στο όγδοο κεφάλαιο). Συνοψίζοντας, η διαδικασία σάρωσης MRS και η αντίστοιχη υλοποίησή της με την κυματιδιακή μέθοδο *LT-WDSS* περιλαμβάνει τα ακόλουθα βήματα:

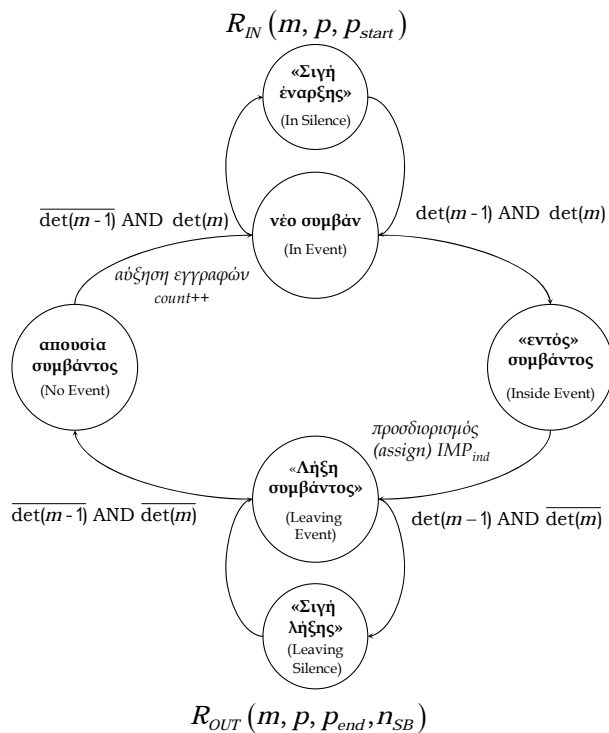
Βήμα 1^ο: Οι καταγραφές «σαρώνονται» με διαδοχικά, μη επικαλυπτόμενα *LB* μπλοκ, στα οποία εφαρμόζεται η κυματιδιακή επεξεργασία αποθορυβοποίησης *WDWF*. Κατόπιν, οι «καθαροί» συντελεστές κυματιδίων εισάγονται στους τελεστές ανίχνευσης σήματος (*det*), ενώ χρησιμοποιούνται και για την εξαγωγή των ενεργειακών παραμέτρων σύννοψης-περίληψης.

Βήμα 2^ο: Τα παράθυρα ανάλυσης *SB* ενεργοποιούνται κάθε φορά που εντοπίζεται ένα νέο συμβάν, με στόχο τον περιορισμό των διαστημάτων σιγής στο πρώτο και τελευταίο παράθυρο *LB*. Έτσι, τα χρονικά όρια των συμβάντων καθορίζονται με μεγαλύτερη ακρίβεια.

Βήμα 3^ο: Οι κυματιδιακές ακολουθίες (εντός ανιχνευμένων συμβάντων) διαμορφώνονται σε ένα πολυδιακριτικό πλέγμα ανάλυσης χρόνου-συχνότητας (multi-resolution, time-frequency bins), βάσει του οποίου πραγματοποιούνται ενεργειακές συγκρίσεις μεταξύ των καναλιών. Έτσι, προκύπτει ο «δείκτης σημαντικότητας καναλιών» (*channel-importance index, IMP_{ind}*), που παρέχει πληροφορίες τοπογραφικής ανάλυσης του παραγόμενου ηχητικού πεδίου [217].

Το «διάγραμμα ροής - εναλλαγής καταστάσεων» (state transition diagram [234]) της ιεραρχικής διαδικασίας σάρωσης MRS, δίνεται στο σχήμα που ακολουθεί [217].

- m : δείκτης *LB*-μπλοκ (current long block)
- p : δείκτης *SB*-μπλοκ (short block indexing)
- R_{IN} : έναρξη συμβάντος (current event start)
- R_{OUT} : λήξη συμβάντος (current event end)
- n_{SB} : αριθμός *SB*-μπλοκ εντός *LB*-μπλοκ (# of short blocks inside long block)
- p_{start} : αριθμός *SB*-μπλοκ σιγής στην αρχή (overhead pause-blocks at starting)
- p_{end} : αριθμός *SB*-μπλοκ σιγής στο τέλος (overhead pause-blocks at ending)
- det : τελεστής εντοπισμού σήματος
- $det(m)$: αποτέλεσμα εντοπισμού “ m μπλοκ”
- AND : Λογικός τελεστής (Boolean AND)
- «...» : Λογικός τελεστής (Boolean NOT)



Σχήμα 4-3. Διάγραμμα «ροής - εναλλαγής καταστάσεων» (state transition diagram) της διαδικασίας MRS.

4.2.3.1 Επεξεργασία αποθορυβοποίησης σε καταγραφές παρατεταμένου χρόνου

Για τις ανάγκες αποθορυβοποίησης των καταγραφών γαστρεντερικής (ακουστικής) δραστηριότητας παρατεταμένου χρόνου, αποφασίστηκε η ενσωμάτωση της μεθόδου *WDWF*. Έχει διαπιστωθεί [179], ότι και οι 4 τεχνικές *WDWF* μπορούν να εφαρμοστούν σε ακολουθίες οσοδήποτε μεγάλου μήκους, χωρίς επιπτώσεις στα αποτελέσματα επεξεργασίας. Βέβαια, αυτό ισχύει για την ανάλυση σύντομου χρόνου (*short-term analysis*), που αφορά ακουστικά συμβάντα με μικρές χρονικές διάρκειες (της τάξης των sec). Στην περίπτωση καταγραφών παρατεταμένης διάρκειας, απαιτείται η χρήση χρονικών παραθύρων επιλογής, όπως περιγράφεται από την εξ. 4-4. Αν συνδυάσουμε τις εξ. 3-48 και 4-4, το πρόβλημα καταστολής θορύβου σε καταγραφές παρατεταμένου χρόνου, ανάγεται στην επεξεργασία όλων των επιλεγμένων παραθύρων $x_m(i)$. Το μόνο ζήτημα που θα μπορούσε να επισημανθεί, σχετίζεται με την ενδεχόμενη δημιουργία ασυνεχειών κατά την έναρξη της επεξεργασίας *WDWF* (κάθε φορά), δηλαδή στον πρώτο συντελεστή, κάθε κυματιδιακής ζώνης (για όλα τα χρονικά παράθυρα). Το πρόβλημα αυτό οφείλεται στην απουσία «μνήμης» κατά την εκκίνηση της διαδικασίας *WDWF*, ενώ συχνά οδηγεί στην εμφάνιση περιοδικών υπολειμμάτων θορύβων στις χρονικές στιγμές έναρξης των παραθύρων [217]. Ωστόσο, το ζήτημα διευθετείται, αν διατηρήσουμε τους τελευταίους (χρονικά) συντελεστές όλων των κυματιδιακών ζωνών (από κάθε παράθυρο επιλογής), ώστε να χρησιμοποιηθούν ως αφητηρία στο αμέσως επόμενο παράθυρο (σχετικά αποτελέσματα δίνονται στο σχ. 4-3).

Αν χρησιμοποιήσουμε το δείκτη " $w=-1$ " για τους συντελεστές αφητηρίας σε κάθε κυματιδιακή ζώνη k και σε κάθε χρονικό παράθυρο m , και εισάγουμε τους ακόλουθους συμβολισμούς:

$$\begin{aligned} P_{x;a(m)}(k, -1) &= P_{x;a(m-1)}(k, W_{xk} - 1) \\ P_{x;a(0)}(k, -1) &= |X_{(1)}(k, 0)|^a \end{aligned} \quad (4-6)$$

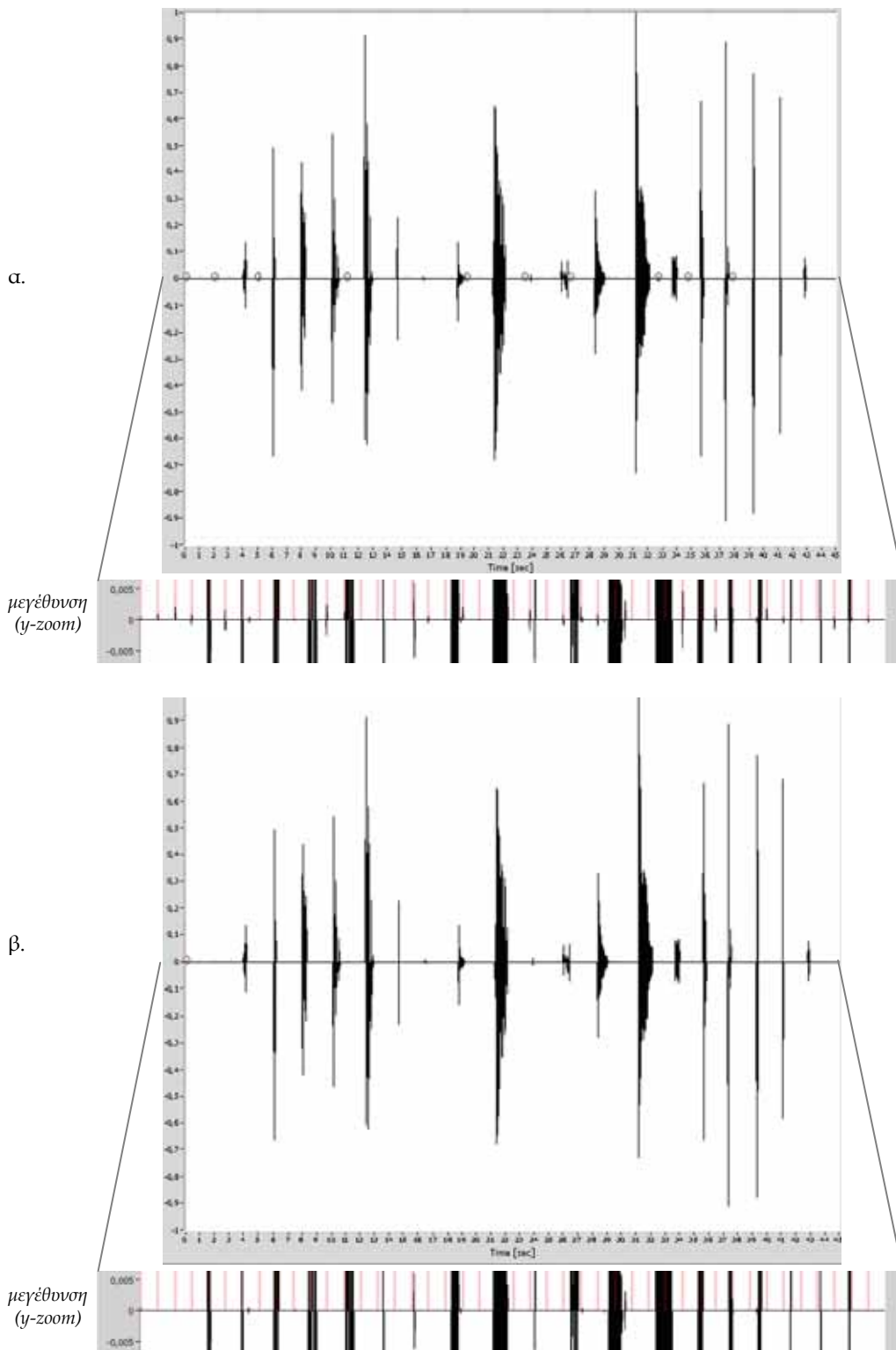
τότε η επεξεργασία *WDWF* με χρήση παραθύρων επιλογής, διαμορφώνεται ως εξής:

$$S_{(m)}^{\sim}(k, w) = \begin{cases} 1 - c \cdot A_{kw} \cdot \left[\frac{P_{nFP;a}(k)}{P_{x;a(m)}(k, w)} \right], \text{ αν } c \cdot A_{kw} \cdot \left[\frac{P_{nFP;a}(k)}{P_{x;a(m)}(k, w)} \right] \leq 1 \\ 0, \text{ διαφορετικά} \end{cases} \quad (4-7)$$

όπου $X_{(m)}(k, w)$ και $S_{(m)}^{\sim}(k, w)$ είναι αντίστοιχα οι ενθόρυβοι και «καθαροί» συντελεστές, ενώ οι εκφράσεις ισχύος $P_{x;a(m)}(k, w)$ μπορούν να αναφέρονται είτε στον τύπο I, είτε στον τύπο II, ανάλογα με τη μέθοδο *WDWF*. Επίσης, η εκτίμηση ισχύος του θορύβου υπολογίζεται μία μόνο φορά, από τους αντίστοιχους συντελεστές $N_{FP}(k, w)$ του διαθέσιμου αποτυπώματος.

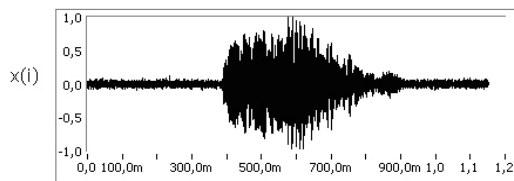
Με την παραπάνω διευθέτηση, η εφαρμογή της μεθόδου *WDWF* σε καταγραφές παρατεταμένου χρόνου διατηρεί όλα τα επιθυμητά στοιχεία (περιορισμένη πολυπλοκότητα, μικρές υπολογιστικές απαιτήσεις, ικανοποιητική καταστολή θορύβου, ελαχιστοποίηση σφαλμάτων και υπολειμμάτων αποθορυβοποίησης) [217]. Επιπρόσθετα, οι τεχνικές *WDWF* έχουν αρκετά πλεονεκτήματα σε σχέση με τις μεθόδους αυτόματης κατωφλίωσης, καθώς: α) το προφίλ του θορύβου *ABN* παραμένει σχεδόν αμετάβλητο κατά μήκος των καταγραφών, ενώ οι όποιες μικρές διακυμάνσεις αντιμετωπίζονται επιτυχώς με τον τύπο *WDWF-II*, β) ο προσδιορισμός του προφίλ θορύβου μπορεί να γίνει με ελεγχόμενο τρόπο («χειρονακτικά» - manually), κι έτσι δεν υπάρχει λόγος πολλαπλών εκτιμήσεων της ίδιας πληροφορίας μέσα από αυτόματες τεχνικές, που συνήθως είναι λιγότερο αποδοτικές γ) πολλές από τις

αυτοματοποιημένες τεχνικές αποθρομβοποίησης λειτουργούν ικανοποιητικά μόνο στην περίπτωση κρουστικών σημάτων, κι έτσι η αξιοποίησή τους σε μη εποικευόμενες διαδικασίες επεξεργασίας είναι ιδιαίτερα επισφαλής, δ) η χρήση μη επικαλυπτόμενων παραθύρων συχνά προκαλεί «απότομες ασυνέχειες», κι έτσι η εφαρμογή των μεθόδων αυτόματης κατωφλίωσης δημιουργεί σημαντικές αλλοιώσεις των σημάτων (σχ. 4-5, 4-6) [179], [217].

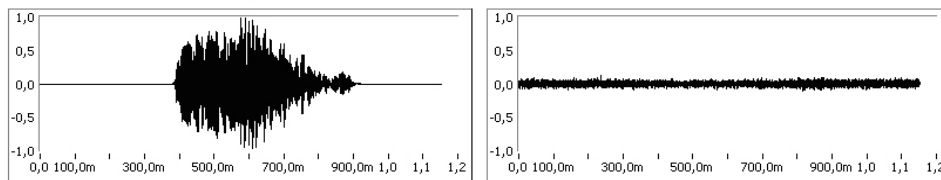


Σχήμα 4-4. Αποτελέσματα της αποθρομβοποίησης WDF με χρήση διαδοχικών παραθύρων επιλογής: α) χωρίς την επαναληπτική διαδικασία (μνήμη) που εισάγει η εξ. 4-6, β) με επιλογή της επαναληπτικής διαδικασίας της εξ. 4-6. Οι ροζ γραμμές δείχνουν τα χρονικά όρια των παραθύρων, ενώ οι κύκλοι τα υπολείμματα θορύβου.

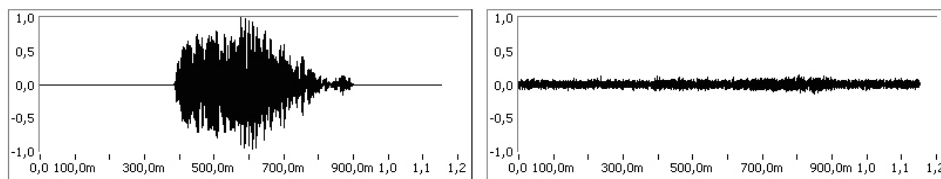
α. Αρχικό σήμα



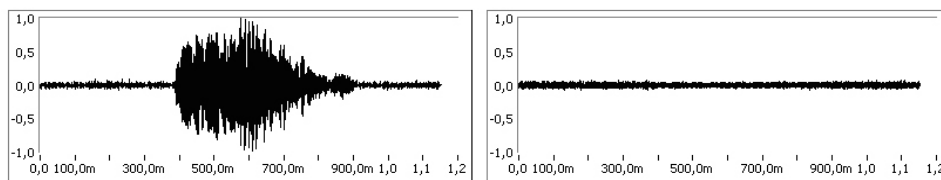
β. WDFW II-6 (DB-6)



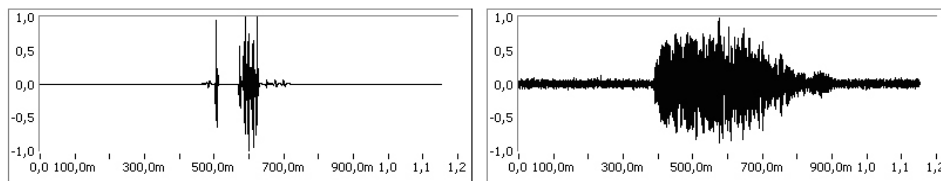
γ. WDFW II-6 (HAAR)



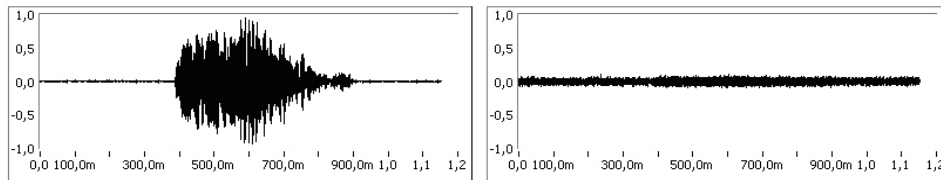
δ. WTST-NST (DB4, Fadj=3)



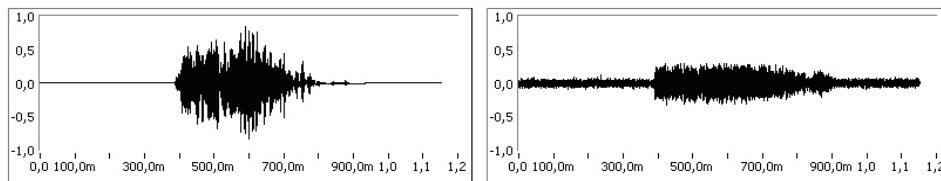
ε. WTST-NST (DB4, Fadj=4)



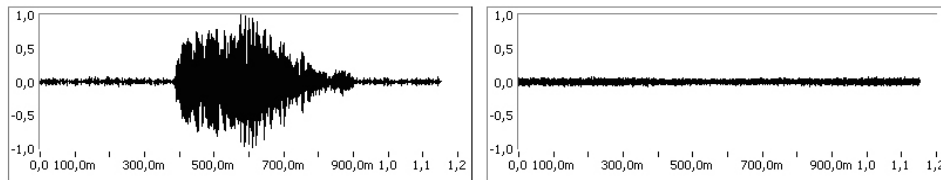
στ. RIGSURE (soft) (DB4)



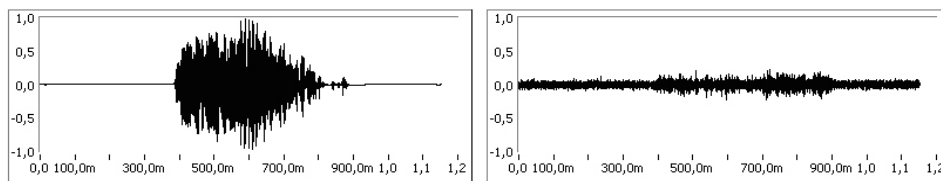
ζ. SQTWOLOG (soft) (DB4)



η. RIGSURE (hard) (DB4)



θ. SQTWOLOG (hard) (DB4)

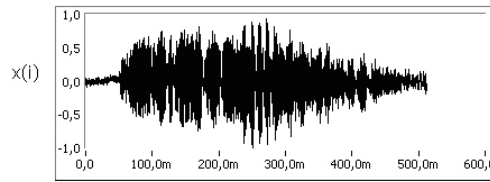


$\tilde{s}(i)$ denoised signal vs time [msec]

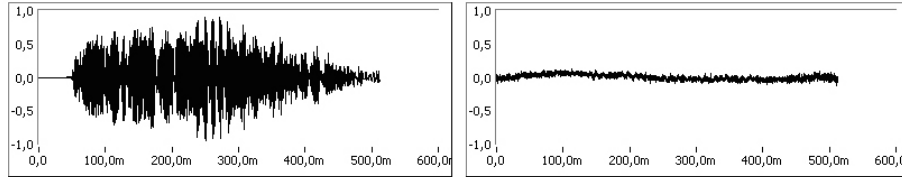
$\tilde{n}(i)$ extracted noise vs time [msec]

Σχήμα 4-5. Σύγκριση των τεχνικών αποθορυβοποίησης WDFW_{II-6} (DB-6, HAAR) με τις WTST-NST ($F_{adj}=3, F_{adj}=4$), RIGSURE και SQTWOLOG, όταν το παράθυρο περιλαμβάνει όλο το σήμα (ανάλυση στο πεδίο χρόνου: στη αριστερή πλευρά απεικονίζεται το αποθορυβοποιημένο σήμα \tilde{s} και στη δεξιά ο εξαγόμενος θόρυβος \tilde{n}).

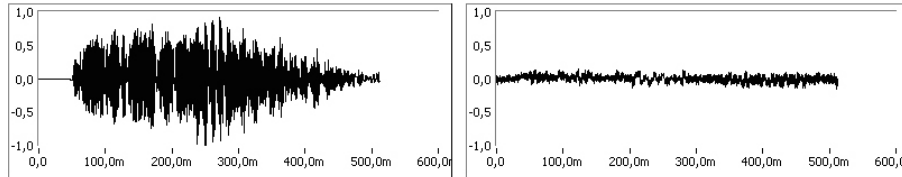
α. Αρχικό σήμα



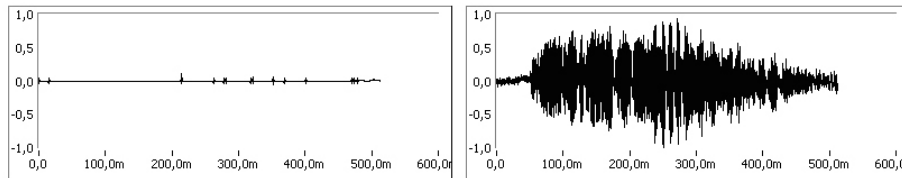
β. WDFW II-6 (DB-6)



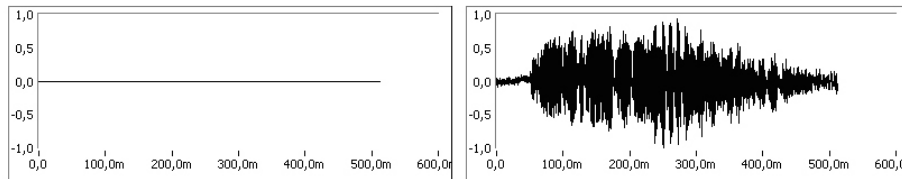
γ. WDFW II-6 (HAAR)



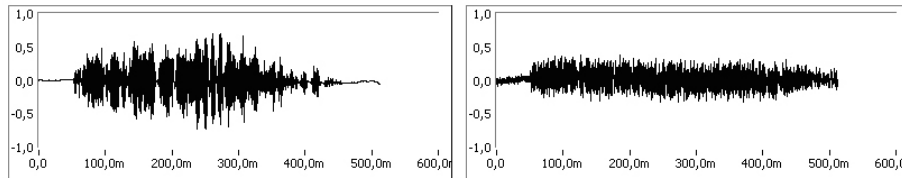
δ. WTST-NST (DB4, Fadj=3)



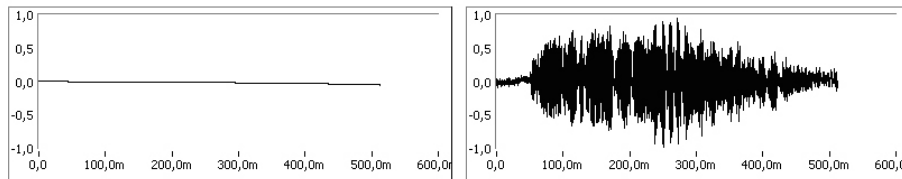
ε. WTST-NST (DB4, Fadj=4)



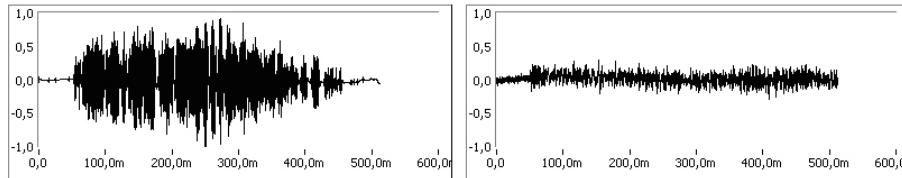
στ. RIGSURE (soft) (DB4)



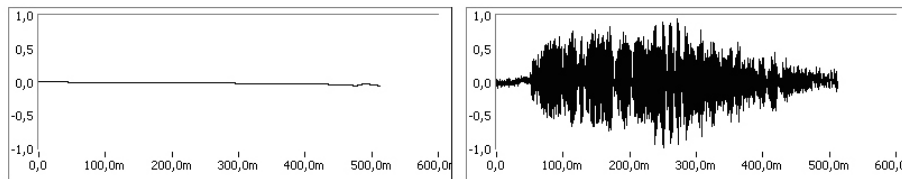
ζ. SQTWOLOG (soft) (DB4)



η. RIGSURE (hard) (DB4)



θ. SQTWOLOG (hard) (DB4)



$\tilde{s}^- (i)$ denoised signal vs time [msec]

$\tilde{n}^- (i)$ extracted noise vs time [msec]

Σχήμα 4-6. Σύγκριση των τεχνικών αποθουροποίησης $WDFW_{II-6}$ (DB-6, HAAR) με τις WTST-NST ($F_{adj}=3, F_{adj}=4$), RIGSURE και SQTWOLOG, όταν το παράθυρο «διακόπτει» το σήμα (ανάλυση στο πεδίο χρόνου: στην αριστερή πλευρά απεικονίζεται το αποθουροποιημένο σήμα s^- και στη δεξιά ο εξαγόμενος θόρυβος n^-).

Για τις ανάγκες αποθορυβοποίησης της διαδικασίας *LT-WDSS*, αποφασίστηκε η χρήση της τεχνικής *WDWF_{II-6}* μαζί με το φίλτρο A_{kw} , καθώς συνδυάζουν την αποτελεσματική απόρριψη των περιόδων σιγής, την ελαχιστοποίηση των υπολειμμάτων θορύβου και το μικρότερο υπολογιστικό κόστος, από όλες τις τεχνικές *WDWF* [179], [217]. Εκτός από την επιλογή του μητρικού κυματιδίου *Daubechies-6 (DB-6)*, έγινε χρήση και του κυματιδίου *Haar*, λόγω των χαρακτηριστικών «απόλυτης συμμετρίας» σχετικά με τα μήκη των κυματιδιακών συντελεστών (το συνολικό τους άθροισμα ισούται με το μήκος της ακολουθίας εισόδου, ενώ το μήκος των συντελεστών σε κάθε κλίμακα είναι ακριβώς το μισό, από το αντίστοιχο της προηγούμενης κλίμακας). Όπως θα δούμε στη συνέχεια, το στοιχείο αυτό έχει ιδιαίτερη σημασία κατά την εφαρμογή των τελεστών ανίχνευσης σιγής *FDPD* στα SB-μπλοκ (σχ.4-3). Αν και το κυματίδιο *DB-6* παρέχει καλύτερα ακουστικά αποτελέσματα από το *Haar* [179], οι διαφορές είναι ελάχιστες, κατά τη σύγκριση με βάση ενεργειακές παραμέτρους και γραφικές απεικονίσεις (σχ. 4-5, 4-6). Έτσι, η χρήση του κυματιδίου *Haar* συγκεντρώνει περισσότερα πλεονεκτήματα κατά την εφαρμογή του στη διαδικασία σάρωσης *LT-WDSS* [217].

Στα σχ. 4-5 και 4-6, παρουσιάζονται δύο «στιγμιότυπα» επεξεργασίας παρατεταμένου χρόνου, με χρήση παραθύρων επιλογής. Στην πρώτη περίπτωση έχει επιλεγεί ολόκληρος ο κύκλος ανόδου-καθόδου του *RS* σήματος, ενώ στη δεύτερη, η εφαρμογή του παραθύρου αποκλείει ένα μικρό τμήμα του σήματος, εμποδίζοντας την ομαλή «ολοκλήρωση» της κυματομορφής (μη ομαλό «σβήσιμο»). Είναι φανερό ότι τα αποτελέσματα καταστολής θορύβου δεν είναι ικανοποιητικά με τις μεθόδους αυτόματης κατωφλίωσης (υπάρχουν έντονα υπολείμματα θορύβου), ενώ παράλληλα προκύπτουν σημαντικές αλλοιώσεις, στην περίπτωση «απότομων διακοπών - ασυνχειών» (σχ. 4-6). Αντίθετα, οι δύο προσεγγίσεις *WDWF (DB-6 και Haar)* δίνουν εξίσου καλά αποτελέσματα και στις δύο περιπτώσεις, ενώ η ικανοποιητική απόρριψη των περιόδων σιγής, διευκολύνει τη διαδικασία εντοπισμού των «σημαντικών βιοακουστικών συμβάντων» [217].

4.2.3.2 Ανίχνευση σημαντικών συμβάντων γαστρεντερικής δραστηριότητας

Σύμφωνα με το μπλοκ διάγραμμα λειτουργίας του σχ. 4-2, η διαδικασία σάρωσης *MRS* περιλαμβάνει κάποια στάδια επεξεργασίας, ανεξάρτητα από τα αποτελέσματα εντοπισμού συμβάντων. Έτσι, όλα τα *LB*-πλαίσια οδηγούνται διαδοχικά στις «ρουτίνες» καταστολής θορύβου, ανανέωσης των ενεργειακών παραμέτρων σύνοψης και «χοντρικής» ανίχνευσης συμβάντων κατά μήκος των *LB*-παραθύρων (*coarse event detection, LB-event detection*). Βέβαια, ο όρος «χοντρικός - συνοπτικός εντοπισμός» (*coarse detection*), που αποδίδεται στη διαδικασία *LB*-ανίχνευσης, είναι κάπως αδόκιμος. Μια πιο σωστή ερμηνεία της παραπάνω έκφρασης είναι η ακόλουθη: «ο χοντρικός εντοπισμός συμβάντων, εντός των πλαισίων *LB*, αποφαινεται για το αν υπάρχει σημαντική δραστηριότητα κατά μήκος του αντίστοιχου μπλοκ, χωρίς ωστόσο να μπορεί να εντοπίσει την ακριβή θέση του συμβάντος που ανιχνεύθηκε» [217].

Το μήκος του *LB*-μπλοκ επιλέχθηκε $N_{LB}=8.192$ δείγματα, που αποτελεί την πλησιέστερη δύναμη του 2 στη διάρκεια 1sec^* , για τη δεδομένη συχνότητα δειγματοληψίας ($f_s=8\text{kHz}$). Η διάρκεια αυτή αποφασίστηκε κατόπιν εμπειρικών παρατηρήσεων [25], [161], [215]-[217], αλλά και με βάση τις συστάσεις της ιατρικής ερευνητικής ομάδας σχετικά με τη διακριτική αναλυτικότητα διαχωρισμού των συμβάντων. Οι τελεστές σύγκρισης για την «απόφαση παρουσίας σήματος» στα *LB*-παραθύρα, βασίστηκαν κατά κύριο λόγο σε αντιληπτικά κριτήρια. Είναι γνωστό [247] ότι, για να γίνει «ικανοποιητικά αντιληπτό» (*noticeable*) ένα

* Όπως και στην *SBS* μέθοδο, η επιλογή «δυνάμεων του 2» για τα μήκη των παραθύρων, αποσκοπεί στην επιτάχυνση των εμπλεκόμενων υπολογισμών και την εξασφάλιση συμμετρίας στα μήκη των κυματιδιακών ακολουθιών.

ακουστικό ερέθισμα, απαιτείται μια διαφορά στάθμης της τάξης των 5dB (σε σχέση με τη «στάθμη φόντου»), ενώ για διαφορές της τάξης των 3dB, το ερέθισμα είναι «οριακά διακριτό» (just perceptible). Αντίστοιχα συμπεράσματα έχουν προκύψει κατά την αξιολόγηση αλγορίθμων καταστολής θορύβου *ABN* [177], [179], [187], όπου διαπιστώθηκε ότι διαφορές στάθμης της τάξης των 6dB, ανάμεσα στα σήματα στενού φασματικού εύρους (narrow-banded signals) και του θορύβου ευρείας ζώνης, αποτελούν ικανοποιητικές συνθήκες για την τέλεια ανακατασκευή του αρχικού -καθαρού- σήματος. Με βάση τα παραπάνω, και επιδιώκοντας την αποφυγή καταχώρησης των σημάτων πολύ χαμηλής στάθμης ως «σημαντικά συμβάντα», αποφασίστηκε η υιοθέτηση της διαφοράς 5dB για τις «πολύ-ζωνικές φασματικές συγκρίσεις» (multi-band comparisons) [217].

Οι προαναφερόμενες «συγκρίσεις εντοπισμού» εφαρμόζονται στους κυματιδιακούς συντελεστές, που προκύπτουν από την ανάλυση *DWT-6* (κατά την επεξεργασία *WDWF_{II-6}*). Από το σχ. 3-15 προκύπτει ότι η συγκεκριμένη κυματιδιακή τοπολογία ορίζει $j=1, \dots, J$ κόμβους φασματικής διαίρεσης. Έτσι, οι φασματικές ζώνες k «κατανέμονται» οκταβικά στην περιοχή $[0, f_s/2]=[0, 4kHz]$, σύμφωνα με τις ακόλουθες σχέσεις [217]:

$j = 1, 2, \dots, J$ κόμβοι κυματιδιακής ανάλυσης (DWT decomposition nodes)

$k = 1, 2, \dots, J + 1$ οκταβικές ζώνες (octave bands)

$$k = \begin{cases} 1, & j = J & \text{συντελεστές κλιμάκωσης (scaling coeff.)} \dots\dots\dots [0 - \frac{f_s}{2^{J+1}}] \\ J - j + 2, & j = J, J - 1, \dots, 1 & \text{συντελεστές κυματιδίων (wavelet coeff.)} \dots\dots\dots [\frac{f_s}{2^{j-1}} - \frac{f_s}{2^j}] \end{cases} \quad (4-8)$$

Οι στάθμες των συνιστωσών σήματος (LS_k) και θορύβου (LN_k) προκύπτουν από τις κυματιδιακές ακολουθίες στην έξοδο του φίλτρου *WDWF_{II-6}*. Από τη σύγκριση της διαφοράς τους με το όριο των 5dB σε κάθε φασματική ζώνη k , προκύπτει η «απόφαση ανίχνευσης σήματος». Δηλαδή, αν σε μια από τις k ζώνες εντοπιστεί σημαντική δραστηριότητα, το αντίστοιχο χρονικό παράθυρο *LB* καταχωρείται ως «ενεργό» [217]. Το μόνο ζήτημα που απομένει να διευθετηθεί για την πλήρη περιγραφή της προσέγγισης «συνοπτικού εντοπισμού» (*MRS-LB*), αφορά στον υπολογισμό των σταθμών σήματος και θορύβου, LS_k και LN_k αντίστοιχα. Δεδομένου ότι ο θόρυβος *ABN* είναι κατά βάση στάσιμη διαδικασία, οι στάθμες LN_k μπορούν να προσδιοριστούν εύκολα με χρήση του τελεστή *rms*:

$$LN_k = 10 \log_{10} \left[\frac{1}{W_{Nk}} \sum_{w=0}^{W_{Nk}-1} N(k, w)^2 + \delta \right], \quad k = 1, \dots, J + 1 \quad (4-9)$$

όπου $N(k, w)$ είναι οι κυματιδιακοί συντελεστές θορύβου, που προκύπτουν είτε από την πράξη αποθορυβοποίησης " $N^-(k, w)$ ", είτε από το διαθέσιμο αποτύπωμα " $N_{FP}(k, w)$ ", W_{Nk} το αντίστοιχο μήκος των κυματιδιακών ακολουθιών σε κάθε ζώνη k , J ο συνολικός αριθμός των κυματιδιακών κόμβων ($J=5$ στην προκειμένη περίπτωση) και δ μια μικρή σταθερά (>0) που εισάγεται για την αποφυγή σφαλμάτων υπολογισμού (*computational overflow*) [217].

Ωστόσο, ο «ορθός» υπολογισμός της ηχητικής στάθμης δεν είναι τόσο απλή υπόθεση στην περίπτωση των γαστρεντερικών βιοακουστικών σημάτων. Αυτό συμβαίνει εξαιτίας των μορφολογικών ιδιαιτεροτήτων των κοιλιακών ήχων, όπου πολλές συνιστώσες παρουσιάζουν έντονο κρουστικό χαρακτήρα με πολύ μικρή διάρκεια. Ως αποτέλεσμα, τα χρήσιμα σήματα δεν παρατηρούνται συνεχώς κατά μήκος των παραθύρων *LB*, αλλά εμφανίζονται σε τυχαίες

χρονικές θέσεις και με τυχαία κυμαινόμενες διάρκειες. Έτσι, ο υπολογισμός της στάθμης τους με απλή εφαρμογή του τελεστή *rms* δεν προσφέρεται, καθώς μπορεί να οδηγήσει σε σφάλματα υπο-εκτίμησης της πραγματικής τιμής, λόγω ενσωμάτωσης των περιόδων σιγής. Αυτός είναι άλλωστε και ο λόγος που κατά την αξιολόγηση των τεχνικών καταστολής θορύβου αξιοποιούνται οι παράμετροι *LSNR* [81], [84]-[85] και *ESNR* [179], [217] αντί των κλασικών σηματοθορυβικών σχέσεων *SNR*. Για τη διευθέτηση παρόμοιων ζητημάτων στα συστήματα ηχομέτρησης (Sound Level Meters –SLM), χρησιμοποιούνται μικρά παράθυρα ολοκλήρωσης κατά τη μέτρηση κρουστικών φαινομένων, όπου έχει τυποποιηθεί η διάρκεια των 35msec (I-ρύθμιση: “I” time-averaging mode- “Impulsive”) [166]. Η επιλογή αυτή έχει γίνει βάσει των αντληπτικών ιδιοτήτων της ανθρώπινης ακοής: “the 35msec time constant was selected as a compromise based on an elementary understanding of the hearing mechanism” [166].

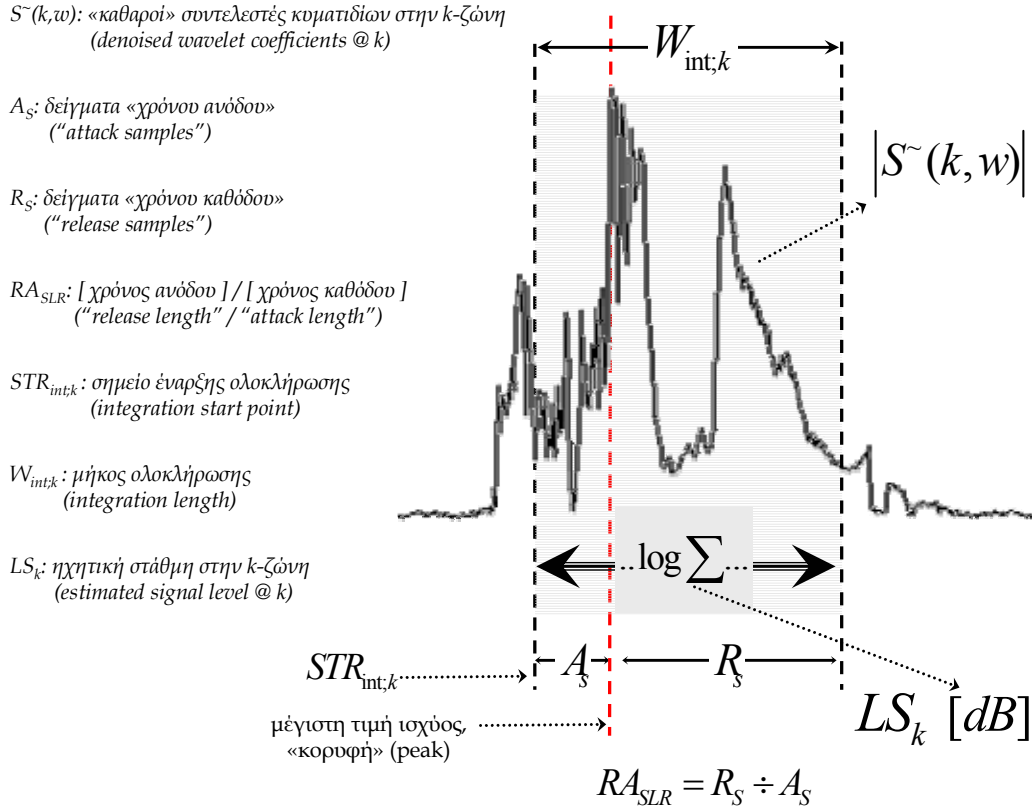
Με βάση τα παραπάνω, αποφασίστηκε η εκτίμηση της ηχητικής στάθμης επιλέγοντας τα διαστήματα ολοκλήρωσης σε θέσεις γύρω από το σημείο «κορυφωσης» (peak-adapted level estimation) [217]. Ως σταθερά ολοκλήρωσης $N_{int;TD}$ (στην περίπτωση απευθείας υπολογισμού της ηχητικής στάθμης στο πεδίο του χρόνου), επιλέχθηκε η δύναμη του 2 που αντιστοιχεί σε ίση ή μεγαλύτερη διάρκεια (*ceiling*), από αυτή των 35msec. Με δεδομένη τη συχνότητα $f_s=8kHz$, η διάρκεια των 35msec αντιστοιχεί σε 280 δείγματα. Η επιλογή της τιμής «οροφής» (*ceiling*) $N_{int;TD}=512$ (2^9), αντί της πλησιέστερης κατώτατης τιμής (*floor*) $N_{int;TD}=256$ (2^8), έγινε για δύο λόγους: α) επειδή δεν έχουν όλοι οι ήχοι κρουστικό χαρακτήρα, και κυρίως β) για να υπάρχει ένας ικανοποιητικός αριθμός συντελεστών ολοκλήρωσης σε όλες τις ζώνες κυματιδιακής ανάλυσης. Το μήκος των 512 δειγμάτων (64msec) διαμορφώνεται κατάλληλα σε όλα τα επίπεδα ανάλυσης k (αναλογικά με τους αντίστοιχους ρυθμούς δειγματοληψίας, εξ. 4-8), κι έτσι προκύπτουν οι σταθερές ολοκλήρωσης $W_{int;k}$ σε κάθε ζώνη [217]:

$$LS_k = 10 \log_{10} \left[\frac{1}{W_{int;k}} \sum_{w=STR_{int;k}}^{STR_{int;k}+W_{int;k}-1} S^-(k, w)^2 + \delta \right], \quad k = 1, \dots, J+1$$

$$W_{int;k} = \frac{N_{int;TD}}{2^j}, \quad j = 1, 2, \dots, J$$

$$STR_{int;k} = \begin{cases} 0, & \left(\arg \max \{ |S^-(k, w)| \} - \left\lfloor \frac{W_{int;k}}{RA_{SLR} + 1} \right\rfloor \right) < 0 \\ [STR_{int;k}], & 0 < \left(\arg \max \{ |S^-(k, w)| \} - \left\lfloor \frac{W_{int;k}}{RA_{SLR} + 1} \right\rfloor \right) \leq W_k - W_{int;k} - 1 \\ W_k - W_{int;k} - 1, & \left(\arg \max \{ |S^-(k, w)| \} - \left\lfloor \frac{W_{int;k}}{RA_{SLR} + 1} \right\rfloor \right) > W_k - W_{int;k} - 1 \end{cases} \quad (4-10)$$

όπου LS_k είναι η εκτιμώμενη στάθμη σήματος (σε κάθε ζώνη), W_k το αντίστοιχο μήκος των κυματιδιακών ακολουθιών, $STR_{int;k}$ είναι το σημείο «έναρξης» της χρονικής ολοκλήρωσης, η παράμετρος RA_{SLR} αποδίδει το λόγο του «χρόνου ανόδου» προς το «χρόνο καθόδου» (Release to Attack Samples' Length Ratio) και χρησιμοποιείται για να καθορίσει την «περιοχή ολοκλήρωσης» γύρω από τη θέση του μεγίστου (στην προκειμένη περίπτωση επιλέχθηκε $RA_{SLR} = 3$), ενώ ο τελεστής “ $\lfloor \cdot \rfloor$ ” δίνει την αμέσως μικρότερη ακέραια τιμή του ορίσματος (*floor operator*). Η διαδικασία υπολογισμού των σταθμών LS_k αποδίδεται γραφικά στο σχ. 4-7.



Σχήμα 4-7. Προοδιορισμός των ηχητικών σταθμών σήματος LS_k σε κάθε ζώνη φασματικής ανάλυσης k .

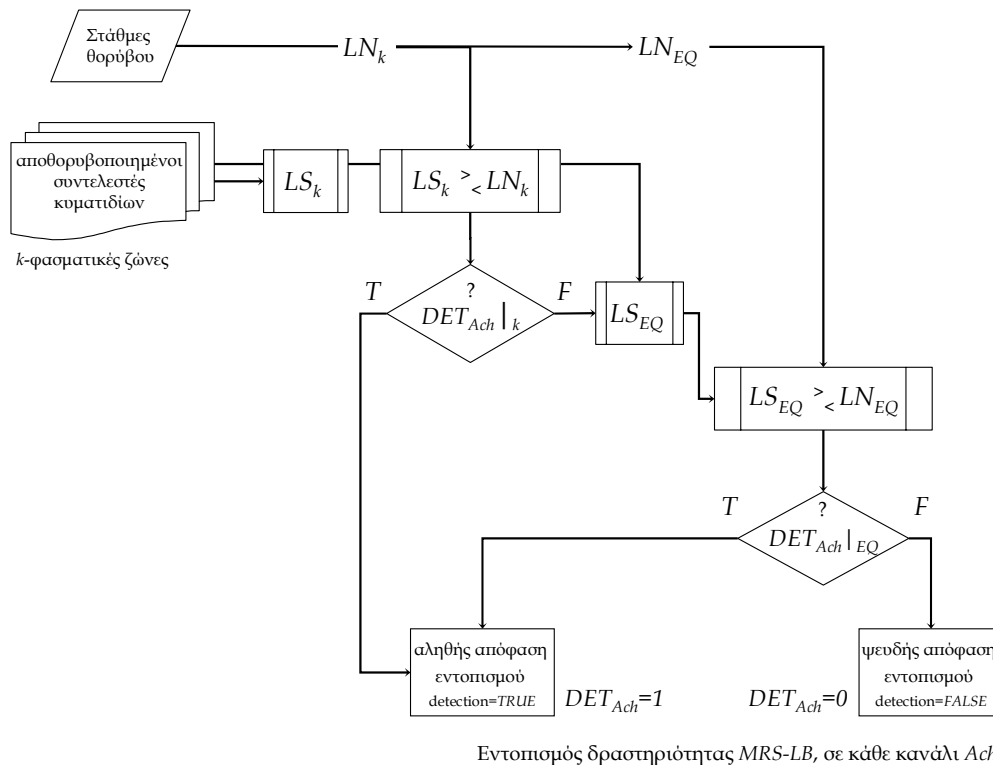
Επιπρόσθετα, για την επίτευξη υψηλής ευαισθησίας ανίχνευσης στην περίπτωση που το σήμα δεν είναι έντονα στενο-ζωνικό, αποφασίστηκε και η ενσωμάτωση της συνολικής ηχητικής στάθμης L_{EQ} (equivalent level). Η τελευταία υπολογίζεται αν θεωρηθεί ότι οι $K=J+1$ φασματικές ζώνες ισοδυναμούν με αντίστοιχο αριθμό ανεξάρτητων -μη συμφασικών (non-coherent)- ηχητικών πηγών [247]:

$$L_{EQ} = 10 \log_{10} \left[\sum_{k=1}^K 10^{\frac{L_k}{10}} \right] \quad (4-11)$$

όπου τα L_{EQ} και L_k μπορούν να αναφέρονται είτε στο σήμα (LS_{EQ} , LS_k) είτε στο θόρυβο (LN_{EQ} , LN_k). Συνδυάζοντας το κριτήριο «σταθμικής διαφοράς 5dB», - σε όλες τις ηχητικές στάθμες (L_{EQ} , L_k) και για όλα τα κανάλια ήχου -"Audio channels" ($Ach = 1, 2, C_n$), - με τον τελεστή «λογικής ένωσης» (Boolean OR), προκύπτει το αποτέλεσμα εντοπισμού $MRS-LB$ [217]:

$$DET_{LB} = OR \left\{ OR \left\{ \begin{array}{l} LS_k \geq LN_k + 5dB, k = 1, \dots, J+1 \\ LS_{EQ} \geq LN_{EQ} + 5dB \end{array} \right\} \right\}_{Ach} \Bigg|_{Ach = 1..C_n} \quad (4-12)$$

όπου C_n είναι ο συνολικός αριθμός των όλων των καναλιών ($C_n=4$ εν προκειμένω). Αυτό σημαίνει ότι, αρκεί ένα από τα κανάλια του τρέχοντος LB -μπλοκ να περιέχει δραστηριότητα ($DET_{Ach}=1$), ώστε η τελική απόφαση παρουσίας σήματος να είναι «αληθής» ($DET_{LB}=1$, $TRUE$). Η διαδικασία ανίχνευσης $MRS-LB$ αποδίδεται γραφικά στο σχ. 4-8 [217].



Σχήμα 4-8. Διαδικασία εντοπισμού δραστηριότητας κατά μήκος των παραθύρων LB ($MRS-LB$), χωριστά για κάθε κανάλι Ach .

4.2.3.3 Οριοθέτηση και αυτόματος διαχωρισμός συμβάντων

Το δεύτερο στάδιο της ιεραρχικής διαδικασίας MRS αφορά στην καλύτερη οριοθέτηση των ανιχνευμένων συμβάντων και την απόρριψη των περιόδων σιγής (silence overheads) στα άκρα των παραθύρων LB . Η διαδικασία αυτή είναι ιδιαίτερα σημαντική, καθώς οι αναλύσεις κατανομής χρόνου-διάρκειας (Time Duration Distribution Analysis -TDDA [25]) είναι πολύ χρήσιμες στη μελέτη της ακουστικής γαστρεντερικής δραστηριότητας [24]-[25], [69], [72], [74]-[75], [86], [88]. Η νέα διαδικασία ανίχνευσης χαρακτηρίζεται από μεγαλύτερη διακριτική ικανότητα, όσον αφορά τις χρονικές θέσεις έναρξης - λήξης των συμβάντων, αλλά και μεγαλύτερη ευαισθησία εντοπισμού των σημάτων χαμηλής ενέργειας. Για το σκοπό αυτό, χρησιμοποιούνται παράθυρα μικρότερης διάρκειας (Short Blocks -SB) εντός των LB -μπλοκ, βάσει των οποίων εκτελείται η νέα διαδικασία ανίχνευσης $MRS-SB$. Η τελευταία ενεργοποιείται κατά τον εντοπισμό «καινούργιων» LB -συμβάντων και εφαρμόζεται μόνο στα LB -μπλοκ έναρξης και λήξης (στην περίπτωση που το αντίστοιχο συμβάν περιλαμβάνει περισσότερα από ένα LB -μπλοκ). Παράλληλα, στη διαδικασία μετέχουν μόνο τα «σημαντικά κανάλια», δηλαδή όσα ικανοποιούν τις συνθήκες « LB -εντοπισμού»: ($DET_{Ach}=1$). Έτσι, η σάρωση $MRS-SB$ εφαρμόζεται σε περιορισμένο αριθμό LB -παραθύρων, με αποτέλεσμα την αύξηση των χαρακτηριστικών ευαισθησίας και αναλυτικότητας της διαδικασίας MRS , χωρίς σημαντική επιβάρυνση του υπολογιστικού της κόστους [217].

Το μήκος των SB -μπλοκ επιλέχθηκε βάσει εμπειρικών παρατηρήσεων, λαμβάνοντας υπόψη και την «ελάχιστη διάρκεια αντίληψης» των 35msec, και ρυθμίστηκε στα $N_{SB}=256$ δείγματα [217]. Η τιμή αυτή αντιστοιχεί σε χρονική διάρκεια 32msec και προσαρμόζεται ανάλογα σε όλες τις ζώνες φασματικής ανάλυσης, υποδιπλασιάζοντας κάθε φορά το

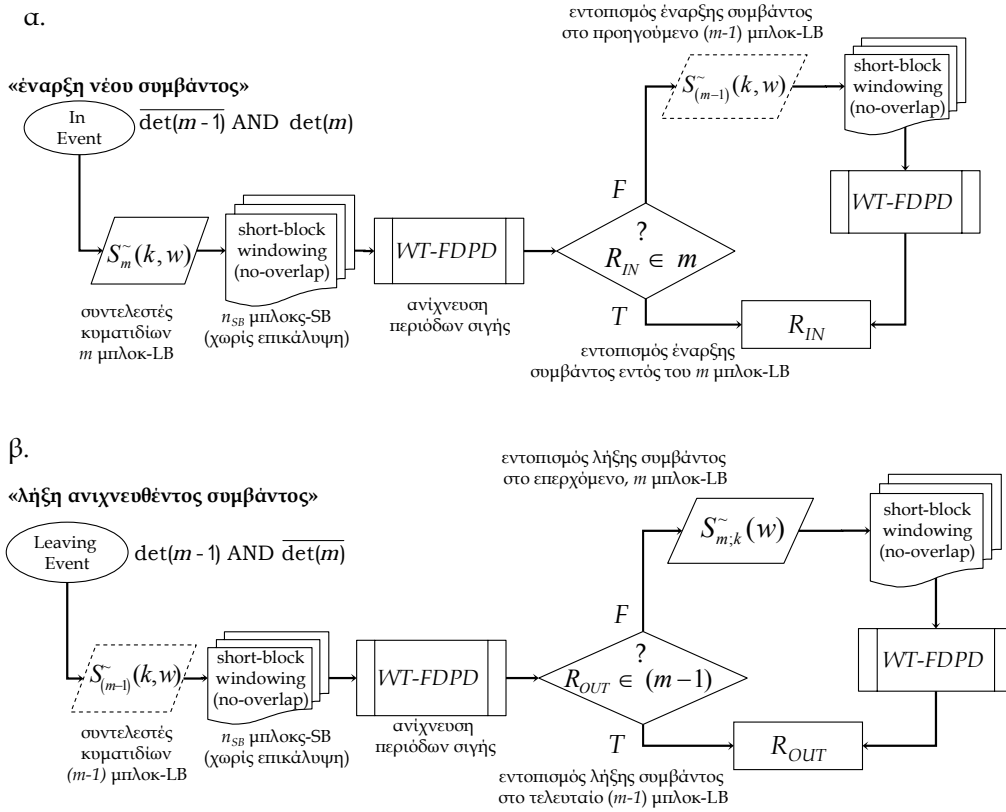
αντίστοιχο μήκος των «κυματιδιακών παραθύρων» (τα μήκη των κυματιδιακών παραθύρων SB διαμορφώνονται στις τιμές 128 ($k=6$), 64, 32, 16, 8 και 8 ($k=1$), δείγματα). Όπως έχει ήδη επισημανθεί, η ανάλυση πολυπλοκότητας κυματομορφών FD και ειδικότερα ο τελεστής ανίχνευσης σιγής FDPD, επιλέχθηκε ανάμεσα σε διάφορες τεχνικές ενεργειακού εντοπισμού, καθώς συγκεντρώνει τα ακόλουθα πλεονεκτήματα: α) είναι πιο ευαίσθητος από τις τεχνικές ενεργειακής σύγκρισης, β) είναι ιδιαίτερα προσαρμοστικός και δεν χρειάζεται την εμπειρική ρύθμιση κατωφλίων, γ) επιδεικνύει εξίσου μεγάλη ακρίβεια - αποτελεσματικότητα όταν τα SB-μπλοκ ανίχνευσης είναι μη επικαλυπτόμενα [217].

Όπως και προκύπτει και από το όνομά του, ο αλγόριθμος FDPD έχει ως στόχο τον εντοπισμό των περιόδων σιγής εντός των «ενεργών» LB-μπλοκ. Αν $s_{p,m}(i)$ είναι το επιλεγμένο τμήμα του «καθαρού» σήματος, που αντιστοιχεί στο p -οστό SB-μπλοκ του m -οστού LB-μπλοκ, τότε οι αντίστοιχες κυματομορφές των περιόδων σιγής θα αντιστοιχούν σε ευθείες γραμμές μηδενικού πλάτους (αφού ο θόρυβος ABN έχει εξαιρεθεί). Δεδομένου ότι η πολυπλοκότητα FD ευθυγράμμων τμημάτων ισούται με τη μονάδα ($FD \text{ of straight lines}=1$), η λογική έκφραση " $FDPD(m,p)>1$ ή $FDPD(p)>1$ ", αποτελεί την ικανή και αναγκαία συνθήκη για την απόφαση παρουσίας σήματος σε κάθε SB-μπλοκ, επιτρέποντας τον εντοπισμό ακόμη και πολύ μικρών διακυμάνσεων [217]. Ο τελεστής FDPD μπορεί να εφαρμοστεί, είτε απευθείας στο πεδίο του χρόνου (TD-FDPD) είτε με τη βοήθεια των συντελεστών κυματιδίων (WT-FDPD). Και στις δύο περιπτώσεις, ορίζονται $n_{SB}=32$ μη επικαλυπτόμενα SB-μπλοκ (τοποθετούνται διαδοχικά εντός των LB-παραθύρων) για καθένα από τα οποία εξετάζεται η συνθήκη " $FDPD(p)>1$ ". Στην περίπτωση του τελεστή WT-FDPD, η ανίχνευση γίνεται ξεχωριστά για καθένα από τις k ζώνες κυματιδιακής - φασματικής ανάλυσης, ενώ τα αντίστοιχα αποτελέσματα εντοπισμού $FDPD_k$ συνδυάζονται με τη συνάρτηση «λογικής σύνθεσης» (Boolean OR) [217]:

$$WT - FDPD(p) = OR\{FDPD_k(p)\}, \quad k = 1, \dots, 6, \quad p = 0, \dots, n_{SB} - 1 \quad (4-13)$$

όπου ο δείκτης p αναφέρεται στα SB-μπλοκ (εντός του αντίστοιχου LB-μπλοκ), ενώ $FDPD_k(p)$ είναι το αποτέλεσμα εντοπισμού στο (p,k) πλαίσιο του σχηματισμένου πλέγματος χρόνου-συχνότητας (32x6 συνολικά πλαίσια: *time-frequency bins*) [217]. Για τις ανάγκες της παρούσης εφαρμογής δοκιμάστηκαν και οι δύο τεχνικές (TD-FDPD, WT-FDPD), οι οποίες αποδείχθηκε ότι παρουσιάζουν παραπλήσια συμπεριφορά, ως προς τα αποτελέσματα ανίχνευσης και τις αντίστοιχες υπολογιστικές απαιτήσεις [217]. Μάλιστα, η μέθοδος WT-FDPD έχει μεγαλύτερη ευαισθησία (ιδιαίτερα για στενοζωνικά σήματα), καθώς εκτελεί τη διαδικασία ανίχνευσης σε πολλές διαφορετικές φασματικές ζώνες. Έτσι, δεδομένης της κυματιδιακής προσέγγισης LT-WDSS, προτιμήθηκε η τεχνική WT-FDPD (αντί της TD-FDPD), επιτρέποντας παράλληλα την αποφυγή του αντίστροφου μετασχηματισμού κυματιδίων [217].

Στην περίπτωση που οι τεχνικές ανίχνευσης FDPD φθάσουν στα όρια του πρώτου (/τελευταίου) LB-μπλοκ χωρίς να εντοπίσουν τα απαραίτητα διαστήματα σιγής έναρξης (/λήξης), η διαδικασία συνεχίζεται στο αμέσως προηγούμενο (/επόμενο) LB-μπλοκ (σχ. 4-9α /σχ. 4-9β). Ιδιαίτερα λειτουργική είναι η εισαγωγή «περιθωρίων σιγής» - «ασφαλείας» (silence overheads) στα όρια αρχής - τέλους των συμβάντων. Στην προκειμένη περίπτωση αποφασίστηκε η ενσωμάτωση δύο «ανεργών» SB-μπλοκ σε κάθε άκρο [217], κυρίως για λόγους καλύτερης γραφικής απεικόνισης και ακουστικής αναπαραγωγής των σημάτων. Η μέθοδος MRS-SB εφαρμόζεται ξεχωριστά σε όλα τα «σημαντικά κανάλια», όπου η μικρότερη χρονική θέση έναρξης και η μεγαλύτερη θέση τερματισμού (μεταξύ των καναλιών), καθορίζουν τα τελικά όρια του συμβάντος [217].



Σχήμα 4-9. Διαδικασία εντοπισμού δραστηριότητας κατά μήκος των παραθύρων SB (MRS-SB): **α)** οριοθέτηση έναρξης στο πρώτο LB-μπλοκ, και **β)** οριοθέτηση λήξης στο τελευταίο LB-μπλοκ.

Αν συμβολίσουμε με r το δείκτη των εντοπισμένων συμβάντων, V το συνολικό αριθμό των LB-μπλοκ (v) που ανήκουν στο r -«χοντρικά εντοπισμένο συμβάν» ($v=r_0, r_1, \dots, r_{V-1}$), p το δείκτη των SB-μπλοκ εντός των LB-παραθύρων ($p=0, 1, \dots, n_{SB}-1$), και p_{start} / p_{end} τα περιθώρια σιγής αρχής-τέλους ($p_{start}=p_{end}=2$), τότε οι θέσεις έναρξης (R_{IN}) /λήξης (R_{OUT}) των συμβάντων, προσδιορίζονται συναρτήσει της αρίθμησης των SB-μπλοκ, με αναφορά το πρώτο (r_0) / τελευταίο (r_{V-1}) LB-μπλοκ, αντίστοιχα [217]:

$$R_{IN} = \begin{cases} \arg \left\{ \sum_{p=0}^{R_{IN} + p_{start} - 1} FDPD(r_0, p) = R_{IN} + p_{start} \right\}, & \text{αν } \sum_{p=0}^{p_{start} - 1} FDPD(r_0, p) = p_{start} \\ \arg \left\{ \sum_{\substack{p=R_{IN} \\ R_{IN} < 0}}^{R_{IN} + p_{start} - 1} FDPD(r_0, p) = p_{start} \right\} & , \text{διαφορετικά} \end{cases} \quad (4-14)$$

$$R_{OUT} = \begin{cases} \arg \left\{ \sum_{p=R_{OUT} - p_{end} + 1}^{n_{SB} - 1} FDPD(r_{V-1}, p) = n_{SB} - (R_{OUT} + p_{end}) \right\}, & \text{αν } \sum_{p=n_{SB} - 1 - p_{end}}^{n_{SB} - 1} FDPD(r_{V-1}, p) = p_{end} \\ \arg \left\{ \sum_{\substack{p=R_{OUT} - p_{end} + 1 \\ R_{out} > n_{SB} - 1}}^{R_{OUT}} FDPD(r_{V-1}, p) = p_{end} \right\} & , \text{διαφορετικά} \end{cases} \quad (4-15)$$

ενώ οι ισχύουν οι ακόλουθες ισοδυναμίες:

$$\begin{aligned} \text{αν } p < 0 &\Leftrightarrow FDDPD(r_0, -|p|) = FDDPD(r_{-1}, n_{SB} - |p|) \\ \text{αν } p > n_{SB} - 1 &\Leftrightarrow FDDPD(r_{V-1}, p) = FDDPD(r_V, p - n_{SB}) \end{aligned} \quad (4-16)$$

όπου “arg” δηλώνει το «όριο», r_{-1} το προηγούμενο από το «πρώτο» LB-μπλοκ (r_0), r_V είναι το επόμενο από το τελευταίο LB-μπλοκ (r_{V-1}), ενώ η παράμετρος $FDDPD(m,p)$ περιγράφει την εφαρμογή του τελεστή ανίχνευσης σιγής $FDDPD$ στο p -οστό SB-μπλοκ του m -οστού LB-μπλοκ ($FDDPD=1 \Leftrightarrow$ “περίοδος σιγής - silent-block”, $FDDPD>1 \Leftrightarrow$ “παρουσία σήματος - signal-block”). Οι αρνητικές τιμές του δείκτη R_{IN} «επισημαίνουν» ότι η έγκυρη θέση έναρξης του συμβάντος βρίσκεται στο προηγούμενο LB-μπλοκ (r_{-1}) και απέχει $|R_{IN}|$ SB-μπλοκ (από την αφετηρία του r_0). Παρόμοια, όταν οι τιμές του δείκτη R_{OUT} είναι μεγαλύτερες από $n_{SB}-1$, αυτό σημαίνει ότι το σημείο τερματισμού του καταχωρημένου συμβάντος βρίσκεται στο επόμενο LB-μπλοκ (r_V), και απέχει $|R_{OUT}|$ LB-μπλοκ (από το σημείο αφετηρίας του r_{V-1}). Σε κάθε περίπτωση, η παράμετρος R_{IN} αντιστοιχεί στο πρώτο SB-μπλοκ του συμβάντος που καταχωρήθηκε, ενώ η R_{OUT} δηλώνει το τελευταίο SB-μπλοκ που ανήκει στο συμβάν [217].

4.2.3.4 Τοπογραφική ερμηνεία συμβάντων

Όπως έχει επισημανθεί, η χρήση πολλαπλών αισθητήρων γίνεται με στόχο την επίτευξη υψηλής ευαισθησίας λήψης, καθώς και την τοπογραφική ανάλυση του παραγόμενου βιοακουστικού πεδίου. Η προσέγγιση αυτή χρησιμοποιήθηκε εκτενώς στις «μανομετρικές» μελέτες [21], [22], [32], [34], αλλά και στις τεχνικές «χωρικής - ακουστικής» ανάλυσης της γαστρεντερικής δραστηριότητας [69], [72], [87]. Και οι δύο ερευνητικές κατευθύνσεις βασίζονται στην αξιοποίηση του πλησιέστερου -στο σημείο διέγερσης- αισθητήρα (Closest Point of Approach - CPA) [87], [158], [248]-[250], ο οποίος προσδιορίζεται βάσει ενεργειακών συγκρίσεων. Ωστόσο, οι τεχνικές ενεργειακού ακουστικού εντοπισμού (Energy Based Localization -EBL) [248]-[250], προϋποθέτουν ότι η ηχητική διέγερση είναι μία (κάθε φορά) και παραμένει ακίνητη (ή μετακινείται με πολύ αργές ταχύτητες). Πρόκειται για το μοντέλο της σημειακής ιστροπικής πηγής, το οποίο αν και έχει χρησιμοποιηθεί σε εφαρμογές ακουστικής χαρτογράφησης της γαστρεντερικής δραστηριότητας [42], [87], πολλοί ερευνητές διατηρούν επιφυλάξεις ως προς την ορθότητα της προσέγγισης [87]. Στο πλαίσιο αυτό, αναζητήθηκε ένα πιο σύνθετο σχήμα περιγραφής για την τοπογραφική ερμηνεία των «βιοακουστικών συμβάντων» και τον χαρακτηρισμό των αντίστοιχων καναλιών [217].

Σύμφωνα με τον Cannon [23], σε πολλές περιπτώσεις οι κοιλιακοί ήχοι παράγονται ως αποτέλεσμα πολλαπλών συσπάσεων που πραγματοποιούνται με μικρές χρονικές αποστάσεις σε διάφορα σημεία του εντέρου. Παράλληλα, συχνά παρατηρείται μετακίνηση (migration) των συστατικών διαταραχών κατά μήκος του γαστρεντερικού σωλήνα, δηλαδή συνεχής μετατόπιση της θέσης ηχητικής διέγερσης* [23], [61]. Άρα, η εύρεση του «CPA-αισθητήρα» δεν είναι πάντα λειτουργική, και δεν καλύπτει πλήρως τις ανάγκες τοπογραφικής ανάλυσης. Έτσι, αποφασίστηκε η ενσωμάτωση του «περιγραφέα σημαντικότητα καναλιών» (channels' importance descriptor - IMP_{ind}), ο οποίος ερμηνεύει την εγγύτητα κάθε καναλιού στο (/στα) σημείο (/σημεία) διέγερσης, επιλέγοντας μία από τις ακόλουθες καταστάσεις: α) **ασήμαντο κανάλι** (*non-significant*) είναι αυτό που δεν ικανοποιεί τις «LB-συνθήκες εντοπισμού», β) **σημαντικό κανάλι** (*significant*) είναι αυτό που ικανοποιεί τις συνθήκες εντοπισμού, τουλάχιστον σε ένα LB-μπλοκ κατά μήκος του συμβάντος, γ) **πρωτεύον κανάλι** (*leading*) είναι αυτό που βρίσκεται πλησιέστερα στο σημείο διέγερσης (CPA), για μία ή περισσότερες

* Σχετικές παρατηρήσεις επιβεβαιώθηκαν πειραματικά με τη βοήθεια φθοροσκοπικών τεχνικών [61].

χρονικές στιγμές κατά μήκος του συμβάντος, και δ) **κυρίαρχο** (*dominant*) είναι το πιο ισχυρό ενεργειακά κανάλι σε σχέση με όλα τα υπόλοιπα, και αποτελεί τη «στατιστικά» εγγύτερη θέση στο σημείο διέγερσης (maximum likelihood CPA) [217].

Από τις προηγούμενες περιγραφές μπορούν να εξαχθούν τα ακόλουθα συμπεράσματα: α) σε κάθε συμβάν είναι δυνατόν περισσότερα από ένα κανάλια να είναι *non-significant*, *significant* ή *leading*, β) ένα μόνο από τα κανάλια χαρακτηρίζεται κάθε φορά ως *dominant*, γ) τα κανάλια που χαρακτηρίζονται ως *leading* ή *dominant* είναι οπωσδήποτε και *significant* (διαφορετικά δεν θα υπήρξε ανίχνευση συμβάντος), δ) εκτός από πολύ σπάνιες περιπτώσεις, το *dominant* κανάλι είναι και *leading*, ε) υπάρχουν περιπτώσεις (κρουστικών συμβάντων, κυρίως) με ένα μόνο *leading* κανάλι (\equiv *dominant*), όπου η αντίστοιχη τοπογραφική ερμηνεία ταυτίζεται με αυτή της προσέγγισης CPA [87], στ) σε κάθε περίπτωση η CPA-προσέγγιση αποτελεί υπό-περίπτωση του προτεινόμενου σχήματος τοπογραφικής περιγραφής [217]. Έτσι, με βάση τα κριτήρια «LB-εντοπισμού», που παρουσιάστηκαν στην § 4.2.3.2 και τους συμβολισμούς των εξ. (4-14) – (4-16), οι αποφάσεις *significant* και *non-significant* για κάθε κανάλι ήχου *Ach* προκύπτουν εύκολα, από τις ακόλουθες σχέσεις [217]:

$$\text{significant}|_{Ach;r} : \sum_{v=r_0}^{r-1} DET_{Ach}|_{LB_v} \geq 1 \quad (4-17)$$

$$\text{non-significant}|_{Ach;r} = \overline{\text{significant}|_{Ach;r}}$$

Σύμφωνα με την ανάλυση που προηγήθηκε, πρωτεύον κανάλι (*leading*) είναι αυτό που για μία ή περισσότερες στιγμές (instances) κατά μήκος του συμβάντος, συγκεντρώνει τη μεγαλύτερη ενέργεια. Για τον ορισμό των «στιγμών» σύγκρισης, αποφασίστηκε η υιοθέτηση ενός πλέγματος χρόνου-συχνότητας (time-frequency bin), που διατηρεί τα χαρακτηριστικά πολύ-διακριτικότητας (multi-resolution) της κυματιδιακής ανάλυσης. Έτσι, αν N_{bin} είναι το «μικρότερο μήκος» των κυματιδιακών ακολουθιών a_j και d_j (εξ. 3-26), τότε, σε κάθε επίπεδο k ορίζεται αντίστοιχος αριθμός $n_{wbin}(k)$ «κυματιδιακών παραθύρων» “*nw*” μήκους N_{bin} , από τις ενεργειακές συγκρίσεις των οποίων προκύπτουν τα πρωτεύοντα (*leading*) κανάλια [217]:

$$n_{wbin}(k) = \left\lfloor \frac{W_{Sk}}{N_{bin}} \right\rfloor = \begin{cases} 1, & k \leq 2 \\ 2^{k-2}, & k > 2 \end{cases}, \quad \text{όπου } N_{bin} = \frac{N_{LB}}{2^J}$$

$$TF_{bin}(nw, k)|_{Ach;v} = \frac{1}{N_{bin}} \sum_{w=(nw-1) \cdot N_{bin}}^{nw \cdot N_{bin}-1} S^-(k, w)^2|_{Ach;v}, \quad k \in [1, J+1], \quad nw \in [1, n_{wbin}(k)] \quad (4-18)$$

$$\text{leading}|_{Ach} : Ach = \underset{\substack{k \in [1, J+1], \\ nw \in [1, n_{wbin}(k)], \\ v \in [r_0, r-1], \\ Ach \in [1, C_n]}}{\arg \max} \left[10 \cdot \log_{10} \left(TF_{bin}(nw, k)|_{Ach;v} + \delta \right) \right]$$

όπου $n_{wbin}(k)$ είναι ο αριθμός των σχηματιζόμενων «χρονικών-πλαισίων» (time-bins) σε κάθε ζώνη k , $W_{Sk} \equiv W_k$ είναι ο αριθμός (μήκος) των συντελεστών κυματιδίων, ενώ η παράμετρος ισχύος $TF_{bin}(nw, k)|_{Ach;v}$ αναπαριστά όλα τα «πλαίσια χρόνου-συχνότητας» (time-frequency bins), που διαμορφώνονται για το *Ach* κανάλι εντός του “*v*” LB-μπλοκ. Στην προκειμένη

περίπτωση, από τις επιλογές των παραθύρων N_{LB} , N_{bin} και το βάθος J του μετασχηματισμού DWT , προκύπτουν $n_{TFbin}=32$ συνολικά πλαίσια χρόνου συχνότητας, εντός κάθε LB -παραθύρου (" $n_{TFbin} = \sum_k [n_{Wbin}(k)]$ "). Η χρήση λογαριθμικών εκφράσεων κατά τη σύγκριση των πλαισίων $TF_{bin}(nw,k)|_{Ach,v}$ (εξ. 4-18γ) γίνεται έτσι, ώστε να μην διαφοροποιούνται τα αποτελέσματα από πολύ μικρές ενεργειακές διακυμάνσεις. Για αντίστοιχους λόγους, προτείνεται ο «ενεργειακός έλεγχος» (gating) όλων των πλαισίων και η εξαίρεσή τους από τη διαδικασία σύγκρισης, εάν έχουν πολύ χαμηλή στάθμη (π.χ. $<1dB$) [217].

Τέλος, ο προσδιορισμός του κυρίαρχου (*dominant*) καναλιού προκύπτει από την «ενεργειακή ολοκλήρωση» όλων των TF_{bin} -πλαisiών κατά μήκος του εκάστοτε συμβάντος, και την αντίστοιχη σύγκριση των ηχητικών καναλιών [217]:

$$\text{dominant}|_{Ach} : Ach = \arg \max_{Ach \in [1, C_n]} \left[10 \cdot \log_{10} \left(\sum_{nw,k,v} \frac{N_{bin} \cdot TF_{bin}(nw,k)|_{Ach,v}}{V \cdot N_{LB}} + \delta \right) \right] \quad (4-19)$$

Αξίζει να σημειωθεί ότι οι συγκρίσεις των εξ. (4-18) και (4-19) πραγματοποιούνται μόνο στα «ενεργά» LB -μπλοκ, σε αντίθεση με την εξ. (4-17), που εφαρμόζεται συνεχώς, παράλληλα με τη διαδικασία LB -εντοπισμού. Ωστόσο, οι στάθμες ισχύος των πλαισίων TF_{bin} (αθροίσματα των εξ. 4-18β και 4-19) υπολογίζονται σε κάθε παράθυρο LB , καθώς χρησιμοποιούνται κατά την εξαγωγή των ενεργειακών παραμέτρων σύνοψης, που εξετάζονται στη συνέχεια [217].

4.2.3.5 Εξαγωγή ενεργειακών παραμέτρων συνοπτικής απεικόνισης

Η πιο απλή προσέγγιση ανάλυσης καταγραφών παρατεταμένου χρόνου είναι η εξαγωγή παραμέτρων σύνοψης, όπως είναι η περιβάλλουσα ισχύος, που απεικονίζει τις ενεργειακές μεταβολές συναρτήσει του χρόνου. Τέτοιου είδους τεχνικές έχουν εφαρμοστεί για την παρακολούθηση της γαστρεντερικής δραστηριότητας στις μανομετρικές μεθόδους και γενικότερα στις καταγραφές ενδοαυλικής πίεσης [21]-[22], [32], [34], [43]-[44], καθώς και στις αντίστοιχες προσεγγίσεις ακουστικής ανάλυσης [25], [68]-[69], [72]-[73], [161]-[162]. Στην προκειμένη περίπτωση αξιοποιήθηκαν οι σταθμικές μεταβολές της περιβάλλουσας ισχύος ($P_{av} [dB]$), σε συνδυασμό με τις οπτικογραφικές εικόνες που «σχηματίζουν» οι στάθμες των κυματιδιακών πλαισίων TF_{bin} (Wavelet Spectrographic Envelope, $WSE [dB]$). Έτσι, με βάση το θεώρημα του Parseval, σχετικά με τη διατήρηση της ενέργειας κατά το μετασχηματισμό κυματιδίων (εξ. 3-26), ο υπολογισμός της στάθμης $P_{av}|_{Ach,m}$ σε κάθε LB -παράθυρο m και για όλα τα κανάλια Ach , υπολογίζεται από την ακόλουθη σχέση [217]:

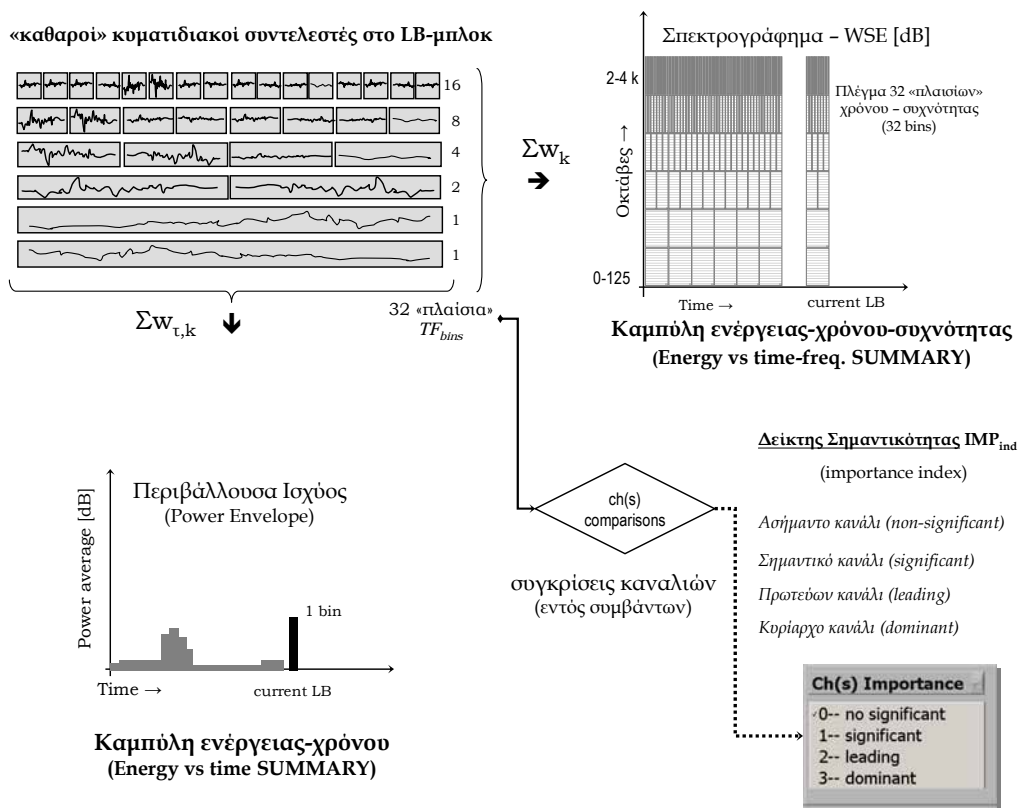
$$P_{av}|_{Ach,m} = 10 \cdot \log \left\{ \left[\frac{1}{N_{LB}} \sum_{k=1}^{J+1} \sum_{w=0}^{W_{sk}-1} S_m(k,w)|_{Ach,m} \right]^2 + \delta \right\} \quad (4-20)$$

Η εξαγωγή οπτικογραφικών εικόνων πραγματοποιείται συνήθως με εφαρμογή του μετασχηματισμού $STFT$, ή οποιασδήποτε άλλης μεθόδου συνδυασμένης ανάλυσης χρόνου - συχνότητας ($JTFA$), συμπεριλαμβανομένων και των κυματιδιακών τεχνικών. Όσον αφορά τις τελευταίες, χαρακτηριστική είναι η χρήση του «κυματιδιακού φάσματος» GWS (Global Wavelet Spectrum) [251], που προκύπτει από την ολοκλήρωση των συντελεστών σε κάθε ζώνη κυματιδιακής ανάλυσης. Έτσι, αν J είναι το βάθος του μετασχηματισμού DWT , προκύπτουν $J+1$ φασματικές - ενεργειακές «εκτιμήσεις» (frequency bins), για κάθε χρονικό παράθυρο m . Αντί αυτού, αποφασίστηκε η εφαρμογή της ενεργειακής ολοκλήρωσης κατά μήκος των n_{TFbin} ($=32$) πλαισίων- TF_{bin} , τα οποία αξιοποιούνται ούτως ή άλλως για την

τοπογραφική ερμηνεία των καναλιών. Με τον τρόπο αυτό εξασφαλίζεται μεγαλύτερη αναλυτικότητα (περισσότερα πλαίσια - “frequency bins”) σε σχέση με την προσέγγιση GWS, ενώ παράλληλα υιοθετείται το πολυδιακριτικό (και οκταβικό) πλέγμα της κομματιδιακής ανάλυσης. Έτσι, οι σπεκτρογραφικές εικόνες WSE ανανεώνονται -για κάθε m , Ach -, με βάση την ακόλουθη σχέση [217]:

$$WSE(nw, k) \Big|_{Ach, m} = 10 \cdot \log \left(TF_{bin}(nw, k) \Big|_{Ach, m} + \delta \right), \quad k \in [1, J + 1], \quad nw \in [1, n_{wbin}(k)] \quad (4-21)$$

Στο σχήμα που ακολουθεί αποδίδεται σχηματικά η διαδικασία ανανέωσης των παραμέτρων σύνοψης, καθώς και η ενεργειακή σύγκριση των καναλιών με στόχο την τοπογραφική ερμηνεία των συμβάντων [217].

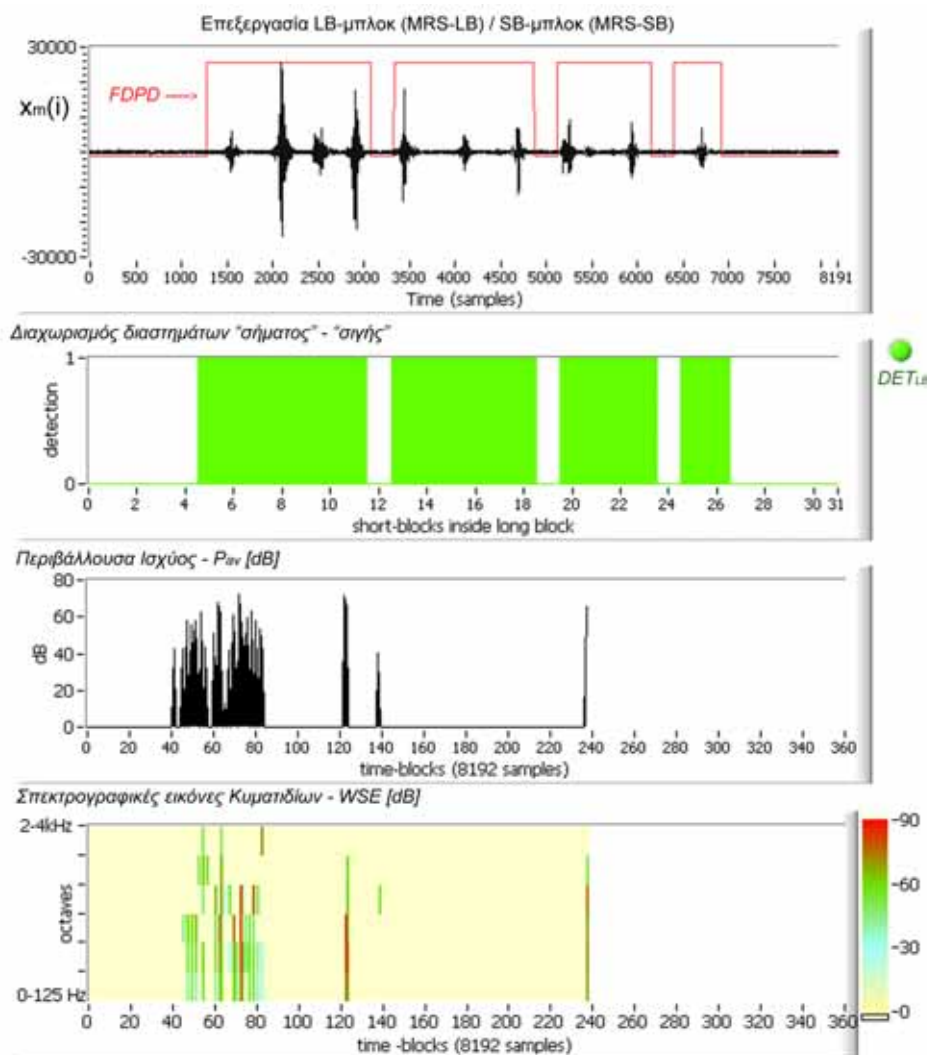


Σχήμα 4-10. Ανανέωση των παραμέτρων σύνοψης (ενέργειας-χρόνου και ενέργειας χρόνου-συχνότητας) και χαρακτηρισμός της σημαντικότητας των καναλιών (IMP_{ind}), κατά την εφαρμογή της μεθόδου LT-WDSS.

4.2.4 Πειραματικά αποτελέσματα αξιολόγησης της μεθόδου LT-WDSS

Για τις ανάγκες σχεδιασμού, «ρύθμισης» (calibration) και τελικής διαμόρφωσης της μεθόδου LT-WDSS, αλλά και όλων των εμπλεκόμενων τεχνικών (MRS-LB, TD-FDPD, WT-FDPD, IMP_{ind}), αξιοποιήθηκαν οι διαθέσιμες πειραματικές καταγραφές παρατεταμένου χρόνου, σε συνδυασμό με πρόσθετα σήματα δοκιμής. Έτσι, κατέστη εφικτή η αξιολόγηση της διαδικασίας πολύ-διακριτικής σάρωσης MRS, βάσει ποιοτικών και ποσοτικών κριτηρίων. Στο σχ. 4-11, παρατίθεται ένα χαρακτηριστικό στιγμιότυπο της διαδικασίας MRS, όπου παρουσιάζονται τα αποτελέσματα LB-εντοπισμού, τα διαστήματα παύσης (SB-εντοπισμός, αποτελέσματα FDPD), και οι γραφικές παραστάσεις των ενεργειακών παραμέτρων σύνοψης P_{av} και WSE. Για την παραστατικότερη παρακολούθηση της λειτουργίας MRS, ο αναγνώστης

παραπέμπεται στα αρχεία κινούμενης εικόνας (βίντεο) *video1.wmv** και *video2.wmv†*, που δημιουργήθηκαν από την αποτύπωση (screen capture) των «περιβαλλόντων λογισμικού», τα οποία αναπτόχθηκαν για τις αντίστοιχες ανάγκες επεξεργασίας. Στο πρώτο παρουσιάζονται μόνο τα αποτελέσματα LB-εντοπισμού, ενώ στο δεύτερο έχουν επιπλέον συμπεριληφθεί οι διαδικασίες «ανίχνευσης σιγής» και ανανέωσης των «γραφημάτων σύνοψης» [217]. Τα συγκεκριμένα παραδείγματα έχουν προκύψει από την εφαρμογή της μεθόδου *LT-WDSS* σε «συνθετικές» καταγραφές (σήματα δοκιμής “*s_{DAS}*”, «μολυσμένα» με GWN θόρυβο).



Σχήμα 4-11. Χαρακτηριστικό στιγμιότυπο (LB-μπλοκ) της διαδικασίας MRS, όπου παρουσιάζονται τα αποτελέσματα LB/SB-εντοπισμού και τα γραφήματα των παραμέτρων ενεργειακής σύνοψης (P_{av} , WSE).

Παράλληλα, δοκιμάστηκαν πρόσθετες τεχνικές αξιολόγησης της μεθόδου *LT-WDSS* και των σχετικών αποτελεσμάτων *LB*- και *SB*-εντοπισμού. Δεδομένου ότι η σάρωση *MRS-LB* βασίζεται κατά κύριο λόγο σε αντιληπτικά κριτήρια («κανόνας διαφοράς 5dB», εξ. 4-12), δεν ενδεικνύεται η αξιοποίηση ποσοτικών δεικτών μέτρησης της αποδοτικότητάς της. Για την αντιμετώπιση αυτού του ζητήματος, εφαρμόστηκε η ακόλουθη πειραματική διαδικασία:

α) Κοιλιακοί ήχοι *s_{DAS}* επιλέχθηκαν από τη βάση δεδομένων που περιγράφηκε στο 3ο κεφάλαιο και χρησιμοποιήθηκαν για τη δημιουργία «συνθετικών καταγραφών δοκιμής».

* http://avlab.ee.auth.gr/LT_WDSS/video-1.wmv

† http://avlab.ee.auth.gr/LT_WDSS/video-2.wmv

β) Διάφορα δείγματα s_{DAS} τοποθετήθηκαν σε τυχαίες θέσεις και διαμορφώθηκαν σε διαφορετικά πλάτη εντός των «συνθετικών καταγραφών», των οποίων η διάρκεια ρυθμίστηκε στα 5 λεπτά.

γ) Οι καταγραφές «μολύνθηκαν» με διαφορετικές επαναλήψεις θορύβου GWN (μηδενικής μέσης τιμής και μοναδιαίας απόκλισης).

ε) Οι ενθόρυβες καταγραφές εξετάστηκαν («ακουστικά» και γραφικά) από μια ομάδα 20 ατόμων (κυρίως ειδικευόμενο προσωπικό και φοιτητές του τμήματος Ιατρικής), τα οποία κλήθηκαν να εντοπίσουν τις θέσεις και τον αριθμό των «σημαντικών» βιοακουστικών συμβάντων γαστρεντερικής δραστηριότητας (αφού πρωτίστως ενημερώθηκαν για την έννοια του όρου «σημαντικό συμβάν», όπως περιγράφηκε στο τρέχων κεφάλαιο).

στ) Από τις συγκρίσεις των παραπάνω αποτελεσμάτων με τα αντίστοιχα της αυτόματης διαδικασίας $MRS-LB$, προέκυψαν τα αποτελέσματα αξιολόγησης [217].

Ειδικότερα, για την έκφραση των αποτελεσμάτων αξιολόγησης με ποσοτικούς δείκτες, επιλέχθηκαν οι ακόλουθοι παράμετροι:

$$DP = \frac{1}{Q} \sum_{q=0}^{Q-1} DP_q, \quad \text{όπου} \quad DP_q = \frac{N_{q;MN} - |N_{q;MN} - N_{q;AU}|}{N_{q;MN}} \cdot 100 \quad (4-22)$$

$$LP = \frac{1}{Q} \sum_{q=0}^{Q-1} LP_q, \quad \text{όπου} \quad LP_q = \frac{N_{q;MN} - \sum_{te=0}^{N_{q;MN}-1} \left[\frac{\min(TL_{MN;q}(te) - [TL_{AU;q}])}{\min(TL_{MN;q}(te) - [TL_{AU;q}]) + 1} \right]}{N_{q;MN}} \cdot 100 \quad (4-23)$$

όπου DP_q είναι το ποσοστό (%) επιτυχούς εντοπισμού (*Detection Performance*) του αριθμού των συμβάντων για κάθε πειραματική καταγραφή q , Q είναι ο συνολικός αριθμός των πειραματικών δοκιμασιών ($Q=20$ εν προκειμένω), DP είναι ο μέσος όρος των αποτελεσμάτων DP_q , $N_{q;AU} / N_{q;MN}$ είναι ο αριθμός των συμβάντων που εντοπίζονται αυτόματα / εμπειρικά σε κάθε πείραμα q , $TL_{MN;q}(te)$ είναι η χρονική θέση εμπειρικού εντοπισμού (εκφρασμένη με τον αύξοντα αριθμό LB -μπλοκ) όλων των συμβάντων δοκιμής (test events $-te$) σε κάθε πείραμα q , και $[TL_{AU;q}]$ είναι η ακολουθία με τις LB -θέσεις των συμβάντων που εντοπίζονται αυτόματα. Έτσι, η μεταβλητή LP_q εκφράζει το ποσοστό των σωστά εντοπισμένων συμβάντων (*Localization Performance*) σε κάθε πείραμα q , ενώ η LP αποτελεί το μέσο όρο όλων των LP_q , δηλαδή αποτελεί το στατιστικά μέσο ποσοστό επιτυχίας εντοπισμού [217].

Κατά τη διαδικασία αυτή, τα εμπειρικά αποτελέσματα εντοπισμού δίνονται αρχικά με τη μορφή πληροφοριών χρονισμού (time-formats), με τον περιορισμό ότι δύο διαδοχικά συμβάντα θα πρέπει να απέχουν μεταξύ τους τουλάχιστον 1sec (≈ 1 LB -μπλοκ). Οι χρονικές θέσεις εντοπισμού μετατρέπονταν στη συνέχεια στην αντίστοιχη LB -αρίθμηση, ώστε να χρησιμοποιηθούν στην εξ. (4-23). Ο ορισμός της παραμέτρου LP_q (εξ. 4-23) έγινε με το εξής σκεπτικό: εάν υπάρχει μία αυτόματα προσδιορισμένη LB -θέση (εντός της ακολουθίας $[TL_{AU;q}]$), που να συμπίπτει με την αντίστοιχη θέση εμπειρικού εντοπισμού $TL_{MN;q}(te)$, τότε ο τελεστής ελάχιστης τιμής “min” προσδίδει στο όρισμα του te -αθροίσματος την τιμή 0. Διαφορετικά, ο τελεστής «οροφής» (ceiling) προσδίδει την τιμή 1. Έτσι, το άθροισμα των αποτελεσμάτων (0 ή 1) κατά μήκος όλων των συμβάντων δοκιμής te , δίνει το συνολικό

αριθμό των συμβάντων για τα οποία σημειώθηκε «αποτυχία εντοπισμού», ενώ η μεταβλητή LP_q εκφράζει το «συμπληρωματικό» ποσοστό (%) «επιτυχούς εντοπισμού» [217].

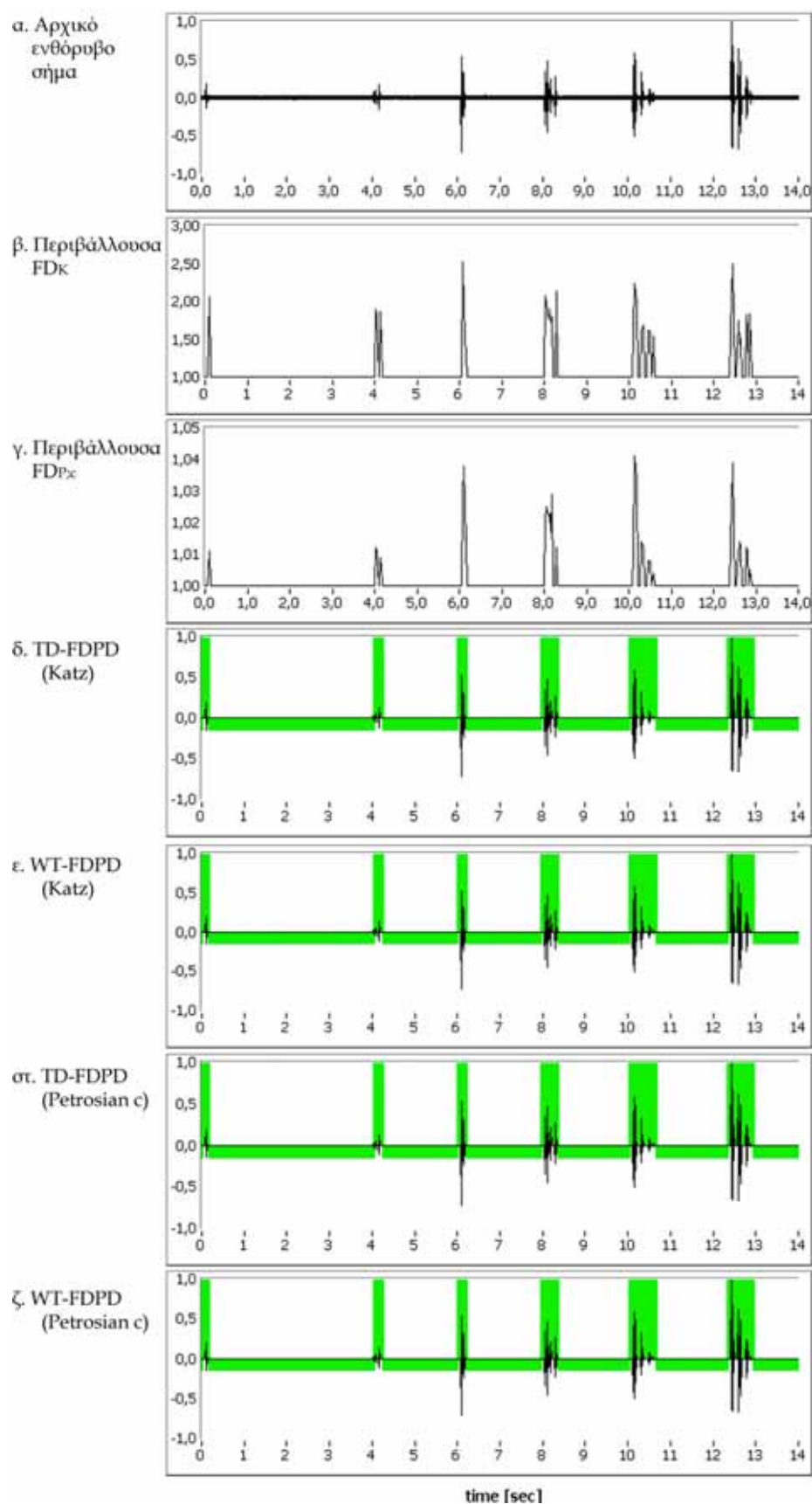
Από τα αποτελέσματα της παραπάνω πειραματικής διαδικασίας προέκυψε ένα μέσο ποσοστό επιτυχούς εντοπισμού του αριθμού των συμβάντων $DP=98,1\%$, όπου η τυπική απόκλιση της ακολουθίας DP_q βρέθηκε ίση με $\sigma_{DP_q}=2,0\%$. Αντίστοιχα, το μέσο ποσοστό επιτυχούς εντοπισμού της θέσης των συμβάντων διαμορφώθηκε σε $LP=93,3\%$, ενώ η αντίστοιχη τυπική απόκλιση της ακολουθίας LP_q υπολογίστηκε σε $\sigma_{LP_q}=3,1\%$. Αξίζει να σημειωθεί ότι τα ποσοστά απόδοσης LP γίνονται ακόμη καλύτερα, αν ληφθούν υπόψη τα αποτελέσματα του SB -εντοπισμού. Έτσι, αν η διαφορά ανάμεσα στις εμπειρικές θέσεις εντοπισμού και τις αντίστοιχες «αυτοματοποιημένες» είναι ίση με 1 παράθυρο- LB (αποτέλεσμα που καταχωρείται ως αποτυχία στην προκειμένη περίπτωση), η εφαρμογή της «οριοθέτησης» $MRS-SB$ επιτρέπει τον επιτυχή προσδιορισμό της θέσης και αυτών των συμβάντων, με αποτέλεσμα τη βελτίωση της συνολικής αποδοτικότητας της μεθόδου [217].

Όσον αφορά την απόδοση της μεθόδου $FDPD$, αυτή σχεδιάστηκε έτσι ώστε να παρέχει μεγαλύτερη ευαισθησία και ακρίβεια εντοπισμού, στις περιοχές γύρω από τα «συνοπτικά», LB -εντοπισμένα, συμβάντα όπου εφαρμόζεται. Στα σχήματα 4-12 και 4-13 παρουσιάζονται τα αποτελέσματα εφαρμογής των τελεστών FD και των μεθόδων $FDPD$, σε κρουστικούς (IB) και παραμένοντες (RS) κοιλιακούς ήχους, αντίστοιχα. Στα διαγράμματα αυτά παρατίθενται και συγκρίσεις ανάμεσα στους FD -τύπους του Katz και του Petrosian (type c). Παρατηρούμε ότι κατά τον υπολογισμό του ενεργειακού φακέλου των σημάτων, η μέθοδος FD_K (σχ. 4-12β, σχ. 4-13β) παρέχει καλύτερα αποτελέσματα από την $FD_{p,c}$ (σχ. 4-12γ, σχ. 4-13γ). Ωστόσο, κατά τη χρήση των δύο τύπων στους αλγορίθμους «ανίχνευσης σιγής» $FDPD$ (σχ. 4-12 ε-στ, σχ. 4-13 ε-στ), προκύπτουν εξίσου ικανοποιητικά αποτελέσματα. Επίσης, παρατηρούμε ότι οι δύο προσεγγίσεις $TD-FDPD$ και $WT-FDPD$ παρουσιάζουν παραπλήσια συμπεριφορά, με τη δεύτερη να είναι ελάχιστα πιο ευαίσθητη από την πρώτη [217].

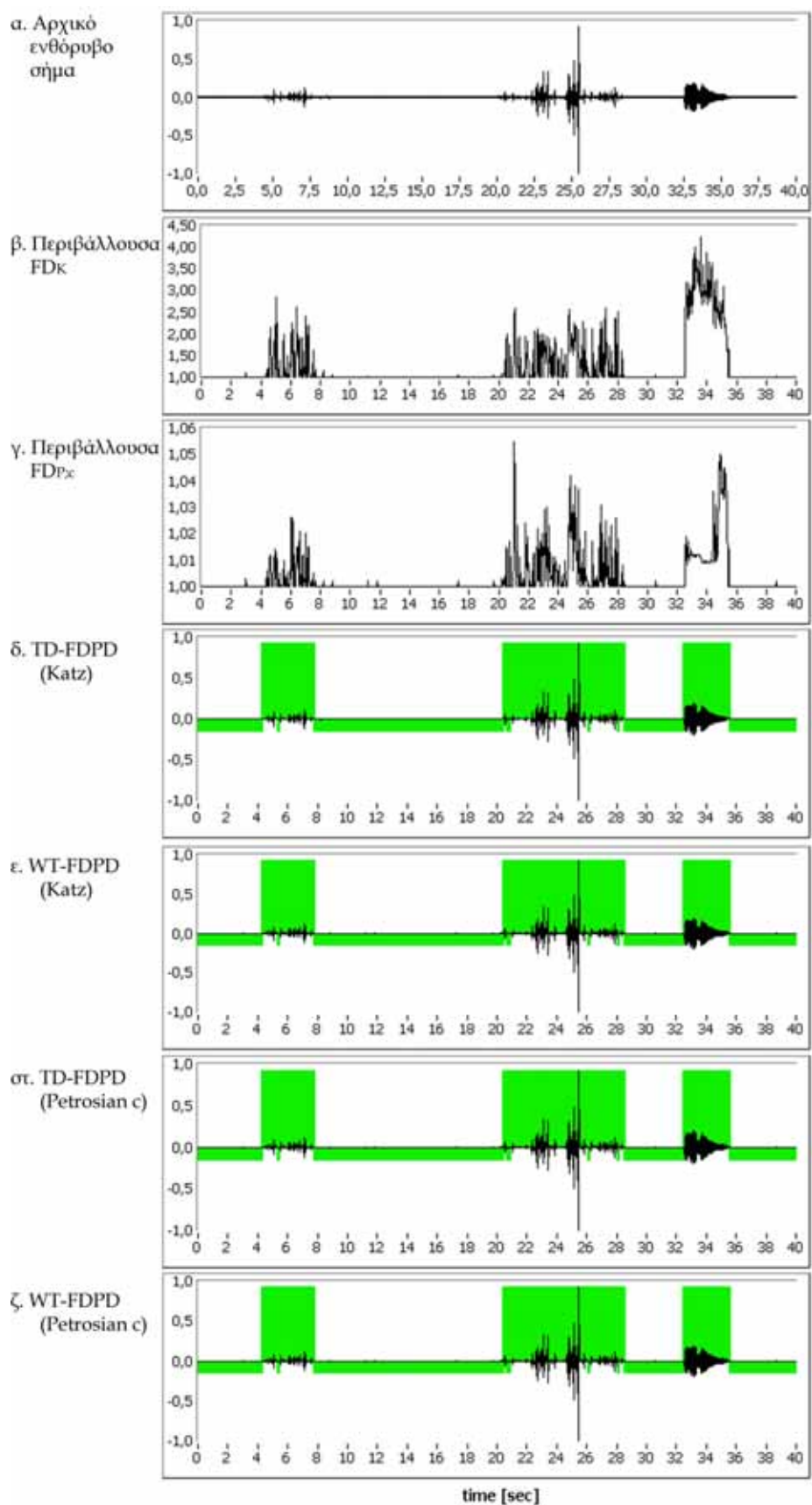
Στο σχ. 4-14, απεικονίζονται τα αποτελέσματα εντοπισμού $MRS-SB$, κατά την εφαρμογή την εφαρμογή των τεχνικών $TD-FDPD$ και $WT-FDPD$ σε σήματα διαφορετικής στάθμης και διαφορετικών μορφολογικών χαρακτηριστικών. Έτσι, χρησιμοποιούνται αντιπροσωπευτικά πρότυπα κοιλιακών ήχων, όπως απομονωμένοι κρουστικοί παλμοί (Solitary CLicks -SCL), επαναλαμβανόμενοι κρουστικοί παλμοί (Repeated CLicks -RCL), «τυχαίες» ακολουθίες RS και IB τμημάτων (Sequences of Irregularly Concatenated Segments -SICS), παραμένοντες -τριζόντες ήχοι (CRepitating SWeeps -CRSW) και παραμένοντες -τονικοί (ή συρίτονες) ήχοι (WhiStling SWeeps -WSSW) [61], [217]-[218] (περισσότερες λεπτομέρειες για τα παραπάνω πρότυπα κοιλιακών ήχων δίνονται στο επόμενο κεφάλαιο). Παράλληλα, τα συγκεκριμένα παραδείγματα έχουν κλιμακούμενες στάθμες ηχητικής εγγραφής, όπως αυτές εκφράζονται από την κανονικοποιημένη παράμετρο ακουστότητας (*Loudness*), για τη δεδομένη κβάντιση (16bit) [217]:

$$Loudness = 20 \cdot \log_{10} \left(\max(|s(i)|) \right) - 20 \cdot \log_{10} \left(2^{16-1} \right) \quad [\text{dB}] \quad (4-24)$$

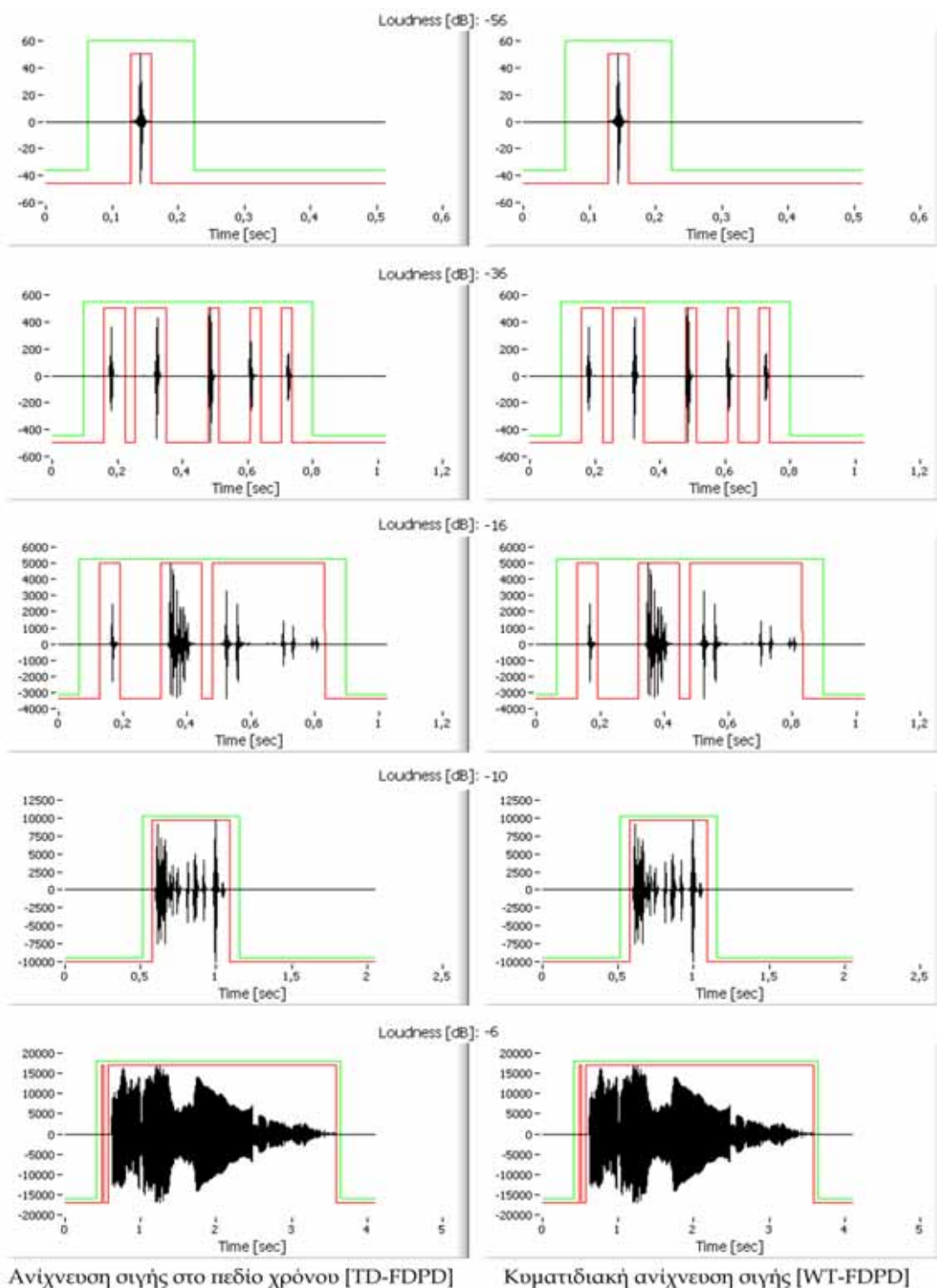
Στο σχ. 4-14 παρουσιάζονται τα αποτελέσματα ανίχνευσης $FDPD$ (κόκκινη σήμανση), καθώς και η τελική οριοθέτηση των συμβάντων με την ενσωμάτωση των περιθωρίων σιγής $p_{start} - p_{end}$ (πράσινη σήμανση). Είναι προφανές ότι, και οι δύο τεχνικές ($TD-FDPD$ και $WT-FDPD$) παρέχουν εξίσου ικανοποιητικά αποτελέσματα ακόμη και σε ιδιαίτερα χαμηλές στάθμες εγγραφής (της τάξης των -44dB), όπου τα σήματα έχουν πολύ μικρό πλάτος (<100) και είναι οριακά ακουστά [217].



Σχήμα 4-12. Αποτελέσματα των τελικών FD και των αντίστοιχων μεθόδων ανίχνευσης σιγής FDDP (TD-FDDP και WT-FDDP) κατά την εφαρμογή τους σε κρουστικούς (IB) κοιλιακούς ήχους (τα «πράσινα-χρωματισμένα πλαίσια σήμανσης» διαχωρίζουν τις «περιόδους σήματος» από τα αντίστοιχα «διαστήματα σιγής»).



Σχήμα 4-13. Αποτελέσματα των τελεστών FD και των αντίστοιχων μεθόδων ανίχνευσης σιγής FDPD (TD-FDPD και WT-FDPD) κατά την εφαρμογή τους σε παραμένοντες (RS) κοιλιακούς ήχους (τα «πράσινα-χρωματισμένα πλαίσια σήμανσης» διαχωρίζουν τις «περιόδους σήματος» από τα αντίστοιχα «διαστήματα σιγής»).



Σχήμα 4-14. Αποτελέσματα ανίχνευσης TD-FDPD/WT-FDPD (αριστερά / δεξιά), σε διαφορετικά πρότυπα κοιλιακών ήχων και για κλιμακούμενες στάθμες (Loudness) ηχητικής εγγραφής («κόκκινη σήμανση»: αρχικά FDPD- όρια εντοπισμού, «πράσινη σήμανση»: τελική οριοθέτηση με ενσωμάτωση των περιθωρίων σιγής).

4.3 Επισκόπηση 4ου κεφαλαίου

Το τέταρτο κεφάλαιο πραγματεύεται τα ζητήματα επεξεργασίας και διαχείρισης των βιοακουστικών καταγραφών παρατεταμένου χρόνου. Στο πλαίσιο αυτό εντάσσονται οι διαδικασίες σάρωσης και ανίχνευσης των συμβάντων σημαντικής δραστηριότητας, η τοπογραφική τους ερμηνεία μέσω του «χαρακτηρισμού» των αντίστοιχων καναλιών και η

εξαγωγή των ενεργειακών παραμέτρων σύνοψης. Παράλληλα, συμπεριλαμβάνονται οι αλγόριθμοι καταστολής θορύβου με χρήση χρονικών παραθύρων, καθώς και οι τεχνικές εντοπισμού των περιόδων παύσης, *FDPD*. Η προτεινόμενη μέθοδος *LT-WDSS*, η οποία περιλαμβάνει όλες τις ρουτίνες επεξεργασίας που εμπλέκονται στη διαδικασία πολυδιακριτικής σάρωσης *MRS*, υλοποιείται εξολοκλήρου σε κυματιδιακό επίπεδο, παρέχοντας νέες δυνατότητες στη μελέτη της γαστρεντερικής κινητικότητας μέσω των μεθόδων γαστρεντερικής φωνογραφίας (*Gastro-Intestinal Phonography -GIP*) [217].

4.3.1 Αποτίμηση προτεινόμενων παρεμβάσεων

Για την αντιμετώπιση των παραπάνω ζητημάτων, αναπτύχθηκαν δύο διαφορετικές στρατηγικές επεξεργασίας: η μέθοδος σάρωσης με χρήση παραθύρων μικρής διάρκειας (*SBS*) και η διαδικασία πολυδιακριτικής σάρωσης (*MRS*). Αν και επιλέχθηκε η προσέγγιση *MRS*, ωστόσο και η πιο κλασσική τεχνική *SBS* υπήρξε αρκετά χρήσιμη, και όπως θα δούμε στο επόμενο κεφάλαιο, αξιοποιήθηκε ως τεχνική προ-επεξεργασίας, κατά την εκπαίδευση ενός «έμπειρου» συστήματος αναγνώρισης προτύπων [215]-[216], [218]. Όσον αφορά την τελικά επιλεγόμενη διαδικασία πολυδιακριτικής σάρωσης *MRS*, αυτή είναι ιδιαίτερα λειτουργική και αποτελεσματική, ενώ συγκεντρώνει αρκετά καινοτόμα στοιχεία. Για την εφαρμογή της αναπτύχθηκε η μέθοδος *LT-WDSS*, η οποία ικανοποιεί πλήρως τις προδιαγραφές που τέθηκαν στις §4.2 και §4.2.1, διευθετώντας όλες τις δυσκολίες της προγενέστερης μεθόδου *SBS*. Έτσι, η χρήση παραθύρων μεγάλης σχετικά διάρκειας (*LB*-μπλοκ) στο μεγαλύτερο τμήμα των καταγραφών επιταχύνει τους υπολογισμούς, ενώ η αξιοποίηση μικρότερων πλαισίων (*SB*-μπλοκ) σε «επιλεκτικές» θέσεις, επιτρέπει την ακριβή οριοθέτηση των συμβάντων. Ως αποτέλεσμα, περιορίζονται οι περιττές διασπάσεις πολύ κοντινών συμβάντων (<1sec), οι οποίες, σύμφωνα με εκτιμήσεις της «ιατρικής» ερευνητική ομάδας, προκαλούν σύγχυση και οδηγούν σε λανθασμένα συμπεράσματα.

4.3.2 Στοιχεία πρωτοτυπίας και ερευνητική συνεισφορά

Το πρώτο και σημαντικότερο στοιχείο πρωτοτυπίας του κεφαλαίου αφορά στη φιλοσοφία της διαδικασίας πολυδιακριτικής σάρωσης *MRS* και τα ιδιαίτερα στοιχεία της μεθόδου *LT-WDSS*. Ειδικότερα, αντιμετωπίζεται επιτυχώς το θέμα της αποθρομβοποίησης καταγραφών παρατεταμένου χρόνου, με τη βοήθεια ορθογωνικών (μη επικαλυπτόμενων) παραθύρων επιλογής (*windowed-WDWF*). Στο πλαίσιο αυτό, επιλέχθηκε και διαμορφώθηκε κατάλληλα η μέθοδος *WDWF_{II-6}*, επιβαρύνοντας κατά το ελάχιστο δυνατό το υπολογιστικό κόστος. Έτσι, η όλη επεξεργασία πραγματοποιείται σε επίπεδο κυματιδιακών συντελεστών (χωρίς την ανάγκη αντίστροφου μετασχηματισμού *IDWT*), απ' όπου προκύπτουν οι «καθαρές από θόρυβο» ενεργειακές παράμετροι σύνοψης και τα αντίστοιχα αποτελέσματα εντοπισμού σήματος. Το νέο σχήμα εντοπισμού διπλής σάρωσης (*dual scanning*), συνδυάζει την ενεργειακή σύγκριση των κυματιδιακών συντελεστών βάσει αντιληπτικών κριτηρίων, με την αξιοποίηση πιο ευαίσθητων μεθόδων οριοθέτησης των ανιχνευμένων συμβάντων, όπως είναι οι τελεστές *Fractal Dimension*. Και στην περίπτωση αυτή, υιοθετήθηκαν οι κυματιδιακές τεχνικές ανίχνευσης σιγής (*WT-FDPD*), ενώ επιλέχθηκε ο προτεινόμενος από τον Petrosian τύπος- ϵ (*FD_{Pc}*), καθώς παρέχει πολύ καλά αποτελέσματα εντοπισμού, με τις μικρότερες υπολογιστικές απαιτήσεις. Εκτός των παραπάνω, μια πρόσθετη καινοτομία αφορά στην οντολογία τοπογραφικής ερμηνείας των συμβάντων, βάσει της παραμέτρου *IMP_{ind}*. Η νέα πρόταση χωρικής ανάλυσης, που στην ουσία εμπεριέχει την προσέγγιση *CPA* που έχει χρησιμοποιηθεί σε σχετικές εφαρμογές στο παρελθόν [69], [72], [87], παρέχει πρόσθετες

πληροφορίες, σχετικά με τη θέση (ή τις θέσεις) και την «τοπογραφική εικόνα» της ηχητικής διέγερσης. Η εφαρμογή της αναμένεται με ιδιαίτερο ενδιαφέρον, τόσο στις ιατρικές ερευνητικές προσεγγίσεις γαστρεντερικής φωνογραφίας, όσο και σε παραπλήσιες μεθόδους βιοακουστικής μελέτης.

Για την καλύτερη κατανόηση της ερευνητικής συνεισφοράς του κεφαλαίου, αρκεί να επισημανθεί ότι τα «ανιχνευμένα σημαντικά συμβάντα» αποτελούν τη βάση για την εφαρμογή όλων των τεχνικών επεξεργασίας που περιγράφονται στη συνέχεια, όπως είναι οι μέθοδοι ανάλυσης προτύπων, οι τεχνικές ακουστικής χαρτογράφησης και οι δομές αρχειοθέτησης και διαχείρισης περιεχομένου. Έτσι, η διαδικασία ανίχνευσης *MRS* και η αντίστοιχη κυματιδιακή μέθοδος *LT-WDSS* αποτελούν ίσως τα πιο κύρια κομβικά σημεία (milestones) της διατριβής, διαμορφώνοντας σε μεγάλο βαθμό το προτεινόμενο πρωτόκολλο ιατρικής μελέτης των κοιλιακών ακουστικών φαινομένων. Ακόμη, ενδιαφέρον παρουσιάζει η αξιοποίησή τους σε παραπλήσιες εφαρμογές βιοακουστικής παρακολούθησης και σε ηχητικές καταγραφές παρατεταμένου χρόνου, γενικότερα. Στο πλαίσιο αυτό, ήδη έχουν ξεκινήσει κάποιες πρώτες πειραματικές προσεγγίσεις για την ανάλυση του ακουστικού θορύβου και τη μελέτη της επίδρασής του στον άνθρωπο [94]-[95], [226].

4.3.3 Προτεινόμενες μελλοντικές εξελίξεις

Οι προτεινόμενες μελλοντικές εξελίξεις αφορούν κυρίως στην αξιοποίηση των μεθόδων σε παραπλήσια επιστημονικά πεδία, όπως είναι η επεξεργασία σημάτων ομιλίας - φωνής, η αυτοματοποίηση διαδικασιών ακουστικών μετρήσεων, η αποτελεσματικότερη διαχείριση μουσικών ηχογραφήσεων, η ανάλυση βιοακουστικών και γενικότερα ψυχο-φυσιολογικών καταγραφών. Μια διαφορετική πρόταση αφορά στην εφαρμογή των αναπτυχθέντων αλγορίθμων σε τεχνικές αυτόματου εντοπισμού κατωφλίων αποθορυβοποίησης για την ανάπτυξη πιο προσαρμοστικών μεθόδων καταστολής θορύβου, με βάση τις απαιτήσεις που παρουσιάστηκαν στο προηγούμενο κεφάλαιο. Σε κάθε περίπτωση, η χρήση των προτεινόμενων μεθόδων στην πράξη, θα αναδείξει τα τυχόν μειονεκτήματά τους (που τουλάχιστον προς το παρόν δεν είναι ορατά) και τις αντίστοιχες ανάγκες εξέλιξης, τόσο ως προς τα αποτελέσματα εντοπισμού σήματος, όσο και ως προς τα συναφή εργαλεία πλοήγησης - διαχείρισης περιεχομένου. Εξάλλου, οι μέθοδοι ανίχνευσης σήματος αποτελούν κοινό παρονομαστή σε πολλά επιστημονικά πεδία, ενώ οι αντίστοιχες τεχνικές επεξεργασίας που έχουν χρησιμοποιηθεί για το σκοπό αυτό, είναι ανεξάντλητες.

ΚΕΦΑΛΑΙΟ 5^ο

5. Τεχνικές αναγνώρισης προτύπων βιοακουστικών φαινομένων

Στο κεφάλαιο αυτό αναλύονται τα ζητήματα ανάλυσης προτύπων των γαστρεντερικών ηχητικών σημάτων. Αρχικά παρατίθεται η βασική θεωρία που διέπει τα συστήματα αναγνώρισης προτύπων, όπως είναι οι τεχνικές στατιστικού διαχωρισμού, οι προσεγγίσεις συντακτικής ταξινόμησης και τα τεχνητά νευρωνικά δίκτυα. Κατόπιν, αναλύεται η υπάρχουσα κατάσταση γύρω από τον προσδιορισμό προτύπων κοιλιακών ήχων και τα κίνητρα που οδήγησαν στην περαιτέρω διερεύνηση αυτού του ζητήματος. Στο πλαίσιο αυτό, παρουσιάζονται δύο νέα σχήματα – λεξικά ταξινόμησης μαζί με τους αντίστοιχους κανόνες περιγραφής – διαχωρισμού των κλάσεων, γεγονός που αποτελεί μία από τις σημαντικότερες καινοτομίες της παρούσης διατριβής. Επίσης, παρουσιάζονται τα αντίστοιχα συστήματα αυτόματης αναγνώρισης προτύπων που αναπτύχθηκαν για κάθε περίπτωση, δίνονται τα αποτελέσματα αξιολόγησής τους και σχολιάζονται τα πλεονεκτήματα της ανάλυσης κοιλιακών ηχητικών προτύπων.

5.1 Εφαρμογές ταξινόμησης και αναγνώρισης προτύπων

Τα τελευταία χρόνια, οι προσεγγίσεις ανάλυσης προτύπων (pattern analysis) και η ανάπτυξη αντίστοιχων εμπειρών συστημάτων (Expert Systems –ES), χρησιμοποιούνται ολοένα και περισσότερο, παρέχοντας νέες δυνατότητες αναζήτησης, εξαγωγής και διαχείρισης της γνώσης - πληροφορίας. Ανάμεσα στις διάφορες εφαρμογές, ξεχωριστή θέση κατέχουν οι μέθοδοι βιοϊατρικής τεχνολογίας [252]-[257], συμπεριλαμβανομένων και των ερευνητικών προσεγγίσεων μελέτης του γαστρεντερικού συστήματος [61], [67], [215]-[218], [258]-[259]. Η διαφορά ανάμεσα στους όρους ταξινόμηση και αναγνώριση προτύπων (pattern classification / pattern recognition) είναι αρκετά δυσδιάκριτη. Αυτό συμβαίνει γιατί οι δύο διαδικασίες συνυπάρχουν και αλληλοσυμπληρώνονται, ενώ ο χαρακτηρισμός ενός συστήματος ως διαδικασία ταξινόμησης ή αναγνώρισης, συχνά, ανάγεται στο είδος της μεθοδολογίας και των αλγορίθμων που επιλέγονται. Η ταξινόμηση προτύπων αναφέρεται κυρίως στην περιγραφή των «κλάσεων» (pattern classes – C_L), στη σύσταση των κανόνων - μηχανισμών σύγκρισης και τη διαμόρφωση του συστήματος αποφάσεων (decision rules). Η αναγνώριση προτύπων είναι μία διαδικασία που έπεται, χρονικά, και αφορά στην αξιοποίηση των κανόνων ταξινόμησης, με στόχο τον καθορισμό του προτύπου ενός οποιοδήποτε τυχαίου «δείγματος» [260]-[261].

Υπάρχουν δύο βασικές προσεγγίσεις σχεδιασμού και υλοποίησης των συστημάτων αναγνώρισης προτύπων: α) οι μέθοδοι εκπαίδευσης με εποπτεία (supervised learning), και β) οι μη εποπτευόμενες μέθοδοι (unsupervised learning) [260]-[263]. Η πρώτη αφορά στην αξιοποίηση «γνωστών δειγμάτων» (labeled samples), με στόχο την εκπαίδευση του εμπειρου συστήματος. Αυτό προϋποθέτει την εκ των προτέρων γνώση του αριθμού των κλάσεων C_L και την εξασφάλιση ενός ικανοποιητικού αριθμού αντιπροσωπευτικών προτύπων, από κάθε κατηγορία. Στην περίπτωση που δεν διατίθενται «ταυτοποιημένα δείγματα εκπαίδευσης», τότε καταφεύγουμε στις τεχνικές αναγνώρισης χωρίς εποπτεία (unsupervised learning), όπου συχνά ούτε ο αριθμός των κλάσεων είναι προκαθορισμένος. Έτσι, η «αναζήτηση

κανόνων ομαδοποίησης» (cluster seeking) γίνεται με βάση τα χαρακτηριστικά των επιλεγμένων παραμέτρων σύγκρισης, τα όρια διαχωρισμού και το μέγιστο επιτρεπτό αριθμό των διαφορετικών κλάσεων. Δύο από τις πιο «δημοφιλείς» αριθμητικές τεχνικές της κατηγορίας αυτής είναι ο αλγόριθμος ομαδοποίησης των k -μέσων (αν αυτή είναι η σωστή απόδοση του όρου “*k-means clustering*”) και η μέθοδος “ISODATA” (Iterative Self-Organizing Data Analysis Technique Algorithm) [260].

5.1.1 Η έννοια του προτύπου

Ο όρος «πρότυπο», ως μετάφραση της αγγλικής λέξης “pattern”, δεν είναι αρκετός από μόνος του, ώστε να περιγράψει πλήρως την έννοια του προτύπου, όπως τη χρησιμοποιούμε στις διάφορες εφαρμογές ταξινόμησης και αυτόματης αναγνώρισης. Στην προκειμένη περίπτωση, τα «πρότυπα» αναφέρονται στη διαδικασία προσδιορισμού αντιπροσωπευτικών δειγμάτων από ένα δειγματοχώρο γεγονότων-φαινομένων, βάσει των οποίων μπορεί να γίνει η ομαδοποίηση των ομοειδών δειγμάτων σε κλάσεις. Έτσι, επιδιώκεται η εύρεση και αξιοποίηση των χαρακτηριστικών εκείνων, που διαχωρίζουν με τον καλύτερο δυνατό τρόπο τα δείγματα διαφορετικών κλάσεων, ενώ παράλληλα παρουσιάζουν μεγάλη ομοιότητα όταν οι περιγραφές αναφέρονται σε πρότυπα της ίδιας κατηγορίας. Επίσης, αναζητούνται οι κανόνες και τα «όρια ομαδοποίησης», βάσει των οποίων θα γίνει η «σύγκριση» των επιλεγμένων χαρακτηριστικών [260]-[261], [264].

Στην πράξη, η υλοποίηση των έμπειρων συστημάτων γίνεται σε διάφορα υπολογιστικά περιβάλλοντα υλικού και λογισμικού, όπου τα αντίστοιχα φυσικά σήματα, αναπαρίστανται με ψηφιακές ακολουθίες (μίας ή περισσότερων διαστάσεων), που προκύπτουν από τη δειγματοληψία και ψηφιοποίηση των φαινομένων πραγματικού κόσμου. Έτσι, οι τεχνικές σύγκρισης και ομαδοποίησης πραγματοποιούνται με βάση χαρακτηριστικές παραμέτρους ανάλυσης (features) των ψηφιακών σημάτων. Η διαδικασία εξαγωγής χαρακτηριστικών παραμέτρων (feature extraction) αποτελεί ίσως τη σημαντικότερη φάση σχεδιασμού και ανάπτυξης των συστημάτων αναγνώρισης προτύπων [260]-[264]. Στόχος της είναι η αξιοποίηση εκείνων μόνο των παραμέτρων που ταυτοποιούν κατά τον καλύτερο δυνατό τρόπο «γεγονότα» της ίδιας ομάδας, παρουσιάζοντας μεγάλη συγκέντρωση των τιμών τους όταν εφαρμόζονται σε δείγματα της ίδιας κλάσης, και αντίστροφα μεγάλες αποκλίσεις όταν εφαρμόζονται σε διαφορετικά πρότυπα. Στο πλαίσιο αυτό, σημαντική θεωρείται η χρήση μικρού αριθμού χαρακτηριστικών παραμέτρων ανάλυσης, στοιχείο που αφενός περιορίζει τις υπολογιστικές απαιτήσεις, αφετέρου καθιστά τη διαδικασία ταξινόμησης πιο λειτουργική (η χρήση πολλών παραμέτρων ενδέχεται να προκαλέσει «σύγχυση» κατά την εκπαίδευση και γενικότερα να δυσχεράνει την αποδοτικότητα του συστήματος) [260]-[264].

Αν S είναι ένα τυχαίο δείγμα του δειγματικού χώρου που περιγράφεται από το διάνυσμα χαρακτηριστικών παραμέτρων (features' vector) $V_S (V_1, V_2, \dots, V_K)$, τότε καθεμιά από τις συνιστώσες $V_k (k=1, \dots, K)$ αντιστοιχεί και σε διαφορετική παράμετρο ανάλυσης, ενώ K είναι η τάξη του γεωμετρικού τόπου που ορίζουν τα διανύσματα V_S . Επομένως, το πρόβλημα ταξινόμησης - αναγνώρισης προτύπων ανάγεται στην εύρεση C_L γεωμετρικών υποχώρων (όσες είναι και οι κλάσεις του «προβλήματος»), καθένας από τους οποίους θα περιλαμβάνει διανύσματα διαφορετικής ομάδας (κλάσης). Οι «υπερ-επιφάνειες» (hyper-surfaces) που χωρίζουν το K -διάστατο χώρο στους C_L υποχώρους των κλάσεων (διάσταση $K-1$) ονομάζονται όρια απόφασης (decision boundaries). Έτσι, η ταξινόμηση προκύπτει από τη σύγκριση των τιμών μιας K -διάστασης «συνάρτησης απόφασης» (decision function $-f_D$), η οποία χρησιμοποιεί τις συναρτήσεις διάκρισης (discriminant functions) $\theta_\lambda (\lambda=1, \dots, C_L)$:

$$f_D(S) \triangleq f_D(\mathbf{V}_S) = f_D|_{\lambda, \xi}(V_1, V_2, \dots, V_K) = \theta_\lambda(S) - \theta_\xi(S) \begin{cases} > 0, & S \in C_\lambda \\ < 0, & S \in C_\xi \end{cases} \quad (5-1)$$

όπου τα λ, ξ αναπαριστούν δύο διαφορετικές κλάσεις C_λ και C_ξ , ενώ οι συναρτήσεις απόφασης θ_λ είναι όσες και ο αριθμός (C_L) των κλάσεων (ορίζεται μία συνάρτηση για κάθε κλάση). Άρα το πρόβλημα της αναγνώρισης ενός τυχαίου δείγματος ανάγεται στη σύγκριση όλων των τιμών $f_D|_{\lambda, \xi}(x)$. Μια πιο χρήσιμη συνάρτηση είναι η συνάρτηση επιλογής (choice function $-f_{CH}$) η οποία ορίζεται ως ακολούθως [260]:

$$f_C(S) \triangleq f_{CH}(\mathbf{V}_S) = \lambda, \quad \text{αν } S \in C_\lambda \Leftrightarrow f_{CH}(S) = \lambda, \quad \text{αν } \theta_\lambda(S) > \theta_\xi(S), \quad \forall \lambda \neq \xi \quad (5-2)$$

5.1.2 Τεχνικές στατιστικής ταξινόμησης προτύπων

Οι τεχνικές στατιστικής ταξινόμησης αποτελούν την απλούστερη και πιο κλασική προσέγγιση διαχωρισμού και αναγνώρισης προτύπων. Στην περίπτωση αυτή αναζητούνται οι κανόνες απόφασης (decision rules) που «εμπεριέχονται» στα γνωστά δείγματα των κλάσεων, δηλαδή οι παράμετροι ανάλυσης που ευνοούν το διαχωρισμό των αντίστοιχων δειγμάτων. Είναι προφανές ότι η επιλογή των χαρακτηριστικών παραμέτρων και η σύσταση των διανυσμάτων V_S καθορίζουν την τάξη του προβλήματος, δηλαδή τη διάσταση (K) του διανυσματικού χώρου, αλλά και όλων των γεωμετρικών τόπων που ορίζονται για κάθε κλάση C_λ . Στόχος της διαδικασίας ομαδοποίησης (clustering) είναι η κατά το δυνατόν αυστηρότερη οριοθέτηση αυτών των υποχώρων, έτσι ώστε οι διαφορετικές κλάσεις να απέχουν μεταξύ τους όσο περισσότερο γίνεται. Οι γεωμετρικοί τόποι των κλάσεων έχουν ακτινική διάταξη με κέντρα (c_λ) τους αντίστοιχους μέσους όρους (*mean*) των γνωστών δειγμάτων εισόδου [260]:

$$c_\lambda(c_{\lambda 1}, \dots, c_{\lambda k}, \dots, c_{\lambda K}) = \text{mean}(\mathcal{S}_\lambda) = \frac{1}{N_\lambda} \sum_{\lambda i=1}^{N_\lambda} S_{\lambda i}(V_1, V_2, \dots, V_K), \quad \lambda = 1, \dots, C_L, \quad \lambda i = 1, \dots, N_\lambda \quad (5-3)$$

όπου N_λ ($\lambda=1, \dots, C_L$) είναι ο αριθμός των γνωστών δειγμάτων $S_\lambda = [S_{\lambda i}]$, $\lambda i=1, \dots, N_\lambda$, κάθε κλάσης C_λ , με κέντρο c_λ . Αν θεωρήσουμε ότι το γεωμετρικό κέντρο των κλάσεων αποτελεί ένα αντιπροσωπευτικό σημείο αναφοράς, τότε ο πιο απλός κανόνας ταξινόμησης είναι αυτός της «πλησιέστερης γειτονιάς» (nearest neighbour), όπου το εκάστοτε δείγμα S κατατάσσεται στην κλάση με τη μικρότερη «ακτινική» απόσταση (Distance $-D$) [260]:

$$S \in C_\lambda, \quad \text{αν } D(S, c_\lambda) < D(S, c_\xi), \quad \forall \lambda, \xi = 1, \dots, C_L, \quad \lambda \neq \xi \quad (5-4)$$

Από την παραπάνω διαδικασία προέκυψε και το όνομα των αντίστοιχων «αλγορίθμων μικρότερης απόστασης» (minimum distance algorithms). Για τον υπολογισμό της απόστασης D χρησιμοποιούνται διάφορες εκφράσεις, με πιο κλασική την Ευκλείδεια (D_E , εξ. 5-5a), ενώ συχνά εφαρμόζεται και η απόσταση Mahalanobis (D_M , εξ. 5-5b) [260], [265]:

$$D_E(\mathbf{V}_S, c_\lambda) = \sqrt{\sum_{\kappa=1}^K (V_\kappa - c_{\lambda \kappa})^2} \quad (5-5)$$

$$D_M(\mathbf{V}_S, c_\lambda) = \sqrt{(\mathbf{V}_S - c_\lambda) \cdot \Sigma_\lambda^{-1} \cdot (\mathbf{V}_S - c_\lambda)}, \quad \text{με } \Sigma_\lambda = \frac{1}{N_\lambda - 1} (\mathcal{S}_\lambda - c_\lambda)(\mathcal{S}_\lambda - c_\lambda)^T$$

όπου $V_S(V_1, \dots, V_{\kappa}, \dots, V_K)$ είναι το διάνυσμα του δείγματος S στον K -διάστατο γεωμετρικό χώρο, Σ_λ είναι ο πίνακας συμμεταβλητότητας (covariance matrix) εντός της κλάσης C_λ , οι εκθέτες «-1» και «T» συμβολίζουν τον αντίστροφο και ανάστροφο πίνακα, ενώ D_E, D_M είναι οι αποστάσεις ανάμεσα στο κέντρο της κλάσης c_λ και το τυχαίο δείγμα S . Η νέα, κανονικοποιημένη απόσταση D_M έχει ως στόχο τη μείωση της επίδρασης των συνιστωσών V_κ για τις οποίες παρουσιάζεται μεγάλη διασπορά εντός των κλάσεων, κι αντίστροφα την μεγαλύτερη επίδραση εκείνων που έχουν μικρότερες διακυμάνσεις. Έτσι, η χρήση της σε εφαρμογές ταξινόμησης είναι πλεονεκτικότερη έναντι της Ευκλείδειας απόστασης, καθώς οι πράξεις συσχέτισης ελέγχουν την ομοιότητα ανάμεσα στο υπό εξέταση δείγμα και το «μέσο προφίλ» κάθε κλάσης C_λ [260], [265].

Σε κάθε περίπτωση, τα διανύσματα V_S αποτελούν κανονικοποιημένες εκφράσεις των αρχικών παραμέτρων ανάλυσης, καθώς οι τελευταίες ενδέχεται να κυμαίνονται σε ένα μεγάλο εύρος τιμών με εντελώς διαφορετικά όρια για κάθε συντεταγμένη. Αυτό σημαίνει ότι η «βαρύτητα» της κάθε παραμέτρου στον υπολογισμό της απόστασης D κι επομένως και στα αποτελέσματα ταξινόμησης, εξαρτάται σε μεγάλο βαθμό από τις διακυμάνσεις των τιμών της, με αρνητικές επιπτώσεις στις τελικές αποφάσεις ομαδοποίησης [260], [265]. Δύο τυπικές προσεγγίσεις κανονικοποίησης (vector scaling) προκύπτουν κατά τη διαίρεση όλων των συνιστωσών V_κ με το μέγιστο παρατηρούμενο εύρος (Range $-R_\kappa$) διακύμανσης, είτε με διαίρεση ως προς τις αντίστοιχες τιμές διασποράς σ_κ^2 [260]:

$$V_\kappa' \leftarrow \frac{V_\kappa}{R_\kappa}, \quad \text{με } R_\kappa = \max_{\lambda, \lambda_i} \{V_\kappa^{\lambda_i}\} - \min_{\lambda, \lambda_i} \{V_\kappa^{\lambda_i}\}, \quad \forall \kappa = 1, \dots, K, \quad \lambda = 1, \dots, C_L, \quad \lambda_i = 1, \dots, N_\lambda$$

$$V_\kappa' \leftarrow \frac{V_\kappa}{\sigma_\kappa}, \quad \text{με } \sigma_\kappa^2 = \frac{1}{\sum_{\lambda=1}^{C_L} N_\lambda} \cdot \sum_{\lambda=1}^{C_L} \sum_{\lambda_i=1}^{N_\lambda} [V_\kappa^{\lambda_i} - \overline{V_\kappa^{\lambda_i}}], \quad \forall \kappa = 1, \dots, K, \quad \lambda = 1, \dots, C_L, \quad \lambda_i = 1, \dots, N_\lambda \quad (5-6)$$

όπου $V_\kappa^{\lambda_i}$ είναι οι συνιστώσες V_κ για όλες τα δείγματα S_{λ_i} και σε όλες τις κλάσεις C_λ , ενώ η παράμετρος $\overline{V_\kappa^{\lambda_i}}$ αντιστοιχεί στο μέσο όρο (mean) όλων των τιμών $V_\kappa^{\lambda_i}$.

Ένα άλλο ζήτημα, αφορά στην επιλογή των καταλληλότερων παραμέτρων ανάλυσης (feature selection), ανάμεσα σε όλες τις μεταβλητές που έχουν «εξαχθεί» ως ικανές για να διαχωρίσουν τις διάφορες κλάσεις (extracted features). Η διαδικασία αυτή αποσκοπεί στον περιορισμό της διάστασης των διανυσμάτων εισόδου (dimensionality reduction), και την αντιμετώπιση των αντιστοιχών προβλημάτων που πηγάζουν από τη χρήση πολλών παραμέτρων ανάλυσης (curse of dimensionality) [262]. Η επιλογή των πιο «εξεχόντων» παραμέτρων (salient features), μπορεί να γίνει με γνώμονα την ελαχιστοποίηση της μέσης απόστασης μεταξύ δειγμάτων ίδιας κλάσης (intra-set distance, D_{intra}) και αντίστοιχα τη μεγιστοποίηση της μέσης απόστασης διαφορετικών κλάσεων (inter-set distance, D_{inter}) [260]:

$$D_{intra}(\lambda) = \frac{1}{N_\lambda \cdot (N_\lambda - 1)} \sum_{\lambda_i=1}^{N_\lambda} \sum_{\lambda_j=1}^{N_\lambda} D^2(S_{\lambda_i}, S_{\lambda_j}), \quad \forall \lambda \in [1, C_L], \quad \lambda_i \neq \lambda_j \quad (5-7)$$

$$D_{inter}(\lambda, \xi) = \frac{1}{N_\lambda \cdot N_\xi} \sum_{\lambda_i=1}^{N_\lambda} \sum_{\xi_j=1}^{N_\xi} D^2(S_{\lambda_i}, S_{\xi_j}), \quad \lambda, \xi \in [1, C_L], \quad \lambda \neq \xi \quad (5-8)$$

Οι παραπάνω εκφράσεις αξιοποιούνται στο σχηματισμό δεικτών ποιότητας (*measures of Quality -Qm*), βάσει των οποίων ελέγχεται η καταλληλότητα των συνιστωσών V_κ [230], [260]:

$$Q_m = \frac{\sum_{\lambda=1}^{C_L} D_{intra}(\lambda)}{\sum_{\lambda=1}^{C_L} \sum_{\xi=1}^{C_L} D_{inter}(\lambda, \xi)} \quad \eta \quad Q_M = \frac{\min_{\lambda, \xi} (D_{inter}(\lambda, \xi))}{\max_{\lambda} (D_{intra}(\lambda))}, \quad \lambda, \xi \in [1, C_L], \quad \lambda \neq \xi \quad (5-9)$$

όπου τα διανύσματα V_S επιλέγονται έτσι ώστε να ελαχιστοποιούν το δείκτη Q_m [260], ή αντίστοιχα να μεγιστοποιούν το δείκτη Q_M για όσο το δυνατό περισσότερους συνδυασμούς κλάσεων C_λ, C_ξ [230]. Ειδικότερα για το δείκτη Q_M , το κριτήριο, ώστε τα διανύσματα $V_S = \{V_\kappa\}$ να συνιστούν αποδοτικούς συνδυασμούς των χαρακτηριστικών παραμέτρων ανάλυσης (ως προς την ικανότητα διαχωρισμού των κλάσεων), εκφράζεται με τη συνθήκη " $Q_M > 1$ " [230]. Εκτός των παραπάνω, μια πρόσθετη διαδικασία που εφαρμόζεται με σκοπό την περιορισμό της διάστασης των διανυσμάτων, αφορά στην αποφυγή χρήσης γραμμικά εξαρτημένων (συσχετισμένων) παραμέτρων ανάλυσης, στοιχείο που συνήθως ελέγχεται με το δείκτη γραμμικής συσχέτισης του *Pearson* (δηλαδή το δείκτη *CCI* της εξ. 3-51) [230], [260]-[262].

5.1.3 Αναγνώριση προτύπων με χρήση νευρωνικών δικτύων

Τα νευρωνικά δίκτυα (*neural networks -NNET*) χρησιμοποιούνται σε μια μεγάλη γκάμα προβλημάτων μηχανικής, συμπεριλαμβανομένων και των εφαρμογών ταξινόμησης και αναγνώρισης προτύπων που προαναφέρθηκαν. Το όνομά τους προήλθε από το γεγονός ότι παρουσιάζουν αρκετές ομοιότητες με τη δομή και τους μηχανισμούς λειτουργίας των φυσικών νευρωνικών δικτύων. Έτσι, αποτελούνται από τεχνητούς νευρώνες, συνάψεις και κόμβους, ενώ η υπολογιστική ισχύς επεξεργασίας και η «ικανότητα λήψης αποφάσεων» αποκτούνται, συνήθως, μέσα από διαδικασίες εκπαίδευσης, παρόμοιες με τους μηχανισμούς απόκτησης γνώσης και εμπειρίας στον άνθρωπο (έμπειρα συστήματα). Ανάμεσα στις κατηγορίες νευρωνικών δικτύων, ιδιαίτερο ενδιαφέρον παρουσιάζουν οι «πολυεπίεδοι τεχνητοί νευρώνες προώθησης προς τα εμπρός» (*Multi-Layer Perceptrons -MLP, Feed-Forward networks*) και οι αντίστοιχες τεχνικές «εποπτευόμενης εκπαίδευσης» (*supervised learning*) με «διάδοση προς τα πίσω» (*Back Propagation -BP*) [263]-[264], [266]-[267], που χρησιμοποιήθηκαν και για τις ανάγκες της παρούσης διατριβής [215]-[216], [218].

Οι τεχνητοί νευρώνες (*TN, perceptrons*) περιλαμβάνουν τις διανυσματικές συνιστώσες εισόδου $\mathbf{V} = \{V_\kappa\}$, τους συντελεστές βαρών $\mathbf{W} = \{w_{\kappa j}\}$, την παράμετρο πόλωσης-κατωφλίου (*bias*) θb και τη συνάρτηση ενεργοποίησης (*activation function- f_{act}*), η οποία δίνει και την τελική έξοδο του νευρώνα (σχ. 5-1α). Οι παράμετροι εισόδου πολλαπλασιάζονται με τα αντίστοιχα διανύσματα βαρών, ενώ το άθροισμά τους αποτελεί την «είσοδο δικτύου» u (*net input*) που οδηγείται στη συνάρτηση ενεργοποίησης. Οι τοπολογίες *MLP* (σχ. 5-1β) σχηματίζονται από πολλούς τέτοιους νευρώνες, οι οποίοι δομούνται σε επίπεδα (*layers -l*) και αντίστοιχους νευρωνικούς κόμβους (*nodes -nd*):

$$\begin{aligned} \mathbf{u}_\ell &= \{u_{\ell, nd(\ell)}\}, \quad \ell = 0, \dots, L_N \\ \mathbf{V}_S = \mathbf{V} &= \{V_\kappa\} = \mathbf{u}|_{\ell=0}, \quad nd(0) = K \\ \theta \mathbf{b}_\ell &= \{\theta b_{\ell j}\}, \quad \ell j = 1, \dots, \text{length}(\mathbf{u}_\ell) \\ \mathbf{W}_\ell &= \{W_{\ell i, \ell j}\}, \quad \ell i = 1, \dots, nd(\ell + 1), \quad \ell j = 1, \dots, \text{length}(\mathbf{u}_\ell) \end{aligned} \quad (5-10)$$

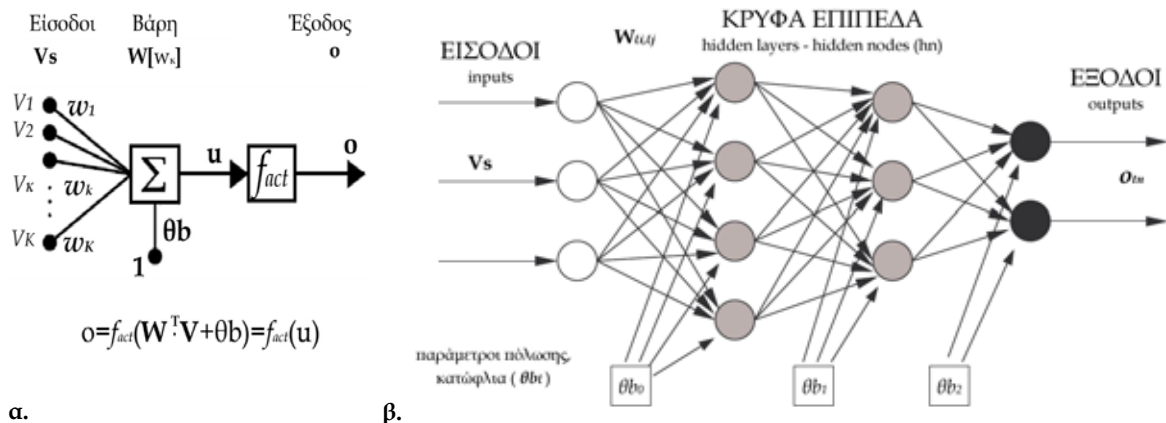
όπου L_N είναι το πλήθος των κρυφών επιπέδων, θb_ℓ οι τιμές πόλωσης - κατωφλίωσης (*bias*) και $\mathbf{W}_\ell = \{W_{\ell i, \ell j}\}$ οι αντίστοιχοι πίνακες των συντελεστές βαρών σε κάθε επίπεδο ℓ . Έτσι, οι

έξοδοι $u_{\ell j}$ των νευρώνων προωθούνται σταδιακά προς τα εμπρός (feed forward) μέχρι το τελικό επίπεδο εξόδου του δικτύου [216], [218], [263]-[264], [266]-[267]:

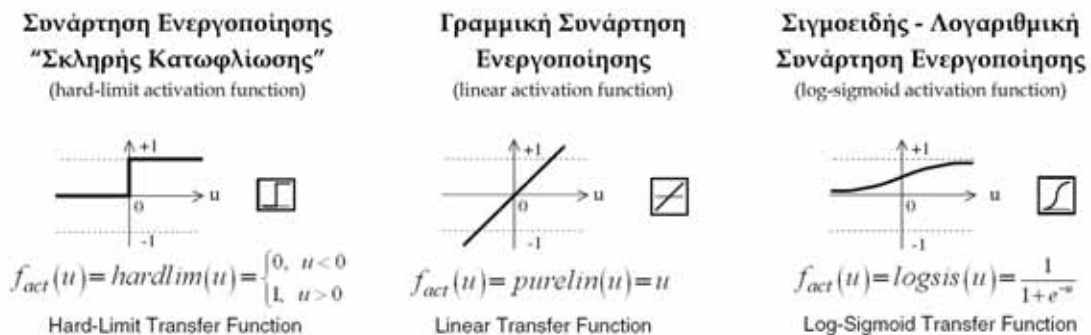
$$\mathbf{u}_{\ell+1} = f_{act}^{\ell}(\mathbf{W}_{\ell} \cdot \mathbf{u}_{\ell} + \theta \mathbf{b}_{\ell}) \Leftrightarrow u_{(\ell+1),\ell i} = f_{act}^{\ell} \left(\sum_{\ell j} W_{\ell i, \ell j} \cdot u_{\ell, \ell j} + \theta b_{\ell i} \right) \quad (5-11)$$

$$\mathbf{o} = \{o_m\} = \mathbf{u}_{(L_N+1)}, \quad m = 1, 2, \dots, T_N$$

Με βάση τα παραπάνω, τα διανύσματα εισόδου (feature vectors) συνδυάζονται με έναν αρκετά περίπλοκο (μη γραμμικό) τρόπο, με στόχο τη διαμόρφωση των T_N συνολικά εξόδων (o_m). Η ισχύς υπολογιστικής επεξεργασίας των δικτύων *MLP* βρίσκεται στους συντελεστές βαρών \mathbf{W}_{ℓ} και στις παραμέτρους πόλωσης-κατώφλιων $\theta \mathbf{b}_{\ell}$, οι οποίοι ρυθμίζονται κατάλληλα κατά τη διαδικασία εκπαίδευσης. Από τις διάφορες συναρτήσεις ενεργοποίησης (σχ. 5-2), συνήθως επιλέγονται σιγμοειδής συναρτήσεις για τα κρυφά επίπεδα, ενώ προτιμώνται γραμμικές συναρτήσεις στο τελευταίο επίπεδο, καθώς επιτρέπουν τη ρύθμιση των εξόδων του δικτύου σε οποιοδήποτε τιμές. Γενικά, δίκτυα *MLP* με ένα κρυφό επίπεδο σιγμοειδών συναρτήσεων και ένα γραμμικό επίπεδο εξόδου, μπορούν να «επιλύσουν» ακόμη και τα πιο δύσκολα προβλήματα, αρκεί βέβαια να επιλεγούν τα κατάλληλα διανύσματα στην είσοδο. Στις εφαρμογές αναγνώρισης προτύπων, ο αριθμός των εξόδων επιλέγεται ίσος με τον αντίστοιχο αριθμό των κλάσεων ($T_N=C_L$), όπου η ενεργή κάθε φορά λογική έξοδος ($o_m=o_{\lambda}=1$) δίνει και το αντίστοιχο αποτέλεσμα ταξινόμησης [216], [218], [263]-[264], [266]-[267].



α. **β.** Σχήμα 5-1. Βασικές τοπολογίες νευρωνικών δικτύων: **α)** ο απλός τεχνητός νευρώνας (TN, perceptron), **β)** τοπολογία MLP με 3 εισόδους, 2 κρυφά επίπεδα (4 και 3 κόμβων) και 2 εξόδους.



Σχήμα 5-2. Τυπικές συναρτήσεις ενεργοποίησης (activation functions, f_{act}) που χρησιμοποιούνται σε πολυεπίπεδα νευρωνικά δίκτυα (MLP).

Κατά τη διαδικασία εκπαίδευσης με εποπτεία, χρησιμοποιούνται ζεύγη γνωστών δειγμάτων (training pairs –“ $tp(nq)=\{V_{in}(nq),dp(nq)\}$, $nq=1,\dots,N_{TP}$ ”), που αποτελούνται από τα N_{TP} διανύσματα εισόδου V_{in} και τα αντίστοιχα «διανύσματα προτύπων» dp (desired patterns). Στην πιο κλασική περίπτωση της διαδικασίας διασταυρούμενου ελέγχου (cross validation), ένα τμήμα των δειγμάτων χρησιμοποιείται κατά τη φάση της «εκμάθησης» (learning part “ $tp_L(nq)$ ”), ενώ το υπόλοιπο χρησιμοποιείται για την αξιολόγηση των αποτελεσμάτων (validation part, “ $tp_V(nq)$ ”). Στόχος της μάθησης είναι να ρυθμιστούν κατάλληλα οι συντελεστές του δικτύου, έτσι ώστε οι έξοδοι του $o(nq)$ να ισοδυναμούν με τα επιθυμητά πρότυπα $dp(nq)$. Στο πλαίσιο αυτό, εκτελούνται πολλοί βρόχοι εκπαίδευσης (*training loops*, tr), σε καθέναν από τους οποίους οι συντελεστές μεταβάλλονται προς την κατεύθυνση μείωσης του μέσου σφάλματος J_{error} . Έτσι, ξεκινώντας από το επίπεδο εξόδου είναι δυνατό να υπολογιστεί το μέσο τετραγωνικό σφάλμα (mean square error, J_{mse}) σε κάθε κόμβο nd και στη συνέχεια να μεταβληθούν όλοι οι συντελεστές $W_{\ell i,\ell j}^{(tr)}$, κατά μία ποσότητα $\Delta W_{\ell i,\ell j}^{(tr)}$, η οποία είναι ανάλογη της αρνητικής κλίσης (gradient) του μέσου σφάλματος $J_{mse}^{(tr)}$. Με τον τρόπο αυτό γίνεται διαδοχικά (προς τα πίσω) η ανανέωση όλων των συντελεστών, στοιχείο από το οποίο προέκυψε και η ονομασία «εκπαίδευση προς τα πίσω» (Back Propagation –BP) και ο αντίστοιχος αλγόριθμος “*Gradient Descent Back Propagation*” (GDBP):

$$J_{mse}^{(tr)} = \frac{1}{N_{TP;L}} \sum_{nq=1}^{N_{TP;L}} J^{(tr)}(nq), \quad \text{με } J^{(tr)}(nq) = \frac{1}{2} \sum_{tn=1}^{T_N} [o^{(tr)}(nq) - dp_L(nq)]^2 \quad (5-12)$$

$$W_{\ell i,\ell j}^{(tr)} = W_{\ell i,\ell j}^{(tr-1)} + n_{LR} \cdot \Delta W_{\ell i,\ell j}^{(tr)} + \mu \cdot W_{\ell i,\ell j}^{(tr-1)}, \quad \text{με } \Delta W_{\ell i,\ell j}^{(tr)} = -\frac{\partial J_{mse}^{(tr)}}{\partial W_{\ell i,\ell j}}$$

όπου n_{LR} είναι ο ρυθμός μάθησης (*Learning Rate*) που καθορίζει την ταχύτητα εκπαίδευσης του δικτύου, ενώ η θετική σταθερά μνήμης μ (momentum term) χρησιμοποιείται έτσι ώστε να εξασφαλιστούν ομαλές μεταβάσεις των συντελεστών και να αποφευχθεί η «παγίδευση» σε τοπικά ακρότατα (local minima) της επιφάνειας σφάλματος J (error surface)· σε περίπτωση που $\mu=0$, η εκπαίδευση εξελίσσεται χωρίς τη χρήση «όρων μνήμης». Βέβαια, αντίστοιχοι κανόνες εφαρμόζονται και για την ανανέωση των συντελεστών πόλωσης θb_ℓ . Η ικανότητα γενίκευσης (generalization) αξιολογείται με τα δείγματα $tp_V(nq)$, τα οποία είναι άγνωστα στο δίκτυο, καθώς δεν χρησιμοποιούνται καθόλου κατά την εκπαίδευση. Γενικά, απαιτείται ένας μεγάλος αριθμός ζευγών εκπαίδευσης (σε σύγκριση με το σύνολο όλων των ρυθμιζόμενων συντελεστών), γιατί διαφορετικά το δίκτυο «υπερ-εκπαιδεύεται» (*overtraining*, *over-fitting*), δηλαδή περιορίζεται στο να «απομνημονεύσει» τα δείγματα εισόδου. Έτσι, υπάρχουν αρκετοί περιορισμοί ως προς το μέγεθος των δικτύων ανάλογα με τη φύση του προβλήματος και το διαθέσιμο αριθμό δειγμάτων εκπαίδευσης [263], [267]-[268], ενώ παράλληλα έχουν προταθεί πιο εξελιγμένες προσεγγίσεις διαμόρφωσης των υποσυνόλων εκπαίδευσης (*k-fold / leave-one-out cross validation*, *jackknife* [267]), καθώς και τεχνικές πρόωρου τερματισμού (*early stopping*) [216], [218], [263]-[264], [266]-[267].

Μια εναλλακτική αντιμετώπιση των παραπάνω προβλημάτων, είναι η αξιοποίηση κανονικοποιημένων εκφράσεων του σφάλματος (regularized mean square error –*msereg*), θέτοντας περιορισμούς (penalty terms) στις επιτρεπές μεταβάσεις μεταξύ διαδοχικών επαναλήψεων [263], [268]-[269]. Όσον αφορά την επιτάχυνση των διαδικασιών εκπαίδευσης, συχνά χρησιμοποιούνται μεταβλητοί ρυθμοί μάθησης (*adaptive learning rate*), οι οποίοι προσαρμόζονται ανάλογα με τις μεταβολές της συνάρτησης κόστους. Παράλληλα, έχουν αναπτυχθεί διάφορες αλγεβρικές τεχνικές γρήγορης εκπαίδευσης (*faster training*), όπως

είναι οι αλγόριθμοι Quasi-Newton και Levenberg-Marquard [254]-[255], [263] και η μέθοδος “resilient propagation” (RPROP) [270]. Η τελευταία χρησιμοποιήθηκε εκτενώς για τις ανάγκες της παρούσης διατριβής, παρέχοντας μεγάλες ταχύτητες εκπαίδευσης και ομαλές διακυμάνσεις κατά την «πτώση» της καμπύλης μέσου σφάλματος [216], [218]. Στη μέθοδο RPROP, οι συντελεστές μεταβάλλονται κατά σταθερές ποσότητες $A_{i,j}^{(tr)}$, με βάση μόνο το πρόσημο και όχι το μέτρο της κλίσης $\partial J/\partial W$ [254], [263], [270]:

$$\Delta W_{i,j}^{(tr)} = -A_{i,j}^{(tr)} \cdot \text{sgn} \left(\frac{\partial J^{(tr)}}{\partial W_{i,j}} \cdot \frac{\partial J^{(tr-1)}}{\partial W_{i,j}} \right), \text{ όπου } \text{sgn} : \text{ η συνάρτηση προσήμου} \quad (5-13)$$

Τα νευρωνικά δίκτυα χρησιμοποιούνται πολύ συχνά σε προβλήματα αναγνώρισης προτύπων, χάρη στην «έμφυτη» ικανότητα προσαρμοστικότητας, που διαθέτουν, και τις, σχετικά, μικρές υπολογιστικές απαιτήσεις. Μια από τις σημαντικότερες παραμέτρους που καθορίζουν την αποτελεσματικότητα των έμπειρων αυτών συστημάτων, εξακολουθεί να είναι η επιλογή των κατάλληλων διανυσματικών χαρακτηριστικών (features vectors), με τις ενδεδειγμένες, κάθε φορά, τοπολογίες δικτύων. Βέβαια, η διαδικασία επιλογής των χαρακτηριστικών παραμέτρων ανάλυσης (feature selection) είναι «λιγότερο κρίσιμη», από ότι στους αλγόριθμους ελάχιστης απόστασης και γενικότερα στις προσεγγίσεις στατιστικής ταξινόμησης. Αυτό συμβαίνει καθώς το δίκτυο μπορεί και προσαρμόζει κατάλληλα τα διανύσματα βαρών στην είσοδο, «ελέγχοντας» την επίδραση κάθε παραμέτρου στο τελικό αποτέλεσμα. Έτσι, αν κάποιες μεταβλητές δεν προσφέρονται για τις εκάστοτε ανάγκες ταξινόμησης, η επίδρασή τους περιορίζεται αυτόματα, δηλαδή οι αντίστοιχες συνάψεις ρυθμίζονται σε χαμηλές τιμές. Ωστόσο, παραμένει υψηλό το υπολογιστικό κόστος, καθώς η τάξη των διανυσμάτων εισόδου είναι μεγαλύτερη απ’ όσο θα έπρεπε. Στο πλαίσιο αυτό, έχουν προταθεί διάφορες τεχνικές προσαρμοστικής ανάπτυξης (grow) και διαγραφής (pruning) νευρώνων σε ένα δίκτυο MLP, ανάλογα με τη συνεισφορά των αντίστοιχων κόμβων - συνάψεων [266]. Παράλληλα, έχουν αναπτυχθεί διάφοροι αλγόριθμοι επιλογής χαρακτηριστικών παραμέτρων (feature selection algorithms), αποκλειστικά για εφαρμογές νευρωνικών δικτύων και ειδικότερα τοπολογιών MLP [271]-[274].

5.1.4 Προσεγγίσεις συντακτικής - ιεραρχικής ταξινόμησης προτύπων

Σε πολλά προβλήματα αναγνώρισης και ιδιαίτερα σε περιπτώσεις περίπλοκων προτύπων, συχνά υιοθετούνται δομές ιεραρχικής - συντακτικής ταξινόμησης, όπου κάθε πρότυπο αντιμετωπίζεται σαν ένα υπερσύνολο, σχηματιζόμενο από πολλά «υπο-πρότυπα» [262], [275], [276]. Τα πρωταρχικά (primitives, elementary) πρότυπα έχουν συνήθως απλούστερη μορφή-σύσταση, ενώ η ταξινόμησή τους μπορεί να γίνει με οποιαδήποτε μέθοδο αναγνώρισης προτύπων, συμπεριλαμβανομένων και των τεχνικών που προαναφέρθηκαν. Τα σύνθετα (complex) πρότυπα υψηλότερου επιπέδου αποτελούνται από συνδυασμούς των βασικών προτύπων, ενώ για τις ανάγκες ταξινόμησής τους υιοθετούνται συντακτικοί κανόνες και λογικές εκφράσεις αποφάσεων. Έτσι, υπάρχει μια αναλογία ανάμεσα στη δομημένη ταξινόμηση προτύπων (structural pattern recognition) και τους αντίστοιχους γλωσσολογικούς κανόνες, όπου τα πρωταρχικά στοιχεία δομούνται σε προτάσεις τυποποιημένης «γραμματικής» και «συντακτικού», έτσι ώστε να σχηματίσουν τα «τελικά πρότυπα» [262].

Το κυριότερο πλεονέκτημα αυτών των προσεγγίσεων (εκτός από τα αποτελέσματα ταξινόμησης) αφορά στην παροχή περιγραφικών σχημάτων σχετικά με τα χαρακτηριστικά

σύστασης και τη μορφολογία των προτύπων [262]. Ένα άλλο θετικό στοιχείο πηγάζει από την ιεραρχική φιλοσοφία αντιμετώπισης, όπου το αρχικό περίπλοκο πρόβλημα αναλύεται σε πολλές και απλούστερες επιμέρους εργασίες, προσδίδοντας μεγαλύτερη λειτουργικότητα και λιγότερες υπολογιστικές απαιτήσεις στη συνολική μεθοδολογία. Ωστόσο, χρειάζεται ιδιαίτερη προσοχή, καθώς τα ενδεχόμενα σφάλματα αναγνώρισης των χαμηλότερων επιπέδων μεταφέρονται αλυσιδωτά έως την κορυφή της πυραμίδας, όπου βρίσκονται τα σύνθετα πρότυπα [275]-[276]. Σε κάθε περίπτωση, η δυνατότητα υιοθέτησης διαφορετικών και ανεξάρτητων εργαλείων ταξινόμησης, σε κάθε επίπεδο της «ιεραρχίας προτύπων», παρέχει πρόσθετη ευελιξία, επιτρέποντας την αποτελεσματική αντιμετώπιση και των πιο σύνθετων προβλημάτων. Για το λόγο αυτό, η συγκεκριμένη προσέγγιση έχει εφαρμοστεί επιτυχώς σε μεθόδους ανάλυσης βιοϊατρικών – βιοακουστικών σημάτων [262], καθώς και σε γενικότερες εφαρμογές αναγνώρισης ηχητικών [275] -[276].

5.2 Ανάλυση προτύπων γαστρεντερικών ακουστικών φαινομένων

Όπως έχει ήδη επισημανθεί, η γαστρεντερική δραστηριότητα παρουσιάζεται σε κύκλους που περιλαμβάνουν τη φάση εφησυχασμού (quiescence phase), την ομαλή φάση (normal phase) ως προς τη μορφολογία και τη συχνότητα επαναλήψεων των συσπάσεων, και την ακανόνιστη φάση τυχαίων συσταλτικών κινήσεων (abnormal phase) [20]-[22]. Τα παραπάνω χαρακτηριστικά, που παρατηρούνται κυρίως στις περιόδους νηστείας, έχουν μελετηθεί με τις μανομετρικές [20]-[22] και ηλεκτρομυογραφικές [26], [31] μεθόδους και συνιστούν τον κύκλο MMC. Παράλληλα, έχουν διαπιστωθεί αντίστοιχες μεταβολές στην αλληλοδιαδοχή των «φάσεων» και με τις βιοακουστικές καταγραφές [24]-[25], [161]-[162], [215]-[218]. Παρόλα αυτά, δεν έχει υιοθετηθεί ως τώρα ένα τυποποιημένο σχήμα ταξινόμησης προτύπων «κοιλιακών ήχων», πέρα από τις γενικές παρατηρήσεις, που αφορούν στα χαρακτηριστικά «κρουστικότητας» και «διάρκειας» των σημάτων [23], [64], [69], [59]. Το στοιχείο αυτό είναι ανασταλτικό στην εξέλιξη των προσεγγίσεων βιοακουστικής διάγνωσης της γαστρεντερικής λειτουργίας, συμπέρασμα που ενισχύεται ακόμη περισσότερο αν αναλογιστούμε ότι, η τυποποίηση αντίστοιχων κανόνων ταξινόμησης είναι στενά συνδεδεμένη με την πρόοδο των καρδιο-αναπνευστικών ακροαστικών μεθόδων [83]-[84], [218].

Γενικά, υπάρχει περιορισμένος αριθμός δημοσιεύσεων στη διεθνή βιβλιογραφία με αντικείμενο την καθιέρωση κανόνων ταξινόμησης των κοιλιακών ήχων. Χαρακτηριστική είναι η εργασία του Arnbjornsson το 1986 [67], ο οποίος επιχείρησε τον διαχωρισμό των «φυσιολογικών» από τους «παθολογικούς» ήχους, καταλήγοντας στο συμπέρασμα ότι μια τέτοια διάκριση δεν είναι εύκολη υπόθεση. Από τα στοιχεία της βιβλιογραφικής αναζήτησης που πραγματοποιήσαμε, δεν έχει διαπιστωθεί πρόοδος προς την κατεύθυνση αυτή, ούτε έχει καταγραφεί κάποια αντίστοιχη ερευνητική προσπάθεια. Παραπλήσια ήταν και η εργασία των Yoshino et al. το 1990 [70], οι οποίοι επιχείρησαν να ταξινομήσουν τους «τύπους» εντερικών ήχων με βάση το φασματικό τους περιεχόμενο. Η εργασία είχε στόχο τη διάγνωση της εντερικής απόφραξης (intestinal obstruction) και κατέληξε στον προσδιορισμό τριών διαφορετικών φασματικών προφίλ - τύπων (types I, II, III). Απ' όσο γνωρίζουμε, η συγκεκριμένη μέθοδος δεν είχε περαιτέρω εξέλιξη, τόσο σε ερευνητικό επίπεδο, όσο και σε επίπεδο κλινικής εφαρμογής, ενώ είναι γνωστό ότι τα περιστατικά απόφραξης μπορούν να διαγνωστούν εύκολα με τις κλινικά αποδεκτές ενδοσκοπικές μεθόδους.

Η προσέγγιση ανάλυσης προτύπων (Abdominal Sound Pattern Analysis -ASPA), που ακολουθήθηκε στο πλαίσιο της παρούσης διατριβής, είναι αρκετά διαφορετική σε σχέση με τις προηγούμενες ερευνητικές προσπάθειες. Στην προκειμένη περίπτωση, η ταξινόμηση

προτύπων αφορά στα απομονωμένα βιοακουστικά συμβάντα μικρής διάρκειας, δηλαδή τις δομικές μονάδες σύστασης του συνολικού μηχανισμού γαστρεντερικής δραστηριότητας, ανεξάρτητα από την κατάσταση (παθολογική ή όχι) των αντίστοιχων υποκειμένων. Η εισαγωγή εργαλείων αναγνώρισης προτύπων στη μεθοδολογία ανάλυσης των καταγραφών παρατεταμένου χρόνου έχει διπλό σκοπό: α) τον εντοπισμό των μη επιθυμητών ηχητικών παρεμβολών και το διαχωρισμό τους από τα χρήσιμα σήματα (δηλαδή την απομάκρυνση του μη στάσιμου προσθετικού θορύβου), και β) την παροχή νέων δυνατοτήτων ανάλυσης, που σχετίζονται με την παρακολούθηση εναλλαγής των προτύπων κατά μήκος των καταγραφών. Στο πλαίσιο αυτό, ενδιαφέρον παρουσιάζει η ανάλυση χρονικής και χωρο-χρονικής κατανομής των προτύπων (Temporal Pattern Distribution Analysis -TPDA, Spatio-Temporal Pattern Distribution Analysis -STPDA), δηλαδή η διερεύνηση του κατά πόσο, ορισμένα πρότυπα εμφανίζονται σε συγκεκριμένες χρονικές στιγμές ή συγκεκριμένες χωρικές θέσεις [217], [218]. Παράλληλα, η καθιέρωση «ορολογίας προτύπων» βοηθά στον καλύτερο σχεδιασμό των διαφόρων τεχνικών επεξεργασίας, για τη σωστή και πλήρη αντιμετώπιση των αντίστοιχων αναγκών (για όλες τις κατηγορίες σήματος). Χαρακτηριστική περίπτωση αποτελεί η διαδικασία αξιολόγησης των τεχνικών *WDWF* (§ 3.2.6.2 και 3.2.6.3), η οποία εξελίχθηκε χωριστά για τα *IB* και *RS* πρότυπα [179].

Η διαδικασία προσδιορισμού των προτύπων βασίστηκε κατά κύριο λόγο σε εμπειρικές - πειραματικές παρατηρήσεις των διαθέσιμων καταγραφών. Όπως έχει ήδη επισημανθεί, εκτός από τις πολυκαναλικές καταγραφές παρατεταμένου χρόνου, χρησιμοποιήθηκαν και άλλες δειγματοληπτικές μετρήσεις μικρότερης διάρκειας (10-30 λεπτά), όπου εκτός από υγιείς συμπεριλήφθηκαν και διάφορες παθολογικές καταστάσεις. Και οι δύο παραπάνω διευθετήσεις έγιναν για λόγους πληρότητας της πειραματικής διαδικασίας [179], [218]. Έτσι, χιλιάδες συμβάντα μικρής διάρκειας απομονώθηκαν από τις πειραματικές καταγραφές και αποτέλεσαν τα αντιπροσωπευτικά δείγματα, τα οποία «εξετάστηκαν» οπτικά και «ακουστικά» με στόχο τον εντοπισμό κοινών χαρακτηριστικών και την ένταξή τους σε ομάδες. Στο πλαίσιο αυτό, υπήρξε στενή συνεργασία ανάμεσα στις δύο ερευνητικές ομάδες (ιατρική και τεχνολογική), ενώ καθοριστική υπήρξε η προτροπή του ιατρικού επιτελείου, για τυποποίηση μικρού σχετικά αριθμού βιοακουστικών προτύπων (τουλάχιστον στα πρώτα στάδια της μελέτης). Επίσης, υπήρξε μέριμνα για την περιγραφή των σημάτων παρεμβολής - θορύβου (με βάση τη σκοπιμότητα που αναλύθηκε πιο πάνω). Για το λόγο αυτό, ο γενικότερος όρος «κοιλιακοί ήχοι» είναι προτιμητέος έναντι των ειδικότερων περιγραφών «εντερικοί - γαστρεντερικοί ήχοι», όταν αναφερόμαστε στον καθορισμό των κλάσεων και των αντίστοιχων προτύπων ταξινόμησης [61], [179], [215]-[218].

Κατά τη σύσταση των λεξικών ταξινόμησης επιχειρήθηκε η ενσωμάτωση όλων των σχετικών παρατηρήσεων και της αντίστοιχης γνώσης, που έχουν προκύψει από παλαιότερες ερευνητικές προσπάθειες. Στο πλαίσιο αυτό, οι φάσεις εφησυχασμού, ομαλής - ρυθμικής δραστηριότητας, και τυχαίων - έντονων «συσταλτικών συμπλεγμάτων» (random and intense clustered contractions) του κύκλου *MMC* που προαναφέρθηκαν, αποτελούν μια πρώτη κατεύθυνση διαχωρισμού. Επίσης, λήφθηκαν υπόψη οι παρατηρήσεις των *Cannon (1905)* [23], *Dalle et al. (1975)* [64], και *Gamer and Ehrenreich (1989)* [69], οι οποίοι διαπίστωσαν: α) την εμφάνιση σύντομων και απότομων κρουστικών σημάτων (*spiky sounds, explosive peaks, intestinal bursts, sudden quick discharges or pops*), που παρατηρούνται απομονωμένα ή σε «επαναλήψεις» (*solitary clicks / repeated multi-clicks*), β) την παρουσία «ομαλών» σημάτων μεγαλύτερης διάρκειας, όπως οι τρίζοντες και οι «μουσικοί» - συρίτονες ήχοι (*gurgling-, rumbling-, rubbing-, crepitating-noises/piping notes*), και γ) το συνδυασμό (μίξη) των παραπάνω

κατηγοριών σε τυχαίες χρονικές ακολουθίες (*combinations of short and prolonged duration oscillating impulses*) [179], [218].

Ειδικότερα, ο προσδιορισμός των διαφόρων προτύπων κοιλιακών ήχων έγινε μέσα από μία πειραματική ιεραρχική διαδικασία δύο επιπέδων, με αποτέλεσμα τη σύσταση δύο αντίστοιχων σχημάτων - λεξικών περιγραφής. Πρόκειται για τις κλάσεις «συνοπτικής - χοντρικής ταξινόμησης» (*Abdominal Sound Coarse Classification, ASCC*) και «λεπτομερούς-αναλυτικής ταξινόμησης» (*Abdominal Sound Fine Classification, ASFC*), που θεμελιώθηκαν διαδοχικά, με την ίδια χρονική σειρά που αναφέρονται. Έτσι, το λεξικό *ASFC* είναι πληρέστερο και αποτελεί επέκταση του αντίστοιχου σχήματος *ASCC*, όπου αξιοποιήθηκε η ήδη αποκτηθείσα εμπειρία, ενώ διευθετήθηκαν και όλα τα ζητήματα που δεν είχαν καλυφθεί ή δεν είχαν αντιμετωπιστεί πλήρως. Όπως έχει ήδη αναφερθεί, στο πλαίσιο των παραπάνω διαδικασιών αξιοποιήθηκαν και οι εικονολήπτες παρακολούθησης με τις αντίστοιχες τεχνικές ανίχνευσης κίνησης, κυρίως για την επαλήθευση (*validation*) των αποτελεσμάτων, σχετικά με την εμφάνιση των σφαλμάτων καταγραφής (*artifacts*) [218].

5.3 Σχήμα συνοπτικής ταξινόμησης κοιλιακών ήχων (*ASCC*)

Το σχήμα *ASCC* αποτελεί την πρώτη προσέγγιση ταξινόμησης των κοιλιακών ήχων και περιλαμβάνει το μικρότερο δυνατό αριθμό χρήσιμων προτύπων. Στην ουσία, πρόκειται για μια πιο επίσημη τυποποίηση του άτυπου διαχωρισμού των εντερικών ήχων σε κρουστικούς (*EBS*) και μη [80]-[85], καθώς και των συχνά παρατηρούμενων «προτύπων παρεμβολής», όπου εκτός των σημάτων φόντου *ABN* (περίοδοι σιγής), συμπεριλαμβάνονται οι τριβοηλεκτρικοί θόρυβοι κίνησης (κυρίως) και οι ήχοι αναπνοής - ροχαλητών (πιο σπάνια). Στο πλαίσιο αυτό, αποφασίστηκε η χρήση των μικρότερων δυνατών (σε διάρκεια) συμβάντων, τα οποία σύμφωνα με την ανάλυση της § 5.1.4, θα αποτελούσαν τα βασικά - πρωταρχικά πρότυπα, δηλαδή τα θεμελιώδη «δομικά συστατικά στοιχεία» του όλου μηχανισμού. Για την επίτευξη του αυστηρού αυτού διαχωρισμού (οριοθέτησης) των συμβάντων, αποφασίστηκε η χρήση της τεχνικής εντοπισμού *SBS*. Το μικρό και ευέλικτο σχήμα ταξινόμησης *ASCC* ενσωματώνει πλήρως τις παρατηρήσεις παλαιότερων ερευνών [23], [64], [69] και κρίθηκε υπεραρκετό για την κάλυψη των λειτουργικών αναγκών, κατά την πρώτη φάση του εγχειρήματος ανάλυσης προτύπων [179], [215]-[216], [218].

5.3.1 Κανόνες περιγραφής των προτύπων συνοπτικής ταξινόμησης *ASCC*

Η παραπάνω διαδικασία οδήγησε στον προσδιορισμό πέντε διαφορετικών κλάσεων, δύο από τις οποίες αντιστοιχούν σε πρότυπα χρήσιμων κοιλιακών ήχων [215]-[216], [218]:

α) **Silent Periods -SP**: πρότυπα συστατικού εφησυχασμού, δηλαδή περιοχές ησυχίας που περιλαμβάνουν μόνο το θόρυβο φόντου και έχουν μικρό ενεργειακό περιεχόμενο. Τα στιγμιότυπα αυτά, αν και έχουν παρόμοια χαρακτηριστικά με το θόρυβο ευρείας ζώνης *ABN*, ενδέχεται να διαφέρουν από το επιλεγμένο αποτόπωμα θορύβου n_{FP} (γι' αυτό και καταχωρούνται - εντοπίζονται ως σημαντικά συμβάντα). Η διαφοροποίησή τους αυτή, έγκειται στις πολύ μικρές κινήσεις των υποκειμένων και των καλωδίων σύνδεσης, σε ακουστικούς θορύβους περιβάλλοντος χαμηλής στάθμης, σε τυχαίες μεταβολές του *ABN*-προφίλ εξαιτίας θερμικών, ηλεκτροστατικών και ηλεκτρομαγνητικών παρεμβολών, ή ακόμη και σε πολύ χαμηλής στάθμης βιοακουστικές συνιστώσες. Ο εντοπισμός τους (ως σημαντικά συμβάντα) συνδέεται κυρίως με τη διαδικασία ανίχνευσης *SBS*, ενώ είναι πιο σπάνιος στη μέθοδο *LT-WDSS* (*MRS* σάρωση), εκτός από τις περιπτώσεις των μη σημαντικών καναλιών.

β) **Sliding Noises – SN**: πρόκειται για θορύβους τριβής που παράγονται από κινήσεις των υποκειμένων ή/και έντονες μετατοπίσεις των αισθητήρων λήψης. Τα σήματα αυτά έχουν έντονες διακυμάνσεις ως προς το ενεργειακό και φασματικό τους περιεχόμενο, στοιχείο που συνδέεται με το βαθμό κίνησης των εξεταζομένων. Το άκουσμά τους παραπέμπει σε σήματα θορύβου, που βέβαια δεν έχουν καμία σχέση με το θόρυβο φόντου *ABN*. Η εισαγωγή της προσαρμοστικής διάταξης *SHIELD* περιόρισε την εμφάνιση αυτών των περιστατικών, ειδικότερα όσων συνδέονται με μικρές κινήσεις. Ωστόσο, τα συμβάντα *SN* εξακολουθούν να παράγονται, ιδιαίτερα κατά τις ανεξέλεγκτες κινήσεις των υποκειμένων σε φάση νυχτερινού ύπνου. Οι παραπάνω «θόρυβοι κίνησης» χαρακτηρίζονται συνήθως από μεγάλη διάρκεια (μερικά δευτερόλεπτα) και αντίστοιχα μεγάλο εύρος δυναμικής περιοχής, που συχνά οδηγεί σε ψαλιδισμούς και έντονες αρμονικές παραμορφώσεις. Από την ενσωμάτωση του συστήματος οπτικοακουστικής παρακολούθησης κατέστη δυνατή η στατιστική ανάλυση εμφάνισης των *SN*-περιστατικών, όπου διαπιστώθηκε η παρουσία τους σε ένα ποσοστό 15,2% (ως προς το συνολικό χρόνο παρακολούθησης) και 11,7% (ως προς το σύνολο των ανιχνευμένων συμβάντων), ενώ παρατηρείται παράλληλη εμφάνιση χρήσιμων κοιλιακών ήχων με θορύβους *SN* σε ένα ποσοστό 0,7% και 5,4%, αντίστοιχα*. Η χρήση της διάταξης *SHIELD* έφερε μείωση των παραπάνω ποσοστών, της τάξης 8-11%, περίπου. Στο πλαίσιο αυτό, υπήρξε σημαντικός προβληματισμός για το αν αρκεί ο εντοπισμός και η απομόνωση των σημάτων *SN*, ή εάν θα έπρεπε να αναπτυχθούν πιο προσαρμοστικές λύσεις καταστολής, για την περίπτωση παράλληλης εμφάνισής τους με χρήσιμες πληροφοριακές συνιστώσες. Τελικά επικράτησε η πρώτη άποψη (τουλάχιστον για την παρούσα φάση), καθώς α) τα ποσοστά συνόπαρξης είναι αρκετά μικρά, β) υπάρχουν σημαντικές αμφιβολίες ως προς τις πιθανότητες επιτυχούς αντιμετώπισης αυτού του προβλήματος (εξαιτίας της φύσης των σημάτων *SN*), και γ) είναι άγνωστο εάν και κατά πόσο οι κινήσεις των υποκειμένων επηρεάζουν την παρουσία κοιλιακών ήχων (πολλοί ιατροί επισήμαναν το ενδεχόμενο παραγωγής βιοακουστικών σημάτων, εξαιτίας παφλασμών του πεπτικού περιεχομένου, λόγω κίνησης). Μια ενδιάμεση λύση αντιμετώπισης δίνεται σε επόμενη ενότητα, κατά την παρουσίαση του συστήματος αναγνώρισης προτύπων *ASFC*.

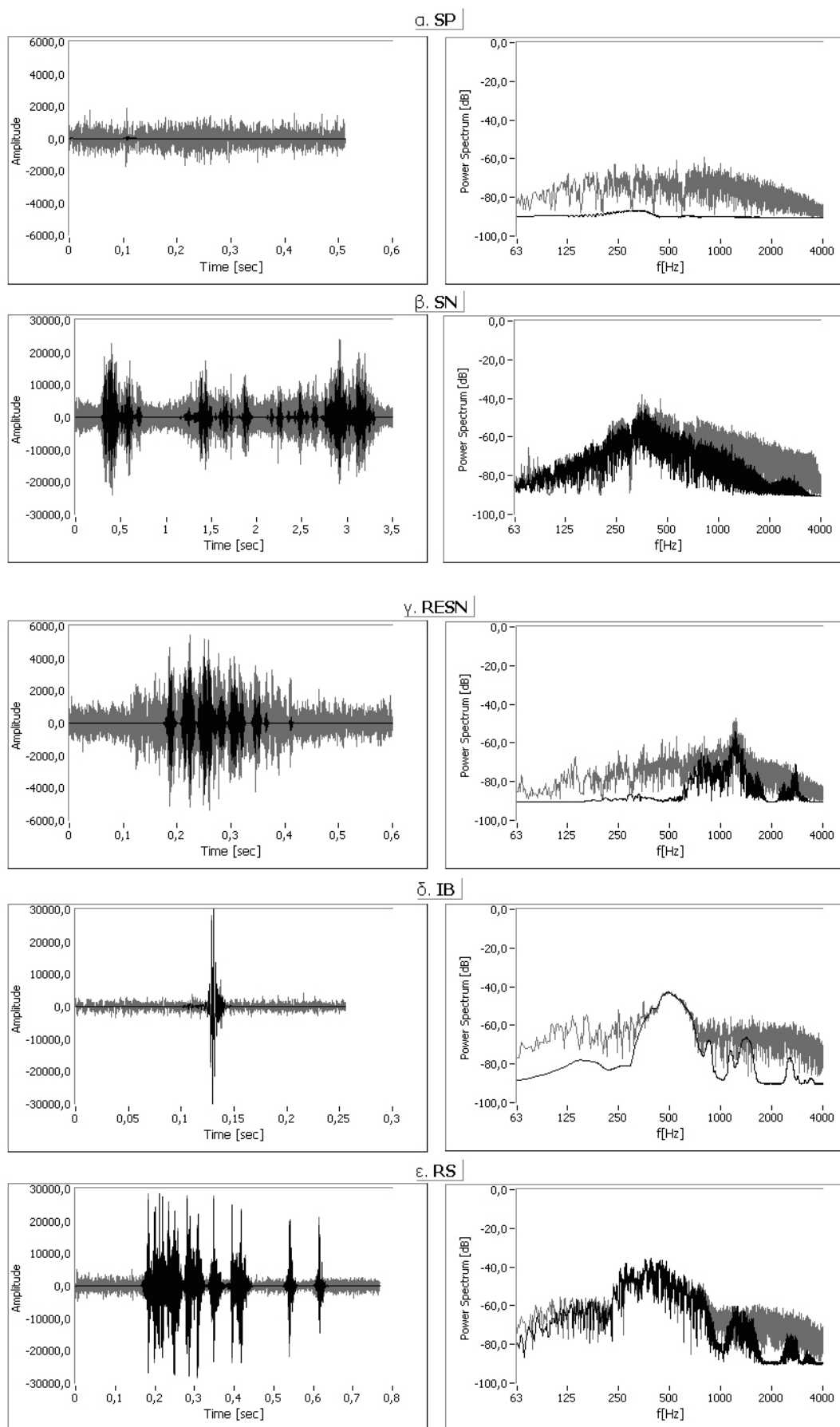
γ) **Respiration and Snoring – RESN**: πρόκειται για ήχους αναπνοής ή/και ροχαλητών που λαμβάνονται στην κοιλιακή περιοχή των υποκειμένων. Το πιο πιθανό είναι ότι οι ήχοι αυτοί οδεύουν μέσω του αέρα (αερόφερτοι θόρυβοι) και λαμβάνονται από τους αισθητήρες στο εξωτερικό μέρος των κοιλιακών τοιχωμάτων. Κατά τη μετάδοσή τους στο εσωτερικό του ανθρώπινου σώματος υπάρχουν πιο μεγάλες απώλειες λόγω της ηχητικής απορρόφησης και των μεγαλύτερων αποστάσεων. Στο ίδιο συμπέρασμα συγκλίνουν και οι οπτικοακουστικές παρατηρήσεις, όπου διαπιστώθηκε ότι η εφαρμογή της διάταξης *SHIELD* ελαχιστοποιεί την εμφάνισή τους. Αξίζει να σημειωθεί ότι τα περιστατικά *RESN* είναι αρκετά σπάνια (δεν εμφανίζονται σε όλες τις καταγραφές), ενώ η παρουσία τους συνδέεται με: α) δυσκολίες αναπνοής ή/και επεισόδια άπνοιας, β) την παρουσία έντονου ροχαλητού, γ) περιοδικές μετατοπίσεις των αισθητήρων από τις θωρακικές - αναπνευστικές κινήσεις. Το τελευταίο οφείλεται κυρίως σε προβλήματα στερέωσης των αισθητήρων και αντιμετωπίστηκε πλήρως με την εισαγωγή της διάταξης *SHIELD*. Οι ήχοι *RESN* έχουν χαμηλή ενέργεια (οριακά μεγαλύτερη από τα πρότυπα *SP*), ενώ το φασματικό τους περιεχόμενο εντοπίζεται σε στενές, σχετικά, περιοχές των μεσαίων συχνοτήτων (γύρω στο 1kHz). Η διάρκειά τους κυμαίνεται από 150 μέχρι 500 msec, περίπου. Χαρακτηριστική είναι η συμμετρία των φασματικών περιοχών ανόδου - καθόδου (symmetrical roll-on / roll-off frequency slopes) γύρω από το «φασματικό κέντρο βάρους» (spectral centroid), καθώς και οι επίσης συμμετρικοί φάκελοι «επίθεσης» και «απόσβεσης» (attack - release envelopes).

* Τα αποτελέσματα αυτά προέκυψαν από μια καταγραφή «νυχτερινού ύπνου» διάρκειας 3 ωρών με μεγάλο αριθμό συμβάντων *SN* (χωρίς τη διάταξη *SHIELD*), υπό την έννοια του χειρόστου σεναρίου ανάλυσης (worst case scenario).

δ) **Intestinal Bursts -IB**: εντερικοί ήχοι μικρής διάρκειας με έντονο κρουστικό χαρακτήρα. Τα σήματα αυτά περιλαμβάνουν έναν κύριο «ωστικό παλμό», εκατέρωθεν του οποίου συχνά εμφανίζονται συνοδευτικές «εξάρσεις» (peaks) μικρότερης στάθμης (pre-echo / post-echo pulses). Οι παλμοί παρουσιάζουν, συνήθως, τονικά χαρακτηριστικά (αντίστοιχα με αυτά των εκθετικά αποσβενδόμενων ημίτονων), οπότε έχουν στενοζωνικό φάσμα, ενώ άλλες φορές έχουν πιο μουντό - ουδέτερο «άκουσμα», οπότε εκτείνονται σε μεγαλύτερο εύρος φάσματος (bandwidth -BW). Έτσι, οι βασικές φασματικές συνιστώσες κυμαίνονται στην περιοχή από 200Hz έως 2,5kHz (περίπου), με ποιο συνηθισμένες τις συχνότητες των 500Hz. Το εύρος φάσματος μεταβάλλεται αντίστοιχα από πολύ χαμηλές έως πολύ μεγάλες τιμές (περίπου 45Hz - 1,2kHz), ενώ το μέσο εύρος BW διαμορφώνεται γύρω στα 200 Hz. Οι διακυμάνσεις πλάτους καλύπτουν όλη τη διαθέσιμη δυναμική περιοχή ($\approx 90\text{dB}/16\text{bit}$), ανάλογα με την ένταση των συστατικών κινήσεων πέψης και την απόσταση από το σημείο διέγερσης. Το κύριο χαρακτηριστικό των προτύπων IB είναι η πολύ μικρή διάρκειά τους (από 30 έως 300 msec περίπου, με μέση τιμή 120 msec), καθώς και οι απότομες μεταβάσεις (transients) των κρουστικών παλμών ($LCF > 15\text{ dB}$). Εμφανίζονται πάρα πολύ συχνά, είτε ως απομονωμένα συμβάντα είτε ως ακολουθίες πολλαπλών «εξάρσεων». Η γένεσή τους συνδέεται με σύντομες συσπάσεις, καθώς και παφλασμούς του πεπτικού χυμού στα εντερικά τοιχώματα.

ε) **Regularly Sustained- RS**: παραμένοντες εντερικοί ήχοι, που έχουν μεγαλύτερη διάρκεια και χαρακτηρίζονται από λιγότερο απότομες καμπύλες «έναρξης» - «απόσβεσης» (attack / release slopes). Στην κατηγορία αυτή ανήκουν πρωτίστως οι τρίζοντες ήχοι, οι οποίοι αποτελούν συνδυασμούς πολλαπλών συσπάσεων (clustered contractions) σαν να πρόκειται για χρονικά-συμπιεσμένους (συμπυκνωμένους) παλμούς IB. Περιλαμβάνονται, επίσης οι «μουσικοί» - συρίτονες ήχοι (πιο σπάνια), καθώς και τυχαίοι συνδυασμοί των παραπάνω, όπου συχνά παρεμβάλλονται κρουστικοί παλμοί. Γενικά, η διάρκειά τους κυμαίνεται από 300msec έως και 3sec, ενώ η πλειονότητα των περιστατικών RS εμφανίζει ένα μέσο μήκος 500msec - 1sec. Για τα ενεργειακά τους χαρακτηριστικά ισχύει ότι και στην περίπτωση των προτύπων IB, δηλαδή υπάρχουν και πάλι μεγάλες διακυμάνσεις πλάτους που καλύπτουν όλο το εύρος δυναμικής περιοχής, αν και τα σήματα RS είναι συνήθως πιο έντονα. Όσον αφορά το φασματικό τους περιεχόμενο, υπάρχουν αρκετές διαφοροποιήσεις, ανάλογα με το είδος των σημάτων και τα ακουστικά χαρακτηριστικά που προαναφέρθηκαν. Σε γενικές γραμμές, η κύρια δραστηριότητά τους εντοπίζεται στις περιοχές 450 - 500 Hz, ενώ πιο σπάνια εκτείνονται από 300 έως 1.700 Hz. Το φασματικό τους εύρος κυμαίνεται από πολύ μικρές τιμές των 30Hz (για τα στενοζωνικά τονικά σήματα) μέχρι και τα 800Hz περίπου (για τους τρίζοντες ήχους και τους συνδυασμούς «τριγμών» με κρουστικούς παλμούς), ενώ το μέσο εύρος BW διαμορφώνεται στα 170Hz, περίπου. Τα συμβάντα RS καταλαμβάνουν μικρότερο ποσοστό στη διάρκεια των καταγραφών παρατεταμένου χρόνου. Η παρουσία τους συνδέεται με πιο «οργανωμένες» συσπάσεις (έντονες - ρυθμικές), οι οποίες δείχνουν να σχετίζονται περισσότερο με τις φάσεις II και III του κύκλου MMC.

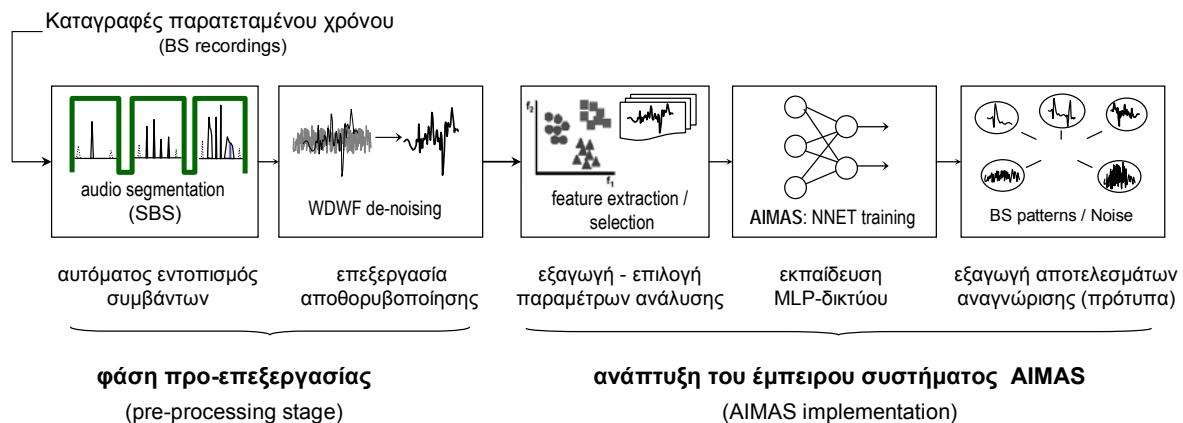
Τα παραπάνω χαρακτηριστικά αποδίδονται με γραφικό τρόπο στο σχ. 5-3, όπου απεικονίζονται αντιπροσωπευτικά πρότυπα των ASCC κλάσεων στο πεδίο του χρόνου και της συχνότητας. Στα διαγράμματα αυτά δίνονται τόσο τα ενθόρυβα σήματα (γκρι χρώμα) όσο και οι αποθορυβοποιημένες τους εκδοχές (μαύρο χρώμα), έτσι ώστε να παρουσιαστούν τα μορφολογικά τους χαρακτηριστικά με τον καλύτερο δυνατό τρόπο. Για τον ίδιο λόγο, οι γραφικές παραστάσεις των προτύπων SP και RESN στο πεδίο του χρόνου έχουν μεγεθυνθεί (auto-scaling), καθώς τα αντίστοιχα σήματα έχουν σχετικά μικρό πλάτος. Τα διαγράμματα φασματικής στάθμης είναι κανονικοποιημένα, όπου η μέγιστη ηχητική στάθμη αντιστοιχεί σε 0 dB [215]-[216], [218].



Σχήμα 5-3. Χαρακτηριστικά «ASCC - πρότυπα» στο πεδίο του χρόνου (αριστερά) και της συχνότητας (δεξιά): με γκρι χρώμα απεικονίζονται τα ενθόρυβα σήματα $x(i)$, $X(k)$ ενώ με μαύρο τα αποθορυβωποιημένα $s^-(i)$, $S^-(k)$.

5.3.2 Αυτόματη αναγνώριση των προτύπων συνοπτικής ταξινόμησης ASCC

Για τις ανάγκες αυτόματης αναγνώρισης των ASCC προτύπων αναζητήθηκαν λύσεις στα νευρωνικά δίκτυα και τις μεθόδους εκπαίδευσης με εποπτεία, που θεωρούνται ιδιαίτερα αποδοτικά, ειδικά όταν οι κλάσεις ταξινόμησης είναι προκαθορισμένες [218], [266]-[267]. Παράλληλα, εξετάστηκε το ενδεχόμενο συνδυαστικής χρήσης νευρωνικών δικτύων και συστημάτων ασαφούς λογικής («νεύρο-ασαφείς» προσεγγίσεις, neural fuzzy), τα οποία επιδεικνύουν αξιοσημείωτη απόδοση σε τέτοιου είδους εφαρμογές [277]-[278]. Ωστόσο, προτιμήθηκαν πιο κλασικές προσεγγίσεις (σε πρώτη φάση), όπως είναι τα πολυεπίπεδα νευρωνικά δίκτυα προώθησης προς τα εμπρός, και η αντίστοιχη εκπαίδευση «πίσω διάδοσης». Στο πλαίσιο αυτό, αξιολογήθηκαν διάφοροι παράμετροι ανάλυσης, ενώ δοκιμάστηκαν πολλές εναλλακτικές τοπολογίες MLP. Στο σχ. 5-4 παρουσιάζονται οι φάσεις ανάπτυξης του δικτύου AIMAS (Autonomous Intestinal Motility Analysis System), όπως ονομάστηκε το έμπειρο σύστημα αναγνώρισης ASCC προτύπων [215]-[216], [218].



Σχήμα 5-4. Οι φάσεις ανάπτυξης του έμπειρου συστήματος AIMAS (αναγνώριση προτύπων ASCC).

Η ανάπτυξη του συστήματος AIMAS έγινε σε κατάλληλο υπολογιστικό περιβάλλον, με τη βοήθεια του λογισμικού πακέτου γραφικού προγραμματισμού *National Instruments LabVIEW* (κατά τη φάση προ-επεξεργασίας) και του λογισμικού πακέτου *Mathworks Matlab 7.0* (κατά τη φάση εκπαίδευσης του δικτύου) [215]-[216], [218].

5.3.2.1 Τεχνικές προ-επεξεργασίας (pre-processing)

Όπως προκύπτει από το παραπάνω σχήμα, κατά τη διαδικασία ανάπτυξης του δικτύου AIMAS, αξιοποιήθηκαν οι διαθέσιμες πειραματικές καταγραφές, ενώ εφαρμόστηκαν και δύο τεχνικές προ-επεξεργασίας: α) η μέθοδος σάρωσης SBS και β) η αποθορυβοποίηση WDWF. Η διαδικασία SBS χρησιμοποιήθηκε με στόχο την εξασφάλιση μεγάλης διακριτικής ικανότητας «τμηματοποίησης» και το διαχωρισμό πολύ κοντινών «τμημάτων» σήματος. Τα «τμήματα» (segments) ορίζονται ως πιο πρωταρχικές δομές από τα «συμβάντα» (events), και αποτελούν συστατικά στοιχεία των τελευταίων (δηλαδή κάθε συμβάν μπορεί να σχηματίζεται από ένα ή περισσότερα τμήματα). Υπό την έννοια αυτή, και δεδομένου ότι η ταξινόμηση ASCC αναφέρεται στα πρωταρχικά πρότυπα (αφορά τις βασικές δομικές μονάδες των τμημάτων), η τμηματοποίηση SBS προτιμήθηκε έναντι των αντιληπτικών κριτηρίων της LT-WDSS [218]. Γενικά, η διαδικασία εντοπισμού μέσω απλών κανόνων ενεργειακής σύγκρισης (όπως συμβαίνει στη μέθοδο SBS) είναι αρκετά ριψοκίνδυνη, και μπορεί να οδηγήσει αφενός στο λανθασμένο εντοπισμό περιόδων σιγής (*False Acceptance -FA*), αφετέρου στη λανθασμένη

απόρριψη τμημάτων σήματος (*False Rejection –FR*) [218]. Έτσι, η εφαρμογή του συστήματος *AIMAS* παρέχει ένα δεύτερο επίπεδο πιο ασφαλούς και έγκυρου διαχωρισμού των περιοχών δραστηριότητας από τις στιγμές εφησυχασμού. Στο πλαίσιο αυτό επιλέχθηκαν παράθυρα μικρού μήκους ($N_W=2.048$ και 4.096 δείγματα) και σχετικά χαμηλά ενεργειακά επίπεδα σύγκρισης [215]-[216], [218].

Το αποτέλεσμα της παραπάνω διαδικασίας είναι ο σχηματισμός των «απομονωμένων» τμημάτων $x_{nq}(i)$, όπου ο δείκτης nq δηλώνει τη χρήση τους ως δείγματα εκπαίδευσης (κατά αντιστοιχία με την § 5.1.3), ενώ ο όρος “ x ” επισημαίνει την παρουσία προσθετικού θορύβου. Αν και οι πρώτες προσπάθειες ταξινόμησης έγιναν στα σήματα $x_{nq}(i)$, γρήγορα διαπιστώθηκε ότι η παρουσία του θορύβου *ABN* δυσχεραίνει τη διαδικασία αναγνώρισης. Τα προβλήματα γίνονται πιο έντονα στις περιπτώσεις που τα χαρακτηριστικά του θορύβου μεταβάλλονται μεταξύ των καταγραφών διαφορετικών υποκειμένων [215]-[216]. Γι’ αυτό, αποφασίστηκε η χρήση των τεχνικών *WDWF*, ενώ τα αποτελέσματα αποθορυβοποίησης αξιοποιήθηκαν και για την εξαγωγή αντίστοιχων παραμέτρων ανάλυσης (περισσότερες λεπτομέρειες δίνονται στη συνέχεια). Έτσι, από την εφαρμογή των δύο παραπάνω τεχνικών προ-επεξεργασίας (*SBS* και *WDWF*) προκύπτουν τα αποθορυβοποιημένα τμήματα σήματος $s_{nq}(i)$ και οι αντίστοιχες κυματιδιακές ακολουθίες $S_{nq}(k,w)$, που στο εξής θα συμβολίζονται $s(i)$ και $S_k(w)$, για λόγους απλότητας [215]-[216], [218].

5.3.2.2 Εξαγωγή χαρακτηριστικών παραμέτρων ανάλυσης (feature extraction)

Όπως έχει ήδη επισημανθεί, η ταυτοποίηση των *ASCC* προτύπων μπορεί να γίνει με τη βοήθεια ενεργειακών και φασματικών παραμέτρων, καθώς επίσης από τη χρονική διάρκεια των σημάτων και τα μορφολογικά στοιχεία των αντίστοιχων κυματομορφών. Στο πλαίσιο αυτό, δοκιμάστηκαν κλασσικές παράμετροι ταξινόμησης ηχητικών σημάτων [228]-[233], [275]-[276], ενώ συστάθηκαν και νέα χαρακτηριστικά μεγέθη, ειδικά για τις ανάγκες της παρούσης εφαρμογής [215]-[216], [218]. Η αξιολόγηση και τελική επιλογή των παραμέτρων, ως προς την καταλληλότητά τους για το διαχωρισμό των *ASCC* κλάσεων, έγινε με βάση τους περιορισμούς της § 5.1.2 (σχέσεις 5-7 έως 5-9), και κυρίως βάσει εμπειρικών παρατηρήσεων και πολλαπλών δοκιμών τύπου “*trial and error*” [215]-[216], [218].

A) Παράμετροι ανάλυσης στο πεδίο του χρόνου (time-domain features, TDF)

Μια απλή ενεργειακή παράμετρος που καθορίζει την ένταση των ακουστικών σημάτων είναι η μέση ισχύς P_{av} , η οποία χρησιμοποιήθηκε και στην εξαγωγή των φακέλων σύνοψης ενέργειας – χρόνου (§ 4.2.3.5, εξ. 4-20). Έτσι, αν N είναι το μήκος του σήματος $s(i)$, τότε:

$$P_{av}(s) = 10 \cdot \log_{10} \left\{ \frac{1}{N} \sum_{i=0}^{N-1} s^2(i) \right\} \quad (5-14)$$

Ένα συμπληρωματικό μέγεθος είναι η διασπορά των μεταβολών ισχύος ως προς το χρόνο “ $var\{P_{av}(ne)\}$ ”, που υπολογίζεται από την εφαρμογή της σχέσης 5-14 σε εσωτερικά παράθυρα επιλογής (ne). Έτσι, εκτιμάται εάν η ενέργεια του σήματος είναι εντοπισμένη σε συγκεκριμένες χρονικές στιγμές (περίπτωση προτύπων *IB*), ή αν «μοιράζεται» σε όλο το μήκος του χρόνου παρατήρησης (περίπτωση προτύπων *RS*). Εναλλακτικά, είναι δυνατόν να χρησιμοποιηθεί το εύρος δυναμικής περιοχής (*Dynamic Range –DR*), που σε συνδυασμό με την παράμετρο P_{av} παρέχει όλες τις απαραίτητες πληροφορίες, σχετικά με τις ενεργειακές

διακομάνσεις του σήματος. Η τελευταία προσέγγιση προτιμήθηκε για λειτουργικούς κυρίως λόγους, όπως είναι ο ευκολότερος υπολογισμός της παραμέτρου DR [215]- [216], [218]:

$$DR = 20 \cdot \log_{10} \left\{ \left\| \max [s(i)] - \min [s(i)] \right\| \right\} \quad (5-15)$$

Εκτός των παραπάνω, δοκιμάστηκαν πρόσθετες παράμετροι ανάλυσης TDF , οι οποίες εξήχθησαν από τις καμπύλες περιβάλλουσας (envelopes) των σημάτων. Για το σκοπό αυτό έγινε χρήση εσωτερικών, μη επικαλυπτόμενων ορθογωνικών παραθύρων (ne) μήκους N_{ENV} , από τη μέση ισχύ των οποίων προκύπτουν οι επιθυμητές καμπύλες φακέλων $S_{ENV}(ne)$:

$$S_{ENV}(ne) = \frac{\sum_{i=ne}^{ne+N_{ENV}-1} s^2(i)}{N_{ENV}}, \quad ne = [0, NE-1], \quad NE = \frac{N}{N_{ENV}} \quad (5-16)$$

Έτσι, συστάθηκαν τρεις νέοι παράμετροι ανάλυσης: α) ο «δείκτης εφησυχασμού» SP_{index} (Silence period index), β) η «μέτρηση των ενεργειακών μεταβάσεων μεταξύ των σημείων αιχμής» PTM (Peak Transition Measure), και γ) ο «δείκτης αιχμής» NP_{index} ("number of peaks" index) [215]-[216], [218]. Ο δείκτης SP_{index} προκύπτει από το λόγο των «σημαντικών» δειγμάτων (δηλαδή εκείνων που ξεπερνούν μια τιμή αναφοράς th_{ENV}) προς το συνολικό μήκος της περιβάλλουσας, και μεταβάλλεται στην περιοχή $[0,1]$:

$$SP_{index} = \frac{\sum_{ne=0}^{NE-1} THR_{[S_{ENV}, th_{ENV}]}(ne)}{NE}, \quad THR_{[S_{ENV}, th_{ENV}]}(ne) = \begin{cases} 1, & S_{ENV}(ne) \geq th_{ENV} \\ 0, & \text{διαφορετικά} \end{cases} \quad (5-17)$$

όπου η τιμή του κατωφλίου σύγκρισης th_{ENV} προσδιορίζεται συνήθως από το διαθέσιμο αποτόπωμα θορύβου, ενώ ο τελεστής THR αντιστοιχεί στη συνάρτηση σκληρής κατωφλίωσης (εξ. 3-27α). Για τον υπολογισμό των NP_{index} και PTM χρησιμοποιούνται οι ακολουθίες $\{PVS_{ENV}, tPVS_{ENV}\}$, οι οποίες καθορίζουν τις θέσεις των σημείων αιχμής στους καρτεσιανούς άξονες $\{x, y\}$. Έτσι, ο δείκτης NP_{index} προκύπτει από το συνολικό αριθμό των κορυφώσεων NP (number of peaks) προς το συνολικό μήκος NE των φακέλων S_{ENV} , και μεταβάλλεται επίσης στο διάστημα $[0,1]$. Η παράμετρος PTM προκύπτει από το άθροισμα των ρυθμών μετάβασης (συντελεστές κλίσης) ανάμεσα σε όλες τις διαδοχικές κορυφές (*summarization of all peaks' alteration in a "gradient-sense"* [218]):

$$\left\{ \begin{array}{l} PVS_{ENV}(np) \\ tPVS_{ENV}(np) \end{array} \right\}_{np \in [0, NP-1]} = \left\{ \begin{array}{l} S_{ENV}(ne), \forall S_{ENV}(ne-1) < S_{ENV}(ne) > S_{ENV}(ne+1) \\ ne \end{array} \right\} \quad (5-18)$$

$$NP_{index} = \frac{NP}{NE}, \quad PTM = \sum_{np=0}^{NP-2} \frac{\|PVS_{ENV}(np+1) - PVS_{ENV}(np)\|}{tPVS_{ENV}(np+1) - tPVS_{ENV}(np)}$$

Μικρές τιμές του δείκτη SP_{index} (κοντά στο μηδέν) συνδέονται με μεγάλες περιόδους εφησυχασμού (πρότυπα SP , $RESN$ και IB). Αντίθετα, τιμές κοντά στη μονάδα αντιστοιχούν σε μεγαλύτερα τμήματα δραστηριότητας (πρότυπα RS , SN). Είναι φανερό ότι στην περίπτωση πολλών κορυφώσεων (χαρακτηριστικά προτύπων RS , SN) οι τιμές του δείκτη NP_{index} πλησιάζουν στη μονάδα, ενώ διαφορετικά τείνουν προς το μηδέν (περιπτώσεις

προτύπων SP , $RESN$, IB). Ανάλογα, όταν υπάρχουν πολλά σημεία αιχμής με μικρές μεταξύ τους αποστάσεις (απότομες μεταβάσεις), τότε η παράμετρος PTM παίρνει μεγάλες τιμές (πρότυπα RS , SN), σε αντίθεση με την περίπτωση λίγων και απομακρυσμένων κορυφώσεων (πρότυπα SP , $RESN$, IB). Δεδομένου ότι τα IB σήματα έχουν (συνήθως) μεγαλύτερες στάθμες από τα επεισόδια $RESN$, τα οποία με τη σειρά τους είναι πιο έντονα από τις περιόδους σιγής (SP), αναμένεται μία κλιμάκωση στις τιμές του δείκτη PTM , η οποία εκφράζεται από την ανισότητα: “ $PTM_{SP} < PTM_{RESN} < PTM_{IB} < PTM_{RS} < PTM_{SN}$ ” [215]-[216].

Μια άλλη κλασική παράμετρος, που χρησιμοποιείται πολύ συχνά για την εκτίμηση της κρουστικότητας των σημάτων, είναι ο «παράγοντας κλίσης» (Crest Factor) [229]-[230]. Στην προκείμενη περίπτωση συστάθηκε ο «αντίστροφος δείκτης» ECF^{-1} (Inverse Envelope Crest Factor), ο οποίος ισούται με το λόγο της μέσης προς τη μέγιστη τιμή του φακέλου S_{ENV} :

$$ECF^{-1} = \frac{\text{mean}(S_{ENV})}{\text{max}(S_{ENV})} \quad (5-19)$$

Είναι προφανές ότι ο όρος ECF^{-1} μεταβάλλεται μεταξύ 0 και 1. Έτσι, στην περίπτωση ομαλών κυματομορφών οι τιμές του πλησιάζουν στη μονάδα (πρότυπα RS , SP , $RESN$), ενώ τείνουν προς το μηδέν όταν εμφανίζονται απότομες κορυφώσεις (πρότυπα IB , SN) [215]-[216], [218].

Οι παραπάνω παράμετροι ανάλυσης χρησιμοποιήθηκαν σε συνδυασμό με φασματικά μεγέθη (περιγράφονται στη συνέχεια) σε διάφορα δίκτυα MLP [215]-[216]. Τα αποτελέσματα ήταν αρκετά ενθαρρυντικά, χωρίς ωστόσο να εξασφαλίζουν την επιθυμητή αποδοτικότητα αναγνώρισης προτύπων (παρατηρήθηκαν ποσοστά επιτυχούς αναγνώρισης 80-85%). Έτσι, αποφασίστηκε η τροποποίηση των παραμέτρων SP_{index} , NP_{index} και PTM , για τους λόγους που αναλύονται ακολούθως. Υπάρχουν σήματα, που ενώ έχουν εντελώς διαφορετικές διάρκειες δραστηριότητας – εφησυχασμού, οι διαφοροποιήσεις τους αυτές αμβλύνονται κατά την εφαρμογή του δείκτη SP_{index} και τον υπολογισμό των «ποσοστών σιγής» [218]. Μάλιστα, τα αποτελέσματα αυτά είναι αρκετά ευαίσθητα σε μεταβολές του μήκους των SBS -παραθύρων, γεγονός που περιορίζει τις δυνατότητες «γενίκευσης» του δικτύου $AIMAS$. Ανάλογα σχόλια μπορούν να διατυπωθούν και για το δείκτη NP_{index} , ο οποίος κατά αντιστοιχία με τον SP_{index} εκφράζει το «ποσοστό» των σημείων αιχμής. Επιπρόσθετα, η παράμετρος PTM επηρεάζεται σε μεγάλο βαθμό από το πλάτος των σημάτων, προκαλώντας αρκετές δυσκολίες ταξινόμησης, που απορρέουν από τη μεγάλη δυναμική περιοχή των κοιλιακών ήχων [218].

Για την αντιμετώπιση των παραπάνω δυσκολιών προτάθηκαν τα ακόλουθα: α) η χρήση του συνολικού αριθμού των κορυφώσεων NP στη θέση του δείκτη NP_{index} , β) η σύσταση μιας νέας κανονικοποιημένης έκφρασης $nPTM$, εκτός από την αρχική παράμετρο PTM , και γ) η εισαγωγή μιας πρόσθετης μεταβλητής που εκφράζει το συνολικό αριθμό των σημαντικών δειγμάτων n_{SS} [218]. Η νέα παράμετρος $nPTM$ υπολογίζεται και πάλι από την εξ. (5-18), με τη διαφορά ότι αυτή τη φορά χρησιμοποιείται μια κανονικοποιημένη εκδοχή του σήματος περιβάλλουσας ($S_{ENV;norm} \in [0,1]$). Η δε εκτίμηση των σημαντικών δειγμάτων n_{SS} γίνεται με τη βοήθεια του μήκους N και του παράγοντα $(1-SP_{index})$: [218]

$$n_{SS} = N \cdot (1 - SP_{index}) \quad (5-20)$$

B) Παράμετροι ανάλυσης στο πεδίο της συχνότητας (frequency-domain features, FDF)

Εκτός των παραπάνω, δοκιμάστηκαν και τελεστές ανάλυσης στο πεδίο της συχνότητας, για την εκτίμηση των συνολικών φασματικών χαρακτηριστικών των σημάτων (stationary spectral analysis [218]). Από τις δοκιμές που έγιναν, τελικά επιλέχθηκαν οι εξής παράμετροι:

α) το φασματικό κέντρο βάρους B (spectral centroid or Brightness) [229]-[230], β) η μέτρηση της φασματικής «επιπεδότητας» SFM (spectral flatness measure) [201], [229]-[230], που ισούται με το λόγο του γεωμετρικού προς τον αριθμητικό μέσο φασματικής ισχύος, γ) το κεντροειδές ισχύος CB_C της ανάλυσης κρισίμων ζωνών CB (σχ. 3-16, 3-17: $WP-17 \cong CB-17$) [218], και δ) η φασματική διασπορά ισχύος CB_V των κρισίμων ζωνών CB :

$$B = df \cdot \frac{\sum_{k=1}^{N/2} \|k \cdot X(k \cdot df)\|}{\sum_{k=1}^{N/2} \|X(k \cdot df)\|} \quad (5-21)$$

$$SFM = 10 \cdot \log_{10} \left\{ \frac{\left[\prod_{i=1}^{N/2} X^2(i \cdot df) \right]^{1/N/2}}{\frac{1}{N/2} \sum_{i=1}^{N/2} X^2(i \cdot df)} \right\} \quad [\text{dB}] \quad (5-22)$$

$$CB_C = \frac{\sum_{k=1}^{17} k \cdot P_{av;k}}{\sum_{k=1}^{17} P_{av;k}}, \quad CB_V = \text{var}\{P_{av;k}\} \quad (5-23)$$

όπου $X(i \cdot df)$ είναι η “ i -” φασματική συνιστώσα στην έξοδο του αλγόριθμο FFT (N σημείων), df το διάστημα δειγματοληψίας (f_s/N), k ο δείκτης κυματιδιακών ζωνών $WP-17$ και $P_{av;k}$ η μέση ισχύς κάθε ζώνης (εφαρμογή της εξ. 5-14: “ $P_{av;k} = P_{av}\{S_k(w)\}$ ”) [215]-[216], [218].

Γ) Κυματιδιακοί παράμετροι ανάλυσης (wavelet-domain features, WDF)

Οι κυματιδιακοί παράμετροι συστάθηκαν με στόχο τη συνδυασμένη ανάλυση χρόνου - συχνότητας (non stationary spectral analysis [218]). Έτσι, αξιοποιήθηκαν παράλληλα με τα χαρακτηριστικά μεγέθη TDF , FDF που προαναφέρθηκαν, αλλά και σε αμιγή νευρωνικά δίκτυα κυματιδίων (wavelet neural networks, $WNNET$) [218], [279], όπου χρησιμοποιούνται αποκλειστικά και μόνο κυματιδιακοί παράμετροι. Η τελευταία προσέγγιση έχει αρκετά λειτουργικά και υπολογιστικά πλεονεκτήματα, που ενισχύονται από το γεγονός ότι κατά την καταστολή θορύβου $WDWF$ εφαρμόζεται ούτως ή άλλως κυματιδιακή ανάλυση. Στο πλαίσιο αυτό, αξιοποιήθηκαν και οι δύο τοπολογίες ανάλυσης ($DWT-6$, $WP-17$), ευνοώντας το σχηματισμό διαφορετικών χαρακτηριστικών παραμέτρων WDF . Έτσι, διαμορφώθηκαν οι κυματιδιακοί δείκτες «μεταβάσεων- PTM », «εφησυχασμού- SP » και «αιχμής- NP », κατά αναλογία με τις αντίστοιχες εκφράσεις στο πεδίο του χρόνου (σχέσεις 5-16 έως 5-18) [215]-[216], [218].

Ειδικότερα, από την εφαρμογή των τελεστών PTM , NP_{index} σε κάθε ζώνη k , προέκυψαν οι νέοι παράμετροι PTM_k και $NP_{k;index}$. Για τη αντιμετώπιση των δυσκολιών που αναφέρθηκαν στην § 5.3.2.2-A, διαμορφώθηκαν οι τροποποιημένοι παράμετροι $nPTM_k$ και NP_k . Επίσης, κατά αντιστοιχία με το δείκτη εφησυχασμού SP_{index} , σχηματίστηκαν οι «λόγοι απόρριψης» των κυματιδιακών ακολουθιών RR_k (k -rejection ratio), αξιοποιώντας τα $WDWF$ -αποτελέσματα κατωφλίωσης. Από τις ακολουθίες RR_k προκύπτει ο αριθμός των σημαντικών συντελεστών κυματιδίων $n_{SS;k}$, ενώ εξάγονται πρόσθετα στατιστικά μεγέθη, όπως είναι η μέση

τιμή RR_M (mean RR_k), το κέντρο βάρους RR_C (centroid RR_k) και η διασπορά RR_V (variance RR_k) των «αποτελεσμάτων απόρριψης» RR_k :

$$RR_k = \frac{\sum_{w=0}^{W_{Sk}-1} THR_{[S_k|0]}(w)}{W_{Sk}}, \quad n_{SS;k} = W_{Sk} \cdot (1 - RR_k), \quad \text{με } THR_{[S_k|0]}(w) = \begin{cases} 1, & |S_k(w)| > 0 \\ 0, & \text{διαφορετικά} \end{cases} \quad (5-24)$$

$$RR_M = \text{mean}\{RR_k\}, \quad RR_C = \frac{\sum_k k \cdot (1 - RR_k)}{\sum_k 1 - RR_k}, \quad RR_V = \text{var}\{RR_k\} \quad (5-25)$$

όπου W_{Sk} είναι το μήκος των κυματιδιακών ακολουθιών σε κάθε ζώνη k , ενώ ο μηδενισμός των συντελεστών $S_k(w)$ οφείλεται στην επεξεργασία $WDWF$. Ο δείκτης k μεταβάλλεται στις περιοχές “1-6” ή “1-17”, ανάλογα με την υιοθετούμενη τοπολογία κυματιδιακής ανάλυσης ($DWT-6$ ή $WP-17$). Για το λόγο αυτό, οι παράμετροι RR_M , RR_C και RR_V θα συμβολίζονται στο εξής RR_{M-6} , RR_{C-6} , RR_{V-6} , και RR_{M-17} , RR_{C-17} , RR_{V-17} , αντίστοιχα [215]-[216], [218].

5.3.2.3 Τοπολογίες δικτύων MLP και αποτελέσματα εκπαίδευσης

Όπως επισημάνθηκε στις § 5.3.1 και 5.3.2, το ζήτημα αναγνώρισης των προτύπων ASCC αντιμετωπίστηκε με τη βοήθεια νευρωνικών δικτύων MLP . Στο πλαίσιο αυτό, αποφασίστηκε η χρήση ενός κρυφού επιπέδου και η επιλογή σιγμοειδών συναρτήσεων ενεργοποίησης σε όλους του κόμβους, εκτός από το επίπεδο εξόδου όπου τέθηκαν γραμμικές συναρτήσεις. Από τις διάφορες δοκιμές σχετικά με το μέγεθος των δικτύων, διαπιστώθηκε ότι τιμές εσωτερικών κόμβων από $hn=10$ έως $hn=20$ είναι αρκετές για τις ανάγκες της τρέχουσας εφαρμογής [215]-[216]. Οι έξοδοι του δικτύου καθορίστηκαν από τον αντίστοιχο αριθμό των ASCC κλάσεων ($C_L=5$). Γενικά, αξιοποιήθηκαν περισσότερα από χίλια δείγματα εκπαίδευσης, τα οποία αφού απομονώθηκαν από τις τεχνικές προ-επεξεργασίας, επιλέχθηκαν και ταυτοποιήθηκαν «χειρωνακτικά» (manually). Δοκιμάστηκαν διάφοροι αλγόριθμοι εκπαίδευσης, όπως είναι η κλασική τεχνική $GDBPM$ (Gradient Descent Back Propagation with Momentum) και κυρίως η μέθοδος $RPROP$. Για το σχηματισμό των υποσυνόλων εκπαίδευσης - αξιολόγησης εφαρμόστηκε η τεχνική “k-fold cross validation”, ενώ προτιμήθηκαν οι κανονικοποιημένες εκφράσεις του μέσου σφάλματος (msereg). Και οι δύο παραπάνω διευθετήσεις έγιναν με στόχο την επίτευξη υψηλών ποσοστών γενίκευσης του δικτύου [215]-[216], [218].

Ειδικότερα, σχηματίστηκαν τρία διαφορετικά διανύσματα εισόδου (\mathbf{V}_1 , \mathbf{V}_2 , \mathbf{V}_3 , εξ. 5-26), χρησιμοποιώντας συνδυασμούς των παραμέτρων που παρουσιάστηκαν στην §5.3.2.2, ενώ ανάλογα ορίστηκαν και τα «διανύσματα-στόχοι» (targets) στην έξοδο του δικτύου (εξ. 5-27):

$$\mathbf{V}_1 = [P_{av} \quad DR \quad SP_{index} \quad NP_{index} \quad PTM \quad ECF^{-1} \quad B \quad SFM \quad CB_c \quad RR_{C-17} \quad RR_{M-17} \quad RR_{V-17}]^T$$

$$\mathbf{V}_2 = [[PTM_k]_{k=1-6} \quad ; \quad [NP_{k:index}]_{k=1-6} \quad ; \quad RR_{C-6} \quad RR_{M-6} \quad RR_{V-6}]^T \quad (5-26)$$

$$\mathbf{V}_3 = [[PTM_k]_{k=1-6} \quad ; \quad [NP_{k:index}]_{k=1-6} \quad ; \quad [RR_k]_{k=1-6}]^T$$

$$SP = [1 \ 0 \ 0 \ 0 \ 0]^T ; SN = [0 \ 1 \ 0 \ 0 \ 0]^T ; RESN = [0 \ 0 \ 1 \ 0 \ 0]^T ; IB = [0 \ 0 \ 0 \ 1 \ 0]^T ; RS = [0 \ 0 \ 0 \ 0 \ 1]^T \quad (5-27)$$

Από την εξ. 5-26 προκύπτει ότι τα διανύσματα \mathbf{V}_2 και \mathbf{V}_3 περιλαμβάνουν μόνο κυματιδιακές παραμέτρους (WDF) και αντιστοιχούν σε αμιγή κυματιδιακά νευρωνικά δίκτυα ($WNNET$). Σε κάθε περίπτωση, οι διανυσματικοί έξοδοι των «εκπαιδευμένων» νευρωνικών δικτύων διαφέρουν από τις «ιδεατές» μορφές (τιμές) της εξ. 5-27. Έτσι, για την εξαγωγή των τελικών αποφάσεων ταξινόμησης απαιτείται μια συνάρτηση μετα-επεξεργασίας (post-processing function, f_{post}). Στην προκειμένη περίπτωση επιλέχθηκε ο τελεστής μεγίστου (max), από την εφαρμογή του οποίου προκύπτουν οι έξοδοι $\{o_\lambda\}$ του συστήματος AIMAS: [218]

$$\mathbf{o}_\lambda = f_{post}(\mathbf{o}_m) = [o_\lambda]_{\lambda=1,\dots,5}^T : o_\lambda = \begin{cases} 1, & tn = \arg \max \{\mathbf{o}_m\} \\ 0, & \text{διαφορετικά} \end{cases}, \quad \forall \lambda, tn \in [1, C_L] \quad (5-28)$$

Σε πρώτη φάση, χρησιμοποιήθηκαν 1.745 δείγματα εκπαίδευσης (αντιπροσωπευτικά πρότυπα κάθε κλάσης), με εισόδους τα διανύσματα \mathbf{V}_1 , \mathbf{V}_2 και \mathbf{V}_3 , όπου τα αντίστοιχα ποσοστά επιτυχούς αναγνώρισης, δεν ξεπέρασαν το 80 με 85% [215]-[216]. Έτσι, έχοντας υπόψη τα σχόλια της § 5.3.2.2, σχετικά με τις ιδιαιτερότητες των παραμέτρων PTM , NP_{index} και SP_{index} , συστάθηκαν τρία νέα σύνολα - διανύσματα χαρακτηριστικών μεγεθών (feature sets: \mathbf{F}_{S1} , \mathbf{F}_{S2} , \mathbf{F}_{S3}) [218]:

$$\begin{aligned} \mathbf{F}_{S1} &= [P_{av} \quad DR \quad SP_{index} \quad NP \quad n_{ss} \quad nPTM \quad ECF^{-1} \quad B \quad SFM \quad CB_c \quad RR_{C-17} \quad RR_{M-17} \quad RR_{V-17}]^T \\ \mathbf{F}_{S2} &= [[PTM_k]_{k=1-6} \quad ; \quad [NP_k]_{k=1-6} \quad ; \quad RR_{C-6} \quad RR_{M-6} \quad RR_{V-6}]^T \\ \mathbf{F}_{S3} &= [[PTM_k]_{k=1-6} \quad ; \quad [NP_k]_{k=1-6} \quad ; \quad [n_{ss;k}]_{k=1-6}]^T \end{aligned} \quad (5-29)$$

Τα διανύσματα \mathbf{F}_{S2} και \mathbf{F}_{S3} περιλαμβάνουν και πάλι μόνο παραμέτρους WDF , ορίζοντας αμιγή κυματιδιακά νευρωνικά δίκτυα $WNNET$. Στην περίπτωση αυτή, δοκιμάστηκαν και οι δύο παράμετροι «μεταβάσεων» (PTM_k και $nPTM_k$), όπου τελικά επιλέχθηκε ο πρώτος, καθώς αποδείχθηκε πιο αποδοτικός. Έτσι, σε αντίθεση με την παράμετρο PTM , η επίδραση του «κυματιδιακού» πλάτους στους δείκτες PTM_k διευκολύνει το διαχωρισμό προτύπων που συνδέονται με συγκεκριμένες φασματικές ζώνες (π.χ. υψηλές τιμές PTM_k σε χαμηλές ζώνες k , αντιστοιχούν συνήθως σε θόρυβο SN). Σε κάθε περίπτωση, οι μεταβολές στην απόδοση των δικτύων, από τη χρήση των δύο διαφορετικών όρων PTM , είναι πολύ μικρές ($\sim 1\%$) [218].

Εκτός από τη σύσταση των νέων παραμέτρων, χρησιμοποιήθηκαν και περισσότερα δείγματα εκπαίδευσης (2.334). Τα τελευταία επιλέχθηκαν με τυχαίο τρόπο από τα συμβάντα των καταγραφών, ανάλογα με τη διαθεσιμότητα των αντίστοιχων προτύπων. Η συμμετοχή κάθε κλάσης στο σύνολο των δειγμάτων εκπαίδευσης, έχει ως εξής: SP (557), IB (555), RS (564), SN (384), $RESN$ (139). Αν εξαιρέσει κανείς το χαμηλό ποσοστό των προτύπων $RESN$, γενικά, υπάρχει μια σχετική ισορροπία ως προς την αντιπροσώπευση των προτύπων στο σύνολο των δειγμάτων εκπαίδευσης [218]. Τα παραπάνω δείγματα μοιράστηκαν σε τρία υποσύνολα, από τα οποία τα δύο χρησιμοποιήθηκαν στην εκπαίδευση και το άλλο στην αξιολόγηση της ικανότητας γενίκευσης του δικτύου (k -fold cross validation). Από όλους τους δυνατούς συνδυασμούς των υποσυνόλων, επιλέχθηκαν τα αποτελέσματα εκπαίδευσης με τα μεγαλύτερα ποσοστά επιτυχούς αναγνώρισης [218]. Λαμβάνοντας, δε, υπόψη τις «οδηγίες» αποφυγής φαινομένων υπερ-εκπαίδευσης της § 5.3.1 (σχετικά με το μέγεθος των δικτύων ως προς τα διαθέσιμα ζεύγη εκπαίδευσης) και με δεδομένες τις διαστάσεις των διανυσμάτων εισόδου \mathbf{F}_{S1} , \mathbf{F}_{S2} , και \mathbf{F}_{S3} , αποφασίστηκε η χρήση 10, 15 και 20 κόμβων hn . Έτσι, συνολικά δοκιμάστηκαν οι τοπολογίες δικτύων $FS1-10$, $FS1-15$, $FS1-20$, $FS2-10$, $FS2-15$, $FS2-20$, $FS3-10$,

FS3-15, FS3-20, όπου ο πρώτος όρος δηλώνει το διάλυμα εισόδου και ο δεύτερος των αριθμό των «κρυφών» νευρώνων hn [218].

Γενικά, η απόδοση αναγνώρισης προτύπων μπορεί να αξιολογηθεί από την έκφραση του μέσου τετραγωνικού σφάλματος mse , ανάμεσα στις πραγματικές εξόδους του δικτύου και τις επιθυμητές τιμές-στόχους των δειγμάτων αξιολόγησης. Ωστόσο, το μέσο σφάλμα mse δεν προτείνεται για την αποτίμηση των ποσοστών επιτυχούς αναγνώρισης (αν και συνολικά εκφράζει τα αποτελέσματα εκπαίδευσης του δικτύου), καθώς δεν συμπεριλαμβάνει την πράξη μετα-επεξεργασίας. Εξάλλου, στόχος του δικτύου AIMAS είναι ο διαχωρισμός των χρήσιμων προτύπων (IB , RS) από τα υπόλοιπα (SP , SN , $RESN$), καθώς και η σωστή ταυτοποίηση των κλάσεων IB και RS . Έτσι, οι τυχόν λανθασμένες ταυτοποιήσεις μεταξύ δειγμάτων SP , SN και $RESN$ είναι ήσσονος σημασίας, και δεν θα έπρεπε να έχουν μεγάλη βαρύτητα στην εκτίμηση της απόδοσης του δικτύου. Με βάση τις παραπάνω παρατηρήσεις συστάθηκαν οι ακόλουθοι δείκτες αξιολόγησης [218]:

$$P = \frac{N_{CS}}{N_{TS}} \cdot 100$$

$$IB_{SCORE} = \frac{N_{CIB}}{N_{TIB}} \cdot 100, \quad RS_{SCORE} = \frac{N_{CRS}}{N_{TRS}} \cdot 100 \quad (5-30)$$

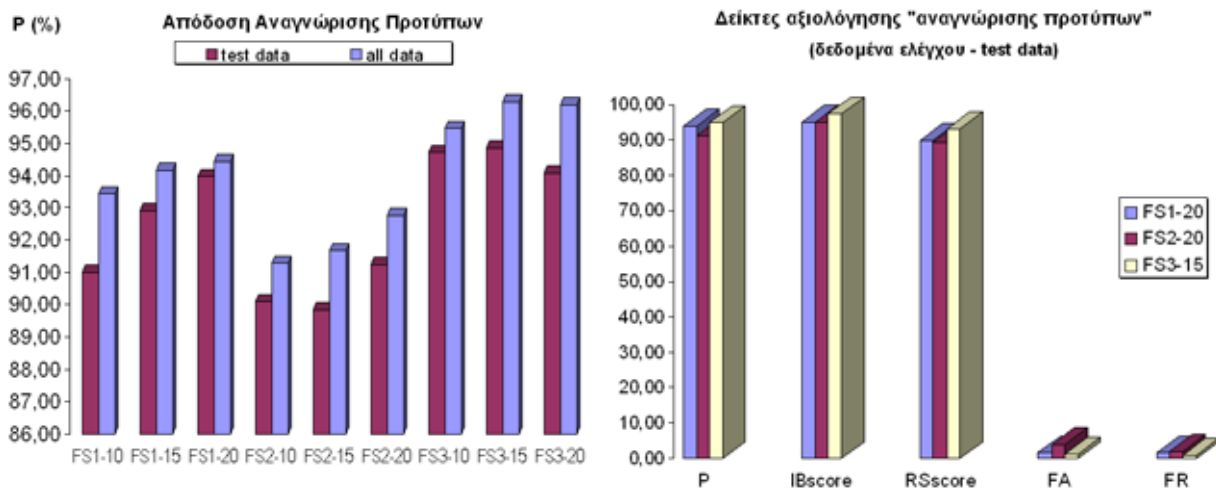
$$FA = \frac{N_{FAS}}{N_{TS}} \cdot 100, \quad FR = \frac{N_{FRS}}{N_{TS}} \cdot 100$$

όπου, P είναι η ποσοστιαία (%) απόδοση του δικτύου (*Performance*), που εκφράζεται από τον αριθμό των σωστά αναγνωρισμένων προτύπων (*Correct Samples*, N_{CS}) προς το συνολικό αριθμό των δειγμάτων (*Total Samples*, N_{TS}), IB_{SCORE} και RS_{SCORE} είναι τα αντίστοιχα ποσοστά επιτυχούς αναγνώρισης των χρήσιμων προτύπων IB , RS (N_{CIB}/N_{CRS} : σωστά αναγνωρισμένα πρότυπα IB/RS , N_{TIB}/N_{TRS} : συνολικός αριθμός δειγμάτων IB/RS), ενώ οι όροι FA/FR (*False Acceptance/ False Rejection*) εκφράζουν τα ποσοστά των δειγμάτων που «κακώς» έγιναν αποδεκτά ή «κακώς» απορρίφθηκαν (N_{FAS} : ο αριθμός των δειγμάτων SP , $RESN$, SN που δεν απομακρύνθηκαν – “*false accepted samples*”, και N_{FRS} : ο αριθμός των δειγμάτων IB , RS , που αποκλείστηκαν ενώ δεν έπρεπε – “*false rejected samples*”). Οι παραπάνω δείκτες μπορεί να εφαρμοστούν στο σύνολο των ζευγών εκπαίδευσης (*training pairs* - N_{TP}) εκφράζοντας τα συνολικά ποσοστά απόδοσης, ή στα δείγματα επαλήθευσης (*validation training pairs* - $N_{TP,V}$) εκτιμώντας την ικανότητα γενίκευσης του δικτύου [218]. Τα δύο αυτά σύνολα δειγμάτων θα αναφέρονται στο εξής με τους όρους “*all data*” και “*test data*”, αντίστοιχα.

Στον πίνακα Π5-1 παρουσιάζονται τα αποτελέσματα αξιολόγησης όλων των δικτύων “FS(1/2/3)-(10/15/20)” που εκπαιδεύτηκαν για την ανάπτυξη του συστήματος AIMAS. Εκτός από τις τιμές των δεικτών απόδοσης και τον αριθμό των ενδιάμεσων νευρώνων hn , δίνονται και οι βρόχοι επανάληψης (*epochs*), που χρειάστηκαν για την εκπαίδευση του δικτύου. Οι γραφικές παραστάσεις των αποτελεσμάτων αυτών δίνονται στο σχ. 5-5. Είναι προφανές ότι το αμιγές κυματιδιακό δίκτυο FS3-15 παρουσιάζει την καλύτερη συμπεριφορά: η συνολική απόδοση διαμορφώνεται στο 94,86%, τα ποσοστά επιτυχούς αναγνώρισης των IB και RS προτύπων είναι 97,3% και 93,09% αντίστοιχα, ενώ οι δείκτες σφάλματος κυμαίνονται σε πολύ χαμηλά επίπεδα ($FA=1,16\%$, $FR=1,03\%$). Μάλιστα, τα ποσοστά αυτά εκφράζουν την ικανότητα γενίκευσης του δικτύου AIMAS, καθώς έχουν προκύψει από τα δείγματα ελέγχου (*test data*). Ιδιαίτερα σημαντικό είναι και το γεγονός ότι το τελικά επλεγμένο δίκτυο FS3-15 αξιοποιεί την ίδια κυματιδιακή δομή ($DWT-6$) με τις τεχνικές $WDWF_{II-6}$ και $LT-WDSS$ [218].

		FS1			FS2			FS3				
		FS1-10 (10.000 βρόχοι-epochs)			FS2-10 (10.000 βρόχοι-epochs)			FS3-10 (15.000 βρόχοι-epochs)				
test data	P	708/	778	91,00 %	701/	778	90,10 %	737/	778	94,73 %	10 hidden neurons (hn)	
	IB _{SCORE}	170/	185	91,89 %	173/	185	93,51 %	177/	185	95,68 %		
	RS _{SCORE}	169/	188	89,89 %	164/	188	87,23 %	174/	188	92,55 %		
	FA	21/	778	2,70 %	33/	778	4,24 %	7/	778	0,90 %		
	FR	16/	778	2,06 %	18/	778	2,31 %	12/	778	1,54 %		
all data	P	2181/	2334	93,44 %	2131/	2334	91,30 %	2228/	2334	95,46 %		
	IB _{SCORE}	528/	555	95,14 %	519/	555	93,51 %	528/	555	95,14 %		
	RS _{SCORE}	516/	564	91,49 %	510/	564	90,43 %	527/	564	93,44 %		
	FA	60/	2334	2,57 %	97/	2334	4,16 %	23/	2334	0,99 %		
	FR	35/	2334	1,50 %	37/	2334	1,59 %	26/	2334	1,11 %		
		FS1-15 (10.000 βρόχοι-epochs)			FS2-15 (10.000 βρόχοι-epochs)			FS3-15 (10.000 βρόχοι-epochs)				
test data	P	723/	778	92,93 %	699/	778	89,85 %	738/	778	94,86 %		15 hidden neurons (hn)
	IB _{SCORE}	171/	185	92,43 %	174/	185	94,05 %	180/	185	97,30 %		
	RS _{SCORE}	173/	188	92,02 %	161/	188	85,64 %	175/	188	93,09 %		
	FA	18/	778	2,31 %	34/	778	4,37 %	9/	778	1,16 %		
	FR	12/	778	1,54 %	23/	778	2,96 %	8/	778	1,03 %		
all data	P	2198/	2334	94,17 %	2140/	2334	91,69 %	2247/	2334	96,27 %		
	IB _{SCORE}	528/	555	95,14 %	519/	555	93,51 %	543/	555	97,84 %		
	RS _{SCORE}	522/	564	92,55 %	516/	564	91,49 %	533/	564	94,50 %		
	FA	53/	2334	2,27 %	89/	2334	3,81 %	23/	2334	0,99 %		
	FR	28/	2334	1,20 %	38/	2334	1,63 %	17/	2334	0,73 %		
		FS1-20 (3.300 βρόχοι-epochs)			FS2-20 (12.500 βρόχοι-epochs)			FS3-20 (5.500 βρόχοι-epochs)				
test data	P	731/	778	93,96 %	699/	778	89,85 %	732/	778	94,09 %	20 hidden neurons (hn)	
	IB _{SCORE}	176/	185	95,14 %	174/	185	94,05 %	180/	185	97,30 %		
	RS _{SCORE}	169/	188	89,89 %	161/	188	85,64 %	175/	188	93,09 %		
	FA	13/	778	1,67 %	34/	778	4,37 %	13/	778	1,67 %		
	FR	13/	778	1,67 %	23/	778	2,96 %	7/	778	0,90 %		
all data	P	2228/	2334	95,46 %	2140/	2334	91,69 %	2245/	2334	96,19 %		
	IB _{SCORE}	528/	555	95,14 %	519/	555	93,51 %	540/	555	97,30 %		
	RS _{SCORE}	527/	564	93,44 %	516/	564	91,49 %	538/	564	95,39 %		
	FA	23/	2334	0,99 %	89/	2334	3,81 %	32/	2334	1,37 %		
	FR	26/	2334	1,11 %	38/	2334	1,63 %	13/	2334	0,56 %		

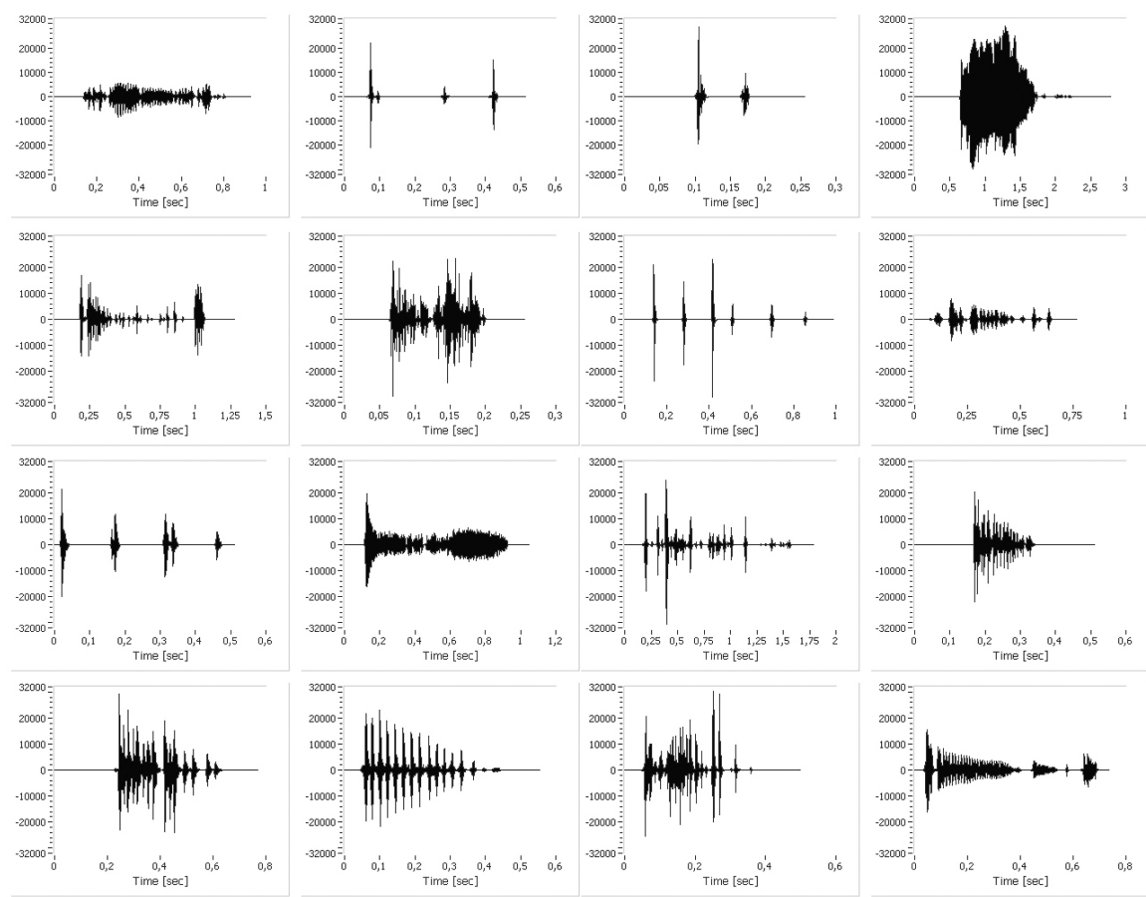
Πίνακας Π5-1. Αποτελέσματα αξιολόγησης (αναγνώρισης προτύπων) των δικτύων MLP που εκπαιδεύτηκαν.



Σχήμα 5-5. Γραφική αναπαράσταση των αποτελεσμάτων αξιολόγησης των δικτύων MLP που εκπαιδεύτηκαν.

5.4 Σχήμα αναλυτικής ταξινόμησης κοιλιακών ήχων (ASFC)

Όπως έχει ήδη επισημανθεί, το σχήμα ταξινόμησης *ASCC* αφορά στα πρωταρχικά (primitive) πρότυπα, τα οποία αποτελούν τα βασικά συστατικά στοιχεία των διαφόρων συμβάντων. Έτσι, αν επιχειρήσουμε να εφαρμόσουμε το λεξικό συνοπτικής ταξινόμησης *ASCC* σε «φυσικά» (manually) απομονωμένους γαστρεντερικούς ήχους, τότε εμφανίζονται αρκετές δυσκολίες (σχ. 5-6). Συγκεκριμένα, υπάρχουν περιπτώσεις όπου τα σήματα περιλαμβάνουν πολλούς κρουστικούς παλμούς, καθώς και τυχαίους συνδυασμούς των τελευταίων με τμήματα ομαλών κυματομορφών, έτσι που η διάκριση *IB/RS* δεν είναι εφαρμόσιμη. Τα ίδια προβλήματα εμφανίζονται και κατά τον αυτόματο εντοπισμό - διαχωρισμό των συμβάντων, εκτός από την περίπτωση εφαρμογής της μεθόδου *SBS* με χρήση πολύ μικρών παραθύρων σάρωσης. Ωστόσο, από την ανάλυση του προηγούμενου κεφαλαίου κατέστη σαφές ότι η τελευταία προσέγγιση δεν είναι ιδιαίτερα λειτουργική, ούτε συμβαδίζει με τις ιατρικές προδιαγραφές μελέτης των γαστρεντερικών ακουστικών καταγραφών παρατεταμένου χρόνου (γι' αυτό και στη θέση της προτάθηκε η μέθοδος *LT-WDSS*) [217]. Επιπρόσθετα, υπάρχει αδυναμία διαχωρισμού των παραμένοντων «τριγμών» από τους συρίτονες ήχους, όπου και οι δύο κατηγορίες ταξινομούνται στην κλάση *RS*. Έτσι, αναζητήθηκαν πιο «εξελιγμένοι» κανόνες διαχωρισμού των κοιλιακών ήχων, με αποτέλεσμα τη σύσταση του λεξικού αναλυτικής ταξινόμησης *ASFC* [61], [217].



Σχήμα 5-6. Σημαντικά συμβάντα γαστρεντερικών ήχων (διαφόρων μορφολογικών χαρακτηριστικών) που έχουν απομονωθεί από τις ακουστικές καταγραφές παρατεταμένου χρόνου. Είναι φανερό ότι η ομαδοποίηση των παραπάνω σημάτων βάσει των κανόνων συνοπτικής ταξινόμησης παρουσιάζει αρκετές δυσκολίες, ενώ το αντίστοιχο λεξικό *ASCC* κρίνεται ανεπαρκές.

5.4.1 Κανόνες περιγραφής των προτύπων αναλυτικής ταξινόμησης ASFC

Η σύσταση των ASFC κλάσεων βασίστηκε και πάλι στη συστηματική ανάλυση ενός μεγάλου αριθμού απομονωμένων συμβάντων, σύμφωνα με τις προδιαγραφές που τέθηκαν στην §5.2. Παράλληλα, υιοθετήθηκε μία φιλοσοφία «ιεραρχικής ταξινόμησης» (§5.1.4) [262], [275]-[276], βάσει της οποίας τα «ASCC-τμήματα» βρίσκονται στο χαμηλότερο επίπεδο ομαδοποίησης (πρωταρχικά πρότυπα), ενώ τα «ψηλότερα» στην ιεραρχία «ASFC-συμβάντα» σχηματίζονται από δομημένες προτάσεις των ASCC-τμημάτων. Με τον τρόπο αυτό, προέκυψαν πέντε ASFC-κλάσεις χρήσιμων κοιλιακών ήχων [61], οι οποίες έχουν ήδη παρουσιαστεί εν μέρει στην §4.2.4 κατά την αξιολόγηση των FDPD-τεχνικών οριοθέτησης των συμβάντων [217]. Στη συνέχεια περιγράφονται τα κύρια χαρακτηριστικά και οι αντίστοιχες ονομασίες των σύνθετων αυτών προτύπων [61] (περίπλοκα πρότυπα - “*complex patterns*”, σύμφωνα με την ορολογία συντακτικής ταξινόμησης [262]):

α) **Solitary CLicks -SCL**: μεμονωμένοι κρουστικοί ήχοι, που αποτελούνται από ένα μοναδικό τμήμα IB, με τις ενδεχόμενες συνοδευτικές «εξάρσεις αρχής - τέλους» (pre-echo/post-echo pulses). Τα χαρακτηριστικά τους (διακυμάνσεις πλάτους, φασματικό περιεχόμενο, διάρκεια, κρουστικότητα, κλπ) είναι όμοια με τα αντίστοιχα των IB-τμημάτων, που παρουσιάστηκαν στην §5.3.1. Από την ανάλυση των πρώτων πειραματικών καταγραφών παρατεταμένου χρόνου διαπιστώθηκε ότι τα συμβάντα SCL έχουν τη μεγαλύτερη συχνότητα παρουσίας από όλα τα πρότυπα ASFC. Έτσι, σε πολλές περιπτώσεις «παρεμβάλλονται» ανάμεσα σε μεγάλες περιόδους σιγής, ενώ εντοπίζονται κυρίως μεταξύ των φάσεων εφησυχασμού και έντονης - ακανόνιστης δραστηριότητας (φάση I προς II του κύκλου MMC), όπου και αποκτούν μεγαλύτερη πυκνότητα εμφάνισης.

β) **Repeated CLicks -RCL**: επαναλαμβανόμενοι κρουστικοί παλμοί, που αποτελούνται από πολλά εσωτερικά τμήματα IB, διαφορετικών συνήθως χαρακτηριστικών. Τα επεισόδια αυτά έχουν πολύ μικρότερα ποσοστά εμφάνισης και συνδέονται κυρίως με την ακανόνιστη φάση II του κύκλου MMC (λίγο πριν αρχίσουν οι ρυθμικές - ομαλές συστολές της φάσης III). Στις περισσότερες περιπτώσεις σχηματίζονται ακολουθίες δύο, τριών και τεσσάρων τμημάτων IB, χωρίς να αποκλείεται και η εμφάνιση πιο μεγάλων «αλυσίδων» (στην παρούσα πιλοτική μελέτη έχουν παρατηρηθεί έως και επτά παλμοί στη σειρά). Η διάρκειά τους κυμαίνεται από 90msec έως 1,5sec, με ποιο συνηθισμένα τα «μήκη» 400msec, περίπου. Το φασματικό τους περιεχόμενο συγκεντρώνεται σε ποιο στενές περιοχές των μεσαίων συχνοτήτων, σε σχέση με τα συμβάντα SCL. Έτσι, το φασματικό κέντρο βάρους διαμορφώνεται γύρω στα 400Hz (με οριακές τιμές από 250Hz έως 1,4kHz, περίπου), ενώ το μέσο εύρος φάσματος BW κυμαίνεται γύρω στα 130-150Hz (με οριακές τιμές από 40 έως 800 Hz, περίπου). Υπάρχουν περιπτώσεις όπου οι εσωτερικοί κρουστικοί παλμοί είναι πάρα πολλοί και πάρα πολύ κοντά μεταξύ τους, οπότε αρχίζουμε να μιλάμε για «συμπλέγματα συστατικών ήχων» (clustered contractions, RS). Άρα, θα πρέπει να υπάρχει μια ελάχιστη απόσταση μεταξύ των επαναλαμβανόμενων παλμών, ούτως ώστε να ερμηνεύονται (οπτικά και ακουστικά) ως διαφορετικές συνιστώσες και να καταχωρούνται ως RCL- εσωτερικά τμήματα. Στην προκειμένη περίπτωση η διαφορά αυτή τέθηκε γύρω στα 30-35msec, λαμβάνοντας υπόψη και τις αντιληπτικές ιδιότητες του ανθρώπου που παρουσιάστηκαν στην § 4.2.3.2.

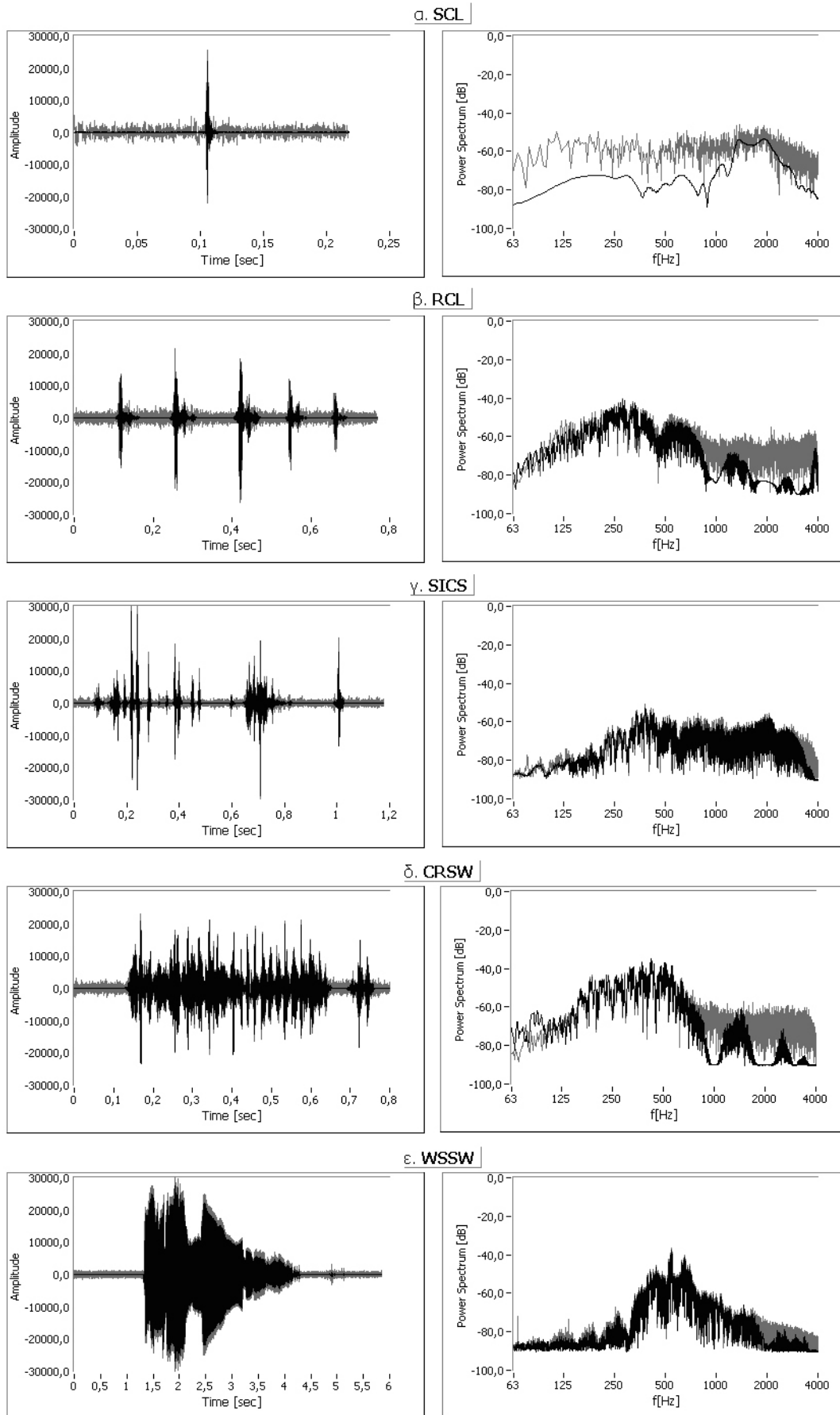
γ) **Sequences of Irregularly Concatenated Segments -SICS**: ακολουθίες ακανόνιστων, τυχαία συνδεδεμένων εσωτερικών τμημάτων IB και RS. Τα επεισόδια SICS είναι αρκετά συχνά κατά μήκος των καταγραφών παρατεταμένου χρόνου (πιο πολύ από τα RCL), ενώ παρατηρούνται κυρίως μεταξύ των φάσεων II και III του κύκλου MMC (όπως τα RCL). Πρόκειται για μια «ημιτελή», κατά κάποιον τρόπο, κλάση, η οποία σχηματίστηκε για να καλύψει το κενό ανάμεσα στους κρουστικούς ήχους (SCL, RCL) και τα ομαλά - παραμένοντα σήματα (CRSW,

WSSW), που περιγράφονται στη συνέχεια. Το βασικό γνώρισμα των συμβάντων *SICS* έχει να κάνει με τη σύστασή τους, δηλαδή με το γεγονός ότι αποτελούνται από «μικτά» τμήματα *IB* και *RS*. Από εκεί κι έπειτα, υπάρχουν μεγάλες διαφορές σχετικά με το πλήθος, τη χρονική σειρά και τα αντίστοιχα χαρακτηριστικά των εσωτερικών *ASCC* τμημάτων. Στο πλαίσιο αυτό, υπήρξε προβληματισμός για το εάν θα έπρεπε να γίνει περαιτέρω διαχωρισμός των προτύπων *SICS*. Τελικά, αποφασίστηκε να μη γίνει κάτι τέτοιο (τουλάχιστον στην παρούσα φάση), καθώς δεν βρέθηκαν αντίστοιχα «ευδιάκριτοι» κανόνες ομαδοποίησης, ενώ σύμφωνα με τις απόψεις της ιατρικής ερευνητικής ομάδας, ο σχηματισμός περισσότερων κλάσεων θα καθιστούσε τη διαδικασία ανάλυσης *ASPA* λιγότερο λειτουργική και αποδοτική. Με βάση τα αποτελέσματα στατιστικής ανάλυσης αντιπροσωπευτικών δειγμάτων, τα πρότυπα *SICS* έχουν τα ακόλουθα χαρακτηριστικά: μέση διάρκεια 1-1,2sec (με εύρος διακυμάνσεων από 170msec έως 2sec, περίπου), τυπικές τιμές φασματικού κέντρου βάρους γύρω στα 550Hz (με οριακές τιμές από 300Hz έως 1,8kHz), μέσο φασματικό εύρος *BW* περίπου 200Hz (με οριακές διακυμάνσεις από 50 έως και 850 Hz).

δ) CRepitating SWeeps -CRSW: πρόκειται για την κύρια κατηγορία των προτύπων *RS*, υπό την έννοια ότι τα αντίστοιχα συμβάντα περιλαμβάνουν ένα μόνο τμήμα παραμένοντων τριγμών, με τις ενδεχόμενες συνοδευτικές «εξάρσεις αρχής - τέλους», μικρότερης έντασης (pre-echo/post-echo pulses). Ο όρος “crepitating” αναφέρεται στους ήχους τριγμών των συμπλεγμάτων συσταλτικής δραστηριότητας που περιγράφηκαν πιο πάνω (clustered contractions) και χρησιμοποιήθηκε πρώτη φορά από τον Cannon [23]. Τα πρότυπα *CRSW* έχουν τα κύρια χαρακτηριστικά των προτύπων *RS*: μέση διάρκεια 800msec (με εύρος διακυμάνσεων από 200msec έως 3sec, περίπου), τυπικές τιμές φασματικού κέντρου βάρους γύρω στα 500Hz (με οριακές τιμές από 250Hz έως 1,7kHz), μέσο φασματικό εύρος *BW* περίπου 240Hz (με οριακές διακυμάνσεις από 35 έως και 750 Hz). Με βάση τις πρώτες πειραματικές αναλύσεις, εκτιμάται ότι τα συμβάντα *CRSW* συνδέονται με τη φάση III του συμπλέγματος *MMC*, καθώς εμφανίζονται μετά το πέρας των «ακανόνιστων φάσεων» (*RCL-SICS*) έχοντας αρκετά έντονο και ρυθμικό χαρακτήρα (συνδέονται με επεισόδια περιοδικού περισταλτισμού «σάρωσης», για την εκκένωση του εντερικού σωλήνα).

ε) WhiStling SWeeps -WSSW: πρόκειται για τους συρίτονες (μουσικούς) κοιλιακούς ήχους, οι οποίοι έχουν έντονο τονικό χαρακτήρα και ομαλές (παραμένουσες) καμπύλες. Σύμφωνα με τον Cannon [23], η γέννησή τους συνδέεται με το σχηματισμό λεπτών στενώσεων του εντερικού σωλήνα κατά τις φάσεις εκκένωσης που περιγράφηκαν πιο πάνω (“piping notes” ήταν ο όρος που χρησιμοποίησε). Ωστόσο, πολλοί ερευνητές δεν αποκλείουν το ενδεχόμενο να παράγονται και από τη γαστρική συσταλτική δραστηριότητα [61], [217]. Τα συμβάντα *WSSW* αποτελούν τη λιγότερο συχνή κλάση από όλα τα *ASFC* πρότυπα. Η εμφάνισή τους παρατηρείται σε χρονικές στιγμές όμοιες με αυτές των συμβάντων *CRSW* (φάση III του *MMC* κύκλου, επεισόδια «σάρωσης»), αλλά σε πολύ μικρότερα ποσοστά. Το φασματικό τους περιεχόμενο είναι εντοπισμένο σε στενο-ζωνικές περιοχές (εξαιτίας του έντονα τονικού τους χαρακτήρα), ενώ δεν αποκλείεται και η εμφάνιση συνοδευτικών κρουστικών παλμών «έναρξης - λήξης», όπως και στην περίπτωση των προτύπων *SCL*, *CRSW*. Με βάση τη στατιστική ανάλυση αντιπροσωπευτικών δειγμάτων, τα πρότυπα *WSSW* έχουν τα εξής χαρακτηριστικά: μέση διάρκεια 600msec (με διακυμάνσεις από 280msec έως 3sec, περίπου), θεμελιώδεις συχνότητες γύρω στα 400-450 Hz (από 240 έως 750 Hz, περίπου), και μέσο φασματικό εύρος $BW \approx 100\text{Hz}$ (από 25 έως και 250 Hz, περίπου).

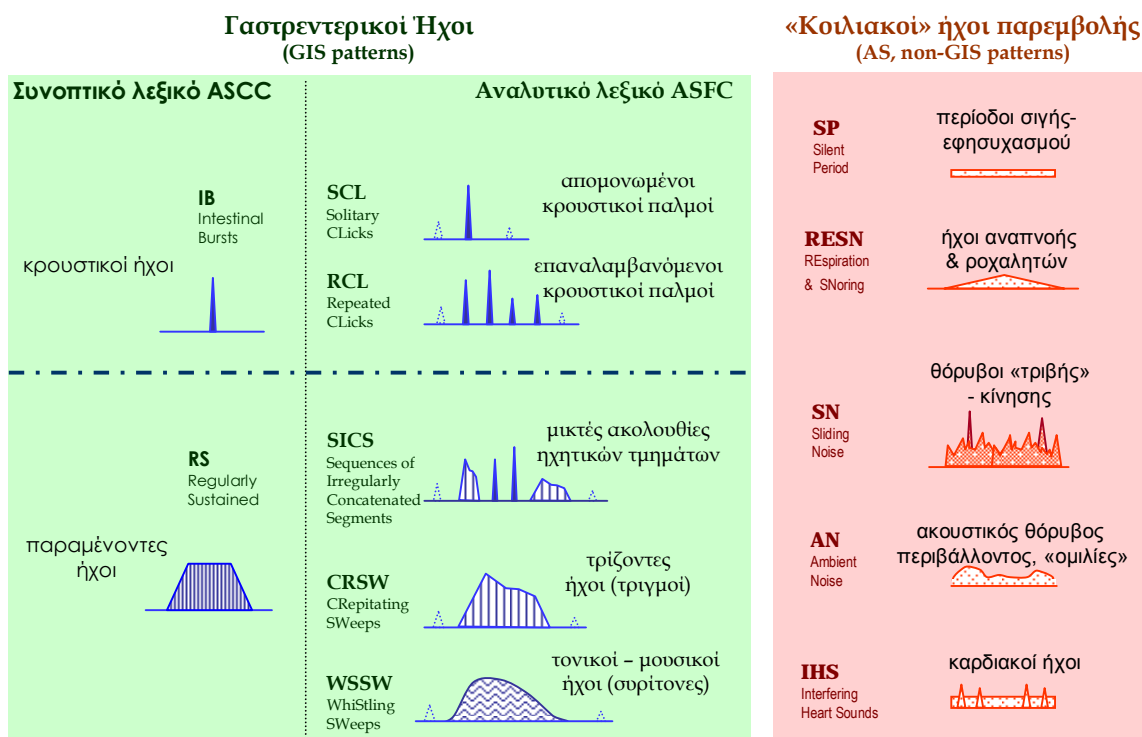
Τα παραπάνω χαρακτηριστικά αποδίδονται με γραφικό τρόπο στο σχ. 5-7, όπου απεικονίζονται αντιπροσωπευτικά πρότυπα των *ASFC* κλάσεων στο πεδίο του χρόνου και της συχνότητας (κατά αντιστοιχία με τα *ASCC* διαγράμματα του σχ. 5-3) [61].



Σχήμα 5-7. Χαρακτηριστικά «ASFC – πρότυπα» στο πεδίο του χρόνου (αριστερά) και της συχνότητας (δεξιά): με γκρι χρώμα απεικονίζονται τα ενθόρυβα σήματα $x(i)$, $X(k)$ ενώ με μαύρο τα αποθορυβοποιημένα $s^-(i)$, $S^-(k)$.

Εκτός των παραπάνω χρήσιμων *ASFC*-προτύπων, συστάθηκαν δύο νέες κλάσεις «θορύβων παρεμβολής» με στόχο τη συμπλήρωση των αντίστοιχων όρων του *ASCC* λεξικού. Πρόκειται για τις περιπτώσεις του περιβαλλοντικού ακουστικού θορύβου (*Ambient Noise – AN*) και των βιοακουστικών σημάτων λειτουργίας της καρδιάς (*Interfering Heart Sounds – IHS*). Στην πρώτη κλάση (*AN*) περιλαμβάνονται κυρίως σήματα ομιλίας (μεταξύ του ιατρικού - νοσηλευτικού προσωπικού και των υποκειμένων), καθώς και ακουστικοί θόρυβοι γειτονικών περιοχών (ανάλογα με τα χαρακτηριστικά ηχομόνωσης των δωματίων εξέτασης). Η γένεση των ανεπιθύμητων, αυτών, παρεμβολών εξαρτάται από τις συνθήκες διεξαγωγής των πειραμάτων, ενώ είναι δυνατός ο απόλυτος έλεγχος και περιορισμός τους. Παράλληλα, έχει διαπιστωθεί ότι η ενσωμάτωση της προσαρμοστικής διάταξης *SHIELD* ελαχιστοποιεί την παρουσία αυτών των φαινομένων, ακόμη και σε πολύ δύσκολες συνθήκες [61], [179], [217]. Όσον αφορά τις παρεμβολές καρδιακών ήχων *IHS*, αυτές είναι σπάνιες σε βιοακουστικές γαστρεντερικές καταγραφές, εκτός από τις περιπτώσεις πειραματόζων [78] ή υποκειμένων με πολύ μικρές φυσικές διαστάσεις (κυρίως βρέφη και νήπια) [59]. Το γεγονός αυτό συμβαίνει καθώς οι παραγόμενοι ήχοι *IHS* εξασθενούν πάρα πολύ κατά τη μετάδοσή τους στο εσωτερικό του ανθρώπινου σώματος (λόγω απόστασης και απορρόφησης). Έτσι, σε καμία από τις καταγραφές της πιλοτικής αυτής φάσης δεν εντοπίστηκαν ήχοι *IHS*, ενώ η ενσωμάτωσή τους έγινε κυρίως για λόγους πληρότητας του *ASFC* λεξικού [61], [179], [217]. Στο σχ. 5-8 δίνονται τα «γραφικά μοντέλα» αναπαράστασης των *ASCC / ASFC* προτύπων, βάσει των αντίστοιχων κανόνων ταξινόμησης των κοιλιακών - γαστρεντερικών ήχων.

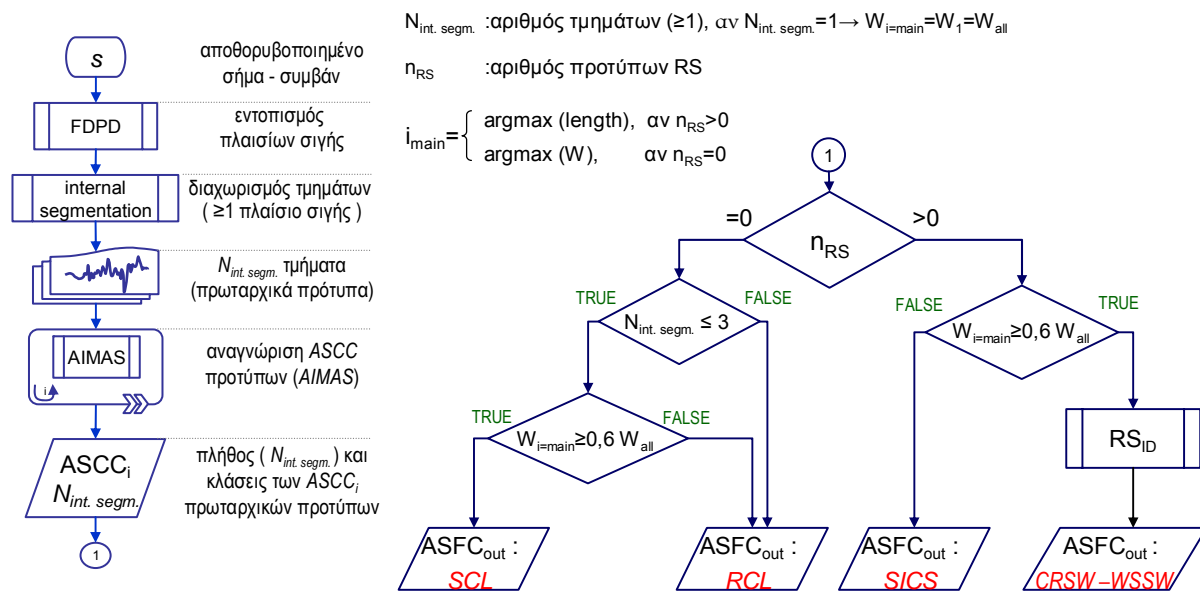
Σχήματα ταξινόμησης κοιλιακών ήχων A. S. P. A. (Abdominal Sounds Pattern Analysis)



Σχήμα 5-8. «Γραφική μοντελοποίηση» των *ASCC/ASFC* προτύπων ταξινόμησης των κοιλιακών ήχων.

5.4.2 Αυτόματη αναγνώριση των προτύπων αναλυτικής ταξινόμησης ASFC

Από την ανάλυση που προηγήθηκε, γίνεται σαφές ότι το σύστημα αναγνώρισης των ASFC-προτύπων (AIMAS+) θα πρέπει να είναι σε θέση να επιτελεί τις ακόλουθες λειτουργίες: **α)** ανίχνευση σιγής (FDPD) των ASFC-συμβάντων με στόχο τον εντοπισμό των εσωτερικών ASCC-τμημάτων (δηλαδή εκείνων των τμημάτων που χωρίζονται από τα «διπλανά» τους με τουλάχιστον ένα «ανενεργό» SB-μπλοκ, ~32msec), **β)** αναγνώριση των ASCC-προτύπων με εφαρμογή του δικτύου AIMAS σε όλα τα εσωτερικά τμήματα ($N_{int.segm.}$), **γ)** απόρριψη των ανεπιθύμητων τμημάτων (SN, RESN, SP) και διατήρηση των χρήσιμων ASCC-κλάσεων (RS, IB), **δ)** εφαρμογή των κανόνων «συντακτικής» ταξινόμησης του ASFC λεξικού και εξαγωγή των τελικών «αναλυτικών προτύπων» (σχ.5-9).



Σχήμα 5-9. Μπλοκ διάγραμμα του συστήματος αναγνώρισης προτύπων AIMAS+, με βάση τους κανόνες αναλυτικής ταξινόμησης (λεξικό ASFC).

Στη διαδικασία αυτή μπορούν να αξιοποιηθούν και οι δύο αλγόριθμοι ανίχνευσης σιγής (TD-FDPD, WT-FDPD), με τον πρώτο να είναι πιο κατάλληλος για την εφαρμογή ολισθαινόντων πλαισίων «σημείο προς σημείο» (point to pint sliding windows). Η διευθέτηση αυτή επιτρέπει την ακριβέστερη οριοθέτηση των εσωτερικών τμημάτων, με το αντίτιμο του μεγαλύτερου υπολογιστικού κόστους. Ωστόσο, οι σχετικές πειραματικές παρατηρήσεις έδειξαν ότι τα αποτελέσματα διαφοροποιούνται ελάχιστα κατά την εφαρμογή μη επικαλυπτόμενων πλαισίων (όπως συμβαίνει και στη διαδικασία MRS-SB, § 4.2.3.3) [217]. Παράλληλα, υιοθετείται η έννοια του βασικού (main) εσωτερικού τμήματος (μπλοκ διάγραμμα του σχ. 5-9β), ώστε να αποφευχθεί η καταχώρηση των πιθανών εξάρσεων «αρχής - τέλους» (pre-echo/post-echo pulses) ως «κανονικών», εσωτερικών τμημάτων. Έτσι, σε κάθε συμβάν «δικαιολογείται» η παρουσία μέχρι και δύο συνοδευτικών παλμών (ένας στην αρχή και ένας στο τέλος), αρκεί η ενέργειά τους να είναι αρκετά μικρότερη από την ενέργεια του βασικού τμήματος (W_{main}). Στην προκειμένη περίπτωση αυτό ελέγχεται με την «εμπειρική» σχέση " $W_{main} \geq 0,6 \cdot W_{all}$ ": η ενέργεια του κύριου τμήματος θα πρέπει είναι μεγαλύτερη ή ίση από το 60% της συνολικής ενέργειας (W_{all}) του συμβάντος, ώστε να μη λαμβάνονται υπόψη οι εκατέρωθεν παλμοί «συνοδείας». Ως βασικό τμήμα καθορίζεται αυτό που έχει τη μεγαλύτερη διάρκεια

(στην περίπτωση που ένα από τα τμήματα ανήκει στην κλάση RS , $n_{RS}>0$), ή αυτό με τη μεγαλύτερη ενέργεια (όταν όλα τα ASCC-τμήματα είναι IB , σχ. 5-9β).

Με βάση τα παραπάνω, ορίζονται οι ακόλουθοι κανόνες συντακτικής ταξινόμησης: α) συμβάντα με έναν κρουστικό παλμό IB ανήκουν στην κλάση SCL , β) συμβάντα με πολλούς εσωτερικούς IB -παλμούς θεωρούνται RCL , γ) συμβάντα που περιέχουν και τα δύο εσωτερικά πρότυπα (IB , RS), καταχωρούνται ως $SICS$, και δ) συμβάντα που έχουν ένα μόνο βασικό τμήμα RS , ανήκουν είτε στην κλάση $CRSW$ είτε στην κλάση $WSSW$. Για το διαχωρισμό $CRSW / WSSW$ αναπτύχθηκε ένα πρόσθετο σύστημα στατιστικής ταξινόμησης RS_{ID} (RS -identification), αξιοποιώντας τις συναρτήσεις ελάχιστης απόστασης (εξ. 5-4, 5-5). Στο πλαίσιο αυτό αξιολογήθηκαν 53 συνολικά παράμετροι ανάλυσης, ανάμεσα στους οποίους και πάρα πολύ περιγραφείς ήχου (Audio Descriptors -D) του πρωτοκόλλου MPEG-7 ("*AudioPower*", "*LogAttackTime*", "*AudioSpectrumCentroid*", "*AudioSpectrumSpread*", "*AudioSpectrumFlatness*", "*AudioSpectrumEnvelope*") [164], [226]-[227], [275], [280]-[283]. Για την επιλογή των βέλτιστων παραμέτρων «ταυτοποίησης» αξιοποιήθηκε το κριτήριο $Q_M>1$ (εξ. 5-9β, §5.1.2), ενώ στη συνέχεια εφαρμόστηκε και ο έλεγχος γραμμικής συσχέτισης (δείκτης CCI , εξ. 3-51). Η τελική σύσταση των χαρακτηριστικών διανυσμάτων ανάλυσης (feature vectors) V_{RS-ID} έχει ως εξής:

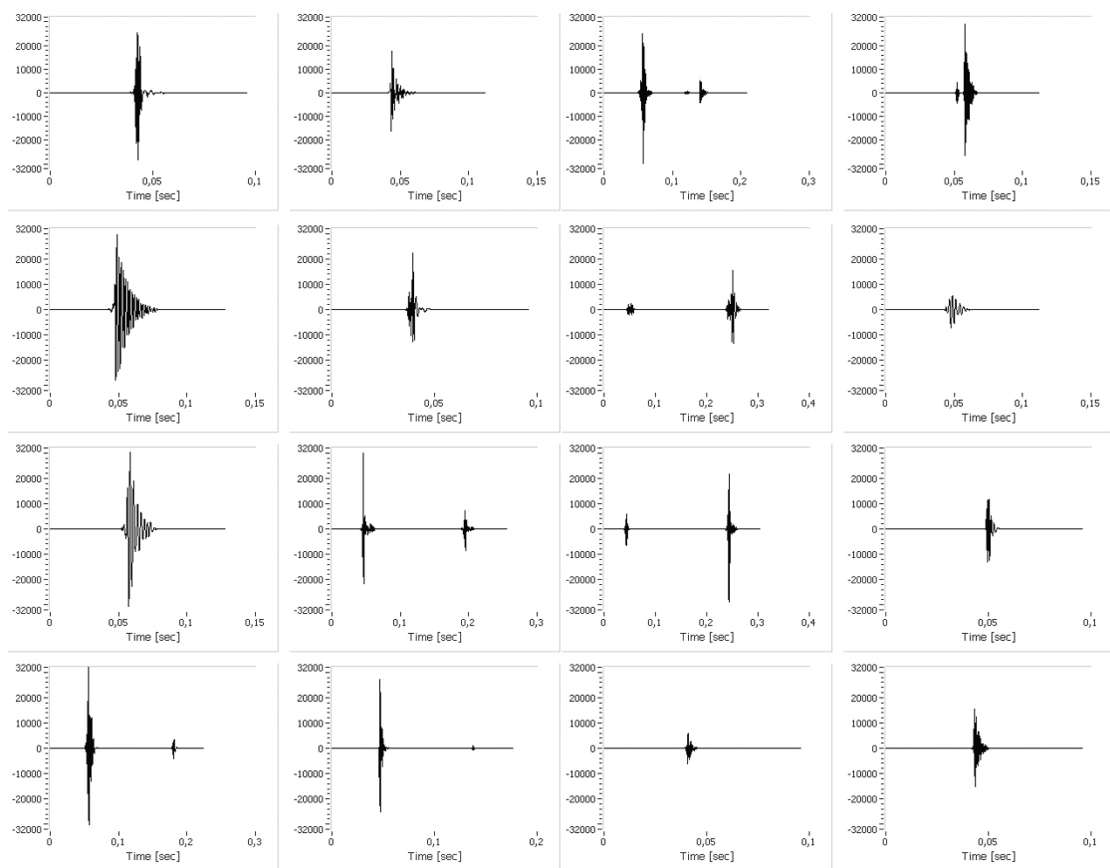
$$V_{RS-ID} = [SS\% \quad nPTM \quad CF \quad SFM], \text{ με } SS\% = \frac{N_{ESNR}}{N} \cdot 100, \quad CF = \frac{\max(s)}{rms(s)} \quad (5-31)$$

όπου $SS\%$ είναι το ποσοστό (%) των σημαντικών δειγμάτων N_{ESNR} (*Significant Samples*, βάσει του $ESNR$ λόγου) προς το συνολικό μήκος N του σήματος, CF είναι ο παράγοντας «κλίσης» (*Crest Factor*), ενώ τα μεγέθη $nPTM$ και SFM είναι τα ίδια που χρησιμοποιήθηκαν και στην εκπαίδευση του δικτύου AIMAS (§ 5.3.2.2, εξ. 5-18 και 5-22).

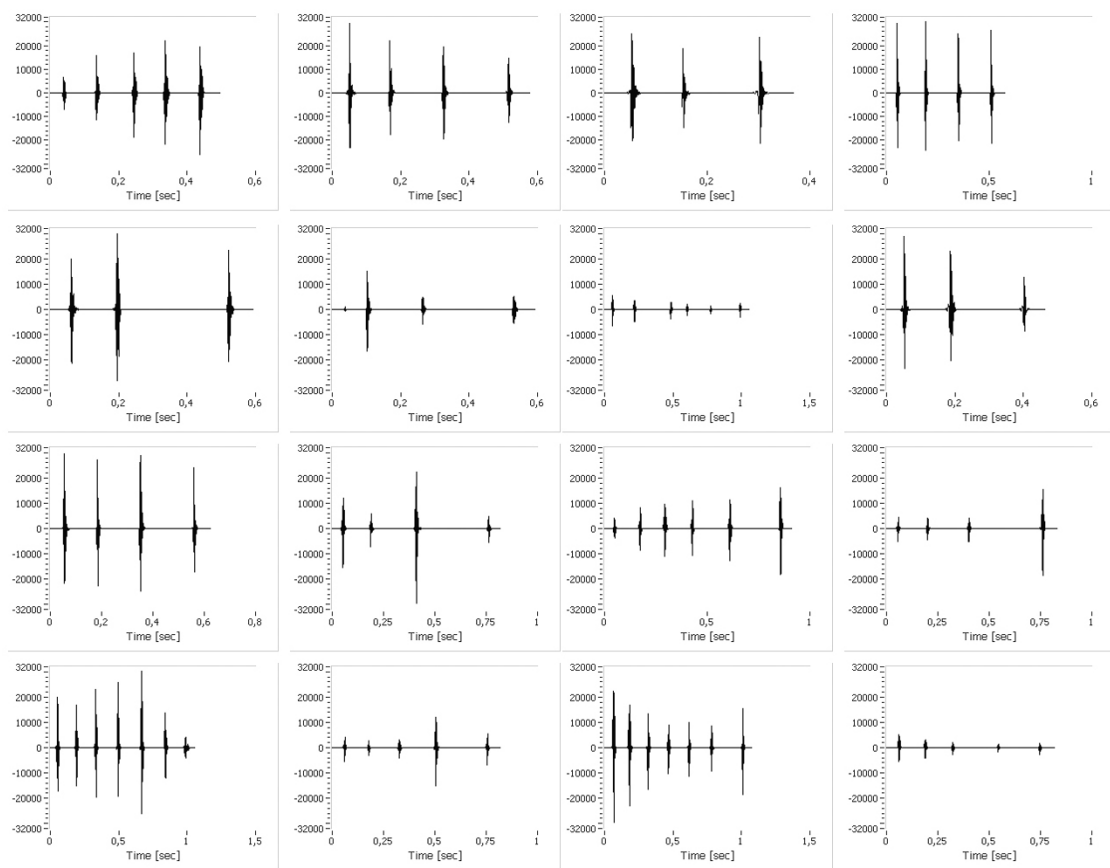
Κατά τη διαδικασία στατιστικής ταξινόμησης αξιοποιήθηκαν 1.176 γνωστά πρότυπα (641 $CRSW$, 534 $WSSW$). Τα δείγματα αυτά ("*all-data*"), «μοιράστηκαν» πολλές φορές σε δύο τυχαία υποσύνολα: το πρώτο (~80%) χρησιμοποιήθηκε (κάθε φορά) στον υπολογισμό του διανυσματικού κέντρου των κλάσεων και το δεύτερο (~20%, "*test-data*") στην αποτίμηση των αντίστοιχων αποτελεσμάτων αναγνώρισης προτύπων. Από την ανάλυση των παραπάνω επαναλήψεων προέκυψαν τα ποσοστά επιτυχούς αναγνώρισης:

$$\begin{aligned} P_{RSID;M}(test-data) &= 93,3\% , & P_{RSID;M}(all-data) &= 94,3\% \\ P_{RSID;E}(test-data) &= 90,6\% , & P_{RSID;E}(all-data) &= 91,1\% \\ P_{CRSW;M}(test-data) &= 97,5\% , & P_{CRSW;M}(all-data) &= 98,6\% \end{aligned} \quad (5-32)$$

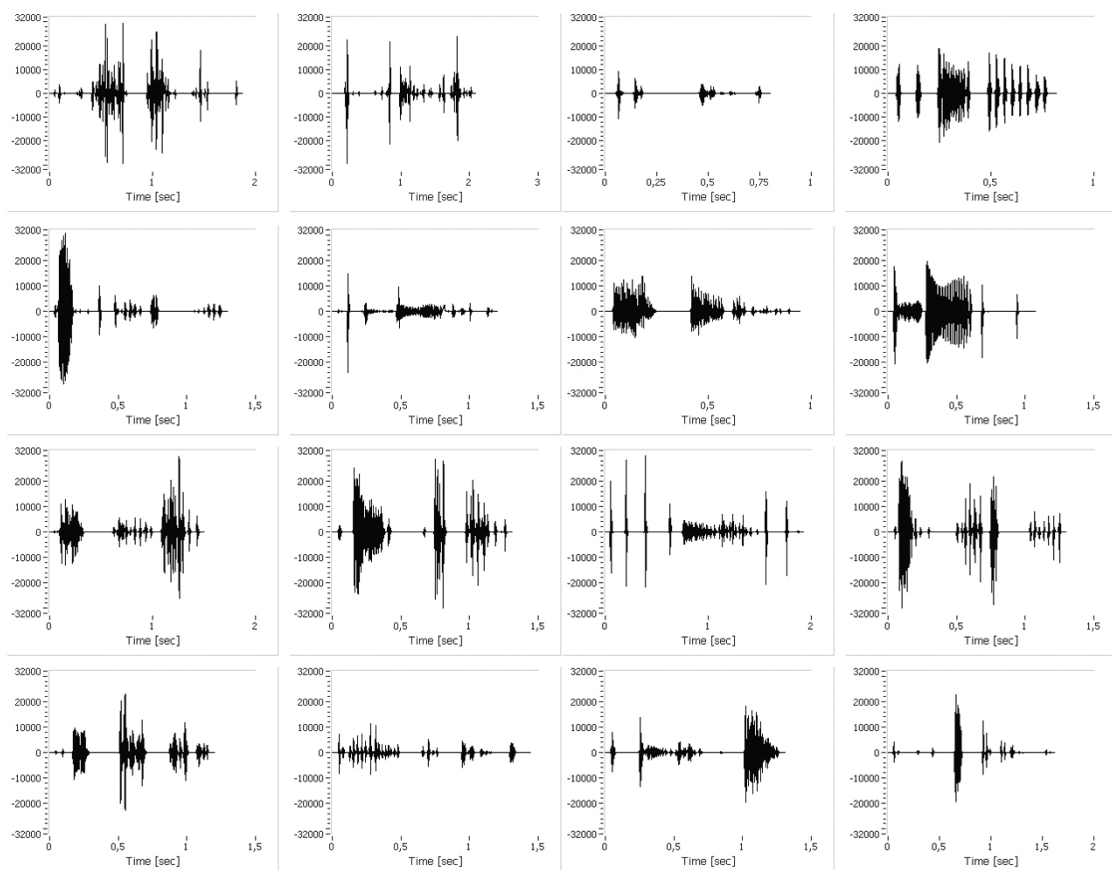
όπου ο όρος P_{RSID} αναφέρεται στην ικανότητα του συστήματος RS_{ID} να αναγνωρίζει σωστά τα RS πρότυπα ($CRSW / WSSW$), ο όρος P_{CRSW} αφορά στην επιτυχή αναγνώριση των $CRSW$ δειγμάτων (κατά αντιστοιχία με τους όρους IB_{SCORE} και RS_{SCORE} , εξ. 5-30), ενώ οι δείκτες M και E δηλώνουν τις αντίστοιχες συναρτήσεις απόστασης D_M , D_E (εξ. 5-5). Είναι φανερό ότι, παρά τη μικρή τάξη των διανυσμάτων V_{RS-ID} ($dim\{V_{RS-ID}\}=4$), το σύστημα RS_{ID} είναι ιδιαίτερα αποδοτικό, ενώ η απόσταση *Mahalanobis* (D_M) δίνει καλύτερα αποτελέσματα (>93%). Τα « $CRSW$ -ποσοστά θετικής αναγνώρισης» είναι ακόμη μεγαλύτερα (>97%), αυξάνοντας έτσι τη συνολική RS_{ID} -απόδοση (λόγω της συχνότερης εμφάνισης των $CRSW$ προτύπων). Στα σχήματα 5-10 έως 5-14 αποτυπώνονται αντιπροσωπευτικά πρότυπα των $ASFC$ κλάσεων, όπως προέκυψαν από τη χρήση του «υβριδικού» συστήματος AIMAS⁺. Από τις χαρακτηριστικές ομοιότητες μεταξύ των σημάτων ίδιων κλάσεων, επιβεβαιώνεται η «εγκυρότητα» και χρηστικότητα του αναλυτικού λεξικού $ASFC$ [61].



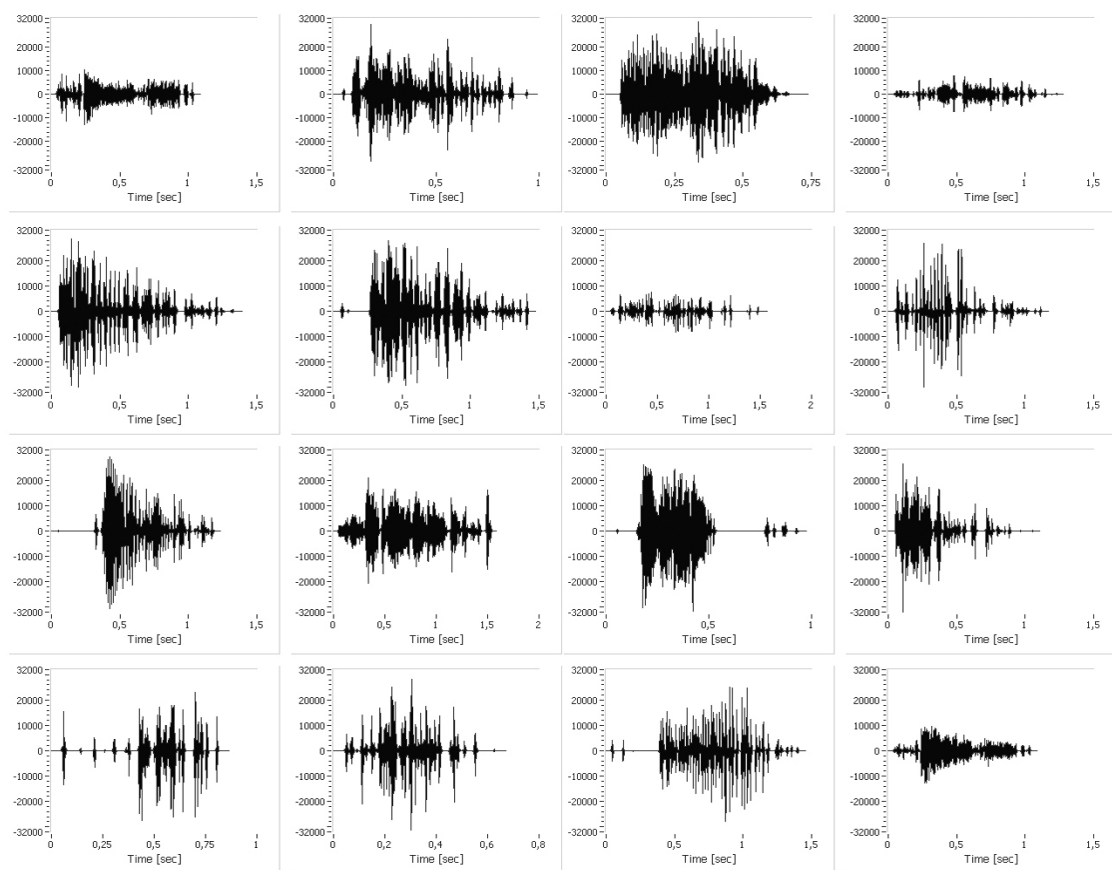
Σχήμα 5-10. Χαρακτηριστικά πρότυπα απομονωμένων κρουστικών παλμών SCL.



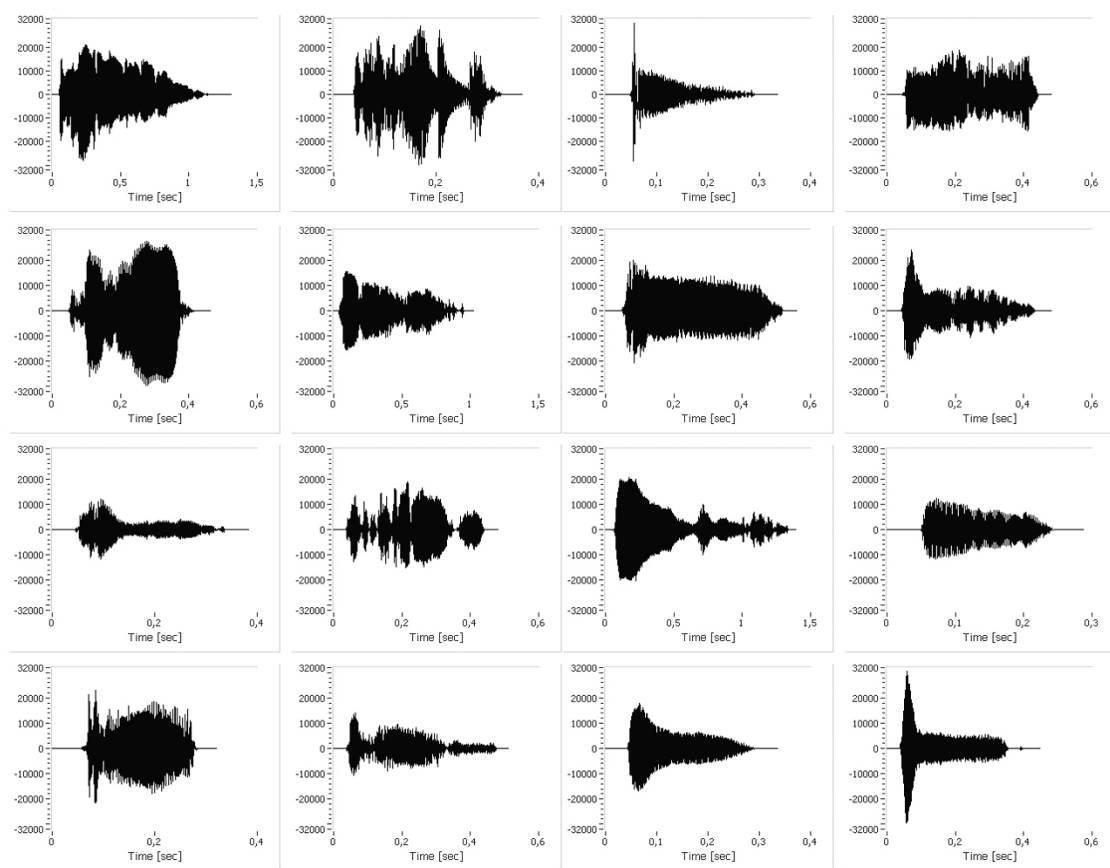
Σχήμα 5-11. Χαρακτηριστικά πρότυπα επαναλαμβανόμενων κρουστικών παλμών RCL.



Σχήμα 5-12. Χαρακτηριστικά πρότυπα «μικτών» κοιλιακών ήχων SICS.



Σχήμα 5-13. Χαρακτηριστικά πρότυπα «τριζόντων» κοιλιακών ήχων CRSW.



Σχήμα 5-14. Χαρακτηριστικά πρότυπα συρίτονων κοιλιακών ήχων WSSW.

5.5 Επισκόπηση 5ου κεφαλαίου

Το κεφάλαιο αυτό ασχολείται με τα θέματα ταξινόμησης και αναγνώρισης προτύπων των βιοακουστικών σημάτων. Αρχικά γίνεται μια σύντομη παρουσίαση των προσεγγίσεων στατιστικού διαχωρισμού και συντακτικής ταξινόμησης, των κανόνων «σχηματισμού» των παραμέτρων ανάλυσης και των «νευρωνικών» τεχνικών, που αξιοποιήθηκαν για τις ανάγκες της παρούσης διατριβής. Το μεγαλύτερο μέρος του κεφαλαίου είναι αφιερωμένο στους κανόνες ομαδοποίησης των κοιλιακών ήχων, τα κίνητρα και τους στόχους ενός τέτοιου εγχειρήματος και τα αντίστοιχα εργαλεία που αναπτύχθηκαν για το σκοπό αυτό.

5.5.1 Αποτίμηση προτεινόμενων παρεμβάσεων

Οι κύριες παρεμβάσεις του κεφαλαίου εντοπίζονται στη δημιουργία δύο λεξικών ταξινόμησης και στην ανάπτυξη των αντίστοιχων εργαλείων αυτόματης αναγνώρισης κοιλιακών ήχων. Έτσι, παρέχονται νέες δυνατότητες μελέτης των ακουστικών καταγραφών γαστρεντερικής κινητικότητας, ενώ παράλληλα απομονώνονται τα επεισόδια μη στάσιμου θορύβου. Τα παραπάνω στοιχεία επιτρέπουν τη σχετική ελευθερία κινήσεων στα υποκείμενα (λιγότερες ενοχλήσεις) κι έτσι η διαδικασία ιατρικής παρακολούθησης καθίσταται πιο αντικειμενική. Η χρήση των νεοσυστατών σχημάτων ταξινόμησης αναμένεται ότι θα συμβάλει στη γρηγορότερη εξέλιξη των ιατρικών μεθόδων γαστρεντερικής ακρόασης, ενώ θα λειτουργήσει και ως «πλότος» κατά το σχεδιασμό σχετικών συστημάτων βιοϊατρικής τεχνολογίας. Κλασικό παράδειγμα αποτελούν οι διαδικασίες αξιολόγησης των αλγορίθμων καταστολής θορύβου *WDWF* και των μεθόδων ανίχνευσης σιγής *FDPD* (3^ο και 4^ο κεφάλαιο,

αντίστοιχα), οι οποίες πραγματοποιήθηκαν ξεχωριστά για τα διάφορα πρότυπα κοιλιακών ήχων [179], [217].

5.5.2 Στοιχεία πρωτοτυπίας & ερευνητική συνεισφορά

Τα στοιχεία πρωτοτυπίας και η ερευνητική συνεισφορά του κεφαλαίου εντοπίζονται, κυρίως, στην ώθηση της ιατρικής έρευνας γύρω από τη μελέτη της γαστρεντερικής κινητικότητας. Με βάση τα αποτελέσματα της σχετικής βιβλιογραφικής αναζήτησης, δεν είναι υπερβολή αν ισχυριστούμε ότι είναι ίσως η πρώτη φορά που εφαρμόζεται ένα τόσο συστηματικό πειραματικό πρωτόκολλο, με στόχο την ταξινόμηση των κοιλιακών ήχων. Οι καινοτομίες της παρούσης εργασίας σχετίζονται με το μέγεθος των «δειγμάτων» που μελετήθηκαν, την ιεραρχική προσέγγιση σύστασης των δύο λεξικών και τη συνεργασία των δύο ερευνητικών ομάδων, όπου εκτός από τις πειραματικές παρατηρήσεις, αξιοποιήθηκε το σύνολο της ιατρικής γνώσης που έχει αποκτηθεί εδώ και έναν, περίπου, αιώνα. Έτσι, χρησιμοποιήθηκαν πάνω από 5.000 απομονωμένα ηχητικά συμβάντα, που προέκυψαν από τριάντα περίπου καταγραφές παρατεταμένου χρόνου και πολύ περισσότερες ηχογραφήσεις μικρότερης διάρκειας, όπου εκτός από υγιείς εθελοντές, συμπεριλήφθηκαν και υποκείμενα με διάφορες παθήσεις. Σε κάθε περίπτωση, στόχος της ταξινόμησης δεν ήταν ο διαχωρισμός των «υγιών» από τους «παθολογικούς ήχους», αλλά η εξεύρεση κοινών χαρακτηριστικών σε επίπεδο «θεμελιωδών προτύπων», δηλαδή η ταξινόμηση των συμβάντων μικρής διάρκειας. Τα έως τώρα αποτελέσματα συνηγορούν στο ότι η «ταυτοποίηση» αυτού του είδους είναι εφαρμόσιμη, ενώ, παράλληλα, η νέα προσέγγιση επιτρέπει τη μελέτη της φυσιολογίας και των λειτουργικών διαταραχών του γαστρεντερικού (που είναι και το ζητούμενο).

Στα θετικά των παρεμβάσεων συγκαταλέγεται και η πολύ καλή απόδοση των δύο συστημάτων αυτόματης αναγνώρισης, *AIMAS* και *AIMAS+*. Η αξιοποίησή τους αναμένεται να συμβάλει στην πρόοδο της ιατρικής έρευνας, όπου ήδη τα πρώτα αποτελέσματα είναι ιδιαίτερα ενθαρρυντικά, καθώς δείχνουν συσχέτιση ανάμεσα στις φάσεις του κύκλου MMC και την αλληλοδιαδοχή των κοιλιακών ηχητικών προτύπων [61]. Ιδιαίτερα βολική είναι, επίσης, η δομή του υβριδικού συστήματος *AIMAS+* και η «φιλοσοφία» των αντίστοιχων συντακτικών κανόνων, καθώς επιτρέπουν τον άμεσο έλεγχο των αποφάσεων ταξινόμησης και την παροχή περιγραφικών στοιχείων για τη μορφή των προτύπων. Έτσι, η σύσταση ακόμη πιο «αναλυτικών» κανόνων ταξινόμησης, στο μέλλον, μπορεί να γίνει εύκολα, με την προσθήκη νέων «ιεραρχικών επιπέδων».

5.5.3 Προτεινόμενες μελλοντικές εξελίξεις

Ο σχεδιασμός μελλοντικής αξιοποίησης προβλέπει την άμεση και συστηματική χρήση των λεξικών ταξινόμησης στις μεθόδους γαστρεντερικής ακρόασης, ενώ η εφαρμογή τους στην «πράξη» θα καθορίσει και την εξέλιξη του όλου εγχειρήματος. Παράλληλα, εξετάζεται το ενδεχόμενο προσθήκης νέων κλάσεων (με πιο πιθανό σενάριο την περαιτέρω ανάλυση των συμβάντων *SICS*), αλλά και η αξιοποίηση πιο σύνθετων ιεραρχικών σχημάτων [275]-[276], [282]. Επιβεβλημένη θεωρείται και η συνεχής εξέλιξη των συστημάτων αναγνώρισης (παρα τις υψηλές αποδόσεις των *AIMAS*, *AIMAS+*), χρησιμοποιώντας συνδυαστικές, νευρο-ασαφείς τεχνικές [277]-[278] και πιο εξελιγμένους αλγορίθμους επιλογής των παραμέτρων ανάλυσης σε νευρωνικά δίκτυα [271]-[274]. Τέλος, εξετάζεται το ενδεχόμενο παράλληλης αξιοποίησης των αποτελεσμάτων «κίνησης βίντεο», μέσα από τη σύσταση αντίστοιχων χαρακτηριστικών παραμέτρων ανάλυσης (features) [218].

ΚΕΦΑΛΑΙΟ 6°

6. Τεχνικές ακουστικής χαρτογράφησης

Στο κεφάλαιο αυτό αναλύονται τα θέματα ακουστικής χαρτογράφησης, δηλαδή τα κίνητρα και οι εφαρμογές των μεθόδων εικονικής αναπαράστασης ηχητικού πεδίου (*sound-field visualization*). Γενικά, οι τεχνικές χωρικής επεξεργασίας (*spatial audio*) είναι ιδιαίτερα δημοφιλείς και αφορούν ένα μεγάλο φάσμα της μηχανικής ήχου, όπως είναι οι εφαρμογές κλασσικής ακουστικής - επικοινωνιών, οι προσεγγίσεις ακουστικής διάγνωσης μηχανικών συστημάτων, οι μέθοδοι βιοακουστικής παρακολούθησης, κ.ά. Στο πλαίσιο αυτό γίνεται μια σύντομη παρουσίαση των σχετικών ερευνητικών προσπαθειών, οι οποίες αποτέλεσαν αντικείμενο μελέτης για την ανάπτυξη ενός νέου συστήματος χαρτογράφησης της βιοακουστικής δραστηριότητας. Παράλληλα, παρατίθενται οι προσεγγίσεις «κοιλιακής χαρτογράφησης» (*abdominal mapping*) που έγιναν στο παρελθόν, όπου διατυπώνονται τα πλεονεκτήματά τους, καθώς και οι δυσκολίες που απορρέουν από τις ιδιαιτερότητες του μέσου διάδοσης, την πολυπλοκότητα του φαινομένου και τους περιορισμούς σχετικά με τον αριθμό και τις θέσεις τοποθέτησης των βιοακουστικών αισθητήρων. Το μεγαλύτερο μέρος του κεφαλαίου είναι αφιερωμένο στην παρουσίαση και αξιολόγηση του νέου συστήματος χωρικής καταγραφής *A-Soundfield 4+3*, το οποίο έχει ήδη περιγραφεί στο δεύτερο κεφάλαιο, χωρίς ωστόσο να παρουσιαστούν οι δυνατότητες τοπογραφικής ανάλυσης, που παρέχει.

Αξίζει να σημειωθεί ότι οι τεχνικές «ακουστικής απεικόνισης» του τρέχοντος κεφαλαίου δεν αποτελούν βασικό αντικείμενο της διατριβής, τουλάχιστον όχι στον ίδιο βαθμό με τις μεθόδους που παρουσιάστηκαν στα τέσσερα προηγούμενα κεφάλαια. Ωστόσο, παρέχουν νέες, εξελιγμένες δυνατότητες ανάλυσης και διαχείρισης περιεχομένου των γαστρεντερικών και, γενικότερα, των βιοακουστικών καταγραφών παρατεταμένου χρόνου. Για τους λόγους αυτούς, καθώς και για λόγους πληρότητας της συγκεκριμένης ερευνητικής προσέγγισης, θεωρήθηκε σκόπιμη η ενσωμάτωση εργαλείων ακουστικής χαρτογράφησης στο προτεινόμενο ιατρικό πρωτόκολλο.

6.1 Χωρική ανάλυση ήχου: θεωρία και εφαρμογές

Υπάρχουν διάφορες προσεγγίσεις χωρικής επεξεργασίας ήχου, των οποίων οι διαφορές έγκεινται τόσο στους εμπλεκόμενους αλγορίθμους, όσο και στα αποτελέσματα ανάλυσης, δηλαδή τους σκοπούς για τους οποίους εφαρμόζεται η διαδικασία χωρικής ανάλυσης. Το κύριο χαρακτηριστικό όλων των μεθόδων είναι ότι χρησιμοποιούν πολλά ηχητικά κανάλια, που προέρχονται από τους αντίστοιχους ηλεκτρακουστικούς μετατροπείς του συστήματος καταγραφής (πολυκαναλικός ήχος, - *multi-channel audio*). Έτσι, διακρίνουμε τις προσεγγίσεις εντοπισμού των ηχητικών πηγών (*Source Location -SL*), τους αλγορίθμους εμπλουτισμού των σημάτων με αντιστάθμιση της χρονικής καθυστέρησης (*Time Delay Compensation - TDC*), τις μεθόδους ακουστικής ολογραφίας εγγύς πεδίου για την ανάλυση «πολύπλοκων» ηχητικών πηγών (*Near-field Acoustic Holography -NAH*) και τις τεχνικές πολυκαναλικής ηχογράφησης - αναπαραγωγής, σύμφωνα με τα πρότυπα στερεοφωνίας και περιφερικού ήχου. Αν και κάθε μια από τις παραπάνω προσεγγίσεις έχει διαφορετικό στόχο, ωστόσο τα όρια διαχωρισμού τους είναι, πολλές φορές, αρκετά δυσδιάκριτα, αφενός γιατί αξιοποιούν

κοινούς αλγορίθμους, αφετέρου γιατί συχνά απαιτείται συνδυασμός τους (π.χ. εντοπισμός ηχητικής πηγής με στόχο τον εμπλουτισμό σήματος ή την απεικόνιση του ηχητικού πεδίου). Σε κάθε περίπτωση, ο πλήρης προσδιορισμός του ηχητικού πεδίου είναι ένα αρκετά δύσκολο εγχείρημα, όπως και όλα τα «αντίστροφα προβλήματα», ενώ παρουσιάζονται και πρόσθετες δυσκολίες που σχετίζονται με τη μοντελοποίηση του μέσου, καθώς και την επιλογή του «σωστού» μηχανισμού διάδοσης των ηχητικών κυμάτων. Έτσι, αξιοποιούνται οι φασικές και ενεργειακές διαφορές ανάμεσα στους ηλεκτρακουστικούς αισθητήρες λήψης (των οποίων οι γεωμετρικές θέσεις είναι γνωστές), εφαρμόζονται αριθμητικές τεχνικές επίλυσης του υπερπροσδιορισμένου αντίστροφου προβλήματος και τελικά επιλέγεται η λύση που οδηγεί στο μικρότερο σφάλμα [158], [160], [248]-[250], [284]-[290].

6.1.1 Τεχνικές εντοπισμού ηχητικών πηγών

Το πιο απλό μοντέλο εντοπισμού της θέσης μιας ηχητικής πηγής αφορά στην περίπτωση του απομακρυσμένου ηχητικού πεδίου (*far field*), όπου η πηγή θεωρείται ότι έχει πλέον πολύ μικρές διαστάσεις (σημειακή πηγή), εκπέμπει «πανκατευθυντικά» (ισοτροπική πηγή), ενώ ο χώρος είναι απαλλαγμένος από εμπόδια (ελεύθερο ηχητικό πεδίο). Έτσι, το μέσο διάδοσης θεωρείται ομογενές και γραμμικό (ο αέρας στην απλούστερη περίπτωση, $c_{\text{sound}} \approx 344 \text{ m/sec}$), ενώ ισχύει ο νόμος του αντιστρόφου τετραγώνου (*Inverse Square Law, ISL: η ηχητική ένταση αποσβένει με το τετράγωνο της απόστασης*) [108]-[109]. Επίσης, με δεδομένη τη μεγάλη απόσταση από το σημείο διέγερσης, τα ηχητικά κύματα θεωρούνται επίπεδα, οπότε η ένταση είναι ανάλογη με το τετράγωνο της ηχητικής πίεσης [108]-[109]. Αν και στην πράξη υπάρχουν μικρές διαφορές από το παραπάνω ιδεατό μοντέλο, ωστόσο τα αποτελέσματα εντοπισμού είναι ιδιαίτερα ικανοποιητικά για τις ανάγκες των περισσότερων εφαρμογών [158], [286]-[287], [291]. Στις περιπτώσεις, δε, όπου το μέσο διάδοσης παρουσιάζει εμπόδια ή αφορά ακόμη και κλειστούς χώρους με πολλές ανακλαστικές επιφάνειες, τότε οι συνιστώσες του αντηχητικού πεδίου μοντελοποιούνται ως ανεπιθύμητος προσθετικός θόρυβος, που δυσχεραίνει τη διαδικασία εντοπισμού [158], [248]-[250], [286]-[287], [291]-[292].

Υπάρχουν περιπτώσεις όπου, ο χωρικός προσδιορισμός της ηχητικής διέγερσης αποτελεί αυτοσκοπό, όπως για παράδειγμα συμβαίνει με τα ραντάρ ηχητικού εντοπισμού [158], [248]-[250], [287]. Σε άλλες πάλι προσεγγίσεις, ο εντοπισμός της ηχητικής θέσης αποτελεί το πρώτο βήμα για μια πιο «βασική» επεξεργασία, όπως είναι ο εμπλουτισμός του σήματος [285]-[287], [291]-[292], καθώς και η απεικόνιση του ηχητικού πεδίου, με προσομοιώσεις διάδοσης προς τα εμπρός (*forward propagation*) [158], [284], [287]-[290]. Χαρακτηριστικές περιπτώσεις της πρώτης κατηγορίας αποτελούν οι συστοιχίες μικροφώνων σε συστήματα συνδιάσκεψης και τηλε-συνδιάσκεψης κλειστών χώρων με στόχο τη βελτίωση «κατανοητότητας του λόγου» (*speech intelligibility*), ενώ υπάρχουν και σχετικές αναφορές επεξεργασίας βιοακουστικών σημάτων [293]-[294]. Κύριος στόχος της δεύτερης κατηγορίας είναι η αξιοποίηση εργαλείων ακουστικής χαρτογράφησης, όπου και πάλι έχουν αναφερθεί ερευνητικές προσπάθειες βιοακουστικής τοπογραφικής απεικόνισης [42], [87].

Οι πιο δημοφιλείς προσεγγίσεις χωρικής επεξεργασίας ήχου αφορούν στις τεχνικές διαμόρφωσης δέσμης (*beam-forming*) [158], [285]-[287], [291]-[292]. Αυτές χρησιμοποιούν γραμμικές ή λογαριθμικές συστοιχίες μικροφώνων, υιοθετώντας τις συνθήκες ελεύθερου και απομακρυσμένου ηχητικού πεδίου που αναφέρθηκαν πιο πάνω. Έτσι, αν υποθέσουμε ότι τα σήματα έχουν στενό φασματικό εύρος, τότε οι μικροφωνικές συνιστώσες παρουσιάζουν φασικές και ενεργειακές διαφορές μεταξύ τους, εξαιτίας της πεπερασμένης ταχύτητας του ήχου και της απόσβεσης λόγω του νόμου του αντιστρόφου τετραγώνου, αντίστοιχα. Οι τιμές

των φασικών και ενεργειακών διαφορών διαμορφώνονται ανάλογα με τα μήκη κύματος και τις αποστάσεις των μικροφώνων μεταξύ τους και από την ηχητική πηγή. Η αντιστάθμιση χρονικής καθυστέρησης *TDC* πραγματοποιείται με τη φασική επεξεργασία κάθε καναλιού χωριστά, δηλαδή την επιβολή χρονικής «υστέρησης». Έτσι, επιτυγχάνεται η συμφασικότητα όλων των συνιστωσών, οι οποίες στη συνέχεια αθροίζονται (*Delay and Sum Beamforming - DSB*). Με δεδομένη, λοιπόν, την τυχαιότητα των ανακλάσεων (προσθετικός θόρυβος) εξασφαλίζεται η βελτίωση της συνολικής σηματοθορυβικής σχέσης (εμπλουτισμός σήματος). Απαραίτητη προϋπόθεση, βέβαια, αποτελεί η γνώση της διεύθυνσης άφιξης (*Direction of Arrival - DOA*). Εάν αυτό δεν είναι γνωστό, τότε ο εντοπισμός της ηχητικής πηγής μπορεί να γίνει οδηγώντας τη δέσμη σε μία διαδικασία γωνιακής σάρωσης, όπου τελικά επιλέγεται η τιμή μεγιστοποίησης της ισχύος στην έξοδο του συστήματος «καθυστέρησης - άθροισης». Στην περίπτωση, δε, που τα σήματα δεν είναι στενοζωνικά, τότε αξιοποιούνται συστοιχίες λογαριθμικής διάταξης, όπου η πολυκαναλική επεξεργασία *DSB* εξελίσσεται παράλληλα σε πολλές φασματικές ζώνες [158], [285]-[287], [291]-[292].

Είναι προφανές ότι η διαδικασία ηχητικού εντοπισμού με την οδηγούμενη διαμόρφωση δέσμης έχει πολλές υπολογιστικές απαιτήσεις και είναι αρκετά χρονοβόρα [285], [287]. Για το λόγο αυτό αξιοποιούνται πιο εξελιγμένες μέθοδοι, όπως είναι οι αλγόριθμοι υπολογισμού των χρονικών καθυστερήσεων (*Time Delay of Arrival - TDA*) μεταξύ των καναλιών. Η φασική διαφορά δύο σημάτων μπορεί να υπολογιστεί με διάφορες μεθόδους, με πιο κλασική τη μεγιστοποίηση της «φάσης φάσματος πεπλεγμένης ισχύος» (*Cross Power Spectrum Phase*), καθώς επίσης με αλγορίθμους ανάλυσης ιδιαζουσών τιμών και μεθόδους στατιστικής ανάλυσης ανώτερης τάξης [97], [249], [285], [287]. Από τις χρονικές διαφορές μεταξύ όλων των ζευγών μικροφώνων προκύπτει ένα «αόριστο» σύστημα εξισώσεων, ενώ η τελική λύση προκύπτει και πάλι από την ελαχιστοποίηση του σφάλματος μέσω αριθμητικών μεθόδων. Από τη στιγμή που θα εντοπιστούν οι συντεταγμένες της ηχητικής πηγής, ο *DSB* εμπλουτισμός σήματος γίνεται, όπως και πιο πάνω, εφαρμόζοντας τη χρονική αντιστάθμιση *TDC*. Οι τεχνικές *TDA* είναι πιο αποδοτικές στην εύρεση της θέσης της ηχητικής πηγής, με το αντίτιμο των μεγαλύτερων απαιτήσεων φασματικού εύρους (ή αντίστοιχα πιο μεγάλου ρυθμού δειγματοληψίας), έτσι ώστε να υπάρχει η απαραίτητη διακριτική ικανότητα για τον εντοπισμό πολύ μικρών χρονικών διαφορών. Στην περίπτωση, δε, που οι αισθητήρες είναι πολύ κοντά στην πηγή (μικρές αποστάσεις), ή όταν η ταχύτητα του μέσου διάδοσης είναι πολύ μεγάλη σε σχέση με το μήκος κύματος των ηχητικών σημάτων, τότε ο υπολογισμός των χρονικών καθυστερήσεων είναι πολύ δύσκολος, έως και αδύνατος [97], [237], [285], [287].

Οι τεχνικές ενεργειακού εντοπισμού αναπτύχθηκαν στην προσπάθεια να ξεπεραστούν τα παραπάνω προβλήματα. Η πιο απλή προσέγγιση ενεργειακού εντοπισμού των ηχητικών πηγών είναι η μέθοδος του «πλησιέστερου αισθητήρα» (*CPA*). Έτσι, αν χρησιμοποιηθεί ένα πλέγμα αισθητήρων λήψης που να καλύπτει τις απαιτήσεις διακριτικής ικανότητας (*spatial resolution*) της εκάστοτε εφαρμογής, τότε η θέση της πηγής εκτιμάται, πολύ απλά, από τις συντεταγμένες του αισθητήρα με τη μεγαλύτερη ενέργεια [158], [248]-[250]. Βέβαια, η χρήση του νόμου του αντιστρόφου τετραγώνου *ISL*, που ισχύει στις περισσότερες περιπτώσεις, επιτρέπει τον προσδιορισμό της ηχητικής θέσης με μεγαλύτερη ακρίβεια και χωρίς την ανάγκη τόσο πολλών μικροφώνων. Η λύση ανάγεται και πάλι στο μηδενισμό της κατάλληλης συνάρτησης σφάλματος, που προκύπτει από το αντίστοιχο σύστημα εξισώσεων. Εάν το μοντέλο ηχητικής απόσβεσης είναι διαφορετικό του νόμου *ISL* (π.χ. στην περίπτωση γραμμικών ηχητικών πηγών, ή σε πολύ μικρές αποστάσεις από την πηγή - παρουσία αντηχητικού πεδίου [108]-[109]), τότε θα πρέπει να προηγηθεί η εκλογή του κατάλληλου

μοντέλου διάδοσης της ηχητικής ενέργειας. Τέλος, στην περίπτωση ταυτόχρονης παρουσίας πολλών ηχητικών πηγών, απαιτούνται διαφοροποιήσεις ως προς τη μοντελοποίηση του προβλήματος και την κατάσταση των εξισώσεων επίλυσης [158], [248]-[250], [284]-[285], ενώ υπάρχουν και εφαρμογές, όπου η διαδικασία εντοπισμού εξελίσσεται παράλληλα σε πολλές περιοχές χρόνου - συχνότητας (time-frequency localization) [295]. Γενικά, οι τεχνικές ενεργειακού εντοπισμού είναι ιδιαίτερα απλές στη σύλληψη και διαθέτουν αρκετά πλεονεκτήματα. Αυτός είναι και ο λόγος που αξιοποιήθηκαν σε σχετικές εργασίες ακουστικής χαρτογράφησης της γαστρεντερικής λειτουργίας [42], [87], ενώ έχουν αξιοποιηθεί και για της ανάγκες της παρούσας διατριβής.

6.1.2 Χωρική καταγραφή και απεικόνιση ηχητικού πεδίου

Υπάρχουν πολλές περιπτώσεις όπου το μοντέλο σημειακής και ισοτροπικής πηγής δεν προσφέρεται για την ανάλυση του ηχητικού πεδίου. Χαρακτηριστικές είναι οι εφαρμογές ακουστικής διάγνωσης μηχανικών και δομικών συστημάτων, οι τεχνικές απεικόνισης της ηχητικής ακτινοβολίας μεγαφώνων-ηχείων, οι προσεγγίσεις χωρικής ακουστικής ανάλυσης πηγών θορύβου, μουσικών οργάνων, κ.ά. [130]-[131], [284], [288]-[290], [296]-[303]. Στις περιπτώσεις αυτές απαιτείται η γνώση τόσο των ενεργειακών όσο και των διανυσματικών χαρακτηριστικών διάδοσης του ηχητικού κύματος, που περιγράφονται πλήρως από την ένταση (Sound Intensity) [296]-[297]. Για το λόγο αυτό, σε πολλές εφαρμογές αυτού του είδους αξιοποιείται μετρητικός εξοπλισμός που επιτρέπει τον προσδιορισμό της ηχητικής έντασης (intensity probes/meters), η οποία παρέχει τη δυνατότητα πλήρους τοπογραφικής ανάλυσης [296]-[297]. Ωστόσο, η ποιο κλασική προσέγγιση χωρικής ανάλυσης ήχου (με τα χαρακτηριστικά που προαναφέρθηκαν) αφορά στη μέθοδο ακουστικής ολογραφίας εγγύς πεδίου (NAH) [298]. Στην περίπτωση αυτή, ορίζεται ένα διοδιάστατο πλέγμα μικροφώνων που τοποθετούνται σε μικρή απόσταση από το ηχογόνο αντικείμενο (X - Y επίπεδο, $Z=0$), από όπου προκύπτει η «εικόνα» ηχητικής ακτινοβολίας για διάφορους κυματαριθμούς (για διάφορες συχνότητες εκπομπής). Έτσι, από τη χρήση του διοδιάστατου μετασχηματισμού *Fourier* είναι δυνατόν να υπολογιστεί ο αντίστοιχος χάρτης ηχητικής κατανομής σε κάθε επίπεδο παράλληλο προς το επίπεδο στερέωσης των αισθητήρων ($Z>0$: απομάκρυνση από την πηγή, $Z<0$: προσέγγιση προς την πηγή). Εξαιτίας του μετασχηματισμού *Fourier*, οι παραπάνω τεχνικές είναι γνωστές με τον όρο "*Fourier Acoustics*" [288]-[290], [298]-[303].

Το κύριο μειονέκτημα των τεχνικών ακουστικής ολογραφίας έγκειται στη χρήση ενός σχετικά μεγάλου αριθμού μικροφώνων, ενώ υπάρχουν και αρκετοί περιορισμοί ως προς το μέγεθος και την αναλυτικότητα του «πλέγματος ηχητικής λήψης», ανάλογα και με τα μήκη κύματος του εκπεμπόμενου ηχητικού πεδίου. Στο πλαίσιο αυτό αναπτύχθηκαν εξελιγμένες αριθμητικές τεχνικές επίλυσης της εξίσωσης ακουστικού κύματος (wave equation), οι οποίες ξεκινούν από μια βασική, παραμετρική μορφή της κυματικής εξίσωσης Helmholtz, και στη συνέχεια προσδιορίζουν τη λύση από τις γνωστές ακουστικές πιέσεις των αισθητήρων (με ελαχιστοποίηση της κατάλληλης συνάρτησης σφάλματος). Ποιο γνωστές είναι οι μέθοδοι *HELS* (Helmholtz Least Squares) [299] και *SONAH* (Statistically Optimal NAH) [302]-[303]. Οι αριθμητικές, αυτές, τεχνικές επιτρέπουν τη χαρτογράφηση αρκετά περίπλοκων ηχητικών πηγών, χωρίς να απαιτείται ιδιαίτερα μεγάλος αριθμός μικροφώνων (μεγαλύτερος, ωστόσο, από τους 4-9 αισθητήρες που δοκιμάστηκαν στην παρούσα εφαρμογή). Επίσης, υπάρχουν απαιτήσεις γνώσης των χαρακτηριστικών του μέσου διάδοσης, που βρίσκεται στην πλευρά του «επιπέδου μέτρησης», το οποίο θεωρείται ομογενές. Στην περίπτωση που τα ακουστικά φαινόμενα δεν είναι στάσιμα κατά την εξέλιξή τους στο χρόνο, είναι δυνατή η εξαγωγή

διαδοχικών πλαισίων χαρτογράφησης (video frames), που περιγράφουν τις χωρο-χρονικές μεταβολές της ηχητικής ενέργειας [300]-[303]. Τα στοιχεία αυτά αποτέλεσαν κίνητρο για την ανάπτυξη αντίστοιχων εργαλείων εικονικής αναπαραγωγής-βίντεο της γαστρεντερικής ακουστικής δραστηριότητας (σχετική ανάλυση γίνεται στη συνέχεια του κεφαλαίου).

6.1.3 Τεχνικές ηχογράφησης και αναπαραγωγής στερεοφωνικού - περιφερικού ήχου

Εκτός των παραπάνω προσεγγίσεων, οι τεχνικές χωρικής καταγραφής και επεξεργασίας ήχου έχουν χρησιμοποιηθεί εκτενώς και στις διάφορες εφαρμογές στέρεο-ηχογράφησης, και τα αντίστοιχα πρότυπα περιφερικού ήχου (surround). Στις περιπτώσεις αυτές δεν ενδιαφέρει τόσο ο εντοπισμός των πηγών και η απεικόνιση της ενεργειακής κατανομής του ηχητικού πεδίου, αλλά κυρίως η επίτευξη υψηλής πιστότητας κατά την αναπαραγωγή (High Fidelity - HiFi). Δηλαδή η δημιουργία συνθηκών ρεαλιστικής «τριδιάστατης ακρόασης», ανάλογα με τις αντίστοιχες συνθήκες που επικρατούσαν κατά την ηχογράφηση. Οι τεχνικές περιφερικού ήχου κερδίζουν συνεχώς έδαφος τις τελευταίες δεκαετίες και αξιοποιούνται σε εφαρμογές ψυχο-ακουστικής, σε διαδραστικά περιβάλλοντα εικονικής πραγματικότητας και κυρίως στη βιομηχανία του οπτικοακουστικού θεάματος [95], [109], [160], [188], [304]-[306].

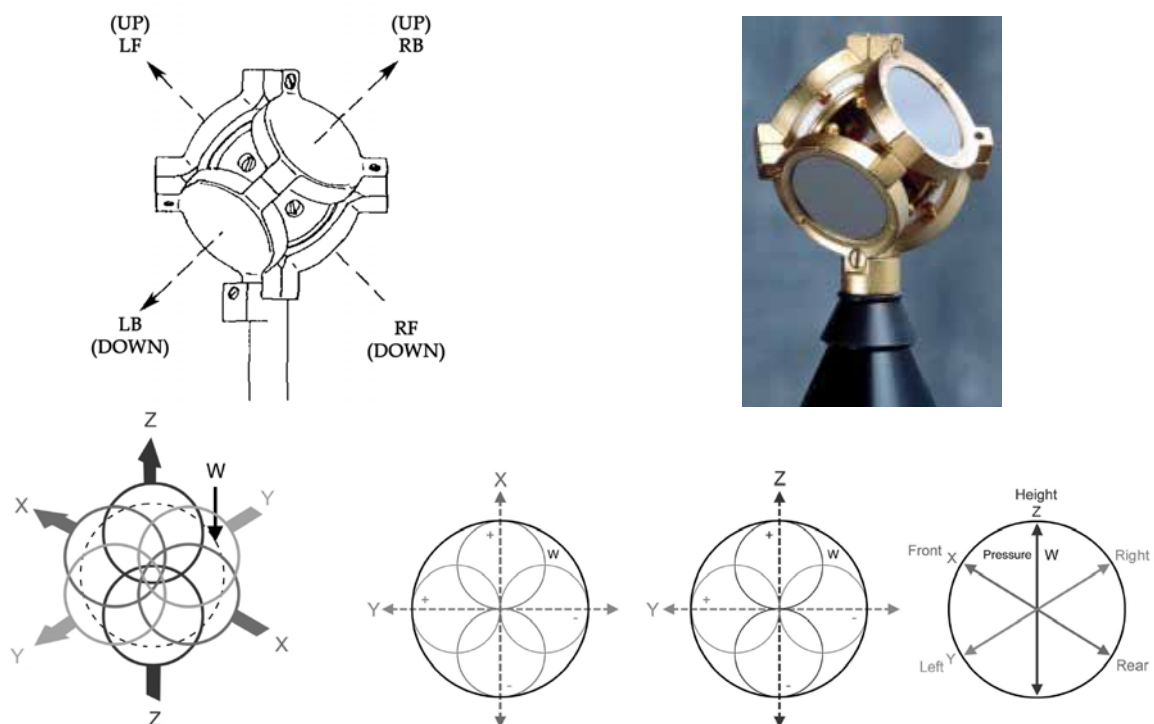
Το πρώτο σύστημα χωρικής καταγραφής ήχου υπήρξε η κλασσική στερεοφωνία, η οποία επιχείρησε να δώσει στον ακροατή την αίσθηση κατανομής του ήχου στο χώρο [160]. Ο όρος στερεοφωνία προκύπτει από τα συνθετικά «στέρεο» και «φωνή», που ετυμολογικά σημαίνει ότι το αντίστοιχο ηχητικό πεδίο έχει τριδιάστατη υπόσταση. Με άλλα λόγια, στόχος της στερεοφωνίας είναι η απόδοση της τριδιάστατης φύσης του ήχου στο χώρο, με τη μεγαλύτερη δυνατή πιστότητα. Σήμερα έχει επικρατήσει να χρησιμοποιούμε τον όρο αυτό για τα δικάναλικά συστήματα ηχογράφησης - αναπαραγωγής, καθώς οι πρώτες ερευνητικές προσπάθειες βασίστηκαν στη χρήση δύο μικροφώνων και δύο ηχείων, αντίστοιχα [160], [305]-[306]. Ο λόγος αξιοποίησης δικάναλων συστημάτων σχετίζεται με την αμφιωτική φύση της ανθρώπινης ακοής, με το γεγονός δηλαδή ότι ο άνθρωπος χρησιμοποιεί δύο αισθητήρια όργανα για την αντίληψη του ήχου στο χώρο [160], [304]-[306].

Ειδικότερα, ο χωρικός εντοπισμός των ακουστικών ερεθισμάτων γίνεται από τις φασικές και σταθμικές διαφορές (Interaural Time Difference -ITD, Interaural Level Difference -ILD) μεταξύ των δύο αυτιών. Έχει διαπιστωθεί ότι ο πρώτος μηχανισμός (ITD) κυριαρχεί σε χαμηλές, σχετικά, συχνότητες (<700 Hz) και ο δεύτερος (ILD) σε υψηλότερες (>700 Hz), κανόνας που είναι γνωστός ως "*Makita Theory*" [160], [307]. Όπως έχει ήδη επισημανθεί στις προηγούμενες παραγράφους, οι χρονικές και ενεργειακές αυτές διαφορές αξιοποιούνται και στις αριθμητικές τεχνικές ηχητικού εντοπισμού. Ωστόσο, η χρήση δύο μόνο αισθητήρων λήψης δεν είναι ιδιαίτερα αποδοτική εν προκειμένω, γεγονός που στην περίπτωση «ηχητικού προσανατολισμού» του ανθρώπου μπορεί να αντισταθμιστεί με τις κινήσεις του κεφαλιού [305]. Για το λόγο αυτό, οι τεχνικές στερεοφωνίας γρήγορα οδηγήθηκαν στη χρήση περισσότερων ηχητικών καναλιών, με αποτέλεσμα τα συστήματα τριφωνίας, τετραφωνίας (Quadraphonic) και τα πιο σύγχρονα πρότυπα περιφερικού ήχου *surround* [160], [305]-[308].

Από τα παραπάνω, οι τετραφωνικές τεχνικές έχουν χρησιμοποιηθεί σε αρκετά μεγάλο βαθμό, κυρίως λόγω της «εύκολης» τετραγωνικής - περιφερειακής διάταξης των μεγαφώνων αναπαραγωγής (Quads), ενώ παράλληλα ιδιαίτερα βολική υπήρξε και η ανάπτυξη των τετρακάναλων συστημάτων μαγνητικής εγγραφής και αναπαραγωγής [160]. Παράλληλα, αναπτύχθηκαν διάφορες τεχνικές μετατροπής των τεσσάρων συνιστωσών σε σύστημα κλασσικής στερεοφωνίας δύο καναλιών (Left-Right) για λόγους συμβατότητας προς τα πίσω (backward compatibility), και αντίστροφα, οι μετατροπές για την κωδικοποίηση δικάναλων

ηχογραφήσεων σε τετρακάναλες. Σε εξέλιξη των παραπάνω συστημάτων σχεδιάστηκαν τα πρότυπα περιφερικού ήχου ακόμη περισσότερων καναλιών, με πιο χαρακτηριστική τη διάταξη “5-1”, η οποία περιλαμβάνει πέντε κανάλια πλήρους ακουστικού φάσματος και ένα κανάλι χαμηλών συχνοτήτων. Κατά το σχεδιασμό και την τυποποίηση όλων αυτών των συστημάτων, ιδιαίτερα μέριμα δόθηκε για την εύκολη μετάβαση από το ένα πολυκαναλικό σύστημα στο άλλο, καθώς και για την εξοικονόμηση αποθηκευτικού χώρου κατά την εγγραφή και μετάδοση των σημάτων. Παράλληλα, τέθηκαν οι κανόνες πολυκαναλικής ηχογράφησης, οι οποίοι προσαρμόστηκαν στις τεχνικές προδιαγραφές των αντίστοιχων προτύπων αναπαραγωγής [160], [304]-[308].

Στο πλαίσιο αυτό, ιδιαίτερο ενδιαφέρον παρουσιάζουν τα μικρόφωνα ηχητικού πεδίου (soundfield microphones) και οι τεχνικές αμβιοφωνίας (ambisonics), που αναπτύχθηκαν από τον Gerzon [307] στην προσπάθεια πλήρους σφαιρικής καταγραφής και αναπαραγωγής του ηχητικού πεδίου, γνωστή και ως «περιφωνία» (*periphony*) [160]. Το μικρόφωνο ηχητικού πεδίου περιλαμβάνει τέσσερις μεμβράνες υπο-καρδιοειδούς πολικού διαγράμματος, που τοποθετούνται στις κορυφές ενός κανονικού τετραέδρου (σχ. 6-1). Έτσι, προκύπτουν οι συνιστώσες W , X , Y , Z , όπου η πρώτη αντιστοιχεί στο σήμα ακουστικής πίεσης (pressure), που θα λάμβανε ένα πανκατευθυντικό μικρόφωνο στο κέντρο συμμετρίας του τετραέδρου, ενώ οι αξονικές X , Y , Z συνιστώσες αντιστοιχούν στα σήματα κλίσης της πίεσης (gradient pressure), που θα λαμβανόταν από ισχυρά κατευθυντικά μικρόφωνα ταχύτητας (με πολικό διάγραμμα “8”). Με τον τρόπο αυτό παρέχονται όλες οι «διανυσματικές» πληροφορίες προέλευσης της ηχητικής διέγερσης, ενώ το σήμα W χρησιμοποιείται ως αναφορά, απ’ όπου προκύπτουν οι πολικότητες των X , Y , Z [160], [306]-[311].



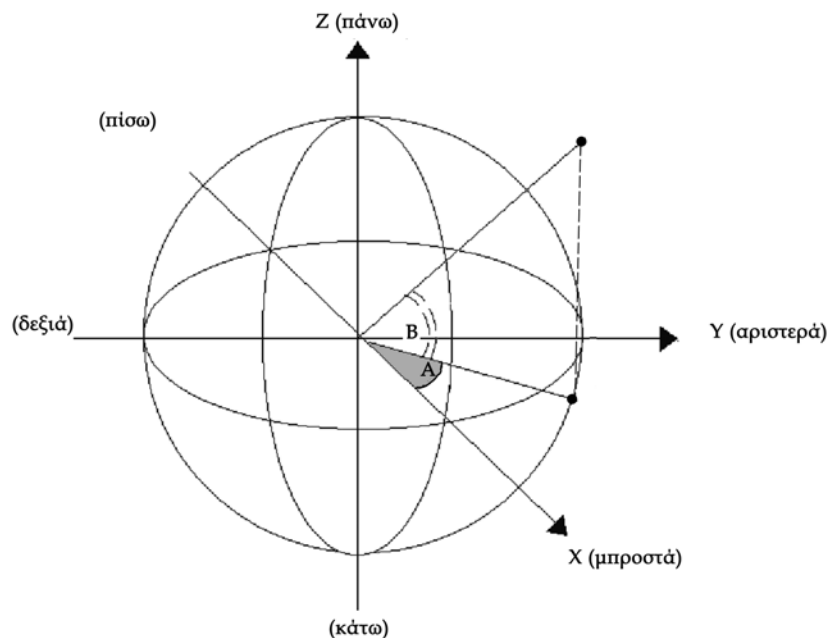
Σχήμα 6-1. Κατασκευή και πολικά διαγράμματα του μικροφώνου ηχητικού πεδίου (soundfield microphone).

Το πρότυπο “B-format”, όπως είναι γνωστό το σύστημα των τεσσάρων συνιστωσών W , X , Y , Z , παρέχει δυνατότητες χωρικής ανάλυσης, χωρίς ωστόσο να ενδιαφέρει τόσο ο εντοπισμός της ηχητικής πηγής, αλλά κυρίως η δημιουργία της κατάλληλης «στερεοφωνικής εικόνας». Στην περίπτωση αυτή, το σύστημα λήψης βρίσκεται στο κέντρο μιας υποτιθέμενης

μοναδιαίας σφαίρας (σχ. 6-2), ενώ από την επεξεργασία των τεσσάρων συνιστωσών είναι δυνατός ο υπολογισμός της διεύθυνσης του ηχητικού πεδίου (χωρίς να είναι δυνατή η εύρεση της ακριβούς θέσης της ηχητικής διέγερσης). Έτσι, αν A είναι η οριζόντια γωνία στο XY επίπεδο (σχ. 6-2) και B η γωνία ανύψωσης (elevation angle), τότε ισχύουν τα ακόλουθα:

$$\begin{aligned} X &= LF - RB + RF - LB = \cos A \cdot \cos B \cdot p_s \\ Y &= LF - RB - RF + LB = \sin A \cdot \cos B \cdot p_s \\ Z &= LF - LB - RB - RF = \sin B \cdot p_s \\ W &= LB + LF + RF + RB = 0,707 \cdot p_s \end{aligned} \quad (6-1)$$

όπου οι δείκτες L (Left), R (Right), F (Front) και B (Back) χρησιμοποιούνται για να δηλώσουν τον εκάστοτε μετατροπέα του μικροφώνου soundfield (αριστερά-δεξιά και μπροστά-πίσω), ο όρος p_s αναφέρεται στην ηχητική πίεση του σήματος εισόδου επάνω στη μοναδιαία σφαίρα, ενώ η πολλαπλασιαστική σταθερά $0,707$ χρησιμοποιείται για την ομοιόμορφη κατανομή της ηχητικής ενέργειας στις τέσσερις συνιστώσες [306]-[307], [309].



Σχήμα 6-2. Χωρική ανάλυση ηχητικού πεδίου με τη βοήθεια σφαιρικών αρμονικών (B-format).

Για τη σωστή αναπαραγωγή του ηχητικού πεδίου από τις συνιστώσες W , X , Y , Z , έχουν προταθεί διάφορες τεχνικές αποκωδικοποίησης και αντίστοιχα σχήματα διάταξης των ηχητικών πηγών [307]-[308]. Παράλληλα, για λόγους συμβατότητας έχουν αναπτυχθεί κανόνες μετατροπής από και προς τα υπόλοιπα πρότυπα στερεοφωνίας (L-R stereo, Quads, surround) [160], [306]-[311]. Ένα πλεονέκτημα του προτύπου B-format είναι η δυνατότητα κυκλικής μετατόπισης του συστήματος συντεταγμένων ή περιστροφής γύρω από κάποιο άξονα, με απλές αλγεβρικές πράξεις. Το στοιχείο αυτό είναι ιδιαίτερα χρήσιμο σε εφαρμογές ακουστικής αναπαραγωγής τρισδιάστατου ήχου (3D audio) και σε περιβάλλοντα εικονικής πραγματικότητας [304], [306]-[307]. Όπως θα δούμε στη συνέχεια, οι τεχνικές αμβιοφωνίας και τα μικρόφωνα ηχητικού πεδίου αποτέλεσαν κίνητρο για την ανάπτυξη μιας νέας μεθόδου χωρικής ανάλυσης της γαστρεντερικής ακουστικής δραστηριότητας, σε συνδυασμό με τα αντίστοιχα εργαλεία «τριδιάστατης» ακρόασης των βιοακουστικών φαινομένων.

6.2 Τοπογραφική ανάλυση γαστρεντερικού βιοακουστικού πεδίου

Η διαδικασία τοπογραφικής ανάλυσης του γαστρεντερικού βιοακουστικού πεδίου είναι ιδιαίτερα σημαντική, καθώς παρέχει πληροφορίες για τις εναλλαγές δραστηριοποίησης των διαφόρων περιοχών του πεπτικού σωλήνα [42], [69], [72]-[73], [87]-[88], [217]. Γι' αυτό και κατά τη μελέτη των γαστρεντερικών βιοακουστικών καταγραφών πολλοί ερευνητές προτείνουν την αξιοποίηση «χωρικών» παραμέτρων (features) [88]. Η πιο απλή προσέγγιση τοπογραφικής ανάλυσης συνίσταται στη χρήση πολλαπλών αισθητήρων και την εφαρμογή του κανόνα πλησιέστερης θέσης -CPA [69], [72], [88]. Στο ίδιο μήκος κύματος συστάθηκε και ο περιγραφέας σημαντικότητας καναλιών IMPind [217], που παρουσιάστηκε στο τέταρτο κεφάλαιο. Παράλληλα, έχουν γίνει προσπάθειες χαρτογράφησης του «κοιλιακού» βιοακουστικού πεδίου, μέσω της εξαγωγής εικόνων χωρικής κατανομής της ηχητικής ενέργειας [42], [87]. Έτσι, από τον εντοπισμό των θέσεων ηχητικής διέγερσης είναι δυνατό να εξαχθούν χρωματικοί χάρτες (*colormaps*), κατά αντιστοιχία με τα γραφήματα ακουστικής ολογραφίας. Οι εικόνες αυτές παρέχουν νέες δυνατότητες μελέτης των βιοακουστικών φαινομένων, τόσο για την «εικονική αποτύπωση» των μεμονωμένων συμβάντων, όσο και για την παρατήρηση των καταγραφών μεγάλης διάρκειας, μέσω τεχνικών χωρικής ολοκλήρωσης - σύνοψης (*video integration – summarization*). Η παρουσίαση, δε, του μηχανισμού γαστρεντερικής λειτουργίας μέσω των χαρτών κατανομής της ηχητικής στάθμης είναι αρκετά λειτουργική, καθώς το ιατρικό προσωπικό είναι περισσότερο εξοικειωμένο με την ανάλυση εικόνων.

Όπως έχει ήδη επισημανθεί, τα πιο σημαντικά προβλήματα των τεχνικών κοιλιακής χαρτογράφησης σχετίζονται με την πολυπλοκότητα του φαινομένου, την εγγύτητα των αισθητήρων στα σημεία ηχητικής διέγερσης και τους άγνωστους μηχανισμούς μετάδοσης των ηχητικών κυμάτων στο εσωτερικό του ανθρώπινου οργανισμού. Από τη βιβλιογραφική αναζήτηση που πραγματοποιήθηκε, διαπιστώθηκε ότι δεν έχουν γίνει πολλές ερευνητικές εργασίες μοντελοποίησης του μυϊκού ιστού και ειδικότερα των κοιλιακών τοιχωμάτων, ως μέσο ακουστικής διάδοσης. Οι περισσότερες μελέτες στον τομέα αυτό εστιάζουν στις τεχνικές υπερηχογράφησης και όχι στα ζητήματα μετάδοσης των ακουστών συχνοτήτων [312]. Αν δεχτούμε ότι η μέση ταχύτητα μετάδοσης του ήχου στους μυϊκούς ιστούς είναι 1.500 m/sec περίπου [87], [312]-[313], τότε τα αντίστοιχα μήκη κύματος ($\lambda=c/f$) του ακουστού φάσματος (20Hz-20kHz) είναι αρκετά μεγάλα, πόσο μάλλον για τις χαμηλές συχνότητες των κοιλιακών ήχων. Έτσι, με δεδομένο το αρκετά λεπτό πάχος των μυϊκών και επιδερμικών στρωμάτων, ίσως να μην έχει μεγάλο νόημα να εξετάζουμε τα κυματικά φαινόμενα ηχητικής μετάδοσης (σκέδαση, απορρόφηση, διάθλαση, κλπ). Στην ορθότητα της υπόθεσης αυτής συνηγορούν και τα αποτελέσματα της εργασίας των Ranta κ.ά. (2003) [87], όπου διαπιστώθηκαν μικρές διαφορές εντοπισμού, από τη χρήση «απορροφητικών» και «μη απορροφητικών» μοντέλων ηχητικής διάδοσης (σχετικές λεπτομέρειες δίνονται στη συνέχεια).

Η πρώτη εργασία χαρτογράφησης της γαστρεντερικής λειτουργίας έγινε από τους Craine κ.ά. (2001) [42]. Η συγκεκριμένη προσέγγιση ήταν ιδιαίτερα απλή στη σύλληψη, θεωρώντας τη δισδιάστατη φύση του προβλήματος (σχ. 6-3α). Έτσι, αν θεωρηθεί ότι η πηγή ηχητικής διέγερσης και οι αισθητήρες λήψης βρίσκονται στο ίδιο επίπεδο, τότε το πρόβλημα εντοπισμού έχει μοναδική λύση που μπορεί να βρεθεί με τη χρήση τριών αισθητήρων [248], από την εφαρμογή τεχνικών «τριγωνοποίησης» (*triangularization*) [42]. Στο πλαίσιο αυτό, θεωρήθηκε ότι το μέσο βιοακουστικής διάδοσης είναι ομογενές, ενώ η μοναδική απώλεια οφείλεται στο «μοίρασμα» ηχητικής ενέργειας, όπως εκφράζεται από το νόμο αντιστρόφου τετραγώνου. Επίσης, οι θέσεις ηχητικής διέγερσης θεωρήθηκαν σημειακές ιστροπικές πηγές, όπως ιστροπικά θεωρήθηκαν και τα κατευθυντικά χαρακτηριστικά των στηθοσκοπιών

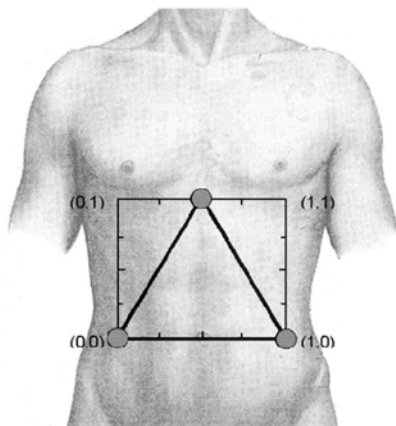
λήψης. Αν και η εργασία αυτή υποστήριξε ότι οι ακουστικοί χάρτες είναι σε θέση να διαχωρίσουν τους «φυσιολογικούς» από τους «δυσλειτουργικούς» μηχανισμούς εντερικής κινητικότητας [42], ωστόσο, όπως σωστά επισημάνθηκε στην εργασία των Ranta κ.ά. [87], δεν υπάρχει δυνατότητα ελέγχου της ορθότητας των αποτελεσμάτων ηχητικού εντοπισμού. Η τελευταία εργασία αποτέλεσε και τη δεύτερη προσπάθεια εξαγωγής «ακουστικών - κοιλιακών χαρτών» (“acoustic - abdominal maps” [87]), η οποία έδωσε μεγαλύτερη έμφαση στην τεχνική, παρά την ιατρική πλευρά του προβλήματος. Έτσι, υπήρξε σκεπτικισμός σχετικά με το μοντέλο ηχητικής διάδοσης, γι’ αυτό και εκτός από το ιδανικό, μη απορροφητικό, μέσο, δοκιμάστηκε και ένα μοντέλο ακουστικής απορρόφησης [87], [313]:

$$\|I_\zeta\| = k_p \cdot \frac{P_{W_s}}{4\pi R_\zeta^2}, \quad \text{μοντέλο χωρίς απορρόφηση (non-absorbent)}$$

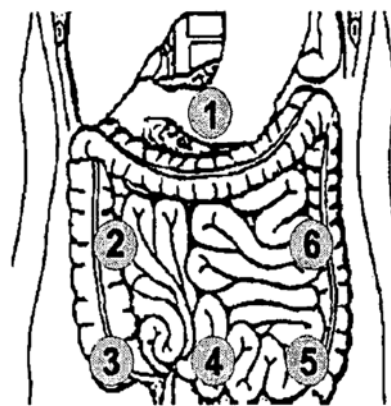
$$\|I_\zeta\| = P_{W_s} \frac{e^{-\mu_d \cdot R_\zeta}}{4\pi R_\zeta^2}, \quad \mu_d = \alpha_d \cdot f^{\beta_d}, \quad \text{μοντέλο με απορρόφηση (absorbent)} \quad (6-2)$$

$$F_{cost}(x, y, z) = \sum_{\substack{\zeta, \xi=1 \\ \zeta \neq \xi}}^6 (P_{W_s, \zeta} - P_{W_s, \xi})^2 = \left(\frac{4\pi}{k_p}\right)^2 \cdot \sum_{\substack{\zeta, \xi=1 \\ \zeta \neq \xi}}^6 (\|I_\zeta\| R_\zeta^2 - \|I_\xi\| R_\xi^2)^2$$

όπου $\|I_\zeta\|$ είναι το μέτρο του διανύσματος ηχητικής έντασης που φθάνει σε κάθε αισθητήρα ζ , P_{W_s} η ακουστική ισχύς της μοναδικής (κάθε φορά) ηχητικής διέγερσης, R_ζ η απόσταση του ζ -αισθητήρα από την πηγή, ενώ οι σταθερές μ_d , α_d και β_d εκφράζουν την ηχητική απόσβεση (decay) λόγω απορρόφησης, συναρτήσει της συχνότητας f [87], [313]. Από την εξ. (6-2γ) προκύπτει ότι τόσο η σταθερά μετάδοσης k_p (propagation constant) όσο και η σταθερά απορρόφησης μ_d δεν επηρεάζουν τη διαδικασία ηχητικού εντοπισμού, η οποία βασίζεται στην ελαχιστοποίηση της συνάρτησης κόστους F_{cost} [87].



α. Craine et al. 2002 [42]



β. Ranta et al. 2003 [87]

Σχήμα 6-3. Τεχνικές ακουστικής χαρτογράφησης του γαστρεντερικού πεδίου με α. τρεις και β. έξι αισθητήρες.

Για τις ανάγκες ηχητικού εντοπισμού στην προκειμένη περίπτωση αξιοποιήθηκαν έξι θέσεις βιοακουστικής λήψης (σχ. 6-3β). Το πρόβλημα εντοπισμού έχει πλέον τριδιάστατη υπόσταση, σε αντίθεση με τη διδιάστατη προσέγγιση των Craine κ.ά. [42], απ’ όπου προκύπτουν και πάλι οι διδιάστατοι χρωματικοί χάρτες χωρικής κατανομής της ηχητικής ενέργειας. Είναι προφανές ότι στην πιο απλή περίπτωση που $k_p=1$, η εξ. (6-2α) εκφράζει το νόμο του αντιστρόφου τετραγώνου *ISL*, ενώ για να ισχύουν οι παραπάνω σχέσεις θα πρέπει

και τα στηθοσκοπία λήψης να έχουν σφαιρικό (πανκατευθυντικό) πολικό διάγραμμα. Η χρήση της συνάρτησης κόστους F_{cost} ερμηνεύεται πολύ απλά από το γεγονός ότι η ακουστική ισχύς $P_{W_s, \zeta}$ που εκτιμάται από κάθε αισθητήρα θα πρέπει να είναι η ίδια (ίση με την πραγματική τιμή P_{W_s}), οπότε οι διαφορές $|P_{W_s, \zeta} - P_{W_s}|$ πρέπει (ιδεατά) να μηδενίζονται. Από τα πειραματικά αποτελέσματα της παραπάνω εργασίας διαπιστώθηκε ότι το μέσο ελάχιστο σφάλμα εντοπισμού παρουσιάζει σχετικά μεγάλες τιμές, με αρκετά έντονες διακυμάνσεις. Οι αιτιάσεις του γεγονότος αυτού εντοπίστηκαν στην αδυναμία των δύο παραπάνω μοντέλων να αποδώσουν την πολυπλοκότητα του φαινομένου, ενώ όπως έχει ήδη αναφερθεί, τα δύο μοντέλα έδωσαν παραπλήσια αποτελέσματα. Έτσι, η προσπάθεια χαρτογράφησης από τη συγκεκριμένη ερευνητική ομάδα εγκαταλείφθηκε ως αρκετά δύσκολο εγχείρημα [87].

6.2.1 Η προτεινόμενη μέθοδος βιοακουστικής χαρτογράφησης

Η προτεινόμενη μέθοδος γαστρεντερικής - ακουστικής - χαρτογράφησης, βασίστηκε στα αποτελέσματα των παραπάνω εργασιών και επιχείρησε να δώσει ένα δεύτερο εργαλείο τοπογραφικής ανάλυσης, εκτός από τον περιγραφέα IMP_{ind} . Στο πλαίσιο αυτό, εκτιμήθηκε ιδιαίτερα η χρησιμότητα των «χαρτών» χωρικής κατανομής της ηχητικής ενέργειας, καθώς και η δυνατότητα οπτικοακουστικής απόδοσης των συμβάντων μέσα από την εναλλαγή εικόνων ακουστικής ολογραφίας (παράλληλα με την αναπαραγωγή ήχου). Από τις παρατηρήσεις των Ranta κ.ά. [87] κατέστη σαφές ότι το εγχείρημα χωρικού εντοπισμού των γαστρεντερικών ηχητικών διεγέρσεων είναι αρκετά δύσκολο. Ωστόσο, το κύριο ερώτημα που θα πρέπει να απαντηθεί αφορά, αφενός στις αιτίες των σφαλμάτων εντοπισμού, αφετέρου στις επιτρεπτές αποκλίσεις, δηλαδή τις απαιτήσεις ακρίβειας.

Σε γενικές γραμμές οι παρατηρήσεις του Ranta [87] είναι σωστές, αν και αξίζει να γίνουν κάποια πρόσθετα σχόλια. Έτσι, η κύρια αιτία σφαλμάτων δεν είναι τόσο η ανομοιογένεια του μέσου (για λόγους που αναφέρθηκαν πιο πάνω), όσο τα χαρακτηριστικά των αισθητήρων λήψης. Όπως αναλύθηκε διεξοδικά στο δεύτερο κεφάλαιο, τα στηθοσκοπία παρουσιάζουν σημαντικές αποκλίσεις από την ιδανική συμπεριφορά του αισθητήρα μηδενικών διαστάσεων, ενώ σε καμιά περίπτωση δεν μπορούμε να ισχυριστούμε ότι λαμβάνουν με την ίδια ευαισθησία τους ήχους από όλες τις διευθύνσεις. Τα προβλήματα αυτά εμφανίζονται και στους πιεζοηλεκτρικούς αισθητήρες επαφής, αν και ο μηχανισμός λειτουργίας τους, σε συνδυασμό με το αρκετά μικρότερο μέγεθός τους, εξασφαλίζουν πολύ καλύτερη απόδοση. Σε κάθε περίπτωση, υπάρχουν διατάξεις υψηλών προδιαγραφών, όπως είναι τα μετρητικά επιταχυνσιόμετρα, που επιτρέπουν την ακριβή γνώση των πολικών διαγραμμάτων λήψης. Μια άλλη πιθανή αιτία σφαλμάτων είναι η χρήση του μοντέλου ISL , που είναι πιθανό να μην ισχύει σε τόσο κοντινές αποστάσεις από την ηχητική διέγερση. Η ομοιόμορφη σφαιρική μετάδοση, ωστόσο, που καθορίζει ο νόμος αντιστρόφου τετραγώνου, είναι η πιο απλή και συνάμα η πιο λογική εκδοχή. Το στοιχείο αυτό ενισχύεται περισσότερο από τα πολύ μεγάλα μήκη κύματος, ή αντίστοιχα από τις μικρές διαστάσεις της ηχητικής πηγής και από την απουσία «κουματικών εμποδίων». Έτσι, με βάση και την ανάλυση που θα ακολουθήσει, τα αναμενόμενα «σφάλματα μοντελοποίησης» είναι αρκετά μικρά και γι' αυτό εκτιμάται ότι βρίσκονται μέσα στα ανεκτά όρια ακρίβειας του όλου εγχειρήματος.

Όσον αφορά στη διακριτική ικανότητα, θα μπορούσε κανείς να ισχυριστεί ότι στις παρατηρήσεις του Ranta υπάρχει μια αντίφαση [87]. Έτσι, ενώ η διαδικασία χαρτογράφησης απορρίπτεται λόγω αμφιβολιών ως προς την ακρίβεια των αποτελεσμάτων, στη θέση της αντιπροτείνεται ο κανόνας *πλησιέστερης θέσης CPA*, που είναι προφανές ότι έχει ακόμη μικρότερη ακρίβεια. Στο πλαίσιο αυτό, αποφασίστηκε η χρήση εργαλείων χωρικού

εντοπισμού και απεικόνιση, λόγω της χρηστικότητας των ενεργειακών-ακουστικών χαρτών, ενώ δόθηκε ιδιαίτερη έμφαση στην εκτίμηση των αναμενόμενων σφαλμάτων. Είναι φανερό ότι η χωρική διακριτικότητα δεν μπορεί να είναι πολύ μεγαλύτερη από τις διαστάσεις των αισθητήρων λήψης, καθώς ακόμη και ο ορισμός της θέσης τους επιδρά στους αλγορίθμους εντοπισμού. Έτσι, είναι σχεδόν παράλογο να περιμένει κανείς την εύρεση της θέσης της ηχητικής πηγής με ακρίβεια της τάξης χιλιοστού και περισσότερο, όταν οι αισθητήρες έχουν διαστάσεις της τάξης εκατοστού, ή των μερικών εκατοστών στην περίπτωση στηθοσκοπιών. Σύμφωνα, δε, και με τις απόψεις της συνεργαζόμενης ιατρικής ομάδας, η επιδίωξη τέτοιων προδιαγραφών ακρίβειας δεν είναι μόνο ουτοπική, αλλά και άνευ σημασίας. Επομένως, η γνώση των προδιαγραφών ακρίβειας εντοπισμού κρίνεται ιδιαίτερα σημαντική, γι' αυτό και απαιτείται η σύσταση κατάλληλων πειραματικών διαδικασιών αξιολόγησης. Το εγχείρημα αυτό παρουσιάζει επίσης πολλές δυσκολίες (σύμφωνα με τον Ranta δεν υπάρχει δυνατότητα επαλήθευσης της εγκυρότητας των αποτελεσμάτων εντοπισμού) και αποτέλεσε ένα ακόμη αντικείμενο της διατριβής, που αντιμετωπίστηκε σε αρκετά ικανοποιητικό βαθμό.

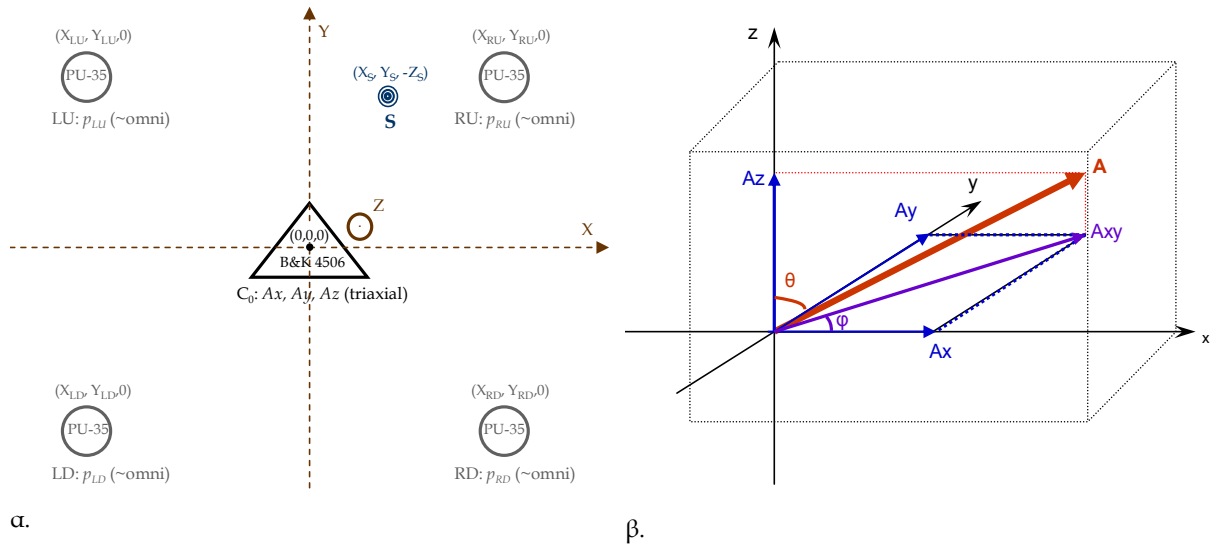
Γενικά, στην ακουστική χαρτογράφηση περίπλοκων φαινομένων επλέγονται τεχνικές αριθμητικής επίλυσης του αντίστροφου προβλήματος (Inverse Numerical Acoustics -INA) [289]-[290], όπως είναι οι μέθοδοι ακουστικής ολογραφίας *NAH*, *SONAH* και *HELIS*. Ωστόσο, οι τεχνικές αυτές απορρίφθηκαν στην προκειμένη περίπτωση για τους ακόλουθους λόγους: α) απαιτούν πολλούς αισθητήρες λήψης (αρκετά περισσότερους από τους τέσσερις έως έξι αισθητήρες που έχουν χρησιμοποιηθεί σε εφαρμογές κοιλιακών ήχων), β) έχουν μεγάλη πολυπλοκότητα και αντίστοιχες υπολογιστικές απαιτήσεις, ενώ η αποτελεσματικότητά τους δεν είναι δεδομένη, καθώς επηρεάζεται από την ελλιπή γνώση του μέσου διάδοσης, γ) οι σχετικά περιορισμένες απαιτήσεις ακρίβειας (εντοπισμού) επιτρέπουν τη χρήση πιο απλών μεθόδων. Με βάση όλα τα παραπάνω, αποφασίστηκε η χρήση του επιταχυνσιομέτρου τριών αξόνων Brüel & Kjaer 4506, με στόχο την εύρεση των διευθύνσεων άφιξης, κατά αντιστοιχία με τα μικρόφωνα ηχητικού πεδίου *Soundfield*. Θα περίμενε κανείς ότι, εκτός από τις αξονικές συνιστώσες X , Y , Z , χρειάζεται και ένας πανκατευθυντικός αισθητήρας (στο ίδιο σημείο) για το σήμα αναφοράς W . Στη θέση του προτιμήθηκαν οι τέσσερις «διαθέσιμοι» περιφερειακοί αισθητήρες του συστήματος *A-Soundfield 4+3*, για λόγους που αναλύονται στη συνέχεια.

Αν (X_s, Y_s, Z_s) είναι οι συντεταγμένες της ηχητικής πηγής και θεωρήσουμε ότι η θέση του επιταχυνσιομέτρου βρίσκεται στο κέντρο των αξόνων $C_0 (0,0,0)$, τότε το επίπεδο $Z=0$ αντιστοιχεί στην κοιλιακή επιφάνεια (σχ. 6-4α). Από τις συνιστώσες επιτάχυνσης A_x , A_y , A_z μπορούν να υπολογιστούν τα μέτρα των γωνιών διεύθυνσης ϕ , θ (σχ. 6-4β):

$$\begin{aligned} A_x &= A \cdot \cos(\phi) \cdot \sin(\theta) \\ A_y &= A \cdot \sin(\phi) \cdot \sin(\theta) \\ A_z &= A \cdot \cos(\theta) \end{aligned} \quad (6-3)$$

$$\left. \begin{aligned} \tan(\phi) &= \frac{A_y}{A_x} = \frac{L_y}{L_x} \\ \tan(\theta) &= \frac{\sqrt{A_x^2 + A_y^2}}{A_z} = \frac{\sqrt{L_x^2 + L_y^2}}{L_z} \end{aligned} \right\} \Rightarrow \begin{aligned} L_x(L_z) &= \frac{\tan(\theta) \cdot L_z}{\sqrt{1 + \tan^2(\phi)}} \\ L_y(L_z) &= \frac{\tan(\theta) \cdot L_z}{\sqrt{1 + \tan^2(\phi)}} \cdot \tan(\phi) \end{aligned} \quad (6-4)$$

όπου A , A_x , A_y , A_z είναι οι ενεργές τιμές των διανυσμάτων A , A_x , A_y , A_z , ενώ τα μήκη L_x , L_y , L_z προκύπτουν από τις συντεταγμένες της πηγής: $\{L_x, L_y, L_z\} = \{|X_s|, |Y_s|, |Z_s|\}$.



Σχήμα 6-4. Σχηματική αναπαράσταση του συστήματος λήψης *A-soundfield 4+3*: α. διάταξη των αισθητήρων (σε κάτοψη), β. γωνίες των διανυσμάτων διεύθυνσης - επιτάχυνσης (σε προοπτική).

Είναι προφανές ότι δεν αρκεί το μέτρο των γωνιών φ , θ αλλά απαιτείται η αλγεβρική τους τιμή που εκτός από τη σχέση των μηκών L_x , L_y ως προς το L_z θα προσδιορίσει και την τελική διεύθυνση άφιξης. Αυτός είναι και ο λόγος που χρησιμοποιείται το σήμα αναφοράς W στο μικρόφωνο *soundfield*, όπου οι φασικές σχέσεις μεταξύ των συνιστωσών καθορίζουν την προέλευση του ηχητικού κύματος. Στην προκειμένη περίπτωση είναι γνωστό ότι η θέση της ηχητικής πηγής βρίσκεται στον ημιχώρο $Z < 0$, δηλαδή το διάνυσμα A_z έχει πάντα θετικό πρόσημο ($A_z > 0$). Έτσι, από τις φασικές συγκρίσεις των σημάτων επιτάχυνσης $A_x(i)$, $A_y(i)$ και $A_z(i)$ είναι δυνατό να υπολογιστούν οι πολικότητες των A_x , A_y :

$$\mathbf{A}_x = \begin{cases} +A_x \cdot \mathbf{e}_x, & \text{αν } mse\{A_{nx}(i) - A_{nz}(i)\} < mse\{A_{nx}(i) - (-A_{nz}(i))\} \\ -A_x \cdot \mathbf{e}_x, & \text{διαφορετικά} \end{cases} \quad (6-5)$$

$$\mathbf{A}_y = \begin{cases} +A_y \cdot \mathbf{e}_y, & \text{αν } mse\{A_{ny}(i) - A_{nz}(i)\} < mse\{A_{ny}(i) - (-A_{nz}(i))\} \\ -A_y \cdot \mathbf{e}_y, & \text{διαφορετικά} \end{cases}$$

όπου $A_{nx}(i)$, $A_{ny}(i)$, $A_{nz}(i)$ είναι κανονικοποιημένες εκδοχές των σημάτων $A_x(i)$, $A_y(i)$, $A_z(i)$, έτσι ώστε να έχουν ίδιο πλάτος και φασικές μόνο διαφορές, $\mathbf{e}_x = (1, 0, 0)$ και $\mathbf{e}_y = (0, 1, 0)$ είναι τα ορθομοναδιαία διανύσματα των αξόνων X , Y , και mse είναι ο τελεστής του μέσου τετραγωνικού σφάλματος (*mean square error*). Η λογική των παραπάνω σχέσεων είναι πολύ απλή, καθώς εξετάζουν εάν τα σήματα $A_x(i)$, $A_y(i)$ είναι πιο πιθανό να έχουν ίδια, ή ανάστροφη, φάση με το $A_z(i)$. Βέβαια, όπως έχει επισημανθεί στο τρίτο κεφάλαιο, υπάρχει πάντα η παρουσία του προσθετικού θορύβου που καθιστά την παραπάνω διαδικασία αρκετά πιο δύσκολη στην πράξη. Έτσι, οι εξ. (6-5) είναι επιρρεπείς σε σφάλματα (ειδικά σε χαμηλές σηματοθορυβικές σχέσεις), ενώ ακόμη και πολύ μικρές φασικές διαφορές, που προκαλεί η αποθορυβοποίηση, ενδέχεται να αλλοιώσουν το τελικό αποτέλεσμα. Στο πλαίσιο αυτό αναζητήθηκε και μια δεύτερη, πιο απλή λύση προσδιορισμού του τεταρτημορίου της ηχητικής διέγερσης (στο XY επίπεδο), χρησιμοποιώντας τα σήματα των περιφερειακών αισθητήρων:

$$\varphi \in \varphi_{quarter} : quarter = \arg \max_{\zeta} (\|x_{\zeta}(i)\|), \quad \zeta = LU, RU, RD, LD \quad (6-6)$$

Όπως θα αναλυθεί στη συνέχεια κατά την παρουσίαση των πειραματικών διαδικασιών αξιολόγησης, η δεύτερη μέθοδος προσδιορισμού της διεύθυνσης του ηχητικού κύματος είναι πιο αποδοτική από την αντίστοιχη της εξ. (6-5), τόσο ως προς τα αποτελέσματα όσο και ως προς το υπολογιστικό κόστος. Έχοντας γνωστές τις γωνίες διεύθυνσης φ , θ , η πιο απλή λύση εύρεσης της θέσης της ηχητικής πηγής είναι να θεωρηθεί ότι όλες οι «σπειρές» του εντέρου, (δηλαδή όλες οι εν δυνάμει ηχητικές πηγές) βρίσκονται κατά προσέγγιση στο ίδιο επίπεδο (π.χ. $Z=-1cm$ ή $Z=-3cm$). Αυτή ήταν και η πρόταση της συνεργαζόμενης ιατρικής ομάδας. Ωστόσο, τελικά αποφασίστηκε και πάλι η αξιοποίηση των περιφερειακών αισθητήρων και η ελαχιστοποίηση της κατάλληλης συνάρτησης κόστους (F_R), κατά αντιστοιχία με την προσέγγιση των Ranta κ.ά. [87]:

$$I_\zeta = \frac{P_{Ws}}{4\pi R^2} = \frac{p_{rms}^2}{\rho_o c} \Rightarrow P_{Ws} = \frac{p_{rms,\zeta}^2}{\rho_o c} 4\pi R_\zeta^2 = const, \quad \zeta = LU, RU, RD, LD \quad (6-7)$$

$$F_R(L_z) = \arg \left(\min \left[\sum_{\zeta, \xi=1, \zeta \neq \xi}^4 \left(p_{rms,\zeta}^2 R_\zeta^2 - p_{rms,\xi}^2 R_\xi^2 \right)^2 \right] \right), \quad \text{όπου } p_{rms,i} = k_s \cdot rms \{ x_\zeta(i) \}$$

όπου R_ζ είναι η απόσταση του ζ-αισθητήρα από την «ελεγχόμενη» θέση ηχητικής διέγερσης, ενώ είναι προφανές ότι, τόσο η σταθερά ευαισθησίας k_s (*sensitivity*) των αισθητήρων, όσο και τα χαρακτηριστικά ($\rho_o c$) του μέσου δεν επηρεάζουν το αποτέλεσμα εντοπισμού. Είναι σαφές ότι στην παραπάνω σχέση υιοθετείται το *ISL* μοντέλο σφαιρικής ηχητικής διάδοσης, όπου το μέτρο της ηχητικής έντασης I αποσβένει με το τετράγωνο της απόστασης και είναι ανάλογο με το τετράγωνο της ακουστικής πίεσης p_{rms} . Επίσης, χρησιμοποιείται το μη απορροφητικό μοντέλο, όπου το μέσο ηχητικής διάδοσης δεν προκαλεί πρόσθετες ενεργειακές απώλειες ή άλλα κυματικά φαινόμενα. Η απόφαση αυτή λήφθηκε έχοντας κατά νου τα αποτελέσματα της εργασίας του Ranta [87] και γενικότερα τα σχόλια των προηγούμενων παραγράφων.

Με βάση τα παραπάνω, η νέα, υβριδική, μέθοδος ηχητικού εντοπισμού αξιοποιεί τόσο τις αξονικές διανυσματικές συνιστώσες για τον υπολογισμό της διεύθυνσης, όσο και τις ενεργειακές στάθμες των περιφερειακών αισθητήρων για την εύρεση της τελικής θέσης της πηγής (**D**irection and **E**nergy **B**ased **L**ocalization –**DEBL**). Έτσι, η αποτελεσματικότητα του αλγορίθμου εξαρτάται πρωτίστως από το σωστό εντοπισμό των σχηματιζόμενων γωνιών και δευτερευόντως από την ελαχιστοποίηση της συνάρτησης κόστους F_R . Αυτό σημαίνει ότι η απόδοση της μεθόδου *DEBL* καθορίζεται κυρίως από την ακρίβεια λήψης των αξονικών συνιστωσών επιτάχυνσης, δηλαδή από τις προδιαγραφές του επιταχυνσιομέτρου Brüel & Kjaer 4506, που όπως παρουσιάστηκε στο δεύτερο κεφάλαιο είναι αρκετά υψηλές. Όπως θα φανεί στην επόμενο παράγραφο, όταν η ηχητική διέγερση βρίσκεται πολύ κοντά στο κέντρο C_0 και άρα η επίδραση του μέσου είναι ακόμη πιο περιορισμένη (λόγω πολύ μικρής απόστασης), τα σφάλματα εντοπισμού ελαχιστοποιούνται.

Όσον αφορά την αριθμητική μέθοδο προσδιορισμού της τελικής θέσης βάσει της εξ. (6-7), υπάρχουν αρκετά πλεονεκτήματα έναντι της προσέγγισης του Ranta, που σχετίζονται τόσο με τον «επικουρικό» χαρακτήρα της τεχνικής, όσο και με τις λιγότερες υπολογιστικές απαιτήσεις. Έτσι, κατά τον έλεγχο των εν δυνάμει θέσεων ηχητικής διέγερσης, η σάρωση δε γίνεται σε όλο το ΧΥ επίπεδο, αλλά μόνο κατά μήκος του άξονα Z , σύμφωνα με την εξ. (6-4). Οι αριθμητικές τεχνικές ηχητικού εντοπισμού έχουν πολλές ομοιότητες με τις μεθόδους εντοπισμού κίνησης των μακρομπλόκ εικόνων (παρουσιάστηκαν στο δεύτερο κεφάλαιο), όπου μπορούν να χρησιμοποιηθούν οι αλγόριθμοι πλήρους αναζήτησης (Full Search –**FS**), λογαριθμικής - πολυδιακριτικής αναζήτησης (Logarithmic Search –**LS**, Multi-resolution

Search -MS), κ.ά. [171]-[172], [248]-[250]. Στην προκειμένη περίπτωση, το υπολογιστικό κόστος της αριθμητικής αναζήτησης είναι αρκετά μικρότερο, καθώς διεξάγεται επάνω στη γνωστή διεύθυνση άφιξης του ηχητικού κύματος. Έτσι, λαμβάνοντας υπόψη τα ανατομικά χαρακτηριστικά της κοιλιακής περιοχής, αποφασίστηκε η ανίχνευση για $-5cm \leq Z \leq 1cm$, με στόχο την ελαχιστοποίηση της συνάρτησης κόστους $F_R (arg-min\{F_R(Z)\})$. Ιδιαίτερη προσοχή χρειάζεται στην επιλογή του κατάλληλου βήματος σάρωσης, ανάλογα με τα χαρακτηριστικά του πεδίου και κυρίως τη διακριτική αναλυτικότητα της μετρητικής διάταξης, δηλαδή το μέγεθος των αισθητήρων. Η επιλογή βημάτων ανίχνευσης ΔZ της τάξης χιλιοστού ($\Delta Z=1mm$) αποδείχθηκε ιδιαίτερα αποδοτική, υπερκαλύπτοντας τις ανάγκες βιοακουστικού εντοπισμού της τρέχουσας εφαρμογής.

6.2.1.1 Ηχητικός εντοπισμός στο «χωροχρόνο» και οπτικοακουστική απόδοση των συμβάντων

Όπως έχει ήδη επισημανθεί στο πρώτο κεφάλαιο, οι θέσεις ηχητικής διέγερσης ενδέχεται να μεταβάλλονται κατά μήκος των κοιλιακών ήχων, ιδιαίτερα σε περιπτώσεις συμβάντων μεγάλης διάρκειας. Τα στοιχεία αυτά προέκυψαν, καταρχήν, από τις πρώτες επιστημονικές παρατηρήσεις του Cannon [23], [36] και είναι σήμερα απολύτως κατανοητά, με βάση τις ιατρικές γνώσεις φυσιολογίας και ανατομίας γύρω από τις συσταλτικές εντερικές κινήσεις. Τα συμπεράσματα αυτά επιβεβαιώθηκαν, επίσης, από τις ενδοσκοπικές και φθοροσκοπικές παρατηρήσεις, που έγιναν δειγματοληπτικά στη φάση διαχωρισμού των προτύπων κοιλιακών ήχων [61]. Για το λόγο αυτό αποφασίστηκε η χρήση σύντομων χρονικών παραθύρων και η εξαγωγή κινούμενων εικόνων ακουστικής χαρτογράφησης, σε αντιστοιχία με τις απεικονίσεις ακουστικής ολογραφίας [290], [300]-[303]. Έτσι, λήφθηκαν και πάλι υπόψη οι αντιληπτικές ικανότητες του ανθρώπου (όρασης αυτή τη φορά), καθώς και οι αντίστοιχες προδιαγραφές των προτύπων κινούμενης εικόνας, όπως είναι η ταχύτητα εναλλαγής 25 εικόνων ανά δευτερόλεπτο (fps) του τηλεοπτικού συστήματος PAL [153]-[154].

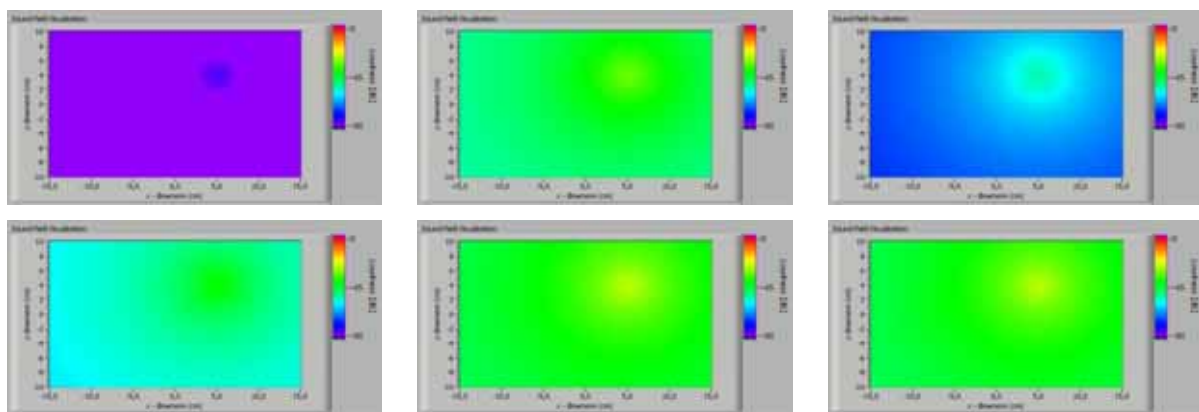
Με δεδομένη τη συχνότητα δειγματοληψίας $f_s=8kHz$, το χρονικό μήκος των εσωτερικών παραθύρων χαρτογράφησης n_{smw} (sound mapping windows) ορίζεται σε $N_{smw}=320$ δείγματα ($=8.000/25$). Παράλληλα, αποφασίστηκε η εφαρμογή επικαλυπτόμενων και μη ορθογωνικών παραθύρων, τα οποία πλέον έχουν τραπεζοειδή μορφή (χρήση 80 δειγμάτων επικάλυψης, 40 στην αρχή και 40 στο τέλος). Και οι δύο αυτές διευθετήσεις έγιναν με στόχο την αποφυγή απότομων μεταβολών μεταξύ των διαδοχικών εικόνων. Για τον ίδιο λόγο απορρίφθηκε η χρήση μεγαλύτερων παραθύρων και κατ' επέκταση μικρότερων ρυθμών fps. Έτσι, ο χωρικός εντοπισμός DEBL (εξ. 6-6, 6-7) εφαρμόζεται σε κάθε παράθυρο n_{smw} (Time Windowed DEBL, TW-DEBL). Μετά τον εντοπισμό της ηχητικής πηγής υπολογίζεται η αντίστοιχη στάθμη, από την εφαρμογή του ISL κανόνα στον πλησιέστερο αισθητήρα. Ακολουθεί η εύρεση όλων των ηχητικών σταθμών σε ένα πλέγμα χωρικής διακριτικότητας ενός χιλιοστού ($\Delta X=\Delta Y=1mm$), στο επίπεδο $Z=0$. Για το σκοπό αυτό διεξάγονται προσομοιώσεις προώθησης προς τα εμπρός (forward propagation) του ISL μοντέλου ηχητικής διάδοσης, απ' όπου προκύπτουν οι ακουστικοί χάρτες ενεργειακής κατανομής (σχ. 6-5).

Η διαδικασία TW-DEBL που περιγράφηκε πιο πάνω, περιορίζει αισθητά τις πιθανότητες ταυτόχρονης εκπομπής περισσότερων από μία ηχητικών πηγών στο μικρό χρονικό διάστημα N_{smw} ($=40msec$), χωρίς ωστόσο να αποκλείεται κάτι τέτοιο. Έτσι, εξετάστηκε το ενδεχόμενο εφαρμογής παραθύρων χρόνου - συχνότητας, όπως συνέβη με τις κυματιδιακές ακολουθίες εντοπισμού και οριοθέτησης των συμβάντων (4^ο κεφάλαιο) [217], φιλοσοφία που έχει χρησιμοποιηθεί και σε εφαρμογές ηχητικού εντοπισμού [295]. Η τεχνική αυτή, αν και θα μείωνε ακόμη περισσότερο την πιθανότητα ταυτόχρονης παρουσίας δύο (ή περισσότερων)

ηχητικών διεγέρσεων (στο ίδιο πλαίσιο χρόνου-συχνότητας), δεν δοκιμάστηκε σε μεγάλο βαθμό, καθώς ξεπερνά τις ανάγκες της διατριβής και τις αντίστοιχες προδιαγραφές της ιατρικής έρευνας (χωρίς να αποκλείεται η εφαρμογή της στο μέλλον). Αξίζει, επίσης, να σημειωθεί ότι και τα επτά σήματα της διάταξης *A-Soundfield 4+3* υπόκεινται στην επίδραση του προσθετικού θορύβου *ABN*, γι' αυτό και προτείνεται η χρήση των τεχνικών καταστολής θορύβου *WDWF*. Στο σχ. 6-5 δίνονται χαρακτηριστικές εικόνες ακουστικής χαρτογράφησης, όπως προέκυψαν από την εφαρμογή της διαδικασίας *TW-DEBL*. Η «λογαριθμική ολοκλήρωση» των εικόνων αυτών εντός κάθε συμβάντος (εφαρμογή της εξ. 4-11 σε όλους τους κόμβους του πλέγματος χωρικής απεικόνισης), έχει ως αποτέλεσμα την εξαγωγή μιας «ισοδύναμης» διδιάστατης εικόνας ακουστικής χαρτογράφησης (*Equivalent Sound Map - ESM*). Επίσης, από την εφαρμογή της τεχνικής *DEBL* σε ολόκληρο το μήκος των συμβάντων προκύπτει η «ισοδύναμη θέση ηχητικής πηγής» (*Equivalent Source Location -ESL*), η οποία ταυτίζεται με την πραγματική θέση διέγερσης στην περίπτωση μιας μόνο πηγής (φαινόμενο αρκετά συχνό σε ήχους σύντομης διάρκειας, όπως είναι τα *SCL* πρότυπα). Τέλος, από τη χωρο-χρονική ολοκλήρωση των *ESM* χαρτών των συμβάντων σε τακτά χρονικά διαστήματα (π.χ. 3 λεπτά), προκύπτουν εικόνες σύνοψης του συνολικού μηχανισμού χωρικής κατανομής της ηχητικής ενέργειας (σε όλο το μήκος των καταγραφών παρατεταμένου χρόνου):

$$L_{SSM}(x, y)|_{SI} = 10 \log \frac{1}{N_{SI}} \sum_{n_{ev}=1}^{N_{EV-SI}} N_{ev}(n_{ev}) \cdot 10^{\frac{L_{ESM}(n_{ev}, x, y)}{10}} \quad (6-8)$$

όπου L_{SSM} είναι η ισοδύναμη εικόνα χαρτογράφησης (*Summarization Sound Map -SSM*) στο διάστημα «σύνοψης» *SI* (*Summarization Interval*), N_{SI} ο συνολικός αριθμός των δειγμάτων χρονικής ολοκλήρωσης, N_{ev} το μήκος του συμβάντος n_{ev} και N_{EV-SI} ο συνολικός αριθμός των συμβάντων στο χρόνο *SI*, ενώ ο όρος $L_{ESM}(n_{ev}, x, y)$ αποδίδει τον *ESM*-χάρτη του συμβάντος n_{ev} για κάθε θέση (X, Y) . Όπως θα δούμε στο επόμενο κεφάλαιο, οι ακουστικοί χάρτες *SSM* είναι ιδιαίτερα χρήσιμοι στη χωρική ανάλυση των καταγραφών παρατεταμένου χρόνου. Η *SSM*-σύνοψη μπορεί να εφαρμοστεί ξεχωριστά για τα διάφορα πρότυπα των ηχητικών γαστρεντερικών συμβάντων, επιτρέποντας τη συνδυασμένη ανάλυση κατανομής των προτύπων στο χωρο-χρόνο (*STPDA*) [217]-[218].



Σχήμα 6-5. Στιγμιότυπα κινούμενων εικόνων ακουστικής χαρτογράφησης, από την εφαρμογή της μεθόδου *DEBL* σε χρονικά παράθυρα n_{smw} (χρονική σειρά: από αριστερά προς τα δεξιά και από πάνω προς τα κάτω).

Εκτός από τα εργαλεία απεικόνισης, ιδιαίτερη έμφαση δόθηκε στην ανάπτυξη τεχνικών στερεοφωνικής ακρόασης, για την ακουστικοποίηση (*auralization*) του παραγόμενου βιο-ακουστικού πεδίου. Η συγκεκριμένη προσέγγιση έγινε με στόχο την πιο ρεαλιστική οπτικοακουστική απόδοση των συμβάντων, με βάση τόσο γραφικά, όσο και ακουστικά -

αντληπτικά χαρακτηριστικά. Έτσι, από το συνδυασμό των κινούμενων εικόνων ακουστικής ολογραφίας και των συνιστωσών πολυκαναλικού ήχου, μπορούν να εξαχθούν οπτικοακουστικά αρχεία “ASF-AVI” (Abdominal SoundField Audio-Visual Information), τα οποία έχουν όλες τις απαραίτητες πληροφορίες των τμημάτων σημαντικής δραστηριότητας. Τα οπτικοακουστικά δεδομένα μπορούν να περιγράψουν πλήρως τα κοιλιακά βιοακουστικά συμβάντα στο χωρο-χρόνο, παρέχοντας νέες δυνατότητες ανάλυσης. Το γεγονός αυτό μπορεί να αξιοποιηθεί, τόσο στις ερευνητικές προσεγγίσεις ιατρικής διάγνωσης «σύντομου χρόνου» [70]-[71], όσο και στην εφαρμογή πρωτοκόλλων μελέτης της κινητικότητας μέσω πολυκαναλικών καταγραφών παρατεταμένης διάρκειας. Έτσι, τα αποτελέσματα χωρικής ανάλυσης και η εναλλαγή των εικόνων ακουστικής ολογραφίας στο χρόνο, μπορούν να αξιοποιηθούν για τον εντοπισμό των διαφόρων φάσεων εντερικής λειτουργίας. Στο πλαίσιο αυτό, ενδιαφέρον παρουσιάζει η μελέτη των «χωρικών προτύπων» και η αναζήτηση αντίστοιχων κανόνων-σχημάτων ταξινόμησης, με βάση χαρακτηριστικές παραμέτρους των εικόνων βιοακουστικής χαρτογράφησης (image features). Αν και τα παραπάνω εγχειρήματα βρίσκονται ακόμη σε νηπιακό στάδιο, ωστόσο, παρέχουν νέες δυνατότητες διαχείρισης, αναζήτησης, σύγκρισης και πλοήγησης μεταξύ των δεδομένων διαφόρων πειραματικών καταγραφών, προωθώντας έτσι τη δι-επιστημονική, ερευνητική συνεργασία [226] (σχετικές λεπτομέρειες δίνονται στο επόμενο κεφάλαιο). Παράλληλα, οι τεχνικές γαστρεντερικής ακρόασης καθίστανται πιο λειτουργικές και «παραστατικές».

Η διαδικασία εξαγωγής των οπτικοακουστικών αρχείων ASF-AVI περιλαμβάνει δύο στάδια: τον υπολογισμό των ακουστικών χαρτών βάσει της μεθόδου TW-DEBL, και την κωδικοποίηση των διαθέσιμων ηχητικών καναλιών στο κατάλληλο πρότυπο περιφερικού ήχου. Από τη διάταξη του συστήματος A-Soundfield 4+3 είναι προφανές ότι οι τέσσερις περιφερειακοί αισθητήρες συνιστούν ένα κλασικό σύστημα τετραφωνίας (*Quadraphonic, Quads*) [160], [308], [314]-[315]. Από τα σήματα αυτά είναι δυνατόν να υπολογιστούν οι δύο «συνολικές» στερεοφωνικές συνιστώσες L_T/R_T (Left/Right Total) για αναπαραγωγή με ένα οποιοδήποτε στερεοφωνικό σύστημα ηχείων ή ακουστικών. Στο πλαίσιο αυτό αποφασίστηκε η χρήση του τετραφωνικού πίνακα κωδικοποίησης Scheiber (“*Quadraphonic Scheiber Matrix Decoder*”), βάσει του οποίου υπολογίζονται τα σήματα L_T/R_T [160], [314]-[315]:

$$\begin{aligned} L_T &= 0,924 \cdot L_F + 0,924 \cdot L_B + 0,383 \cdot R_F - 0,383 \cdot R_B \\ R_T &= 0,924 \cdot R_F + 0,924 \cdot R_B + 0,383 \cdot L_F - 0,383 \cdot L_B \end{aligned} \quad (6-9)$$

$\mu\epsilon \ L_F \equiv LU, \ L_B \equiv LD, \ R_F \equiv RU, \ R_B \equiv RD$

όπου οι συνιστώσες L_F, L_B, R_F και R_B αναφέρονται στα αντίστοιχα σήματα των τεσσάρων περιφερειακών αισθητήρων $x_\zeta(i)$, $\zeta = \{LU, LD, RU, RD\}$. Είναι προφανές ότι από τις σχέσεις της εξ. (6-9) μπορούν να εξαχθούν τα ισοδύναμα με τα αρχικά σήματα L'_F, L'_B, R'_F και R'_B , εφαρμόζοντας τον αντίστοιχο πίνακα αποκωδικοποίησης [160], [314]-[315]. Έτσι, εκτός από την κλασική στερεοφωνία δύο καναλιών μπορούν να αξιοποιηθούν πιο σύνθετες διατάξεις ηχείων ή/και ακουστικά περιφερικού ήχου “*Dolby Headphones*” [316], με τελικό αποτέλεσμα τη μεγαλύτερη πιστότητα χωρικής αναπαραγωγής (αφού προηγηθεί η μετατροπή στο κατάλληλο πρότυπο *surround*) [160], [308], [314]-[316].

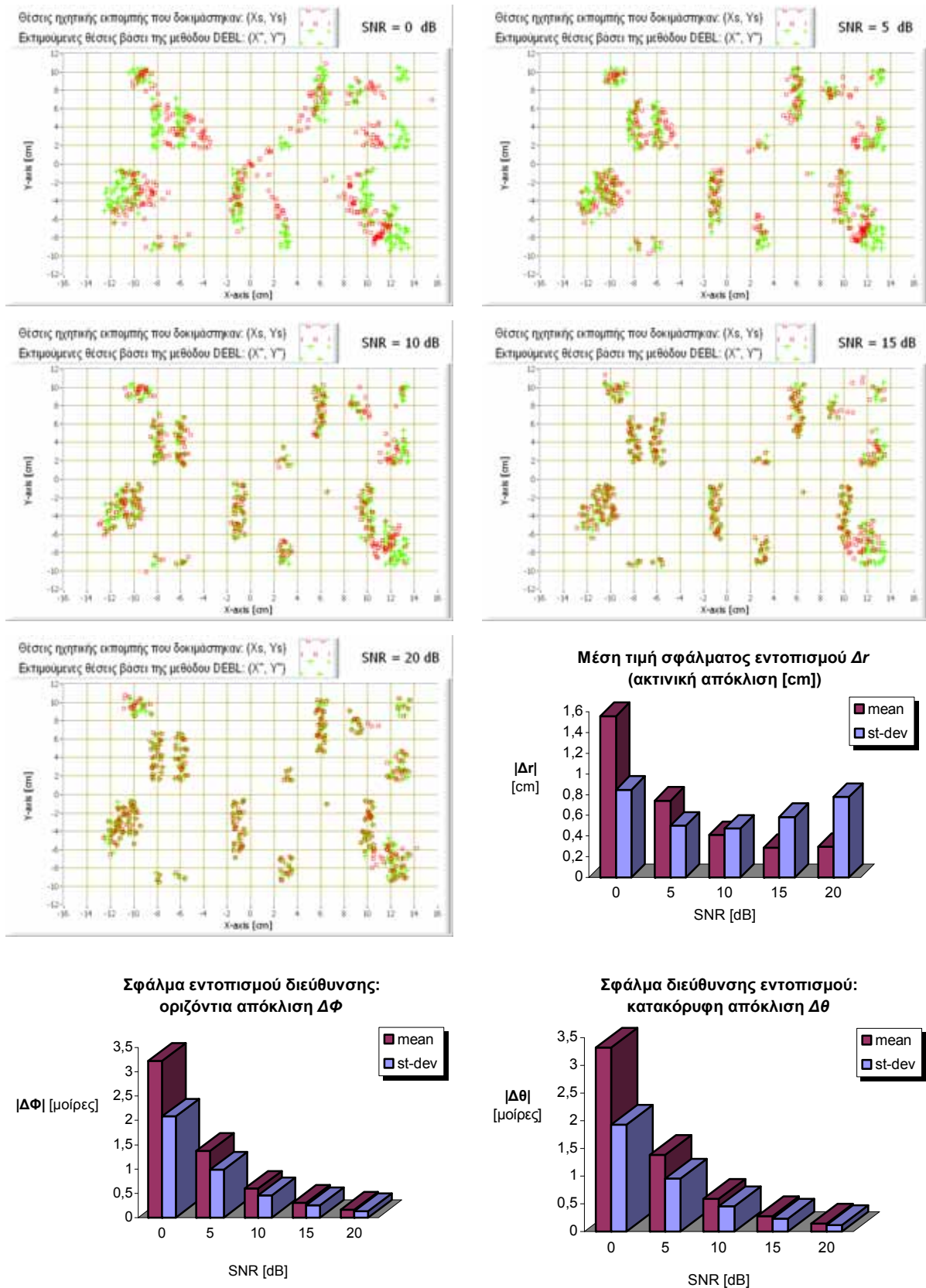
6.2.1.2 Αξιολόγηση της μεθόδου DEBL μέσω προσομοιώσεων σε περιβάλλον λογισμικού

Μια από τις κύριες αιτιάσεις των Ranta κ.ά. [87], για την απόρριψη των προσεγγίσεων βιοακουστικής χαρτογράφησης [41], αφορά στις αμφιβολίες αποτελεσματικότητας και την

αδυναμία αξιολόγησης των αντίστοιχων αλγορίθμων. Για το λόγο αυτό, στην προκειμένη περίπτωση δόθηκε ιδιαίτερη έμφαση για την αναζήτηση κανόνων αξιολόγησης, μέσα από τη σύσταση κατάλληλων πειραματικών διαδικασιών. Η πιο απλή προσέγγιση για την εκτίμηση της αποδοτικότητας ηχητικού εντοπισμού είναι η διεξαγωγή προσομοιώσεων ηχητικής διάδοσης σε περιβάλλοντα λογισμικού [248]-[249]. Οι διαδικασίες αυτού του είδους έχουν το πλεονέκτημα ότι δεν απαιτείται εκτέλεση πειραμάτων σε πραγματικές συνθήκες, στοιχείο ιδιαίτερα χρήσιμο στην περίπτωση που δεν είναι διαθέσιμος ο απαιτούμενος εξοπλισμός στο σύνολό του, ή όταν υπάρχουν δυσκολίες που σχετίζονται με την αδυναμία ελέγχου των ηχητικών πηγών και τις ιδιαιτερότητες του μέσου διάδοσης (όπως συμβαίνει και με τη συγκεκριμένη εφαρμογή). Στον αντίποδα, τα μοντέλα αριθμητικής προσομοίωσης συχνά εξιδανικεύουν τη συμπεριφορά του μέσου και τους αντίστοιχους κανόνες κυματικής διάδοσης, παρουσιάζοντας έτσι αποκλίσεις με τα πραγματικά αποτελέσματα [248]-[249]. Για το σκοπό αυτό, συνήθως εισάγονται συνιστώσες θορύβου, που προσεγγίζουν τις «ατέλειες» του μέσου, αλλά και την πραγματική παρουσία θορύβου *ABN* στους αισθητήρες [248]-[249].

Έχοντας υπόψη τα συμπεράσματα των Ranta κ.ά. [87] και τις αντίστοιχες παρατηρήσεις της § 6.2, αποφασίστηκε η υιοθέτηση του ισοτροπικού-ομογενούς μέσου και του *ISL* κανόνα (ομοιόμορφη - σφαιρική διάδοση του βιοακουστικού πεδίου), παράλληλα με τη χρήση λευκού γκαουσιανού θορύβου *GWN*. Έτσι, επιλέχθηκαν τυχαίες θέσεις ηχητικής διέγερσης στο χώρο (X, Y, Z : $Z < 0$), όπου διεξήχθησαν πολλές προσομοιώσεις και για διάφορους λόγους σήματος προς θόρυβο *SNR* (0, 5, 10, 15 και 20dB), ακολουθώντας μια πάγια διαδικασία αξιολόγησης των αποτελεσμάτων ακουστικού εντοπισμού [248]-[249]. Ως σήματα διέγερσης χρησιμοποιήθηκαν τόνοι διαφόρων συχνοτήτων, όπου όλες οι συνιστώσες της υποτιθέμενης μετρητικής διάταξης μολύνθηκαν με τυχαίο θόρυβο ίσης ενέργειας (οι λόγοι *SNR* ορίζονται βάσει του ασθενέστερου, κάθε φορά, αισθητήρα). Η χρήση τονικών σημάτων έγινε κυρίως στη φιλοσοφία γενικότερης αξιολόγησης της μεθόδου *DEBL*, ανεξάρτητα από τη φύση του συγκεκριμένου προβλήματος, αλλά και για πρακτικούς λόγους, όπως ο εύκολος καθορισμός των σχέσεων *SNR*. Εξάλλου, ο εντοπισμός *DEBL* βασίζεται μόνο σε ενεργειακά κριτήρια και δεν επηρεάζεται από τα μορφολογικά χαρακτηριστικά των σημάτων.

Αξίζει να σημειωθεί ότι, υπάρχουν αρκετές επιφυλάξεις σχετικά με το πόσο οι ατέλειες του μέσου αποδίδονται σωστά από τον προσθετικό θόρυβο *GWN*, ή εάν λαμβάνονται υπόψη τα πραγματικά κατευθυντικά χαρακτηριστικά των αισθητήρων. Η έντονη παρουσία θορύβου, από την άλλη πλευρά, είναι πολύ πιθανό να δημιουργεί μεγαλύτερα προβλήματα εντοπισμού, σε σχέση με τις πραγματικές ιδιαιτερότητες του μέσου και των αισθητήρων. Σε κάθε περίπτωση, η προτεινόμενη ανάλυση μοντελοποίησης παρέχει χρήσιμες πληροφορίες, τόσο για τις δυνατότητες εντοπισμού της κυματικής διεύθυνσης από τις αξονικές συνιστώσες επιτάχυνσης, όσο και για την τελική αριθμητική επίλυση βάσει της εξ. (6-7). Εξάλλου οι προσομοιώσεις αυτές είχαν επικουρικό χαρακτήρα και διεξήχθησαν σε συνδυασμό με μια αντίστοιχη πειραματική διαδικασία φυσικής μοντελοποίησης (περιγράφεται στην επόμενη παράγραφο). Στο σχ. 6-6 δίνεται η στατιστική ανάλυση των σφαλμάτων εντοπισμού για διάφορες θέσεις διέγερσης και λόγους *SNR*. Είναι προφανές ότι η μέθοδος *DEBL* έχει πολύ καλή απόδοση ακόμη και σε ιδιαίτερα χαμηλές σηματοθορυβικές σχέσεις, παρέχοντας πολύ μεγαλύτερη αναλυτικότητα από τον κανόνα *CPA*. Παρατηρούμε, επίσης, ότι οι αποκλίσεις εντοπισμού Δr οφείλονται κυρίως στην «αριθμητική μέθοδο επίλυσης» και όχι στην εύρεση της διεύθυνσης (οι γωνιακές αποκλίσεις $\Delta\Phi$, $\Delta\theta$ είναι πολύ μικρές). Έτσι, η χρήση ενός πρόσθετου 3-αξονικού επιταχυνσιόμετρου, για τον υπολογισμό της διεύθυνσης ως προς ένα δεύτερο σημείο, εκτιμάται ότι θα δώσει ακόμη καλύτερα αποτελέσματα εντοπισμού.



Σχήμα 6-6. Αποτελέσματα εντοπισμού-DEBL (προσομοιώσεις για διάφορους λόγους SNR): (α)-(ε) διαγράμματα χωρικής διασποράς, (στ)-(η) μέσο σφάλμα - τυπική απόκλιση της (στ) ακτινικής απόκλισης Δr , (ζ) της απόκλισης στην οριζόντια διεύθυνση $\Delta \Phi$ και (η) της απόκλισης στην κατακόρυφη διεύθυνση $\Delta \Theta$.

6.2.1.3 Αξιολόγηση της μεθόδου DEBL μέσω φυσικών πειραματικών μοντέλων

Η πιο αξιόπιστη διαδικασία αξιολόγησης των μεθόδων ακουστικού εντοπισμού αφορά στη διεξαγωγή πειραμάτων, όπου οι ηχητικές πηγές τοποθετούνται σε διάφορες γνωστές θέσεις και κατόπιν προσδιορίζεται το σφάλμα εντοπισμού [249], [284], [287]. Είναι προφανές ότι κάτι τέτοιο δεν είναι εύκολο να γίνει στην προκειμένη περίπτωση, τουλάχιστον όχι μέσω ανώδυνων και μη επιβαρυντικών στον ανθρώπινο οργανισμό επεμβάσεων. Για το σκοπό αυτό δημιουργήθηκαν φυσικά μοντέλα προσομοίωσης των χαρακτηριστικών των κοιλιακών τοιχωμάτων (ως μέσα βιοακουστικής μετάδοσης), με τα οποία κατέστη δυνατή η διενέργεια πειραμάτων ελεγχόμενης ηχητικής εκπομπής. Κίνητρο για την προσέγγιση αυτή αποτέλεσε η εργασία των Mansy κ.ά. (1999) [317], που χρησιμοποίησαν αντίστοιχα φυσικά μοντέλα προσομοίωσης του ανθρώπινου μυϊκού ιστού, για την αξιολόγηση διαφόρων αισθητήρων βιοακουστικής καταγραφής. Πρόσφατα έχει δημοσιευτεί μια ακόμη εργασία (Kraman κ.ά. 2006) [318], με αντικείμενο το σχεδιασμό, την κατασκευή και την αξιολόγηση ενός πρότυπου φυσικού μοντέλου, που προτείνεται ειδικά για «δοκιμές» ηλεκτρακουστικών μετατροπέων αναπνευστικών ήχων: “*bioacoustic transducer testing (BATT) system for respiratory sounds*”. Για τη δημιουργία των «τεχνητών» μέσων βιοακουστικής μετάδοσης, και στις δύο παραπάνω εργασίες, χρησιμοποιείται παχύρρευστο υλικό GEL, καθώς και στρώματα ζωικού λίπους, ενώ ο έλεγχος της ηχητικής εκπομπής γίνεται μέσω μικρών ηχείων και ακουστικών (*headphones – earphone transducers*) [317]-[318].

Στο πλαίσιο αυτό, αποφασίστηκε η κατασκευή τεχνητών ομοιωμάτων (μοντέλων) του μέσου μετάδοσης των κοιλιακών ήχων (*Abdominal Sound-Propagation Media Models, ASPMM*). Ειδικότερα, δημιουργήθηκαν δύο μοντέλα ASPMM (σχ. 6-7). Το πρώτο αποτελείται από μια πλαστική θερμοφώρα, στο εσωτερικό της οποίας τοποθετήθηκε παχύρρευστο μίγμα λαδιού, και μία συνθετική δερμάτινη επιφάνεια. Η προσαρμογή και συγκράτηση των δύο αυτών «υλικών» έγινε με τη βοήθεια επιστρώσεων λίπους και ισχυρών κολλητικών ταινιών διπλής όψης (σχ. 6-7α). Για την υποτιθέμενη βιοακουστική διέγερση χρησιμοποιήθηκε η γεννήτρια δονήσεων “*Vibration Exciter - B&K Type 4809*” στην οποία προσαρμόστηκε μια ακίδα μικρού πάχους (σχ. 6-7α). Παρά τις υψηλές προδιαγραφές της γεννήτριας δονήσεων B&K 4809, το συγκεκριμένο πειραματικό μοντέλο εγκαταλείφθηκε για δύο βασικούς λόγους: α) υπήρξαν προβλήματα στήριξης του ομοιώματος ASPMM, το οποίο είχε, ανεπιθύμητα, ευμετάβλητο σχήμα, και β) η επαφή της ακίδας διέγερσης στο τεχνητό βιοακουστικό μέσο διαφέρει αρκετά από τον πραγματικό μηχανισμό συσταλτικής δραστηριότητας, παραγωγής και μετάδοσης των κοιλιακών ηχητικών κυμάτων. Έτσι, το μοντέλο αυτό δε χρησιμοποιήθηκε στα πειράματα χαρτογράφησης, αλλά περιορίστηκε στις μετρήσεις αξιολόγησης των αισθητήρων και του συστήματος καταγραφής, που παρουσιάστηκαν στο δεύτερο κεφάλαιο.

Το δεύτερο ομοίωμα ASPMM δημιουργήθηκε σε μια προσπάθεια αντιμετώπισης των παραπάνω προβλημάτων (σχ. 6-7β). Χρησιμοποιήθηκε και πάλι η επιφάνεια δέρματος, η οποία αυτή του φορά επικολλήθηκε σε δύο έντονα απορροφητικά υλικά τα οποία είχαν εμποτιστεί με μίγμα παχύρρευστου λαδιού και είχαν περιτυλιχθεί με διαφανή μεμβράνη. Η προσαρμογή και συγκράτηση των δύο παραπάνω «στρωμάτων ανομοιογενούς ρευστού» με το έντονα αναδιπλούμενο δερμάτινο ύφασμα έγινε με επαλείψεις επεξεργασμένου λίπους. Για τις ανάγκες ηχητικής διέγερσης χρησιμοποιήθηκαν τα ηλεκτροδυναμικά ακουστικά *Sony MDR-E819V earphones*, τα οποία έχουν πολύ καλά χαρακτηριστικά απόκρισης (ευαισθησία 108 dB/mW, συχνοτική περιοχή 12-22.000 Hz, ισχύς 50mW) και πολύ μικρό μέγεθος. Η τοποθέτησή τους έγινε πάνω στο ηχοαπορροφητικό υλικό του σχ. 6-7β, επιτρέποντας την καλύτερη προσέγγιση των γαστρεντερικών συσταλτικών μηχανισμών βιοακουστικής

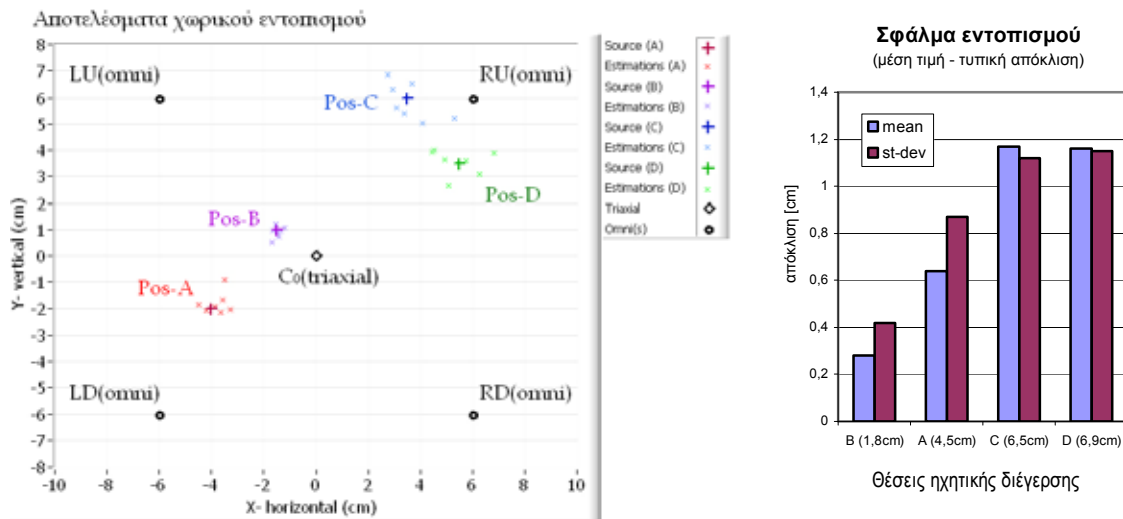
μετάδοσης, αλλά και την αποφυγή ηχητικών ανακλάσεων στη βάση στερέωσης. Για την καλύτερη επαφή και μηχανική προσαρμογή των αισθητήρων στη μετρούμενη επιφάνεια, καθώς και για την απομόνωση των ακουστικών θορύβων περιβάλλοντος κατά τη διεξαγωγή των πειραμάτων, χρησιμοποιήθηκε το γιλέκο *WAAV* (σχ. 6-7β). Θα μπορούσε κανείς να παρατηρήσει ότι το προτεινόμενο μοντέλο *ASPMM* είναι αρκετά περίπλοκο, σε βαθμό που να μην προσεγγίζει τόσο καλά το μυϊκό ιστό, όπως συμβαίνει με τα στρώματα *GEL* των εργασιών που προαναφέρθηκαν [317]-[318]. Ωστόσο, αυτό αποτελεί περισσότερο πλεονέκτημα παρά μειονέκτημα, καθώς επιτρέπει την αξιολόγηση της μεθόδου *DEBL* σε ακόμη πιο δύσκολες συνθήκες. Εξάλλου, το πραγματικό μέσο μετάδοσης των γαστρεντερικών βιοακουστικών κυμάτων είναι αρκετά πιο περίπλοκο, όπου εκτός από το μυϊκό ιστό των κοιλιακών τοιχωμάτων είναι προφανές ότι περιλαμβάνονται και τμήματα του εντέρου.



Σχήμα 6-7. Πειραματικές διατάξεις αξιολόγησης της μεθόδου ηχητικού εντοπισμού *DEBL*.

Χρησιμοποιώντας το μοντέλο *ASPMM* του σχ. 6-7β διεξήχθησαν πρόσθετα πειράματα αξιολόγησης των *DEBL*-αποτελεσμάτων εντοπισμού. Στην περίπτωση αυτή αξιοποιήθηκαν αποθρονοποιημένοι κοιλιακοί ήχοι $s_{DAS}(i)$ ως σήματα διέγερσης, όπως συνέβη και με την ποσοτική αξιολόγηση των τεχνικών *WDWF* στο τρίτο κεφάλαιο. Οι δοκιμές έγιναν για διάφορες θέσεις εκπομπής (όχι βέβαια στον ίδιο βαθμό με την έκταση των προσομοιώσεων), αφού πρωτίστως τα σήματα $s_{DAS}(i)$ κανονικοποιήθηκαν στη στάθμη -3dB . Από την παρουσία του προσθετικού θορύβου *ABN*, λόγω των χαμηλών ακουστικών σταθμών και των μεγάλων τιμών ενίσχυσης, προέκυψαν καταγραφές με *ESNR*-σηματοθορυβικές σχέσεις της τάξης των $3\text{-}12\text{dB}$. Έτσι, χρησιμοποιήθηκε η μέθοδος $WDWF_{L17}$, η οποία, βάσει της ποσοτικής ανάλυσης του τρίτου κεφαλαίου, είναι η πιο αποδοτική από τις *WDWF*-τεχνικές αποθρονοποίησης. Στο σχ. 6-8 παρουσιάζονται τα αποτελέσματα εντοπισμού για τέσσερις ενδεικτικές θέσεις ηχητικής διέγερσης. Παρατηρούμε ότι τα σφάλματα απόκλισης είναι ιδιαίτερα χαμηλά, ειδικά στις περιπτώσεις που η πηγή βρίσκεται πολύ κοντά στο κέντρο συμμετρίας της μετρητικής διάταξης *A-Soundfield 4+3* (θέσεις *A* και *B*). Το στοιχείο αυτό ερμηνεύεται από το γεγονός ότι το ηχητικό κύμα διανύει μικρή απόσταση, κι έτσι η επίδραση του μέσου, λόγω

των μη ιδεατών του χαρακτηριστικών, είναι ακόμη ποιο περιορισμένη. Είναι φανερό ότι η χρήση περισσότερων 3-αξονικών επιταχυνσιομέτρων και η εφαρμογή της μεθόδου *DEBL*, με βάση τον πλησιέστερο κάθε φορά *CPA*-αισθητήρα, μπορεί να οδηγήσει σε ακόμη καλύτερα αποτελέσματα. Η διαπίστωση αυτή σε συνδυασμό με τις επισημάνσεις της προηγούμενης παραγράφου αποτελούν τη βάση σχεδιασμού των μελλοντικών εξελίξεων, που αναλύονται στη συνέχεια. Σε κάθε περίπτωση, οι διαδικασίες πειραματικής αξιολόγησης αποδεικνύουν ότι τα αποτελέσματα ηχητικού εντοπισμού της μεθόδου *DEBL* είναι πολύ ικανοποιητικά, καλύπτοντας σε μεγάλο βαθμό τις προδιαγραφές της συγκεκριμένης εφαρμογής.



Σχήμα 6-8. Πειραματικά αποτελέσματα ηχητικού εντοπισμού *DEBL* και στατιστική ανάλυση σφάλματος.

6.3 Επισκόπηση βου κεφαλαίου

Το κεφάλαιο αυτό εστιάζει στα ζητήματα τοπογραφικής βιοακουστικής ανάλυσης, μέσα από μεθόδους ηχητικού εντοπισμού και αντίστοιχα εργαλεία ακουστικής χαρτογράφησης. Ξεκινώντας από τις εργασίες των Craine [42] και Ranta [87], αναζητήθηκαν νέες δυνατότητες μελέτης της γαστρεντερικής κινητικότητας, μέσω των εικόνων χωρικής κατανομής της βιοακουστικής ενέργειας. Έτσι, αναπτύχθηκε μια νέα τεχνική ηχητικού εντοπισμού, της οποίας κίνητρο αποτέλεσαν τα μικρόφωνα ηχητικού πεδίου *SoundField*. Η μέθοδος *DEBL* αξιοποιεί στοιχεία των τεχνικών εύρεσης της κυματικής διεύθυνσης (*DOA*) και των μεθόδων ενεργειακού εντοπισμού (*EBL*), ενώ, για πρώτη φορά σε τέτοιου είδους εφαρμογές, γίνεται χρήση 3-αξονικού επιταχυνσιομέτρου. Λαμβάνοντας υπόψη τις παρατηρήσεις του Ranta [87], σχετικά με την αδυναμία αξιολόγησης των τεχνικών βιοακουστικού εντοπισμού, εκτελέστηκαν δύο πειραματικές διαδικασίες, τα αποτελέσματα των οποίων επαληθεύουν την αποδοτικότητα της μεθόδου *DEBL*. Παράλληλα, αναπτύχθηκαν νέα, ευέλικτα εργαλεία πλοήγησης, διαχείρισης περιεχομένου και ανάλυσης, όπως οι ακουστικοί χάρτες σύνοψης *SSM* για τη μελέτη των καταγραφών παρατεταμένου χρόνου, και τα αρχεία *ASF-AVI* για την οπτικοακουστική απόδοση των μεμονωμένων συμβάντων.

Σε ότι αφορά τις προτεινόμενες μελλοντικές εξελίξεις, ήδη έχει αναφερθεί η δυνατότητα τροποποίησης της μεθόδου με χρήση περισσότερων επιταχυνσιομέτρων τριών αξόνων, ή/και βελτιώσεις των αριθμητικών αλγορίθμων. Ενδιαφέρον παρουσιάζει, επίσης, η εφαρμογή σχημάτων χρόνου - συχνότητας *JTFA* με στόχο την αύξηση της χωρο-χρονικής ανάλυσης εντοπισμού, καθώς και το ενδεχόμενο αξιολόγησης της μεθόδου σε υποκείμενα, με την κατάποση μικροσκοπικών στοιχείων ηχητικής εκπομπής (ηλεκτρακουστικές κάψουλες).

ΚΕΦΑΛΑΙΟ 7^ο

7. Τεχνικές διαχείρισης και ανάλυσης περιεχομένου

Το κεφάλαιο αυτό αφορά στο σχεδιασμό μιας ευέλικτης μεθοδολογίας διαχείρισης πολυμεσικού περιεχομένου, που ενσωματώνει όλες τις πληροφορίες σε μια τυποποιημένη δομή αρχειοθέτησης και μια αντίστοιχη οντολογία ανάλυσης. Η προσπάθεια αυτή έχει σαν στόχο την αποτελεσματική καταχώρηση, πλοήγηση και μελέτη όλων των πειραματικών δεδομένων. Οι προδιαγραφές του συγκεκριμένου εγχειρήματος έχουν αρκετές ομοιότητες με απαιτητικές εφαρμογές περιγραφής οπτικοακουστικού περιεχομένου, για τις ανάγκες των οποίων προτάθηκε και το πρότυπο MPEG-7. Έτσι, η προτεινόμενη οντολογία ανάλυσης σχεδιάστηκε υιοθετώντας τη βασική φιλοσοφία του προτύπου MPEG-7, στοιχείο που επιτρέπει την τυποποίηση των «παρατηρήσεων», τη μελλοντική διεπιστημονική συνεργασία και την προώθηση των σχετικών ερευνητικών προσεγγίσεων βιοϊατρικής τεχνολογίας. Στο πλαίσιο αυτό αναλύονται τα βασικά χαρακτηριστικά του προτύπου MPEG-7, με έμφαση στα στοιχεία που αξιοποιήθηκαν για τις ανάγκες της παρούσης εφαρμογής. Το κύριο μέρος του κεφαλαίου εστιάζει στην παρουσίαση της προτεινόμενης οντολογίας και τις καινοτομίες που εισάγονται για τη μελέτη των βιοακουστικών φαινομένων παρατεταμένου χρόνου. Επίσης, διατυπώνονται συμπεράσματα σχετικά με τη χρηστικότητα της νέας μεθοδολογίας, καθώς και για τις δυνατότητες μελλοντικής εξέλιξης.

7.1 Το πρότυπο MPEG-7

Η ραγδαία ανάπτυξη της τεχνολογίας, τα τελευταία χρόνια, έχει πυροδοτήσει την αντίστοιχη πρόοδο στο πεδίο του ήχου και της εικόνας. Καθώς οι ρυθμοί παραγωγής οπτικοακουστικής πληροφορίας αυξάνονται συνεχώς, ανάγκες όπως είναι η αναζήτηση, η κατηγοριοποίηση, η σύνοψη του ψηφιοποιημένου υλικού και άλλες παρόμοιες ενέργειες αποκτούν, αν όχι μεγαλύτερο, εξίσου σημαντικό ρόλο στην αξιοποίηση του υλικού. Την ανάγκη αυτή έρχεται να καλύψει το πρότυπο MPEG-7, του οποίου ο ρόλος είναι να παρέχει εννοιολογικές περιγραφές που σχετίζονται με ένα κομμάτι οπτικοακουστικού υλικού. Σε αντιδιαστολή με τα πρότυπα MPEG-1, MPEG-2, MPEG-4, το MPEG-7 δεν ασχολείται με την κωδικοποιημένη αναπαράσταση του υλικού. Ο κύριος ρόλος των προτύπων MPEG-1/2/4 ήταν η προτυποποιημένη αποθήκευση της πληροφορίας, εφαρμόζοντας βέλτιστες μεθόδους συμπίεσης και αυξάνοντας έτσι το λόγο όγκου αποθηκευμένου υλικού προς απαιτούμενο χώρο αποθήκευσης (για μια υποκειμενικά ή αντικειμενικά δεδομένη ποιότητα). Το MPEG-7 έχει σαν στόχο τη περιγραφή του αποθηκευμένου υλικού. Σχηματικά, η φιλοσοφία του συνοψίζεται στη φράση “bits about the bits”, δηλαδή πληροφορίες για το περιεχόμενο και όχι το ίδιο το περιεχόμενο. Έτσι, μπορεί να χρησιμοποιηθεί ανεξαρτήτως του μέσου και του τρόπου αποθήκευσης του οπτικοακουστικού υλικού [164]-[165], [226], [280], [283].

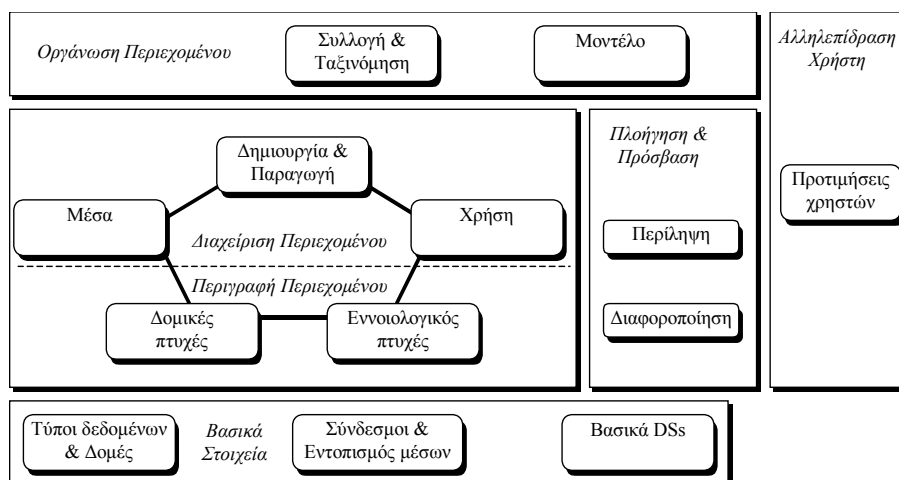
Ο βασικότερος στόχος του προτύπου είναι η παροχή ενός συνόλου μεθόδων και εργαλείων για τις διαφορετικές κλάσεις περιγραφής πολυμεσικού περιεχομένου. Ο όρος «κλάσεις περιγραφής» σημαίνει διαφορετικές πιθανές οπτικές γωνίες, από τις οποίες θα καλυφθεί κάποιο οπτικοακουστικό περιεχόμενο. Βασικό στοιχείο αποτελεί το γεγονός ότι κάθε οντότητα μπορεί να περιγραφεί με πολλαπλούς τρόπους, ανάλογα με τη χρήση της. Οι

Βασικές κλάσεις είναι οι εξής: καταγραφής (transcriptive), φυσική (physical), αντίληψης (perceptual), βασισμένη στο μέσο (medium-based). Πάνω από αυτές τις δομές βρίσκεται μία περιγραφή αρχιτεκτονικής (architectural), η οποία δίνει τη σχέση μεταξύ μεγάλων τμημάτων των δεδομένων και τη σχέση μεταξύ των υποκείμενων κλάσεων περιγραφής. Ακόμη πιο πάνω βρίσκεται η περιγραφή σχολιασμού (annotation), που περιλαμβάνει κάθε είδους σχόλια πάνω στο περιεχόμενο [164]-[165], [226], [280], [283].

Το πρότυπο MPEG-7 στοχεύει στο να εξασφαλίσει ένα πλαίσιο εργασίας που θα επιτρέπει αναφορές σε κομμάτια ενός αντικειμένου, σε ένα ολόκληρο αντικείμενο ή σε σύνολο αντικειμένων. Η ευελιξία σχετίζεται με το αν είναι δυνατόν να περιγραφεί το περιεχόμενο ενός κομματιού του υλικού, με τρόπο τέτοιο, ώστε να είναι δυνατή η συναγωγή και η συνεπαγωγή ενός είδους πληροφορίας από ένα άλλο (π.χ. εξαγωγή οπτικών από ηχητικές πληροφορίες και το αντίστροφο). Οι περιγραφές μπορούν να εγκατασταθούν με φυσικό τρόπο στο αντίστοιχο οπτικοακουστικό υλικό, μέσα στην ίδια ροή δεδομένων ή στο ίδιο σύστημα αποθήκευσης, μπορούν όμως να υπάρχουν ανεξάρτητα σε τελείως διαφορετικό σημείο. Για παράδειγμα το MPEG-7 μπορεί να χρησιμοποιηθεί για την περιγραφή μιας ταινίας (π.χ. διαίρεση σε σκηνές, διάλογοι, περιγραφή μουσικής επένδυσης κ.α.) είτε αυτή είναι αποθηκευμένη σε κινηματογραφικό φιλμ, είτε σε κάποιο ψηφιακό μέσο (π.χ. DVD/MPEG-2) [164]-[165], [226], [280], [283].

7.1.1 Τα μέρη του MPEG-7

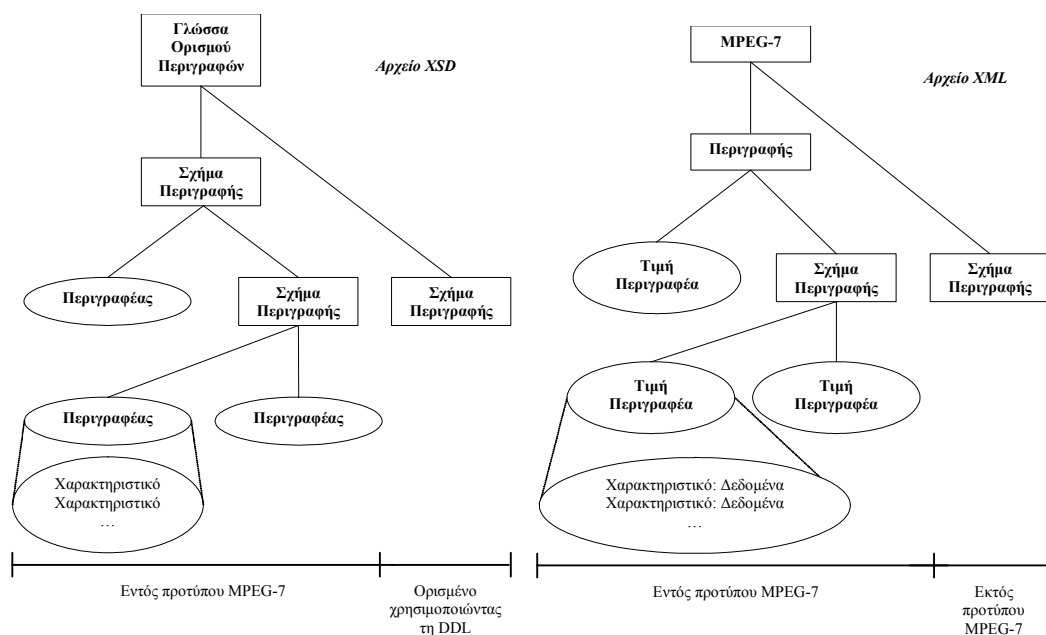
Οι πρωτεύουσες λειτουργίες παρέχονται από μια ομάδα τμημάτων (parts) του προτύπου MPEG-7. Στα τμήματα αυτά περιλαμβάνονται τα *Συστήματα (MPEG-7 Systems)*, η *Γλώσσα Ορισμού Περιγραφών (MPEG-7 Description Definition Language, DDL)*, το *Οπτικό μέρος (MPEG-7 Visual)*, το *Ηχητικό μέρος (MPEG-7 Audio)*, τα *Σχήματα Περιγραφής Πολυμέσων (MPEG-7 Multimedia Description Schemes - MDS)*, το *Λογισμικό Αναφοράς (MPEG-7 Reference Software)*, η *Συμμόρφωση με το MPEG-7 (MPEG-7 Conformance)*, η *Εξαγωγή και Χρήση Περιγραφών (MPEG-7 Extraction and Use of Descriptions)*, η *Εξαγωγή και Χρήση των MPEG-7 Περιγραφών (Extraction and use of MPEG-7 descriptions)*, τα *Προφίλ και τα Επίπεδα (Profiles and levels)*, οι *Ορισμοί Σχημάτων (Schema definition)* και τα *MPEG-7 προφίλ σχημάτων (MPEG-7 profile schemas)*. Συγκεκριμένα, οι «δομές» *MDS* περιλαμβάνουν σετ των εργαλείων περιγραφής (*Descriptors, Description Schemes*) των οντοτήτων γενικής χρήσης και των πολυμεσικών οντοτήτων [226], [280]. Χρησιμοποιούνται για την περιγραφή υλικού που, στη γενικότερη περίπτωση, είναι αποθηκευμένο σε πολλά μέσα, και κατηγοριοποιούνται σύμφωνα με το παρακάτω σχήμα.



Σχήμα 7-1. Δομή σχημάτων περιγραφής πολυμέσων MDS (Multimedia Description Schemes).

7.1.2 Δομή του προτύπου MPEG-7

Το πρότυπο MPEG-7 (Multimedia Content Description Interface) στοχεύει στην παροχή ενός προτυποποιημένου πυρήνα τεχνολογιών, που να επιτρέπουν την περιγραφή οπτικο-ακουστικού περιεχομένου σε πολυμεσικά περιβάλλοντα. Δεν αποτελεί μια ξεχωριστή γλώσσα προγραμματισμού αλλά μια δομή βασισμένη σε XML. Για την επίτευξη του στόχου αυτού, το MPEG-7 προτυποποιεί τα παρακάτω συστατικά, εμφανιζόμενα από το μερικότερο στο γενικότερο: Δεδομένα (Data), Χαρακτηριστικό (Feature), Περιγραφέας (Descriptor), Τιμή Περιγραφέα (Descriptor Value), Σχήμα Περιγραφής (Description Scheme), Περιγραφή (Description), Κωδικοποιημένη Περιγραφή (Coded Description), Γλώσσα Ορισμού Περιγραφών (Description Definition Language) [164]-[165], [226], [280], [283].



Σχήμα 7-2. Δόμηση προτύπου MPEG-7 (XSD/XML).

Όπως φαίνεται και στο σχ. 7-2, τα προαναφερθέντα εργαλεία χρησιμοποιούνται αρχικά για να οριστούν οι προδιαγραφές μιας περιγραφής και κατόπιν η ίδια η περιγραφή. Επίσης, το πρότυπο παρέχει ελευθερία προσθήκης νέων περιγραφών (πέρα από τις προτεινόμενες), αυξάνοντας την ευελιξία και προσαρμοστικότητα του. Η επιλογή της γλώσσας XML, πάνω στην οποία βασίστηκε το πρότυπο, έγινε καθώς, τόσο η διαδεδομένη χρήση της, όσο και η απλότητά της, ικανοποιούν σε πολύ μεγάλο βαθμό τις απαιτήσεις για διαλειτουργικότητα και ευρύτητα χρήσης του προτύπου [164]-[165], [226], [280], [283].

7.1.3 Περιγραφείς και σχήματα περιγραφής

Σύμφωνα με όλα τα παραπάνω, μια MPEG-7 περιγραφή δομείται ιεραρχικά από τους περιγραφείς χαμηλού επιπέδου *Ds*, που αποτελούν χαρακτηριστικές παραμέτρους ανάλυσης (features). Οι περιγραφείς *Ds* συγκροτούνται σε σχήματα περιγραφής *DSs*, τα οποία με τη σειρά τους συνδυάζονται με άλλα δομικά στοιχεία για να ορίσουν τις τελικές προδιαγραφές των περιγραφών. Έτσι, στόχος των διαφόρων MPEG-7 εφαρμογών είναι η επιλογή των καταλληλότερων, ανά περίπτωση, στοιχείων *D*, *DS* για τον ορισμό και την εξαγωγή των τελικών περιγραφών. Στη συνέχεια παρατίθενται οι κυριότεροι περιγραφείς *Ds* και τα

αντίστοιχα σχήματα *DSs*, που βρίσκονται στον πυρήνα του προτύπου MPEG-7, αλλά και θεωρούνται κατάλληλοι για την προτεινόμενη οντολογία διαχείρισης περιεχομένου.

Όπως προκύπτει από το σχ. 7-1, ανάμεσα στα βασικά στοιχεία συμπεριλαμβάνονται τα εργαλεία αναφοράς και εντοπισμού των μέσων. Στο πλαίσιο αυτό προτείνονται τα σχήματα “MediaURL(D)” και “MediaLocator(DS)”, που στην απλούστερη περίπτωση παρέχουν τη «διεύθυνση» του μέσου (διαδικτυακή διεύθυνση URL, ή διαδρομή αρχείου filepath). Επίσης συμπεριλαμβάνονται οι περιγραφείς χρονισμού “Time (DS)” και “MediaTime (DS)”, που ορίζουν το χρόνο στα φαινόμενα «πραγματικού κόσμου», αλλά και στο εσωτερικό των οπτικοακουστικών καταγραφών. Από το συνδυασμό των παραπάνω προκύπτουν τα σχήματα “Segment(DS)”, “VideoSegment(DS)” και “AudioSegment(DS)”, που καθορίζουν τμήματα (segments) του καταγεγραμμένου υλικού. Ακόμη, υπάρχουν σχήματα περιγραφής των μέσων (“MediaIdentification(DS)”, “MediaCoding(DS)”, “MediaFormat(DS)”, κλπ), που παρέχουν πληροφορίες για τις συνθήκες δημιουργίας του περιεχομένου, την ποιότητα και τα χαρακτηριστικά του υλικού (συμπίεση, κωδικοποίηση, κλπ), καθώς και τους τρόπους αξιοποίησής του. Ιδιαίτερα χρήσιμα είναι τα «σχήματα σχολιασμού», τα οποία περιέχουν περιγραφές «ελεύθερου κειμένου» (Textual Description), καθώς και «δομημένες περιγραφές κειμένου» (Structured Annotation). Στο πλαίσιο αυτό, έχουν προτυποποιηθεί σχήματα που προορίζονται για την περιγραφή τόπων “Place(DS)”, προσώπων “Person(DS)”, οργανισμών “Organization(DS)”, κ.ά. [280], [319]-[320].

Εκτός των παραπάνω, υπάρχουν διάφορα σχήματα περιγραφής, που προκύπτουν από το συνδυασμό χαρακτηριστικών παραμέτρων ήχου - εικόνας (audio-video features) και τα οποία χρησιμοποιούνται για την περιγραφή επιλεγμένων τμημάτων των οπτικοακουστικών καταγραφών. Με τον τρόπο αυτό διευκολύνονται οι διαδικασίες ανάλυσης, πλοήγησης και αναζήτησης, καθώς πλέον οι περιγραφές συντάσσονται με βάση πηγαία χαρακτηριστικά του περιεχομένου και όχι με επιπρόσθετες πληροφορίες κειμένου. Μία από τις πιο σημαντικές παραμέτρους περιγραφής των αρχείων κινούμενης εικόνας, είναι η ένταση κίνησης. Έτσι, έχουν προταθεί διάφοροι MPEG-7 περιγραφείς κίνησης εικόνας [321], μεταξύ των οποίων και το σχήμα περιγραφής “MotionActivity(DS)”. Το σχήμα αυτό περιέχει τέσσερις μεταβλητές (attributes): την ένταση κίνησης (*Intensity of Activity*), την κατεύθυνση (*Direction of Activity*), τη χωρική και χρονική κατανομή (*Spatial Distribution of Activity*, *Temporal Distribution of Activity*). Η ένταση, που αποτελεί και την πιο βασική παράμετρο, υπολογίζεται από τη διασπορά των διανυσμάτων κίνησης (motion vectors), και λαμβάνει από πολύ χαμηλές έως πολύ υψηλές τιμές που κλιμακώνονται σε πέντε επίπεδα δομημένου κειμένου: {*Very Low Intensity*, *Low intensity*, *Medium Intensity*, *High Intensity*, *Very High Intensity*}. Η τυποποίηση του περιγραφέα έγινε έπειτα από πειραματικές παρατηρήσεις σχετικά με το πώς ο άνθρωπος προσλαμβάνει τη «δραστηριότητα κίνησης βίντεο» [321]-[322].

Παράλληλα, έχουν τυποποιηθεί αντίστοιχα σχήματα για την περιγραφή των ηχητικών πληροφοριών. Οι περιγραφείς αυτοί καλούνται να παρουσιάσουν και να ταυτοποιήσουν τα διάφορα ηχητικά σήματα με βάση τα ενεργειακά και φασματικά τους χαρακτηριστικά, καθώς και με πρόσθετα αντιληπτικά κριτήρια. Έτσι, μερικοί από τους MPEG-7 περιγραφείς ήχου χαμηλού επιπέδου είναι οι ακόλουθοι [226]-[227], [280]-[283]:

Λογαριθμικός χρόνος ανόδου σήματος (LogAttackTime D): πρόκειται για τη λογαριθμική έκφραση του χρόνου που απαιτείται για τη μετάβαση του σήματος από ένα προκαθορισμένο κατώφλι χαμηλής στάθμης στη μέγιστη τιμή. Δίνει ένα μέτρο της βιαιότητας του ήχου, που επιτρέπει το διαχωρισμό των κρουστικών σημάτων από τα πιο ομαλά.

Ηχητική Ισχύς (AudioPower D): περιγράφει τη στιγμιαία τιμή της ηχητικής ενέργειας και εφαρμόζεται σε επλεγμένα τμήματα, των οποίων η διάρκεια είναι ακέραια πολλαπλάσια των 10 msec. Προκύπτει από το άθροισμα όλων των δειγμάτων στιγμιαίας ισχύος, και είναι χρήσιμος για το διαχωρισμό σημάτων που διαφέρουν ενεργειακά.

Περιβάλλουσα Ηχητικού Φάσματος Ισχύος (AudioSpectrumEnvelope D): πρόκειται για ένα διάνυσμα που περιγράφει το βραχυπρόθεσμο φάσμα ενέργειας ενός ηχητικού σήματος, που είναι χωρισμένο σε διαστήματα οκτάβας (λογαριθμικό - οκταβικό σχήμα). Εκτός των τιμών φασματικής ισχύος, παρέχονται η μικρότερη και η μεγαλύτερη συχνότητα των περιοχών. Ο υπολογισμός του μπορεί να γίνει με χρήση του βραχυπρόθεσμου μετασχηματισμού Fourier, αλλά και με τράπεζες φίλτρων. Χρησιμοποιείται για την περιγραφή ηχητικών σημάτων με βάση φασματικά χαρακτηριστικά, καθώς και για την εξαγωγή σπεκτρογραφημάτων.

Κέντρο βάρους ηχητικού φάσματος (AudioSpectrumCentroid D): είναι το κέντρο βαρύτητας του λογαριθμικού φάσματος ισχύος και υπολογίζεται από τον παρακάτω τύπο:

$$ASC = \frac{\sum_{k=1}^{N/2} \log_2(f[k]/1000) \cdot P_x(k)}{\sum_{k=1}^{N/2} P_x(k)} \quad (7-1)$$

όπου $P_x[k]$ είναι η ισχύς του σήματος εισόδου x , στη συχνότητα $f[k]$. Η παραπάνω έκφραση περιλαμβάνει αντιληπτικά κριτήρια που σχετίζονται με τις ιδιότητες της ανθρώπινης ακοής στην περιοχή του 1kHz. Μπορεί να υπολογιστεί με τη βοήθεια του μετασχηματισμού FFT, καθώς και από τις τιμές του διανύσματος AudioSpectrumEnvelope(D).

Διασπορά ηχητικού φάσματος (AudioSpectrumSpread D): είναι ένα μέτρο διασποράς του φάσματος ισχύος γύρω από το φασματικό κέντρο βάρους ASC και υπολογίζεται ως εξής:

$$ASS = \sqrt{\frac{\sum_{k=1}^{N/2} [\log_2(f[k]/1000) - ASC]^2 \cdot P_x(k)}{\sum_{k=1}^{N/2} P_x(k)}} \quad (7-2)$$

Ο περιγραφέας ASS είναι κατάλληλος για την ταυτοποίηση στενοζωνικών σημάτων, όπως είναι οι τονικοί ήχοι, σε αντίθεση με τους κρουστικούς ήχους και τα σήματα θορύβου.

7.1.4 Προτεινόμενες εφαρμογές MPEG-7 στη βιοϊατρική τεχνολογία

Όπως προκύπτει και από το κείμενο εργασίας (working document) που αφορά στις προτεινόμενες MPEG-7 εφαρμογές [323], ο τομέας της βιοϊατρικής τεχνολογίας αποτελεί ένα από τα πεδία για τα οποία προορίζεται η τεχνολογία MPEG-7. Είναι γεγονός ότι, οι νέες βιοϊατρικές ερευνητικές κατευθύνσεις και οι προσπάθειες διεπιστημονικής συνεργασίας γύρω από τα θέματα των επιστημών υγείας, απαιτούν τη συστηματικοποίηση καταχώρησης και διαχείρισης των ιατρικών παρατηρήσεων, μέσα από τη σύσταση κατάλληλων οντολογιών [90]-[93]. Στις πληροφορίες προς διαχείριση, εκτός από τις κλινικές παρατηρήσεις, θα πρέπει να υπάρχει δυνατότητα πρόσβασης σε «αρχεία ιατρικών εξετάσεων», μεταξύ των οποίων

περιλαμβάνονται, ως επί το πλείστον, οπτικοακουστικά δεδομένα [92]. Στο πλαίσιο αυτό, το πρότυπο MPEG-7 αποτελεί μια ραγδαία εξελισσόμενη τεχνολογία, που αφενός υπόσχεται να ικανοποιήσει τις ανάγκες των σύνθετων βιοϊατρικών οντολογιών, αφετέρου είναι σε θέση να εξασφαλίσει την απαραίτητη τυποποίηση, συγχρόνως, με τη διαλειτουργικότητα και την εύκολη επεκτασιμότητα [93], [226], [323].

Η ιατρική είναι μια περιοχή στην οποία, η οπτική αναγνώριση είναι συχνά μια σημαντική τεχνική για τη διάγνωση. Στην ιατρική βιβλιογραφία υπάρχει πλήθος από φωτογραφίες, τρισδιάστατες εικόνες και άτλαντες, που απεικονίζουν, υπό διαφορετικές, συχνά, κλίμακες, τις φυσιολογικές και παθολογικές μορφές των διαφόρων ανθρώπινων οργάνων. Σε πολλές περιπτώσεις, η αποτελεσματική διάγνωση προκύπτει ως αποτέλεσμα σύγκρισης μιας δεδομένης εικόνας με τις αντίστοιχες των διαθέσιμων βιβλιοθηκών. Η ποσότητα πληροφοριών που έχει καταχωρηθεί μέχρι σήμερα σε τέτοιους άτλαντες είναι ήδη πολύ μεγάλη και συνεχίζει να αυξάνεται, ενώ η περιγραφή και διαχείριση του περιεχομένου αυτού, μόνο με τη χρήση δομών κειμένου, είναι ιδιαίτερα δύσκολη. Επομένως, υπάρχει μια αυξανόμενη ζήτηση για τις μηχανές αναζήτησης να μπορούν να αποκριθούν σε ερωτήσεις (queries) με βάση τα πηγαιία χαρακτηριστικά των εικόνων, αλλά και όλων των βιοϊατρικών πληροφοριών, γενικότερα. Παράλληλα, έχοντας υπόψη τη ραγδαία εξέλιξη των τεχνικών τρισδιάστατης-απεικόνισης (εικόνες μαγνητικής τομογραφίας, απεικονίσεις εγκεφαλικής δραστηριότητας, κλπ), η τυποποίηση κατάλληλων περιγραφών και αντίστοιχων σχημάτων περιγραφής στην περίπτωση αυτή θα οδηγήσει στην αποτελεσματικότερη διαχείριση και ενοποίηση των εργαλείων ανάλυσης, σε ακολουθίες δεδομένων δύο, τριών και περισσότερων διαστάσεων. Επιπρόσθετα, προτείνεται η διασύνδεση των εικόνων με σήματα βιοακουστικής καταγραφής αλλά και κλινικές παρατηρήσεις, με στόχο την διευκόλυνση της διαγνωστικής διαδικασίας (π.χ. συνδυασμένη ανάλυση ενός ηχητικού σήματος «βήχα» με μια θωρακική ακτινογραφία) [323].

Στο πλαίσιο αυτό, έχουν ήδη ξεκινήσει οι πρώτες ερευνητικές προσπάθειες αξιοποίησης του προτύπου MPEG-7 στη βιοϊατρική τεχνολογία, με στόχο την καλύτερη διαχείριση των ιατρικών οπτικοακουστικών δεδομένων, τη βέλτιστη αξιοποίηση του ιατρικού ανθρώπινου δυναμικού, την υποβοήθηση της διαγνωστικής διαδικασίας μέσα από εμπειρα συστήματα ηλεκτρονικής ιατρικής και την αξιοποίηση δομών τηλεϊατρικής [92]. Παράλληλα, έχουν προταθεί MPEG-7 σχήματα περιγραφής για την οπτικοακουστική εποπτεία των διαδικασιών ιατρικής εξέτασης, όπως συμβαίνει στην περίπτωση εγκεφαλογραφικών καταγραφών [150], αλλά και στην παρούσα εφαρμογή, για την τεχνικά υποβοηθούμενη παρακολούθηση της πειραματικής διαδικασίας [226]. Τελευταία έχει δημοσιευθεί και μια ακόμη εργασία στον τομέα της γαστρεντερολογίας [324], που αφορά στη χρήση του πρωτοκόλλου MPEG-7, σε συνδυασμό με μικροσκοπικές κάψουλες ενδοσκόπησης του γαστρεντερικού συστήματος. Η σημαντικότητα της εργασίας αυτής έγκειται, τόσο στην τυποποίηση κατάλληλων σχημάτων περιγραφής για την αυτοματοποίηση της συγκεκριμένης διαγνωστικής εξέτασης, όσο και στη μελλοντική της αξιοποίηση σε συνδυασμό με άλλες βιοϊατρικές μεθόδους και εναλλακτικά σχήματα περιγραφής, σαν κι αυτά που τυποποιούνται στην τρέχουσα εργασία.

7.2 Διαχείριση περιεχομένου και ανάλυση των πειραματικών καταγραφών

Η προτεινόμενη οντολογία διαχείρισης περιεχομένου και ανάλυσης έχει σαν στόχο την ενσωμάτωση όλων των πειραματικών παρατηρήσεων σε ένα ευέλικτο σχήμα περιγραφής. Όπως έχει ήδη επισημανθεί, ο συγκεκριμένος σχεδιασμός έγινε με βάση τη φιλοσοφία του προτύπου MPEG-7, απόφαση που λήφθηκε εξαιτίας πολλών παραγόντων. Καταρχήν, η

διαδικασία πολυδιακριτικής σάρωσης *MRS* και η καταχώρηση των «αναφορών χρονισμού» των εντοπισμένων συμβάντων στα ομώνυμα αρχεία (*.mrs), παρουσιάζει αρκετές ομοιότητες με τις δομές των σχημάτων περιγραφής που προτείνονται από το MPEG-7. Επίσης, οι δείκτες σημαντικότητας καναλιών (*IMP_{ind}*) και η πληροφορία προτύπου μπορούν εύκολα να αποδοθούν από περιγραφείς δομημένου κειμένου. Μάλιστα, υπάρχουν πολλά κοινά σημεία ανάμεσα στην ιεραρχική ταξινόμηση προτύπων, που προτείνεται από το αναλυτικό λεξικό *ASFC*, και τη φιλοσοφία των MPEG-7 σχημάτων ταξινόμησης ηχητικών προτύπων [227], [275]-[276]. Παράλληλα, το MPEG-7 παρέχει τη δυνατότητα ενσωμάτωσης και πρόσθετων στοιχείων, όπως είναι τα μετα-δεδομένα περιγραφής των πειραματικών διαδικασιών και των υποκειμένων, τα σχήματα περιγραφής των οπτικοακουστικών καταγραφών εποπτείας, καθώς επίσης οι εξαγόμενες πληροφορίες χωρικής απεικόνισης και ηχητικής απόδοσης του παραγόμενου βιοακουστικού πεδίου (ASF-AVI).

Έτσι, η κύρια εργασία που απομένει για την πλήρη συμμόρφωση με το πρότυπο, αφορά στην εξαγωγή των δομημένων κειμένων (XML), δηλαδή στην παρεμβολή ενός MPEG-7 «μεταφραστή» που δέχεται στην είσοδο τα δυαδικά (binary) αρχεία *.mrs και δίνει στην έξοδο τα αρχεία *.xml των MPEG-7 περιγραφών. Αν και κάτι τέτοιο δεν μας απασχόλησε ιδιαίτερα στην παρούσα φάση, ωστόσο, η μελλοντική του διευθέτηση κρίνεται ιδιαίτερα σημαντική, καθώς θα επιτρέψει την πλήρη τυποποίηση και διαλειτουργικότητα του προτεινόμενου πρωτοκόλλου ιατρικής μελέτης. Το στοιχείο αυτό, που αποτελεί ίσως το πιο σημαντικό κριτήριο επιλογής του προτύπου MPEG-7 (αντί της υιοθέτησης μιας καινούργιας φιλοσοφίας διαχείρισης δεδομένων), αναμένεται ότι θα συμβάλει στην προώθηση της διεπιστημονικής συνεργασίας και στην εξέλιξη όλων των ερευνητικών προσεγγίσεων βιοακουστικής παρακολούθησης της γαστρεντερικής κινητικότητας.

7.2.1 Σχεδιασμός εννοιολογικής δομής περιγραφής περιεχομένου

Όπως έχει ήδη επισημανθεί από το πρώτο κίονας κεφάλαιο του παρόντος, στόχος της προτεινόμενης μεθοδολογίας είναι η αξιοποίηση ευέλικτων δομών διαχείρισης περιεχομένου και ανάλυσης, με στόχο την αντιμετώπιση των προβλημάτων που προκύπτουν εξαιτίας της πολύ μεγάλης διάρκειας των βιοακουστικών καταγραφών γαστρεντερικής κινητικότητας και των ιδιομορφιών της μη εποπτευόμενης πειραματικής διαδικασίας. Αυτός ήταν και ένας από τους βασικούς λόγους που εμπόδισε την εξέλιξη των σχετικών ερευνητικών προσεγγίσεων, καθώς κύριο χαρακτηριστικό των καταγραφών είναι ο μεγάλος όγκος πρωτογενούς πληροφορίας, με τα αντίστοιχα προβλήματα επεξεργασίας και διαχείρισης. Για την αντιμετώπιση τους έχουν προταθεί τεχνικές μακροσκοπικής ανάλυσης, μέσα από σήματα περιβάλλουσας και εργαλεία συνοπτικής απεικόνισης [64], [69], [72], καθώς και η μείωση του χρόνου παρακολούθησης λαμβάνοντας αντιπροσωπευτικά δείγματα του συνολικού φαινομένου [42], [74]-[75]. Στην παρούσα προσέγγιση επιχειρείται συγκερασμός των παραπάνω τεχνικών, εισάγοντας την έννοια των «συμβάντων» [217], [226].

Τα συμβάντα αποτελούν περιόδους που σηματοδοτούν την εξέλιξη του φαινομένου, δηλαδή τα δομικά στοιχεία του σήματος που απαιτείται να έχουν τη μέγιστη δυνατή αναλυτικότητα, καθώς αποτελούν φορείς της πληροφορίας σε επίπεδο σύντομου χρόνου. Διέπονται από ορισμένα χαρακτηριστικά τα οποία μπορούν να αξιολογηθούν, τόσο σε επίπεδο ανίχνευσης, όσο και σε επίπεδο περιγραφής [226]. Έτσι, εκτός από τις «καμπύλες σύννοψης», που αναπαριστούν τη συνολική εξέλιξη του φαινομένου, η αλληλοδιαδοχή των συμβάντων και τα αντίστοιχα σχήματα περιγραφής παρέχουν μια διαφορετική προσέγγιση παρατήρησης του συνολικού φαινομένου. Στο πλαίσιο αυτό, εκτός από τους σχολιασμούς

που αναφέρθηκαν στα προηγούμενα κεφάλαια (δείκτης σημαντικότητας καναλιών, είδος προτύπου, κλπ), υπάρχει δυνατότητα αξιοποίησης των οπτικοακουστικών αρχείων ASF-AVI για την παρουσίαση των μεταβολών του βιοακουστικού πεδίου, τόσο σε επίπεδο συμβάντος, όσο και συνολικά, μέσα από τεχνικές «υπέρθεσης» ηχητικών σταθμών (εξ. 6-8). Γενικά, τα συμβάντα παρέχουν τη δυνατότητα απομόνωσης των ανεπιθύμητων παρεμβολών, αλλά και την ολοκλήρωση όλων των πληροφοριών ως προς το χρόνο, παρέχοντας νέα, συνδυαστικά σχήματα σύννοησης [226].

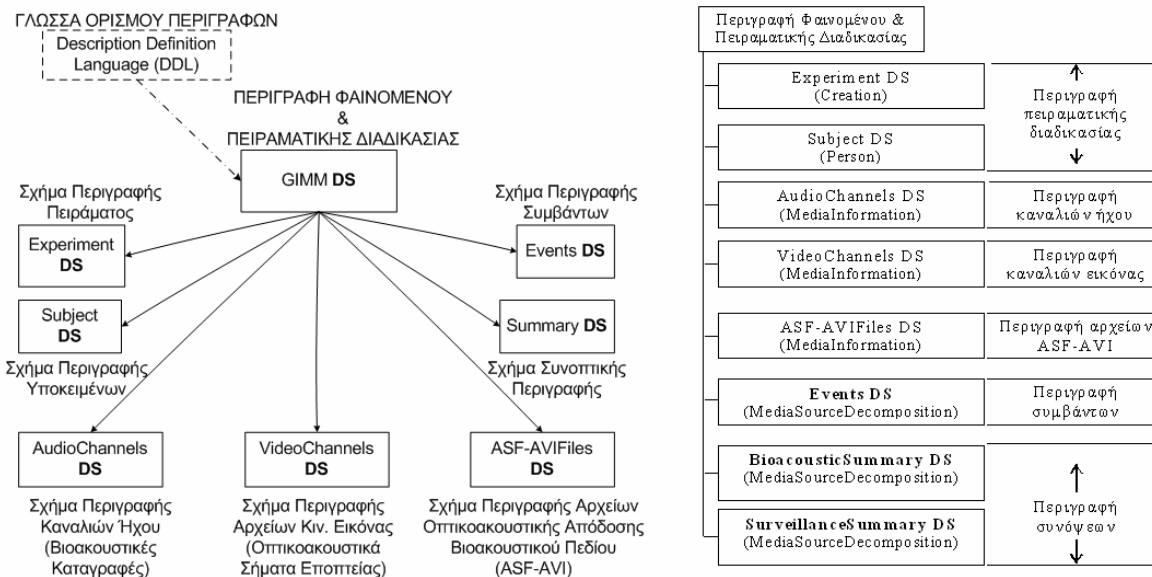
Σημαντική, βέβαια, παράμετρος στην παρατήρηση και την ανάλυση των φαινομένων παρατεταμένου χρόνου είναι και η περιγραφή των συνθηκών εξέλιξης, η οποία αναφέρεται στα «ευρήματα» των συστημάτων αυτόματης και χειροκίνητης εποπτείας. Δηλαδή, η κρίση των συνθηκών από τον παρατηρητή γίνεται μέσω οπτικών και ακουστικών ερεθισμάτων που λαμβάνει περιφερειακά του φαινομένου, με τη βοήθεια του συστήματος οπτικοακουστικής παρακολούθησης. Παράλληλα, ιδιαίτερα χρήσιμες θεωρούνται και οι πληροφορίες σχετικά με το ιατρικό ιστορικό των υποκειμένων, αλλά και όλες τις μεταβλητές της πειραματικής διαδικασίας (κατάσταση ύπνου, φαρμακευτική αγωγή, περίοδος νηστείας ή μεταγευματική φάση, είδος γεύματος, κλπ). Τα στοιχεία αυτά συμβάλλουν στην περαιτέρω ομαδοποίηση των παρατηρήσεων, ανάλογα με τις συνθήκες διεξαγωγής των ψυχο-φυσιολογικών μετρήσεων και τα κλινικά ευρήματα του ιατρικού προσωπικού [226].

Με βάση τα παραπάνω, τα προτεινόμενα σχήματα περιγραφής διακρίνονται ανάλογα με τα οπτικοακουστικά αρχεία στα οποία αναφέρονται, αλλά και με βάση τις πληροφορίες που παρέχουν, δηλαδή τους σκοπούς που καλούνται να εξυπηρετήσουν. Σχετικά με την πρώτη διάκριση, έχουμε τις περιγραφές που αφορούν στο οπτικοακουστικό σύστημα εποπτείας, στα επιμέρους κανάλια των πολυκάναλων βιοακουστικών καταγραφών, καθώς και στα αρχεία οπτικοακουστικής απόδοσης του κοιλιακού ηχητικού πεδίου ASF-AVI. Όσον αφορά τη διάκριση με βάση το περιεχόμενο, έχουμε τα σχήματα περιγραφής των συνθηκών διεξαγωγής της πειραματικής διαδικασίας, τα σχήματα συνοπτικής περιγραφής του προς εξέταση φαινομένου και τα σχήματα περιγραφής των επιμέρους συμβάντων. Απαραίτητοι είναι, επίσης, οι περιγραφές που αναφέρονται στα εμπλεκόμενα αρχεία, καθορίζοντας την ακριβή τους θέση (διαδρομή αρχείου *filepath*), το είδος της παρεχόμενης πληροφορίας, τους οπτικοακουστικούς μετατροπείς λήψης που χρησιμοποιήθηκαν και τις χρονικές στιγμές έναρξης και λήξης της πειραματικής διαδικασίας [226]. Η προτεινόμενη εννοιολογική δομή παρέχει τη δυνατότητα πλοήγησης, σύγκρισης, αναζήτησης και απεικόνισης των διαφόρων στιγμιότυπων των καταγραφών, αξιοποιώντας αποκλειστικά και μόνο τις πληροφορίες που ενσωματώνονται στην τελική περιγραφή. Παράλληλα, υπάρχει η δυνατότητα αξιοποίησης των σχημάτων περιγραφής για την εφαρμογή πρόσθετων τεχνικών μετα-επεξεργασίας, που επιτρέπουν την ανάπτυξη εναλλακτικών εργαλείων ανάλυσης και την υιοθέτηση νέων μεθοδολογικών προσεγγίσεων (σχετικές λεπτομέρειες δίνονται στη συνέχεια) [226].

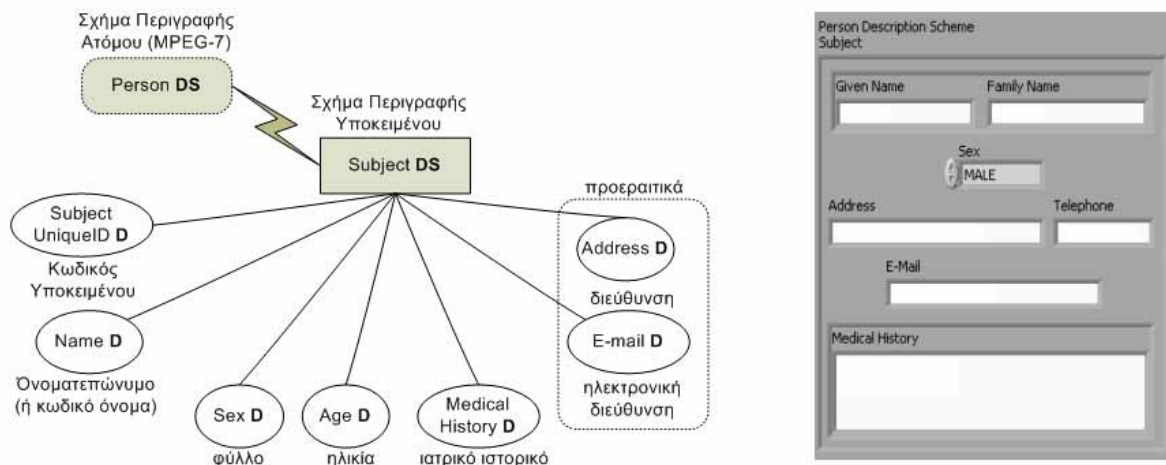
7.2.2 Προτεινόμενα σχήματα περιγραφής των πειραματικών καταγραφών

Η προτεινόμενη οντολογία διαχείρισης οπτικοακουστικού περιεχομένου, για τη μελέτη της γαστρεντερικής κινητικότητας (*Gastro-Intestinal Motility Monitoring Description Scheme - GIMM DS*), σχεδιάστηκε σύμφωνα με τις προδιαγραφές που αναλύθηκαν πιο πάνω και αποτελείται από επτά επιμέρους οντότητες (σχ. 7-3). Ανάμεσα σ' αυτές είναι τα σχήματα περιγραφής της πειραματικής διαδικασίας (*Experiment DS*), των υποκειμένων (*Subject DS*) και όλων των εμπλεκόμενων οπτικοακουστικών καναλιών (αρχείων). Τα τελευταία αφορούν στις βιοακουστικές καταγραφές (*AudioChannels DS*), τα σήματα οπτικής παρακολούθησης

(VideoChannels DS) και τα οπτικοακουστικά αρχεία απόδοσης του βιοακουστικού πεδίου ASF-AVI (ASF-AVIFiles DS). Ακόμη, υπάρχουν τα σχήματα σύνοψης (Summary DS) και περιγραφής των συμβάντων (Events DS), όπου το πρώτο επιχειρεί να εξάγει συνοπτικές παραμέτρους παρακολούθησης των φαινομένων, ενώ το δεύτερο εστιάζει στην αναλυτική παρουσίαση όλων των συμβάντων. Το σχήμα σύνοψης Summary DS αναφέρεται στις βιοακουστικές καταγραφές (BioacousticSummary DS), αλλά και στα σήματα οπτικοακουστικής εποπτείας (SurveillanceSummary DS). Στο σχ. 7.3β δίνεται η σχέση των επιμέρους σχημάτων με τα στοιχεία DSs και Ds που προτείνονται από το MPEG-7 [226].



Σχήμα 7-3. Η προτεινόμενη οντολογία ανάλυσης (Gastro-Intestinal Motility Monitoring DS -GIMM DS).



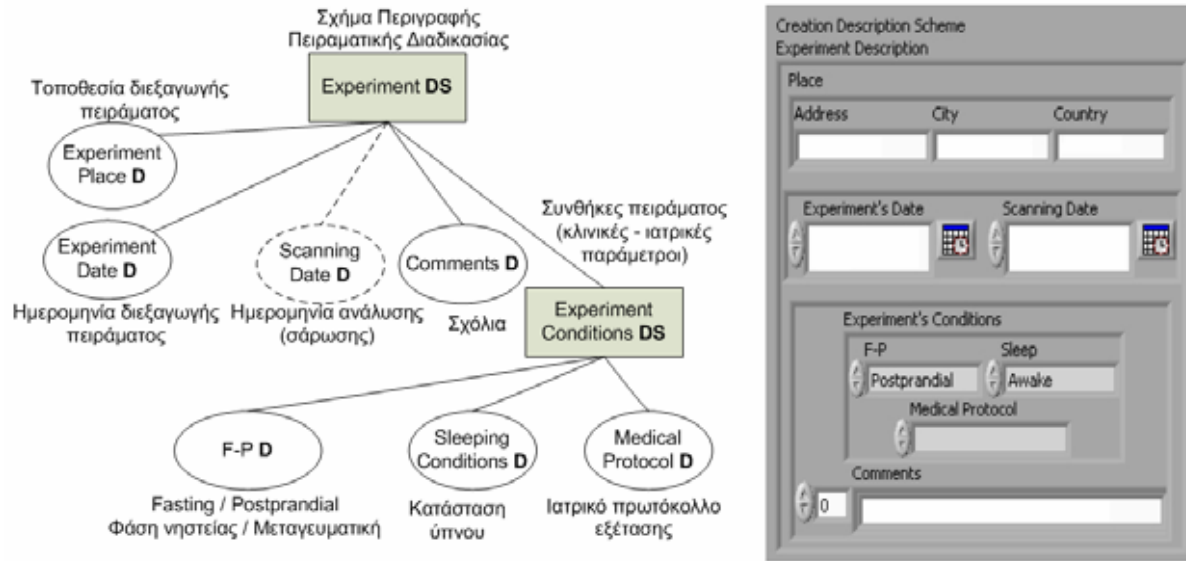
Σχήμα 7-4. Σχήμα περιγραφής υποκειμένων: α. διάγραμμα, β. περιβάλλον διαχείρισης.

Ειδικότερα, στην περίπτωση του σχήματος περιγραφής των υποκειμένων (Subject DS) χρησιμοποιείται το προτεινόμενο, από το MPEG-7, σχήμα περιγραφής ατόμων (Person DS), με τις απαραίτητες, βέβαια, τροποποιήσεις για την προσαρμογή στις ανάγκες της τρέχουσας εφαρμογής (σχ. 7-4). Έτσι, ορίζονται οι περιγραφείς κωδικού (SubjectUniqueID D), ονόματος (Name D), φύλλου (Sex D), ηλικίας (Age D) και, προαιρετικά, οι περιγραφείς διεύθυνσης (Address D) και ηλεκτρονικής διεύθυνσης (E-mail D). Επίσης, εισάγεται ο

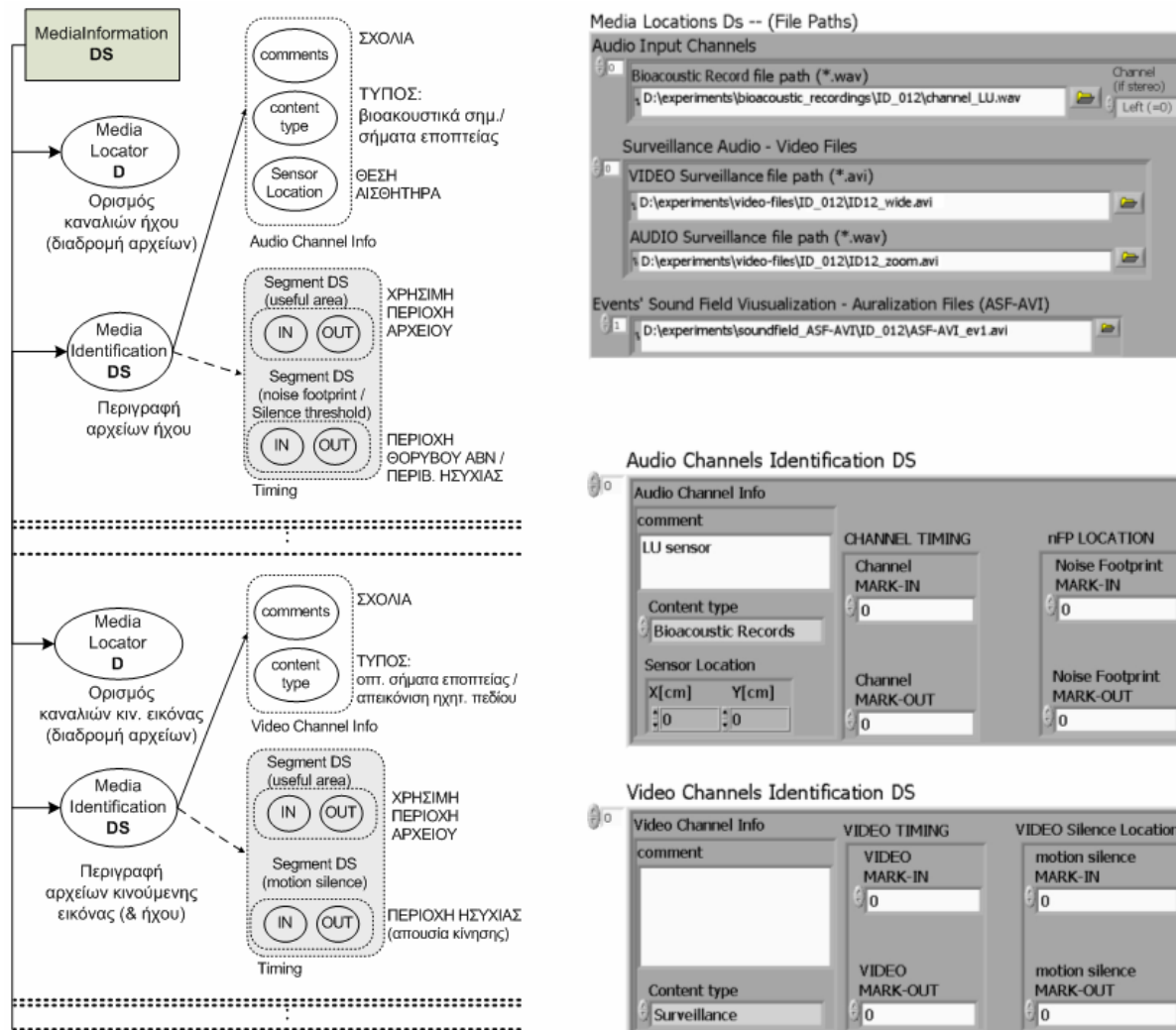
περιγραφείας ιατρικού ιστορικού (MedicalHistory D), ο οποίος έχει τη μορφή σχολιασμών ελεύθερου κειμένου (TextAnnotation D) και παρέχει τη δυνατότητα εισαγωγής χρησιμών στοιχείων που μπορούν να αξιολογηθούν κατά την ιατρική μελέτη. Παρόμοιας φύσης πληροφορίες καταχωρούνται και στο σχήμα περιγραφής των πειραμάτων (Experiment DS, σχ. 7-5). Διακρίνουμε τους «υπο-περιγραφείς» τοποθεσίας και ημερομηνίας διεξαγωγής της πειραματικής διαδικασίας (ExperimentPlace D και ExperimentDate D, αντίστοιχα), καθώς και τον προαιρετικό περιγραφέα της ημερομηνίας ανάλυσης (σάρωση MRS), οπότε και έγινε η καταχώρηση των πληροφοριών (ScanningDate D). Εδώ υπάρχει ένα ακόμη σχήμα για την περιγραφή των «ιατρικών» συνθηκών της πειραματικής διαδικασίας (ExperimentConditions DS). Περιέχει δύο περιγραφείς δομημένου κειμένου (StructuralAnnotation D) που αφορούν στην «κατάσταση διατροφής» (F-P D, fasting / postprandial) και στην «κατάσταση ύπνου» (SleepingConditions D) των υποκειμένων, έναν περιγραφέα για σχόλια (Comments D) και άλλον ένα για το ιατρικό πρωτόκολλο εξέτασης (MedicalProtocol D), αν υπάρχει κάποιο.

Το σχήμα περιγραφής των οπτικοακουστικών μέσων (MediaInformation DS) παρέχει όλες τις απαραίτητες πληροφορίες σχετικά με τη θέση των αντίστοιχων αρχείων, τη φύση του περιεχομένου τους, τα χαρακτηριστικά των οπτικοακουστικών μετατροπέων λήψης, καθώς και κάποια συνοδευτικά σχόλια με πληροφορίες χρονισμού (σχ. 7-6). Στην περίπτωση του σχήματος σύννοψης, χρησιμοποιούνται τα σχήματα ορισμού των τμημάτων (Segment DS) και κατόπιν εφαρμόζονται οι περιγραφείς AudioPower D, AudioSpectrumEnvelope D και VideoMotionActivity D (σχ. 7-7). Για την «αποσύνθεση» (Media Decomposition) των βιοακουστικών καταγραφών (BioacousticSummary DS) χρησιμοποιούνται μόνο οι δύο πρώτοι περιγραφείς, σε αντίθεση με την περίπτωση των οπτικοακουστικών σημάτων εποπτείας (SurveillanceSummary DS) όπου χρησιμοποιούνται και οι τρεις περιγραφείς. Οι τιμές τους μπορούν να υπολογιστούν από την περιβάλλουσα ακουστικής ισχύος P_{av} (εξ. 4-20, § 4.2.3.5), του κυματιδιακού φασματικού φακέλου WSE (εξ. 4-21, § 4.2.3.5) και της καμπύλης έντασης κίνησης I_{mot} (εξ. 2-6, § 2.2.2.3.). Τα τελευταία προκύπτουν από τη διαδικασία σάρωσης MRS (σχ. 4-11) και τις αντίστοιχες τεχνικές ανίχνευσης κίνησης βίντεο (σχ. 2-18). Βέβαια, απαιτούνται και κάποιες μικρές τροποποιήσεις προκειμένου να εξασφαλιστεί η πλήρης συμμόρφωση με τους προτεινόμενους, από το MPEG-7 πρότυπο, περιγραφείς. Έτσι, οι τιμές του περιγραφέα ακουστικής ισχύος προσδιορίζεται για κάθε LB -μπλοκ, ενώ ο περιγραφέας περιβάλλουσας φάσματος προσδιορίζεται από τις ακολουθίες WSE , για δεκαέξι εσωτερικά πλαίσια σε κάθε LB -παράθυρο. Τέλος, για την καταχώρηση της έντασης κίνησης θα πρέπει να γίνει μια κλιμάκωση της καμπύλης I_{mot} με βάση τα πέντε επίπεδα που έχουν τυποποιηθεί στον αντίστοιχο MPEG-7 περιγραφέα.

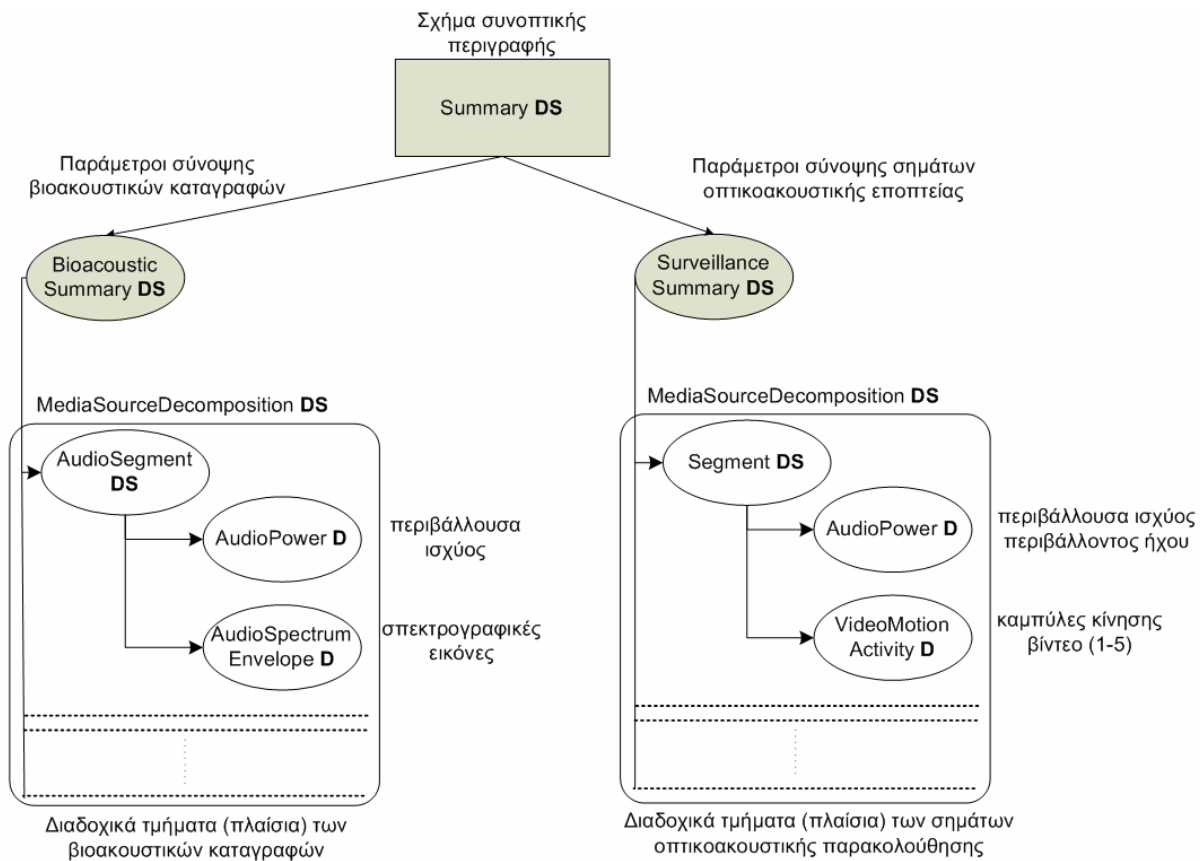
Όσον αφορά το σχήμα περιγραφής των συμβάντων (Events DS), εδώ διακρίνουμε τις συνολικές πληροφορίες, όπως είναι η χρονική διάρκεια και η εκτιμώμενη ισοδύναμη θέση της ηχητικής πηγής (ESL), και τους επιμέρους περιγραφείς που εφαρμόζονται χωριστά σε κάθε βιοακουστικό κανάλι (σχ. 7-8). Για τον καθορισμό της ESL θέσης χρησιμοποιείται ο περιγραφέας συντεταγμένων (Spatial2DCoordinates D). Ακόμη ορίζονται οι περιγραφείς προτύπου (Pattern D) και σημαντικότητας καναλιών (IMPind D), οι οποίοι έχουν μορφή δομημένου κειμένου. Τέλος, χρησιμοποιούνται και κάποιοι χαμηλού επιπέδου περιγραφείς (LLDs), που καλούνται να «εκφράσουν» τα κύρια χαρακτηριστικά των συμβάντων, όπως η ενέργεια, το φασματικό περιεχόμενο και η κρουστικότητα. Οι περιγραφείς αυτοί μπορούν να αξιοποιηθούν για τις ανάγκες αναζήτησης, σύννοψης και πλοήγησης στα συμβάντα ίδιων ή διαφορετικών καταγραφών, ενώ υπάρχει δυνατότητα σύνδεσης με τα αρχεία ASF-AVI και με τις τεχνικές μετα-επεξεργασίας, που παρέχουν πρόσθετα εργαλεία ανάλυσης.



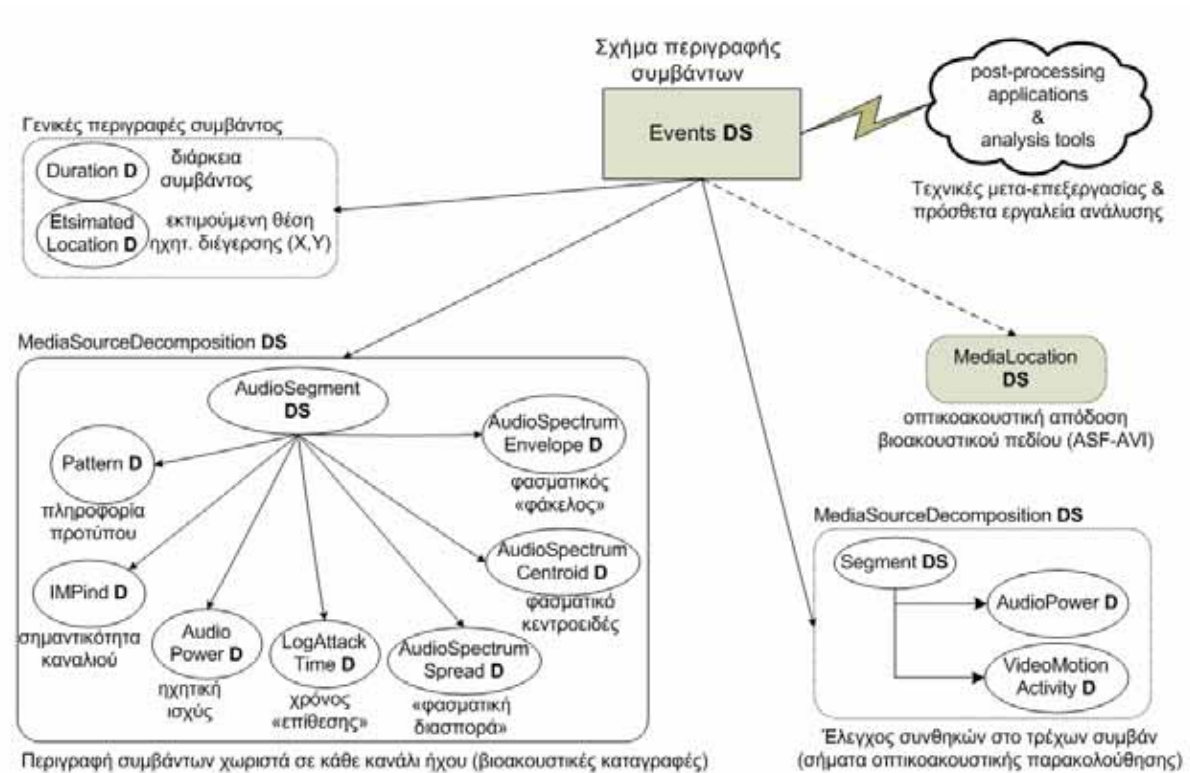
Σχήμα 7-5. Σχήμα περιγραφής πειραματικής διαδικασίας: α. διάγραμμα, β. περιβάλλον διαχείρισης.



Σχήμα 7-6. Σχήμα περιγραφής οπτικοακουστικών μέσων (αρχείων): α. διάγραμμα, β. περιβάλλον διαχείρισης.



Σχήμα 7-7. Σχήμα «σύνοψης» βιοακουστικών καταγραφών & σημάτων οπτικοακουστικής εποπτείας.



Σχήμα 7-8. Σχήμα περιγραφής επιμέρους συμβάντων.

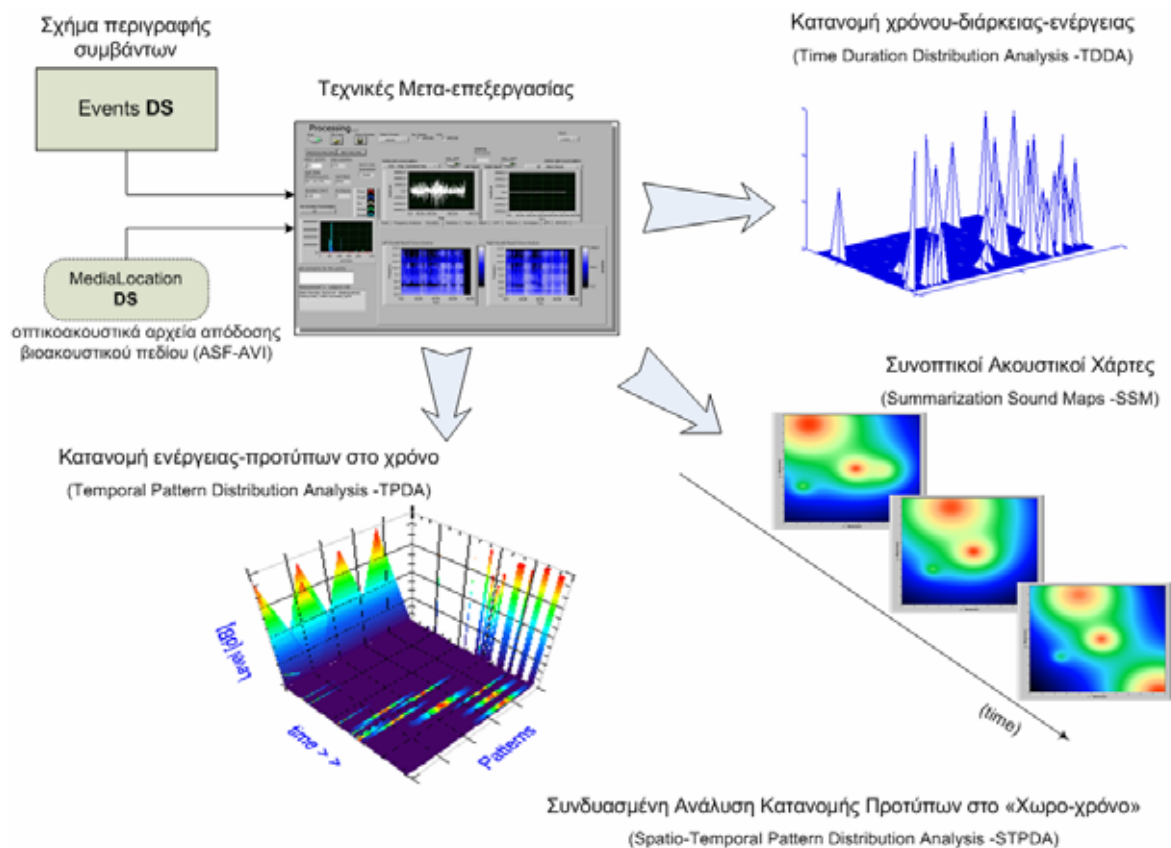
7.2.3 Τεχνικές μετα-επεξεργασίας και εργαλεία ανάλυσης

Το σχήμα περιγραφής που αναλύθηκε πιο πάνω επιτρέπει την αξιοποίηση πρόσθετων εργαλείων μετα-επεξεργασίας, τόσο για τη μελέτη των μεμονωμένων επεισοδίων, όσο και για την παρακολούθηση των καταγραφών, συνολικά. Έτσι, είναι δυνατή η ομαδοποίηση των συμβάντων με βάση τις τιμές των περιγραφέων χαμηλού επιπέδου, των πιθανών κλινικών παρατηρήσεων και του ιατρικού ιστορικού των υποκειμένων, όπως δηλαδή συμβαίνει στις προσεγγίσεις βιοακουστικής ανάλυσης σύντομου χρόνου, για τη διάγνωση διαφόρων παθολογιών [70]-[71]. Η σημαντικότερη, όμως, καινοτομία που εισάγει η νέα μεθοδολογία περιγραφής περιεχομένου, αφορά στη μελέτη των καταγραφών παρατεταμένου χρόνου, με τη βοήθεια του σχήματος περιγραφής των συμβάντων. Τα τελευταία, ως δομικά στοιχεία εξέλιξης του συνολικού φαινομένου, παρέχουν νέες δυνατότητες παρατήρησης και εξαγωγής πρόσθετων παραμέτρων σύνοψης. Αν, λοιπόν, απομονωθούν όλα τα πρότυπα που δεν αποτελούν χρήσιμους κοιλιακούς ήχους, και θεωρηθεί ότι στις χρονικές στιγμές απουσίας συμβάντων επικρατεί απόλυτη σιγή, μπορούν να εξαχθούν καμπύλες περιβάλλουσας και σπεκτρογραφικές εικόνες, σαν κι αυτές του σχ. 4-11 που προκύπτουν από την ανάλυση των πηγαίων καταγραφών. Επιπλέον, από τη χρήση του δείκτη σημαντικότητας IMP_{ind} (επιλογή των κυρίαρχων, κάθε φορά, καναλιών), είναι δυνατή η εξαγωγή πολλαπλών γραφημάτων, για την παρακολούθηση των ενεργειακών μεταβολών ως προς το χρόνο, σε κάθε θέση λήψης και με μικρή επίδραση των φαινομένων συνακρόασης.

Σε επέκταση των παραπάνω, από τη γνώση της διάρκειας των διαφόρων συμβάντων μπορούν να προκύψουν τρισδιάστατα γραφήματα κατανομής ενέργειας - διάρκειας (Time Duration Distribution Analysis -TDDA, σχ. 7-9), τα οποία σύμφωνα με πολλούς ερευνητές είναι ιδιαίτερα περιγραφικά [24]-[25], [72], [161]. Επιπρόσθετα, είναι δυνατή η αξιοποίηση περισσότερων ανεξάρτητων μεταβλητών (π.χ. ενέργεια, διάρκεια, πληροφορία προτύπου, τοπογραφική πληροφορία, κοκ), παρέχοντας συνδυαστικά γραφήματα χωρικής ανάλυσης και εξέλιξης των προτύπων στο χρόνο (Temporal Pattern Distribution Analysis -TPDA, Spatio-Temporal Pattern Distribution Analysis -STPDA) [217], [226]DA) [217], [226]. Σχετικά παραδείγματα απεικόνισης δίνονται στο σχ. 7-9, όπου εκτός από τις καμπύλες περιβάλλουσας ισχύος και τις σπεκτρογραφικές εικόνες του σχ. 4-11, πλέον χρησιμοποιούνται πρόσθετοι χρωματικοί χάρτες τριών διαστάσεων, αλλά και γραφήματα TDDA για διάφορες κατηγορίες προτύπων και σε διάφορες θέσεις λήψης.

Ιδιαίτερα λειτουργικά κρίνονται και τα διαγράμματα ακουστικής χαρτογράφησης, που προκύπτουν από τη σύνδεση των συμβάντων με τα αρχεία οπτικοακουστικής απόδοσης του βιοακουστικού πεδίου (ASF-AVI). Σύμφωνα δε, με τις τεχνικές χωρο-χρονικής ολοκλήρωσης που παρουσιάστηκαν στο προηγούμενο κεφάλαιο (εξ. 6-8), είναι δυνατό να δημιουργηθούν οι εικόνες συνοπτικής χαρτογράφησης SSM, που απεικονίζουν τη χωρική κατανομή της βιοακουστικής ενέργειας σε διάφορα χρονικά διαστήματα (σχ. 7-9). Η χρησιμότητα αυτών των εργαλείων είναι πολύ μεγάλη, στοιχείο που ενισχύεται ακόμη περισσότερο από το γεγονός ότι το ιατρικό προσωπικό είναι αρκετά εξοικειωμένο με την παρατήρηση εικόνων. Αν, λοιπόν, «ολοκληρώσουμε» τις εικόνες χωρικής κατανομής της ηχητικής στάθμης (εξ. 6-8) με διαστήματα ολοκλήρωσης 3 λεπτών, τότε μια τυπική καταγραφή διάρκειας 2 ωρών μπορεί να μελετηθεί με τη βοήθεια 40, μόλις, εικόνων ακουστικής χαρτογράφησης, που αντιστοιχούν σε ένα αρχείο κινούμενης εικόνας διάρκειας 1,6 sec. Μάλιστα, οι χάρτες αυτοί μπορούν να εξαχθούν χωριστά για τις διάφορες κλάσεις ηχητικών προτύπων, παρέχοντας εναλλακτικές δυνατότητες μελέτης της χωρο-χρονικής κατανομής τους (STPDA). Πιο σημαντικό απ' όλα θεωρείται το γεγονός ότι τα παραπάνω εργαλεία ανάλυσης είναι

παραμετροποίηση, ανάλογα με τις επιλογές του χρήστη (modular analysis), ενώ για την εφαρμογή τους απαιτούνται μόνο οι περιγραφές της προτεινόμενης οντολογίας (GIMM DS) και η πρόσβαση στα ASF-AVI αρχεία οπτικοακουστικής απόδοσης των συμβάντων.



Σχήμα 7-9. Τεχνικές μετα-επεξεργασίας και πρόσθετα εργαλεία ανάλυσης.

7.3 Επισκόπηση 7ου κεφαλαίου

Το κεφάλαιο αυτό έχει αντικείμενο το σχεδιασμό μια νέας οντολογίας διαχείρισης περιεχομένου και ανάλυσης, για τη μελέτη όλων των πειραματικών δεδομένων βιοϊατρικής παρακολούθησης. Έτσι, αφού παρουσιάστηκαν οι ιδιαιτερότητες του εγχειρήματος και οι αντίστοιχες τεχνικές προδιαγραφές της ιατρικής ερευνητικής διαδικασίας, αποφασίστηκε η υιοθέτηση της φιλοσοφίας του προτύπου MPEG-7, που αποτελεί τεχνολογία αιχμής και προτείνεται για ιδιαίτερα απαιτητικές εφαρμογές περιγραφής πολυμεσικού περιεχομένου. Στο πλαίσιο αυτό, αναλήθηκαν τα βασικά στοιχεία του προτύπου MPEG-7, με έμφαση στους περιγραφείς και τα σχήματα περιγραφής που χρησιμοποιήθηκαν για τις ανάγκες της παρούσας εφαρμογής. Η προτεινόμενη μεθοδολογία και το αντίστοιχο σχήμα περιγραφής παρέχουν νέες λειτουργικές δυνατότητες στην ανάλυση της γαστρεντερικής βιοακουστικής δραστηριότητας και μπορούν να αξιοποιηθούν, τόσο στις προσεγγίσεις ανάλυσης σύντομου χρόνου, όσο και στις καταγραφές παρατεταμένου χρόνου. Η πλήρης συμμόρφωση με τις προδιαγραφές του προτύπου αναμένεται ότι θα εξασφαλίσει την απαραίτητη τυποποίηση και διαλειτουργικότητα για τη θεμελίωση μιας κλινικά αποδεκτής διαγνωστικής μεθόδου, για τη μελέτη των μηχανισμών γαστρεντερικής κινητικότητας. Παράλληλα, εξετάζεται το ενδεχόμενο εξέλιξης των σχημάτων περιγραφής, η μελλοντική ενσωμάτωσή τους σε συνδυαστικές μεθόδους ψυχοφυσιολογικής παρακολούθησης και, γενικά, η χρήση τους σε εφαρμογές διαχείρισης οπτικοακουστικού περιεχομένου με παρόμοια χαρακτηριστικά.

ΚΕΦΑΛΑΙΟ 8^ο

8. Συμπεράσματα - Προοπτικές

Η παρούσα διατριβή έχει ως αντικείμενο την ανάπτυξη νέων τεχνικών επεξεργασίας και διαχείρισης πολυκαναλικών, οπτικοακουστικών καταγραφών, με εφαρμογή στη μελέτη βιοακουστικών φαινομένων παρατεταμένου χρόνου. Κίνητρο για την ανάπτυξη των τεχνικών αυτών αποτέλεσαν οι ιατρικές προσεγγίσεις παρακολούθησης των μηχανισμών γαστρεντερικής κινητικότητας (μέσω των παραγόμενων κοιλιακών ήχων), με έμφαση στη μελέτη της φυσιολογίας και τη διάγνωση των λειτουργικών διαταραχών. Πρόκειται για ένα αρκετά δύσκολο εγχείρημα, καθώς η ιατρική γνώση στον τομέα αυτό είναι αρκετά περιορισμένη. Έτσι, αν και οι πρώτες σχετικές ερευνητικές προσπάθειες ξεκίνησαν πριν από έναν αιώνα περίπου, και παρά το γεγονός ότι οι λειτουργικές διαταραχές του γαστρεντερικού συστήματος απασχολούν σήμερα ένα μεγάλο μέρος του ανθρώπινου πληθυσμού, δεν υπάρχει μια κλινικά αποδεκτή μέθοδος για την αντικειμενική αξιολόγηση της γαστρεντερικής κινητικότητας. Η συγκεκριμένη ερευνητική προσέγγιση βιοακουστικής συγκεντρώνει πολλές ιδιαιτερότητες, που την καθιστούν μια απαιτητική «περίπτωση μελέτης» των πειραματικών διαδικασιών ψυχο-φυσιολογικής παρακολούθησης.

Στόχος, λοιπόν, της διατριβής ήταν η τεχνολογική υποστήριξη της ιατρικής έρευνας. Στο πλαίσιο αυτό αναπτύχθηκαν μια σειρά από εργαλεία, που σχετίζονται με τις ανάγκες καταγραφής, επεξεργασίας, διαχείρισης περιεχομένου και ανάλυσης των πολυκαναλικών βιοακουστικών καταγραφών, αλλά και των σημάτων οπτικοακουστικής παρακολούθησης της πειραματικής διαδικασίας. Η συνεισφορά της εργασίας εντοπίζεται στην εισαγωγή νέων μεθοδολογικών προσεγγίσεων και αντίστοιχων αλγορίθμων επεξεργασίας σήματος, που μπορούν να αξιοποιηθούν, τόσο για τις ανάγκες του συγκεκριμένου συστήματος βιοϊατρικής τεχνολογίας, όσο και σε παραπλήσιες εφαρμογές των τεχνολογιών ήχου και εικόνας, για την επεξεργασία και διαχείριση οπτικοακουστικού περιεχομένου.

8.1 Ερευνητική συνεισφορά & πρωτοτυπία της διατριβής

Από την ανάλυση που προηγήθηκε, γίνεται σαφές ότι η ερευνητική συνεισφορά και η πρωτοτυπία της διατριβής αφορά δύο επιμέρους τομείς. Η προώθηση της ιατρικής έρευνας γύρω από τη μελέτη της γαστρεντερικής κινητικότητας και η υλοποίηση αντίστοιχων συστημάτων βιοϊατρικής παρακολούθησης, είναι ο ένας από αυτούς. Ο σχεδιασμός και η ανάπτυξη νέων αλγορίθμων επεξεργασίας και διαχείρισης οπτικοακουστικών καταγραφών αποτελεί το δεύτερο πεδίο ερευνητικής συνεισφοράς. Έτσι, τα περισσότερα από τα εργαλεία που αναπτύχθηκαν μπορούν να χρησιμοποιηθούν σε σύνθετα πρωτόκολλα ψυχοφυσιολογικής παρακολούθησης, αλλά και σε εντελώς διαφορετικές εφαρμογές, όπως είναι οι τεχνικές ανάλυσης ακουστικού θορύβου, η επεξεργασία σημάτων ομιλίας και μουσικής, οι μέθοδοι ακουστικής χαρτογράφησης, τα συστήματα παρακολούθησης - ελέγχου ασφαλείας, οι «μηχανές» περιγραφής και διαχείρισης πολυμεσικού περιεχομένου.

8.1.1 Συνεισφορά στην έρευνα «βιοϊατρικής τεχνολογίας»

Η πιο σημαντική συνεισφορά στη επιστήμη της ιατρικής αφορά στην τυποποίηση μιας νέας μεθοδολογία μελέτης της γαστρεντερικής ακουστικής δραστηριότητας, και για μεγάλο χρόνο παρακολούθησης. Το προτεινόμενο βιοϊατρικό σύστημα περιλαμβάνει, καταρχάς, μια νέα διάταξη πολυκαναλικής καταγραφής των βιοακουστικών σημάτων, η οποία διαθέτει μεγάλη χωρική ευαισθησία λήψης και δυνατότητες τοπογραφικής ανάλυσης. Ανάμεσα σε όλα τα άλλα, ιδιαίτερη έμφαση δόθηκε στον περιορισμό της ενόχλησης και στην ασφάλεια των υποκειμένων κατά τη διεξαγωγή της πειραματικής διαδικασίας. Έτσι, εκτός των ηλεκτρονικών κυκλωμάτων προενίσχυσης και των μονάδων τροφοδοσίας, συστάθηκαν δύο νέες διατάξεις (*FSA*, *WAAV*) για την καλύτερη στερέωση και μηχανική προσαρμογή των βιοακουστικών αισθητήρων, στην περιοχή των κοιλιακών τοιχωμάτων. Η αξιοποίηση του συστήματος οπτικοακουστικής εποπτείας έδωσε τη δυνατότητα αξιολόγησης της πρωτότυπης αυτής κατασκευής *SHIELD* (συνδυασμός των δύο διατάξεων *FSA*, *WAAV*). Ως αποτέλεσμα, παρατηρήθηκε βελτίωση της απόδοσης του συστήματος λήψης, περιορισμός των ακουστικών θορύβων περιβάλλοντος, μικρότερη ενόχληση στα υποκείμενα και καλύτερα λειτουργικά χαρακτηριστικά, που σχετίζονται με την προετοιμασία της πειραματικής διαδικασίας και των υποκειμένων. Τα στοιχεία αυτά μπορούν να χρησιμοποιηθούν ως πιλότος για την τυποποίηση μιας αντίστοιχης διάταξης σε επίπεδο βιομηχανικού σχεδιασμού, που θα μπορεί να έχει εφαρμογή σε πολλές παραπλήσιες βιοακουστικές μεθόδους.

Παράλληλα, σημαντική κρίνεται και η συνεισφορά του συστήματος οπτικοακουστικής παρακολούθησης, που επιτρέπει τον έλεγχο της πειραματικής διαδικασίας και των υποκειμένων. Στο πλαίσιο αυτό δοκιμάστηκαν διάφορα πρότυπα συμπίεσης εικόνας για να επιλεγεί τελικά ο βέλτιστος κωδικοποιητής. Ακόμη, αναπτύχθηκαν τεχνικές ανίχνευσης κίνησης εικόνας, με στόχο την καλύτερη διαχείριση του περιεχομένου και τον πιο αποτελεσματικό έλεγχο των σημάτων οπτικοακουστικής εποπτείας. Η αξιοποίησή τους συνίσταται στην ικανοποίηση διαφόρων απαιτήσεων της τρέχουσας εφαρμογής, καθώς και στη μελλοντική εφαρμογή πιο σύνθετων πρωτοκόλλων ψυχοφυσιολογικής μελέτης.

Όσον αφορά στην επεξεργασία των κοιλιακών ηχητικών σημάτων, αναπτύχθηκαν δύο νέοι αλγόριθμοι αποθορυβοποίησης, συνδυάζοντας τα χαρακτηριστικά των φίλτρων Wiener με τις τεχνικές κυματιδιακής επεξεργασίας. Οι προτεινόμενες *WDWF* μέθοδοι έχουν πολύ καλή απόδοση σε όλες τις κατηγορίες γαστρεντερικών ήχων (και χωρίς μεγάλο υπολογιστικό κόστος), σε αντίθεση με τις προϋπάρχουσες τεχνικές, που προτείνονται για την επεξεργασία κρουστικών βιοακουστικών σημάτων. Τα αποτελέσματα αυτά επιβεβαιώθηκαν μέσα από διαδικασίες ποιοτικής και ποσοτικής αξιολόγησης, που εφαρμόστηκαν, τόσο σε φυσικές καταγραφές κοιλιακών ήχων, όσο και σε συνθετικά σήματα δοκιμής. Η χρησιμότητα των νέων αλγορίθμων εντοπίζεται στην υποβοήθηση της ακροαστική ανάλυσης από το ιατρικό προσωπικό, καθώς και σε μεθόδους αυτοματοποιημένης ανάλυσης, όπως είναι η ανίχνευση σήματος, τα συστήματα αναγνώρισης προτύπων και τα εργαλεία χαρτογράφησης.

Για τις ανάγκες επεξεργασίας και διαχείρισης των καταγραφών παρατεταμένου χρόνου, αναπτύχθηκε μια νέα μέθοδος πολύ-διακριτικής σάρωσης (*MRS*). Η υλοποίηση του σχετικού αλγόριθμου *LT-WDSS* έγινε σε επίπεδο κυματιδιακών συντελεστών, όπου αξιοποιήθηκαν αντιληπτικά κριτήρια και τεχνικές ανάλυσης της πολυπλοκότητας κυματομορφών (*fractal dimension*). Η νέα μέθοδος παρέχει τη δυνατότητα εντοπισμού των τμημάτων σημαντικής δραστηριότητας, ενώ επιτρέπει την τοπογραφική ερμηνεία των συμβάντων και την εξαγωγή ενεργειακών παραμέτρων σύνοψης, για την απεικόνιση του συνολικού φαινομένου. Με βάση τα αποτελέσματα ποιοτικής και ποσοτικής αξιολόγησης, η νέα τεχνική έχει πολύ καλή

απόδοση και μεγάλη ακρίβεια εντοπισμού, χωρίς ιδιαίτερες υπολογιστικές απαιτήσεις. Τα στοιχεία αυτά επιβεβαιώθηκαν και από τα μέλη της ιατρικής ερευνητικής ομάδας, που μετείχαν στις σχετικές πειραματικές διαδικασίες.

Μια από τις καινοτομίες της διατριβής είναι και η εισαγωγή του επιταχυνσιομέτρου τριών αξόνων στο σύστημα καταγραφής, που χρησιμοποιείται πρώτη φορά στην ανάλυση κοιλιακών ήχων. Το στοιχείο αυτό επέτρεψε την ανάπτυξη νέων, ιδιαίτερα αποτελεσματικών μεθόδων βιοακουστικού εντοπισμού και χαρτογράφησης, ξεπερνώντας πολλά από τα προβλήματα των προγενέστερων μεθόδων. Έτσι, υπήρξε δυνατότητα αξιολόγησης της νέας μεθόδου με την εκτέλεση προσομοιώσεων λογισμικού και τη σύσταση κατάλληλων πειραματικών διατάξεων. Ως αποτέλεσμα, παρέχονται νέα εργαλεία απεικόνισης και ακουστικής αναπαραγωγής του παραγόμενου βιοακουστικού πεδίου.

Η μεγαλύτερη, ίσως, συνεισφορά στην προώθηση της ιατρικής έρευνας, αφορά στις τεχνικές ταξινόμησης και αναγνώρισης προτύπων των κοιλιακών ήχων. Η σπουδαιότητα του εγχειρήματος αυτού έγκειται στο γεγονός ότι δεν έχουν, έως τώρα, προσδιοριστεί σχετικοί κανόνες περιγραφής των κλάσεων, σε αντίθεση με τις περιπτώσεις των καρδιακών και αναπνευστικών ήχων, όπου η σύσταση των λεξικών ταξινόμησης και των αντίστοιχων όρων περιγραφής είναι στενά συνδεδεμένη με την πρόοδο των μεθόδων διάγνωσης. Στην προκειμένη περίπτωση, επιχειρήθηκε η τυποποίηση της υπάρχουσας γνώσης και η εξερεύνηση νέων κανόνων ομαδοποίησης και περιγραφής των προτύπων, μέσα από μια συστηματική πειραματική διαδικασία μελέτης ενός μεγάλου αριθμού δειγμάτων, που διεξήχθη σε συνεργασία με την ιατρική ερευνητική ομάδα. Ως αποτέλεσμα συστάθηκαν δύο λεξικά ιεραρχικής ταξινόμησης των κοιλιακών ήχων, ενώ αναπτύχθηκαν και δύο αντίστοιχα συστήματα αυτόματης αναγνώρισης προτύπων. Για το σκοπό αυτό, χρησιμοποιήθηκαν τεχνικές στατιστικές και συντακτικής ταξινόμησης, καθώς και νευρωνικά δίκτυα. Η ανάλυση προτύπων παρέχει νέες, πιο περιγραφικές δομές μελέτης της γαστρεντερικής κινητικότητας, ενώ, παράλληλα, επιτρέπει την απομόνωση των κλάσεων που αντιστοιχούν σε σήματα θορύβου. Σύμφωνα με ιατρικές εκτιμήσεις, τα πρώτα αποτελέσματα ανάλυσης προτύπων είναι ιδιαίτερα ενθαρρυντικά, καθώς φαίνεται να υπάρχουν αρκετές ομοιότητες ανάμεσα στην αλληλοδιαδοχή τους και τις φάσεις του κύκλου MMC.

Ένα από τα μεγαλύτερα εμπόδια στην εξέλιξη των μεθόδων ανάλυσης των ακουστικών γαστρεντερικών καταγραφών παρατεταμένου χρόνου, αφορά στις δυσκολίες επεξεργασίας και διαχείρισης αυτών των καταγραφών. Για το λόγο αυτό, ιδιαίτερη βαρύτητα δόθηκε στο σχεδιασμό ευέλικτων δομών περιγραφής, διαχείρισης και ανάλυσης του οπτικοακουστικού υλικού. Έτσι, σχεδιάστηκε ένα νέο σχήμα περιγραφής περιεχομένου και μια αντίστοιχη οντολογία ανάλυσης, που είναι πλήρως συμβατά με το πρότυπο MPEG-7 και παρέχουν νέες δυνατότητες παρακολούθησης όλων των πειραματικών δεδομένων. Η νέα προσέγγιση αποτελεί σημαντική καινοτομία της διατριβής, ενώ τα χαρακτηριστικά τυποποίησης και διαλειτουργικότητας της τεχνολογία MPEG-7 αναμένεται να συμβάλουν σημαντικά στην προώθηση όλων των ερευνητικών μεθόδων μελέτης των γαστρεντερικών ήχων.

Στην παρούσα, λοιπόν, κατάσταση έχει αναπτυχθεί ένας πυρήνας εργαλείων και αλγορίθμων επεξεργασίας για τη μελέτη της γαστρεντερικής λειτουργίας, μέσω βιοακουστικών καταγραφών παρατεταμένης διάρκειας. Εκτιμάται, ότι η εφαρμογή του προτεινόμενου πρωτοκόλλου ανάλυσης θα βοηθήσει σημαντικά τη συστηματική μελέτη των κοιλιακών ήχων, τόσο για ηχητικές ακολουθίες μικρής διάρκειας, όσο και για καταγραφές παρατεταμένου χρόνου. Επίσης, θα διευκολύνει το σχεδιασμό αντίστοιχων πρωτοκόλλων

ανάλυσης σε άλλα πεδία βιοϊατρικής παρακολούθησης, καθώς και τη συνδυασμένη μελέτη των κοιλιακών ήχων με άλλες ψυχο-φυσιολογικές παραμέτρους.

8.1.2 Συνεισφορά στην έρευνα «μηχανικής»

Όσον αφορά στη γενικότερη ερευνητική συνεισφορά της διατριβής, πολλά από τα εργαλεία που αναπτύχθηκαν μπορούν να αξιοποιηθούν σε πλήθος άλλων εφαρμογών. Ως παράδειγμα, μπορούμε να αναφέρουμε τη χρήση των αλγορίθμων καταστολής θορύβου *WDWF* σε διάφορα άλλα πεδία, όπως είναι η επεξεργασία ομιλίας και γενικότερα η αποκατάσταση αλλοιωμένων ηχογραφήσεων. Οι προτεινόμενες τεχνικές έχουν εφαρμοστεί σε σήματα δοκιμής, αλλά και σε καταγραφές συνδιασκέψεων, παρέχοντας ιδιαίτερα ικανοποιητικά αποτελέσματα. Παράλληλα, έχουν δρομολογηθεί τροποποιήσεις με βάση τις ανάγκες των αντίστοιχων εφαρμογών, για την αποκατάσταση ενθόρυβων σημάτων ομιλίας και αρχαιακού υλικού μουσικών ηχογραφήσεων.

Εξίσου χρήσιμη θεωρείται και η μέθοδος ακουστικής χαρτογράφησης που αναπτύχθηκε, η οποία μπορεί να επεκταθεί, τόσο σε μετρητικές διατάξεις δονήσεων, όσο και σε κλασσικές ακουστικές εφαρμογές. Έτσι, κατά αντιστοιχία με τα επιταχυνσιόμετρα τριών αξόνων, προτείνεται ο συνδυασμός δύο ή περισσότερων μικροφώνων ηχητικού πεδίου (*SoundField*), με στόχο τον εντοπισμό των ηχητικών πηγών, την επιλεκτική λήψη και ενίσχυση των σημάτων με την επιβολή καθυστερήσεων χρονικής αντιστάθμισης. Για το σκοπό αυτό έχουν, ήδη, πραγματοποιηθεί κάποιες πρώτες διαδικασίες προσομοίωσης, όπου τα αποτελέσματα είναι ιδιαίτερα ικανοποιητικά.

Σε ένα σχετικά μεγάλο εύρος εφαρμογών μπορεί να αξιοποιηθεί και η νέα μέθοδος πολυδιακριτικής σάρωσης *MRS*, η οποία αναπτύχθηκε για την ανάλυση πολυκαναλικών ηχητικών καταγραφών παρατεταμένου χρόνου. Ο συγκεκριμένος αλγόριθμος επεξεργασίας *LT-WDSS* έχει δοκιμαστεί επιτυχώς και σε καταγραφές ανάλυσης ακουστικού θορύβου. Παράλληλα, με ενδιαφέρον αναμένεται η αξιοποίησή του και σε παραπλήσιες εφαρμογές, όπως είναι η αυτοματοποίηση των διαδικασιών παρακολούθησης σε συστήματα ασφαλείας, σε εφαρμογές διαχείρισης πολυκαναλικών μουσικών ηχογραφήσεων, σε περιβάλλοντα οπτικοακουστικής συναρμογής, κ.ά. Ιδιαίτερα χρήσιμα είναι και τα συμπεράσματα σχετικά με την ανάπτυξη των τεχνικών ανίχνευσης κίνησης βίντεο. Έτσι, εξετάζεται το ενδεχόμενο εξέλιξης των μεθόδων που αναπτύχθηκαν, καθώς και η αξιοποίησή τους σε πολλές από τις εφαρμογές που προαναφέρθηκαν.

Στο ίδιο μήκος κύματος είναι και οι προτάσεις για τη μελλοντική αξιοποίηση της προτεινόμενης οντολογίας ανάλυσης, η οποία σχεδιάστηκε με βάση το πρότυπο *MPEG-7*. Επιπλέον, έχουν δρομολογηθεί οι εργασίες ανάπτυξης ενός περιβάλλοντος πλοήγησης, που να είναι πλήρως συμβατό με το πρότυπο (*MPEG-7 browser*), και το οποίο θα μπορεί να αξιοποιηθεί, τόσο για τις ανάγκες της παρούσας βιοϊατρικής μεθόδου, όσο και για διάφορες άλλες οπτικοακουστικές εφαρμογές. Το εγχείρημα αυτό είναι ιδιαίτερα σημαντικό καθώς δεν έχουν, ακόμη, υλοποιηθεί σε μεγάλο βαθμό τέτοιου είδους εργαλεία.

Τέλος, για κάθε μια από τις μεθόδους και τους αλγορίθμους που αναπτύχθηκαν, υπάρχει δυνατότητα μελλοντικής εξέλιξης, τόσο για την καλύτερη εξυπηρέτηση της ιατρικής έρευνας, όσο και για την αποτελεσματικότερη εφαρμογή τους σε παραπλήσια επιστημονικά πεδία. Οι προτάσεις αυτές αναλύθηκαν ξεχωριστά σε κάθε κεφάλαιο και αφορούν στις μεθόδους καταστολής θορύβου, τις τεχνικές ανίχνευσης σήματος, τα συστήματα αναγνώρισης προτύπων, τα εργαλεία ακουστικής χαρτογράφησης και απεικόνισης ηχητικού πεδίου, το *MPEG-7* σχήμα περιγραφής και διαχείρισης του οπτικοακουστικού περιεχομένου.

9. Βιβλιογραφικές αναφορές

- [1]. Pauri F., Pierelli F., Chatrian G-E. and Erdly W.W., "Long-term EEG-video-audio monitoring: computer detection of focal EEG seizure patterns", *Electroencephalography and clinical Neurophysiology*, vol. 82, pp. 1-9, 1992.
- [2]. Yen T.Y., Radwin R.G., "A video-based system for acquiring biomechanical data synchronized with arbitrary events and activities", *IEEE Trans. Biomed. Eng.*, vol. 42, no. 9, pp. 944-948, June 1995.
- [3]. Josefsson T., Nordh E., and Eriksson P-O., "A flexible high-precision video system for digital recording of motor acts through lightweight reflex markers", *Computer Methods and Programs in Biomedicine*, Elsevier, vol. 49, no 2, pp. 119-184, March 1996.
- [4]. Guerri J.C., Esteve M., Palau C., Monfort M., and Sarti M.A., "A software tool to acquire, synchronise and playback multimedia data: an application in kinesiology", *Computer Methods and Programs in Biomedicine*, vol. 62, no. 1, pp. 51-58, 2000.
- [5]. Zeng S., Powers J.R., and Hsiao H., "A New Video-Synchronized Multichannel Biomedical Data Acquisition System", *IEEE Trans Biomed Eng.*, vol. 47, no. 3, pp. 412 - 419, March 2000.
- [6]. Zhanjian Li, da Silva A.M., Cunha, J.P.S., "Movement quantification in epileptic seizures: a new approach to video-EEG analysis", *IEEE Trans. Biomed. Eng.*, vol. 49, no. 6, pp. 565-573, June 2002.
- [7]. Pavlidis I., "Continuous physiological monitoring", in *Proceedings of the 25th IEEE EMBS Conference*, Cancun, Mexico, pp. 1084-1087, Sept. 2003.
- [8]. Keogh E., Celly B., Ratanamahatana C.A., and Zordan V., "A Novel Technique for Indexing Video Surveillance Data", In *proceedings of ACM SIGMM 2003 Workshop on Video Surveillance IWVS'03*, in conjunction with eleventh ACM International Conference on Multimedia, Berkeley, CA, November 2003.
- [9]. Murthy R., Pavlidis I. and Tsiamyrtzis P., "Touchless monitoring of breathing function", in *Proceedings of the 26th IEEE EMBS Conference*, San Fransisco, CA, pp. 1196-1199, Sept. 2004.
- [10]. Murthy R. and Pavlidis I., "Noncontact measurement of breathing function - A novel method using infrared imaging and advanced statistics", *Engineering in Medicine and Biology Magazine IEEE*, vol. 25, no. 3, pp. 57-67, May-June 2006.
- [11]. Karayiannis N.B., Sami A., Frost J.D., Wise M.S., Mizrahi E.M., "Automated extraction of temporal motor activity signals from video recordings of neonatal seizures based on adaptive block matching", *IEEE Trans. Biomed. Eng.*, vol. 52, no. 4, pp. 676-686, Apr. 2005.
- [12]. Coyle M.A., Keenan D.B., Henderson L. S., Watkins M.L., Haumann B.K., Mayleben D.W., and Wilson M.G., "Evaluation of an ambulatory system for the quantification of cough frequency in patients with chronic obstructive pulmonary disease", *Cough*, vol. 1, no. 3, pp. 1-7, 2005.
- [13]. Davey M.J., "Investigation of sleep disorders", *J. Paediatr. Child Health*, vol. 41, pp. 16-20, 2005.
- [14]. Hensley MJ, Hillman DR, McEvoy RD, Neill AM, Solin P, Teichtahl H, Thompson BR, Tolhurst S, Thornton AT, Worsnop CJ, "Guidelines for sleep studies in adults", Prepared for the Australasian Sleep Association & Thoracic Society of Australia and New Zealand, pp. 1-38, Oct. 2005.
- [15]. Drossman D.A., "The functional gastrointestinal disorders and the Rome II process", *Gut*; vol. 45 (Suppl II), pp. II1-II5, 1999.
- [16]. Talley N.J. et al., "B. Functional Gastroduodenal Disorders", in: D. A. Drossman (Ed.), *ROME II: The functional Gastrointestinal Disorders*, 2nd ed. Degnon Associates, chapter 6, pp. 299 - 350, 2000.
- [17]. Thompson W.G., "C. Functional Bowel Disorders and D. Functional Abdominal Pain", in: D. A. Drossman (Ed.), *ROME II: The functional Gastrointestinal Disorders*, 2nd ed. Degnon Associates, chapter 7, pp. 351-432, 2000.

- [18]. Holtmann G. and Enck P., "Stress and gastrointestinal motility in humans: a review of the literature", *J Gastrointest Motil*, vol. 3, no. 4, pp. 245-254, 1991.
- [19]. Farthing M.J.G. "Fortnightly Review: Irritable bowel, irritable body, or irritable brain?", *BMJ*, vol. 310, pp. 171-175, Jan. 1995.
- [20]. Weisbrodt N. W., "Motility of the Small Intestine", in: L. R. Johnson (Ed.), *Physiology of the Gastrointestinal Tract*, 2nd ed, New York: Raven, vol. 1, pp. 631-663, 1987.
- [21]. Sharna S., "Intestinal Manometry to Evaluate Intestinal Motor Function", in *Gastrointestinal Motility: Which test?*, Read N. W. (Ed.), Wrighton Biomedical Publishing Ltd., vol. 18, pp. 157-166, 1989.
- [22]. Kellow J. E., Gill R. C. and Wingate D. L., "Prolonged Ambulant Recordings of Small Motility Demonstrate Abnormalities in the Irritable Bowel Syndrome", *Gastroenterology*, vol. 98, no. 5, pp. 1208-1218, May 1990.
- [23]. Cannon W.B., "Auscultation of the rhythmic sounds produced by the stomach and intestines", *Am. J. Physiology*, vol. 14, pp. 339- 353, 1905.
- [24]. Δημούλας Χ., "Καταγραφή και ψηφιακή επεξεργασία γαστρεντερικών ήχων", Διπλωματική διατριβή, Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, Οκτώβριος 1997.
- [25]. Pasiadis, C., Papanikolaou, G., Dimoulas, Ch. & Kalampakas A., "Intestinal Motility Recording and Analysis", *Proceedings of the VIII Mediterranean Conference on Medical and Biological Engineering and Computing MEDICON '98*, Lemesos, Cyprus, 1998.
- [26]. Bardakjian B. L., "Gastrointestinal System", in: J. D. Bronzino (Ed.), *The Biomedical Engineering Handbook*, IEEE Press and CRC Press Inc, vol. 6, pp. 57-69, 1995.
- [27]. Leicester Medical School, available online: <http://www.gastroatlas.com>.
- [28]. Current Medicine Group, online: <http://www.le.ac.uk/pathology/teach/va/anatomy/frmst.html>.
- [29]. Schulze K., "Imaging and modelling of digestion in the stomach and the duodenum", *Neurogastroenterol Motil*, vol. 18, pp. 172-183, 2006.
- [30]. Dooley CP, DiLorenzo C, Valenzuela JE., "Variability of migrating motor complex in humans", *Dig Dis Sci*, vol. 37, pp. 723-728, 1992.
- [31]. Chen J., Schirmer B.D., and McCallum R.W., "Measurement of electric activity of the human small intestine using surface electrodes", *IEEE Trans. Biomed. Eng.*, vol. 40, no. 6, pp. 598-602, Jun. 1993.
- [32]. Shibata Y., Toyoda S., Nimura Y., Miyati M., "Patterns of intestinal motility recovery during the early stage following abdominal surgery: clinical and manometric study", *World J. Surg.*, vol. 21, no. 8, pp. 806-809, Oct. 1997.
- [33]. Luiking Y.C., VanDen Reijden A.C., Van Berge Henegouwen G.P., Akkermans L.M.A., "Migrating motor complex cycle duration is determined by gastric or duodenal origin of phase III", *Am J Physiol*, 275, pp. G1246- G1251, 1998.
- [34]. Miedema B. W., Schillie S., Simons J. W., Burgess S. V., Liem T., and Silver D., "Small bowel motility and transit after aortic surgery", *J. Vasc. Surg.*, vol. 36, pp. 19-24, 2002.
- [35]. Vantrappen G., Janssens J., Coremans G., and Jian R., "Gastrointestinal Motility Disorders, Digestive Diseases and Sciences", vol. 31, no. 9, pp. 5S-25S, Sept. 1986.
- [36]. Cannon W.B., "The movements of the intestines studied by means of the roentgen rays", *Am. J. Physiology*, vol. 6, pp. 251- 277, 1902.
- [37]. Maglinte D.D.T., Kelvin F.M., O'Connor K., Lappas J.C., Chernish S.M., "Current status of small bowel radiography", *Abdom Imaging*, vol. 21, pp. 247-257, 1996.
- [38]. Merc & CO., INC, available online: <http://www.merck.com>.

- [39]. Kerlin P., Phillips S., "Variability of motility of the ileum and jejunum in healthy humans", *Gastroenterology* 82, pp. 694-700, 1982.
- [40]. Kellow J.E., Borody T.J., Phillips S.F., Tucker R.L., Haddad A.C., "Human interdigestive motility: variation in patterns from esophagus to colon", *Gastroenterology* 91, pp. 386-395, 1986.
- [41]. Garcia-Casado J., Martinez-de-Juan J.L., and Ponce J.L., "Non-invasive measurement and analysis of intestinal myoelectrical activity using surface electrodes, *IEEE Trans. Biomed. Eng.*, vol. 52, no. 6, pp. 983-991, 2005.
- [42]. Craine BL, Silpa ML, O'Toole CJ. "Two-dimensional positional mapping of gastrointestinal sounds in control and functional bowel syndrome patients", *Dig Dis Sci*, vol. 47, pp. 1290-1296, 2002.
- [43]. Thompson D.G., Wingate D.L., Archer L., Benson M.J., Green W.J., and Hardy R.J., "Normal patterns of human upper small bowel motor activity recorded by prolonged radiotelemetry", *Gut*, vol. 21, pp. 500-506, 1980.
- [44]. Wen-Xing Wang, Guo-Zheng Yan, Fang Sun, Ping-Ping Jiang, Wen-Qiang Zhang, Gen-Fu Zhang, "A non-invasive method for gastrointestinal parameter monitoring", *World J Gastroenterol*, Vol. 11, no 4, pp. 521-524, 2005.
- [45]. Iddan G, Meron G, Glukhovskiy A, Swain P., "Wireless capsule endoscopy", *Nature*, vol. 25, pp. 405-417, May 2000.
- [46]. Lewis B.S., Swain P., "Capsule endoscopy in the evaluation of patients with suspected small intestinal bleeding: results of a pilot study", *Gastrointestinal Endoscopy*, vol. 56, no. 3, pp. 349-353, 2002.
- [47]. Napierkowski JJ, Maydonovitch CL, Belle LS, Brand WT Jr, Holtzmuller KC, "Wireless capsule endoscopy in a community gastroenterology practice", *J. Clin Gastroenterol.*, Vol. 39, no 1, pp.36-41, Jan 2005.
- [48]. Huisman TA, Jacot P, Benoit CH, et al., "Crohn's disease: evaluation with computed tomography and magnetic resonance imaging", *Schweiz Rundsch Med Prax*, vol. 88, no. 43, pp. 1759-1764, 1999.
- [49]. Gregersen H., Gilja O.H., Hausken T., et al., "Mechanical properties in the human gastric antrum using B-mode ultrasonography and antral distension", *Am J Physiol Gastrointest Liver Physiol*, vol. 283, pp. G368- G375, 2002.
- [50]. Pallotta N., "Ultrasonography in the assessment of gallbladder motor activity", *Dig Liver Dis*, vol. 35, pp. S67- S69, 2003.
- [51]. Pallotta N., Calabrese E., Alghisi F., Corazziari E., "Ultrasonography in the assessment of gastric and gallbladder motor function", *Diabetes Nutr Metab*, vol. 17, pp. 37-42, 2004.
- [52]. Swain P. and Fritscher-Ravens A., "Role of video endoscopy in managing small bowel disease", *Gut*, vol. 53, pp. 1866-1875, 2004.
- [53]. Dowling K., "Limbless Locomotion: Learning to Crawl with a Snake Robot", Doctoral dissertation, tech. report CMU-RI-TR-97-48, Robotics Institute, Carnegie Mellon University, December, 1997.
- [54]. Hoeg H.D., Slatkin A.B., Burdick J.W., and Grundfest W.S., "Biomechanical modeling of the small intestine as required for the design and operation of a robotic endoscope", *IEEE Int. Conf. on Robotics and Automation*, San Francisco, CA, 2000.
- [55]. Funaki B., Zaleski G.X., Lorenz J., Menocci P.B., Funaki A.N., Rosenblum J.D., Straus C., Leef J.A., "Radiologic Gastrostomy Placement: Pigtail- Versus Mushroom - Retained Catheters", *Am. J. Roentology*, vol. 175, pp. 375-379, 2000.
- [56]. Phee L., Accoto D., Mencias A., Stefanini C., Carrozza M.C., and Dario P., "Analysis and development of locomotion devices for the gastrointestinal track", *IEEE Trans. Biomed. Eng.*, vol. 49, no. 6, pp. 613-616, Jun. 2002.
- [57]. Rogalla P, Werner Rustner M, Huitema A, et al. "Virtual endoscopy of the small bowel: phantom study and preliminary clinical results", *Eur Radiol*, vol. 8, no. 4, pp. 563-567, 1998.

- [58]. Schreyer A.G., Herfarth H., Kikinis R., Seitz J., Schölmerich J., Geissler A., Feuerbach S., "3-D modeling and virtual endoscopy of the small bowel based on MR Imaging in patients with inflammatory bowel disease", *Investigative Radiology*, vol. 39, no 9, pp. 528-533, 2002.
- [59]. Hadjileontiadis L.J., Liatsos C.N., Mavrogiannis C.C., Rokkas T.A., and Panas S.M., "Enhancement of bowel sounds by wavelet-based filtering", *IEEE Trans. Biomed. Eng.*, vol. 47, no. 7, pp. 876-886, Jul. 2000.
- [60]. Yuki M., Adachi K., Fujishiro H., Uchida Y., Miyaoka Y., Yoshino N., Yuki T., Ono M., Kinoshita Y., "Is a Computerized Bowel Sound Auscultation System Useful for the Detection of Increased Bowel Motility?", *Am J Gastroenterol*, vol. 97, no. 7, pp. 1846-1847, 2002.
- [61]. Kalampakas A., Dimoulas C.A., Kalliris G.M., Papanikolaou G.V., "A systematic approach of abdominal sounds' classification. A novel, computer assisted, analysis' effort of prolonged recordings", *Gastroenterology*, (to be submitted), 2006.
- [62]. Lifespan, available online: <http://www.lifespan.org/>.
- [63]. MedlinePlus, available online: <http://www.nlm.nih.gov/medlineplus/>.
- [64]. Dalle D., Devroede G., Thibault R., and Perrault J., "Computer analysis of Bowel Sounds", *Comput. Biol. Med.*, vol. 4, pp. 247-256, 1975.
- [65]. Vasseur C, Devroede G, Dalle D, Van Houtte N, Bastin E, Thibault R., "Postprandial bowel sounds", *IEEE Trans Biomed Eng.*, vol. 22, no. 5, pp.443-448, Sept. 1975.
- [66]. Politzer JP, Devroede G, Vasseur C, Gerard J, Thibault R., "The genesis of bowel sounds: influence of viscus and gastrointestinal content", *Gastroenterology*, vol. 71, no. 2, pp. 282-285, Aug. 1976.
- [67]. Arnbjornsson E., "Normal and pathological bowel sound patterns", *Ann Chir Gynaecol.*, vol. 75, no. 6, pp. 314-318, 1986.
- [68]. Campbell F.C., Storey B.E., Cullen P.T., "Surface vibration analysis (SVA): A new non-invasive monitor of gastrointestinal activity", *Gut*, vol. 30, pp. 39-45, 1989.
- [69]. Gamer C. and Ehrenreich H., "Non invasive topographic analysis of intestinal activity in man on the basis of acoustic phenomena", *Res. Exp. Med. (Bed.)*, vol. 189, no. 2, pp. 129-140, 1989.
- [70]. Yoshino H., Abe Y., Yoshino T., Oshato K., "Clinical application of spectral analysis of bowel sounds in intestinal obstruction", *Dis. Col. & Rect.*, vol. 33, no 9, pp. 753-757, Sept. 1990.
- [71]. Liatsos C, Hadjileontiadis LJ, Mavrogiannis C, Patch D, Panas SM, Burroughs AK. "Bowel sounds analysis: A novel non-invasive method for diagnosis of small-volume ascites", *Dig Dis Sci*, vol. 48, pp. 1630-1636, 2003.
- [72]. Tomomasa T, Morikawa A, Sandler RH, Mansy HA, Koneko H, Masahiko T, Hyman PE, Itoh Z., "Gastrointestinal sounds and migrating motor complex in fasted humans", *Am J Gastroenterol*, vol. 94, no. 2, pp. 374-381, 1999.
- [73]. Tomomasa T., Takahashi A., Yasushi Nako, Hiroaki Kaneko, Masahiko Tabata, Yoshiaki Tsuchida, and Akihiro Morikawa, "Analysis of gastrointestinal sounds in infants with pyloric stenosis before and after pyloromyotomy", *PEDIATRICS* vol. 104, no. 5, p. e60, Nov. 1999.
- [74]. Craine BL, Silpa ML, O'Toole CJ, "Computerized Auscultation Applied to Irritable Bowel Syndrome", *Dig Dis Sci*, vol. 44, no. 9, pp. 1887- 1892, 1999.
- [75]. Craine BL, Silpa ML, O'Toole CJ, "Enterotachogram analysis to distinguish irritable bowel syndrome from Crohn's disease," *Dig Dis Sci*, vol. 46, no. 9, pp. 1974-1979, 2001.
- [76]. Radnitz C.L., and Blanchard E.B., "Bowel sound biofeedback as a treatment Irritable Bowel Syndrome", *Biofeedback and Self-Regulation*, vol. 13, no. 2, pp. 169-179, 1988.
- [77]. Radnitz C.L., and Blanchard E.B., "A 1- and 2-Year Follow-Up Study of Bowel Sound Biofeedback as a Treatment for Irritable Bowel Syndrome", *Biofeedback and Self-Regulation*, vol. 14, no. 4, pp. 333-338, 1989.

- [78]. Mansy HA, Sandler RH, "Bowel-sound signal enhancement using adaptive filtering", IEEE Eng. Med. Biol. Mag., vol. 16, no. 6, pp. 105-117, Nov.-Dec. 1997.
- [79]. Bray D., Reilly R.B., Haskin L., McCormack B., "Assessing motility through abdominal sound monitoring", Proc. 19th International Conference - IEEE/EMBS, Chicago, IL. USA, pp. 2398-2400, 1997.
- [80]. Hadjileontiadis L.J. and Panas, S. M., "On modeling impulsive bioacoustic signals using a-stable distributions: application in discontinuous adventitious sounds and explosive bowel sounds", Proc. of 20th IEEE EMBS 98, Hong Kong, vol. 20, No. 1, pp. 13-16, 1998.
- [81]. Hadjileontiadis L.J. and Rekanos I.T., "Enhancement of explosive bowel sounds using kurtosis-based filtering", Proc. of 25th IEEE EMBS 2003, Cancun, Mexico, pp. 2479-2482, Sept. 2003.
- [82]. Hadjileontiadis L.J. and Rekanos I. T., "Detection of explosive lung and bowel sounds by means of fractal dimension", IEEE Signal Processing Lett., vol. 10, no. 10, pp. 311-314, Oct. 2003.
- [83]. Hadjileontiadis L.J., "Wavelet-Based Enhancement of Lung and Bowel Sounds Using Fractal Dimension Thresholding – Part I: Methodology", IEEE Trans. Biomed. Eng., vol. 52, no. 6, pp. 1143-1148, Jun. 2005.
- [84]. Hadjileontiadis L.J., "Wavelet-Based Enhancement of Lung and Bowel Sounds Using Fractal Dimension Thresholding – Part II: Application Results", IEEE Trans. Biomed. Eng., vol. 52, no. 6, pp. 1050-1064, Jun. 2005.
- [85]. Rekanos I.T. and Hadjileontiadis L.J., An iterative kurtosis-based technique for the detection of nonstationary bioacoustic signals, Signal Processing, Elsevier (article in press), 2006.
- [86]. Ranta R., Louis-Dorr V., Heinrich Ch., Wolf D., and Guillemin F., "Wavelet-based bowel sounds denoising, segmentation and characterization", in Proc. of the 23th IEEE-EMBS Conference, Istanbul, Turkey, pp. 1903-1906, 2001.
- [87]. Ranta R., Louis-Dorr V., Heinrich Ch., Wolf D., and Guillemin F., "Towards an acoustic map of abdominal activity", Proc. of the 25th IEEE-EMBS Conference, EMBS'03, Cancun, Mexico, pp. 2769-2772, 2003.
- [88]. Ranta R., Louis-Dorr V., Heinrich Ch., Wolf D., and Guillemin F., "Principal component analysis and interpretation of bowel sounds", Proc. of the 26th IEEE-EMBS Conference, EMBS'04, San Francisco, CA, USA, pp. 227-230, 2004.
- [89]. Liatsos C.N., Hadjileontiadis L.J., Mavrogiannis C.C., Rokkas T.A. and Panas S.M., "Enhanced De-Noising of Bowel Sounds Using a Wavelet-Based Filter", Proceedings of the VIII Mediterranean Conference on Medical and Biological Engineering and Computing MEDICON '98, Lemesos, Cyprus, 1998.
- [90]. Kashyap V., "Emergent Semantics: An organizing principle for Biomedical Informatics and Knowledge Management", AIS SIGSEMIS Bulletin, vol. 1, no. 2, pp. 46-49, July 2004.
- [91]. Golbreich C., "Challenges for Knowledge Management in the Biomedical Domain", AIS SIGSEMIS Bulletin, vol. 1, no. 2, pp. 43-45, July 2004.
- [92]. Cuggia M., Mougin F., Le Beux P., "Indexing method of digital audiovisual medical resources with semantic Web integration", International Journal of Medical Informatics, Elsevier, vol. 74, pp. 169-177, 2005.
- [93]. Bodenreider O. and Burgun A., "Biomedical ontologies", in: Chen H., Fuller S., Hersh W.R., Friedman C. (Ed.), Medical informatics: Advances in knowledge management and data mining in biomedicine: Springer-Verlag; chapter 2.5, pp. 211-236, 2005.
- [94]. Χρηστίδης Δ.Α., Δημούλας Χ., Καλλίρης Γ., Σεβαστιάδης Χ., Κουντουράς Ν. και Παπανικολάου Γ., "Σχεδιασμός και κατασκευή ψυχοφυσιολογικής-βιοιατρικής εφαρμογής για την αντικειμενική μέτρηση της επίδρασης του θορύβου πόλης στον ανθρώπινο οργανισμό", 1ο Πανελλήνιο Συνέδριο της Ελληνικής Εταιρείας Ψυχοφυσιολογίας, Θεσσαλονίκη, 7-10 Ιουν. 2001.
- [95]. Dimoulas C., Kalliris G., Sevastiadis C., Papanikolaou G., Christidis D., "Development of an engineering application for subjective evaluation of human response to noise", AES 110th Convention, Preprint Number: 5408, April 2001.

- [96]. Χατζηλεοντιάδης Λ., "Ανάλυση και επεξεργασία των αναπνευστικών ήχων με τη χρήση της στατιστικής - φάσματος ανώτερης τάξης και του μετασχηματισμού κυματιδίων", Διδακτορική διατριβή, Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, Απρίλιος 1997.
- [97]. Χατζηλεοντιάδης Λ., "Μέθοδοι επεξεργασίας βιοϊατρικών σημάτων", Σημειώσεις, Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, Μάιος 2000.
- [98]. Shung K.K. and Cloutier G., "Medical diagnosis with acoustics", in: Crocker M. (Ed.), Encyclopedia of acoustics, ISBN 0-471-80465-7, John Wiley & Sons Inc, vol. 3, ch. 143, pp. 1739-1750, 1997.
- [99]. Goovaerts H.G., Rempelman O., van Geijn H.P., "A transducer for detection of fetal breathing movements", IEEE Transactions on Biomedical Engineering, vol. 36, no. 4, pp. 471 - 478, April 1989.
- [100]. Kovács F., "A Rule-Based Phonocardiographic Method for Long-Term Fetal Heart Rate Monitoring", IEEE Trans. Biomed. Eng., vol. 47, no. 1, pp. 124-130, Jan 2000.
- [101]. Sebald D.J., Bahr D.E., and Kahn A.R., "Narrowband Auscultatory Blood Pressure Measurement", IEEE Trans. Biomed. Eng., vol. 49, no. 9, pp. 1038-1044, Sept. 2002.
- [102]. Thompson B.M. and Andrews S.R., "The emerging field of sound training", IEEE Engineering in Medicine and Biology Magazine, vol. 18, no. 2, pp. 89 - 96, March-April 1999.
- [103]. Παστιάδης Κ., "Σύγχρονες τεχνικές ανάλυσης φωνής ατόμων με προβλήματα ομιλίας", Διδακτορική διατριβή, Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, Οκτώβριος 2002.
- [104]. Umaphathy K., Krishnan S., Parsa V., and Jamieson D.G., "Discrimination of pathological voices using a time-frequency approach", IEEE Trans. Biomed. Eng., vol. 52, no. 3, pp. 421-430, Mar. 2005.
- [105]. Wikipedia, online encyclopedia: <http://www.wikipedia.org>.
- [106]. Panoulas K., Kitsas I., Georgiadis P., Zoumpoulidis T., Fotiou I., and Hadjileontiadis L.J., "BlueSteth: a Bluetooth-based stethoscope for healthcare and educational applications", IEEE Computer Society, CSIDC 2002 Final Report, pp. 1-30, 2002.
- [107]. Ballou G., "Microphones", in: G. Ballou (Ed.), Handbook for Sound Engineers: The new audio cyclopedia, Howard W. Sans & Co., ISBN 0-672-2-21983-2, part 3, ch. 13, pp. 315-404, 1987.
- [108]. Σκαρλάτος Δ., «Εφαρμοσμένη Ακουστική», Φιλομάθεια (Εκδ.), ISBN 960-87710-1-3, Πάτρα, Οκτώβριος 2003.
- [109]. Παπανικολάου Γ., «Ηλεκτρακουστική», 2η έκδοση, University Studio Press, ISBN 960-12-1394-5, Θεσσαλονίκη 2005.
- [110]. Brüel & Kjær, "Microphone Handbook, vol. 1: theory", www.bk.dk (available online), Technical Documentation, BE 1447 -11, July 1996.
- [111]. Brüel & Kjær, "Measuring Microphones", www.bk.dk (available online), Lecture Note BA 7216-15, 1998.
- [112]. Clifford M., "Microphones", Tab Books Inc., ISBN: 0830604758, Dec. 1986.
- [113]. Vannuccini L., Earis J.E., Helistö P., Cheetham B.M.G., Rossi M., Sovijärvi A.R.A., Vanderschoot J., "Capturing and preprocessing of respiratory sounds", Eur Respir Rev, vol. 10, no. 77, pp. 616-620, 2000.
- [114]. Möser M., "Vibration measurements and instrumentation", in: Crocker M. (Ed.), Encyclopedia of acoustics, ISBN 0-471-80465-7, John Wiley & Sons Inc, vol. 2, ch. 72, pp. 857-868, 1997.
- [115]. NTD resource center, "Piezoelectric Transducers", available online: www.ndt-ed.org.
- [116]. Aszkler C., "The Principles of Acceleration, Shock, and Vibration Sensors", Sensors Magazine Online, www.sensorsmag.com, May 2005.
- [117]. Wilson J., "A Practical Approach to Vibration Detection and Measurement Part 1: Physical Principles and Detection Techniques", Sensors Magazine Online - Editorial, www.sensorsmag.com, Feb. 1999.

- [118]. Tustin W., "An Introduction to Classical Sinusoidal Vibration", Sensors Magazine Online, www.sensorsmag.com, May 2004.
- [119]. Wilson J., "A Practical Approach to Vibration Detection and Measurement Part 2: Dynamic and Environmental Effects of Performance", Sensors Magazine Online, www.sensorsmag.com, March 1999.
- [120]. Kulwanoski G. and Schnellinger J., "The Principles of Piezoelectric Accelerometers", Sensors Magazine Online - Editorial, www.sensorsmag.com, Feb. 2004.
- [121]. Brüel & Kjær, "Vibration transducers and signal conditioning", www.bk.dk (available online), Lecture Note BA 7675-12, 1998.
- [122]. Hansen K.S., "A New Design Principle for Triaxial Piezoelectric Accelerometers", www.bk.dk (available online), in: H.K. Zaveri (Ed.), Brüel & Kjær Technical Review, no. 1, 2002.
- [123]. Ballou G., "Filters and Equalizers", in: G. Ballou (Ed.), Handbook for Sound Engineers: The new audio cyclopedia, Howard W. Sans & Co., ISBN 0-672-2-21983-2, part 4, ch. 17, pp. 569-612, 1987.
- [124]. Patronis G.Jr. and Barkhard M., "Amplifiers", in: G. Ballou (Ed.), Handbook for Sound Engineers: The new audio cyclopedia, Howard W. Sans & Co., ISBN 0-672-2-21983-2, part 4, ch. 15, pp. 489-544, 1987.
- [125]. Παπανικολάου Γ., Σεργιάδης Γ., Καλλιρης Γ., Παστιάδης Κ., Δημούλας Χ., "Ψηφιακή Επεξεργασία Ήχου", Εκπαιδευτικό υλικό στο πλαίσιο του προγράμματος: "Εκπαίδευση από απόσταση στο ΑΠΘ", multimedia CD-ROM, Θεσσαλονίκη, 1999.
- [126]. Manquen D., "Magnetic recording and playback", in: G. Ballou (Ed.), Handbook for Sound Engineers: The new audio cyclopedia, Howard W. Sans & Co., ISBN 0-672-2-21983-2, part 5, ch. 24, pp. 933-976, 1987.
- [127]. Manquen D., "Digital recording and playback", in: G. Ballou (Ed.), Handbook for Sound Engineers: The new audio cyclopedia, Howard W. Sans & Co., ISBN 0-672-2-21983-2, part 5, ch. 25, pp. 977-994, 1987.
- [128]. Mainardi L.T., Bianchi A.M., and Cerutti S., "Digital Biomedical Signal Acquisition and Processing", in: J. D. Bronzino (Ed.), The Biomedical Engineering Handbook, Second Edition, CRC Press LLC, ISBN 0-8493-0461-X, vol. 1, ch. 53, 2000.
- [129]. Kurumbalapatiya D. and Hoole, S.R.H., "Data Acquisition", in: Dorf R.C. (Ed.), The Electrical Engineering Handbook, second edition, ISBN: 0-8493-8574-1, CRC Press LLC, ch. 84, 2000.
- [130]. Mitchell J. S., "Vibration analysis - its evolution and use in machinery health monitoring", Endevco tech. paper TP 265, www.endevco.com (available online), 1975.
- [131]. Krishnappa G., "Machinery condition monitoring", in: Crocker M. (Ed.), Encyclopedia of acoustics, ISBN 0-471-80465-7, John Wiley & Sons Inc, vol. 2, ch. 73, pp. 869-879, 1997.
- [132]. Hadjileontiadis L.J. and Panas S.M., "Adaptive Reduction of Heart Sounds from Lung Sounds Using Fourth-Order Statistics", IEEE Transactions on Biomedical Engineering, vol. 44, no. 7, pp. 642-648, July 1997.
- [133]. Hadjileontiadis L.J. and Panas S.M., "Separation of Discontinuous Adventitious Sounds from Vesicular Sounds, Using a Wavelet-based Filter", IEEE Transactions on Biomedical Engineering, vol. 44, no. 12, pp. 1269-1281, Dec. 1997.
- [134]. Kaniusas E., Pftzner H. and Saletu B., "Acoustical signal properties for cardiac/respiratory activity and apneas", IEEE Trans. Biomed. Eng., vol. 52, no. 11, pp. 1812 - 1822, Nov. 2005.
- [135]. Jennings D., Flint A., Turton B.C.H. and Nokes L.D.M., "Introduction to medical electronics applications", ISBN: 0-340-61457-9, Edward Arnold, London, 1995.
- [136]. Σεργιάδης Γ., «Βιοϊατρική Τεχνολογία», Σημειώσεις, Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, 1992.
- [137]. Pepperell J.C.T., Davies R.J.O. and Stradling J.R., "Sleep studies for sleep apnoea", Physiol. Meas. 23 R39-R74, 2002.

- [138]. Nakajima K., Matsumoto Y., and Tamura T., "Development of real-time image sequence analysis for evaluating posture change and respiratory rate of a subject in bed", *Physiological Measurement*, vol. 22, pp. N21-N28, 2001.
- [139]. Chang Y. D., Lee D.G., Ryoo S.Y., Lee D.S., "Detection of patient movement for video EEG monitoring", *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS 1998)*, vol. 20, no. 2, pp. 959-962, 1998.
- [140]. Bye A., Kok D., Ferenschild F. and Vles J., "Paroxysmal non-epileptic events in children: A retrospective study over a period of 10 years", *J. Paediatr. Child Health*, vol. 36, pp. 244-248, 2000.
- [141]. Sun M., Liu Q., Scheuer M.L., Sclabassi, R.J., "Assessment of object-based video compression for epilepsy monitoring", *Proceedings of the Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society (EMBS/BMES 2002)*, Vol. 2, pp. 1045-1046, 23-26 Oct. 2002.
- [142]. Liu Q., Sclabassi, R.J., Sun M., "A new change detection method and its application to epilepsy monitoring video", *Proceedings of the IEEE 30th Annual Northeast Bioengineering Conference*, pp. 59-60, 17-18 Apr. 2004.
- [143]. Liu Q., Sclabassi, R.J., Sun M., "Change detection in epilepsy monitoring video based on Markov random field theory", *Proceedings of the 2004 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS 2004)*, pp. 63-66, 18-19 Nov. 2004.
- [144]. Karayiannis N.B., Xiong Y. Sami A., Frost J.D., Wise M.S., Mizrahi E.M., "Improving the accuracy and reliability of motion tracking methods used for extracting temporal motor activity signals from video recordings of neonatal seizures", *IEEE Trans. Biomed. Eng.*, vol. 52, no. 4, pp. 747-749, Apr. 2005.
- [145]. Karayiannis N.B., Xiong Y. Sami A., Frost J.D., Wise M.S., Mizrahi E.M., "Quantifying motion in video recordings of neonatal seizures by robust motion trackers based on block motion models", *IEEE Trans. Biomed. Eng.*, vol. 52, no. 6, pp. 1065-1077, June 2005.
- [146]. Karayiannis N.B., Tao G., "An improved procedure for the extraction of temporal motion strength signals from video recordings of neonatal seizures", *Image and Vision Computing*, Elsevier, vol. 24, pp. 27-40, 2006.
- [147]. Liu Q., Sun M., Sclabassi, R.J., "An application of MAP to change detection in moving video", *Fourth International Symposium on Uncertainty Modeling and Analysis (ISUMA 2003)*, pp. 318-323, 2003.
- [148]. Liu Q., Sclabassi R.J., Scheuer M.L. and Sun M., "A two-step method for compression of medical monitoring video", *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS 2003)*, Vol. 1, pp. 845-848, 17-21 Sept. 2003.
- [149]. Jian Xu, Hakki Ilgin, Liu Q., Sun M., Sclabassi R.J., Chaparro L.F., "Multi-channel video for patient monitoring based on DCT compositing", *Proceedings of the IEEE 30th Annual Northeast Bioengineering Conference*, pp. 63-64, 17-18 Apr. 2004.
- [150]. Pon L-S, Sclabassi R.J., Qiang Liu, Scheuer, M., and Mingui Sun, "A medical EEG/video multimedia content description system", *Proceedings of the International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS 04)*, pp. 63-66, 592-595 Nov. 2004.
- [151]. Young I.T., Gerbrands J.J., and Lucas J. van Vliet, "Image Processing Fundamentals", in: Madisetti K.V. and Williams D.B. (Ed.), *Digital Signal Processing Handbook*, ISBN 0849385725, CRC Press LLC, ch. 51, 1999.
- [152]. Hecht J., Watkins L.S., Becker R.A., "Optoelectronics", in: Dorf R.C. (Ed.), *The Electrical Engineering Handbook*, second edition, ISBN: 0-8493-8574-1, CRC Press LLC, ch. 31, 2000.
- [153]. Παπανικολάου Γ., Μπλιάτσοι Χ., Βεγίρης Χ., "Τηλεοπτικά Συστήματα", Σημειώσεις, Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, 2003.
- [154]. Ντελόπουλος Α., "Συστήματα πολυμέσων στην εικονική πραγματικότητα", Σημειώσεις, Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, 2004.

- [155]. McGowan J.F., "AVI Overview", www.jmcgowan.com/avi/, (available online).
- [156]. National Instruments , "IMAQ Vision for LabVIEW™ User Manual", Part Number 371007A-01, www.ni.com (available online), Aug. 2004.
- [157]. National Instruments , "IMAQ Vision Concepts Manual", Part Number 372916C 01, www.ni.com (available online), Aug. 2004.
- [158]. Chen J.C., Yao K., and Hudson R.E., "Source localization and beamforming", IEEE Signal Processing Magazine, vol. 19, no. 2, pp. 30-39, March 2002.
- [159]. Brüel & Kjær, "Miniature Triaxial DeltaTron® Accelerometers - Types 4506, 4506 B, 4506 B002, 4506 B003", www.bk.dk (available online), Product Data BP 1838-16, 2004.
- [160]. Streicher R., and Everest A.F., "The new stereo soundbook", second edition, ISBN 0 9665162 0 6, Audio Engineering Associates, Pasadena, California, 1998.
- [161]. Dimoulas Ch., Papanikolaou G., Kalliris G. Pasiadis C., "Computer aided systems for prolonged recording and analysis of human bowel sounds", The Journal of the Acoustical Society of America, vol. 105, no 2, p. 1102 (abstract), Feb. 1999.
- [162]. Dimoulas C., Kalliris G., Papanikolaou G., Kalampakas A., "Prolonged recording, data manipulation and analysis of human bowel sounds using multi-channel, continuous acquisition and streaming to disk techniques", National Instruments World Wide Conference "NI-Days 2002-2003", Thessaloniki 2003.
- [163]. Srinivasan S., Pohsiang (John) Hsu, Holcomb T., Mukerjee K., Regunathan S.L., Lin B., Liang J., Lee M-C and Ribas-Corbera J., "Windows Media Video 9: overview and applications", Signal Processing: Image Communication, vol. 19, no. 9, pp. 851-875, Oct. 2004.
- [164]. "Special issue on MPEG-7", IEEE Transactions on Circuits and Systems for video Technology, vol. 11, no 6, pp. 685-772, June 2001.
- [165]. "Special issue on MPEG-7", Signal Processing: Image Communication, Elsevier, vol. 16, no. 1-2, pp. 1-293, Sept. 2000.
- [166]. Brüel P. V., Pope J., Zaveri H.K., "Acoustical Measurements and Instrumentation: Introduction", in: Crocker M. (Ed.), Encyclopedia of acoustics, ISBN 0-471-80465-7, John Wiley & Sons Inc, vol. 4, ch. 154, pp. 1839-1844, 1997.
- [167]. Jeannin S., Jasinschi R., She A., Naveen T., Mory B. and Tabatabai A., "Motion descriptors for content-based video representation", Signal Processing: Image Communication, Elsevier, vol. 16, no. 1-2, pp. 59-85, Sept. 2000.
- [168]. Jeannin S. and Divakaran A., "MPEG-7 visual motion descriptors", IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 6, pp. 720-724, June 2001.
- [169]. Fonseca P.M. and Pereira F., "Automatic video summarization based on MPEG-7 descriptions", Signal Processing: Image Communication, Elsevier, vol. 19, no. 8, pp. 685-699, Sept. 2004.
- [170]. Peker K.A. and Divakaran A., "Framework for measurement of the intensity of motion activity of video segments", Journal of Visual Communication and Image Representation, vol. 15, no. 3, pp. 265-284, Sept. 2004.
- [171]. Jo Yew Tham, Ranganath S., Ranganath M., and Kassim A.A. "A novel unrestricted center-biased diamond search algorithm for block motion estimation", IEEE Transactions on Circuits and Systems for Video Technology, vol. 8, no. 4, pp. 369 - 377, Aug. 1998.
- [172]. Zhu S. and Ma K-K, "A new diamond search algorithm for fast block-matching motion estimation", IEEE Transactions on Image Processing, vol. 9, no. 2, pp. 287 - 290, Feb. 2000.
- [173]. Collins R.T., Lipton A. J., and Kanade T., "Introduction to the Special Section on Video Surveillance", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 8, pp. 745-746, Aug. 2000.
- [174]. Collins R.T., Lipton A.J., Kanade T., Fujiyoshi H., Duggins D., Tsin Y., Tolliver D., Enomoto N., Hasegawa O., Burt P., Wixson L., "A system for video surveillance and monitoring: VSAM final report, Technical Report CMURI-R-00-12, Carnegie Mellon University, 2000.

- [175]. López M.T., Fernández-Caballero A., Fernández M.A., Mira J. and Delgado A.E., "Motion features to enhance scene segmentation in active visual attention", *Pattern Recognition Letters*, Elsevier, vol. 27, no. 5, pp. 469-478, Apr. 2006.
- [176]. Ugur Töreyn B., Enis Çetin A., Aksay A. and Bilgay Akhan M., "Moving object detection in wavelet compressed video", *Signal Processing: Image Communication*, Elsevier, vol. 20, no. 3, pp. 255-264, March 2005.
- [177]. Καλλιρής Γ., "Νέες τεχνικές αποκατάστασης αλλοιωμένων ακουστικών σημάτων ομιλίας και μουσικής", Διδακτορική διατριβή, Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, 1995.
- [178]. Vaseghi S. V., *Advanced Digital Signal Processing and Noise Reduction*, ISBN: 0-471-62692-9, Second Edition, John Wiley & Sons Ltd, New York, USA, 2000.
- [179]. Dimoulas C., Kalliris G., Papanikolaou G., Kalampakas A., "Novel wavelet domain Wiener filtering denoising techniques: application to bowel sounds captured by means of abdominal surface vibrations", *Biomedical Signal Processing and Control* (2006), doi:10.1016/j.bspc.2006.08.004, Elsevier, 2006 (in press).
- [180]. Boll S., "Suppression of Acoustic Noise in Speech Using Spectral Subtraction", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-27, no. 2, pp. 113-120, April 1979.
- [181]. Virag N., "Single channel speech enhancement based on masking properties of the human auditory system", *IEEE Transactions on Speech and Signal Processing*, vol. 7, no. 2, pp. 126-137, March 1999.
- [182]. Spriet A., Moonen M., Wouters J., "Spatially pre-processed speech distortion weighted multi-channel Wiener filtering for noise reduction", *Signal Processing*, Elsevier, vol. 84, no. 12, pp. 2367-2387, 2004.
- [183]. Spriet A., Moonen M., Wouters J., "The impact of speech detection errors on the noise reduction performance of multi-channel filtering and Generalized Sidelob Cancellation", *Signal Processing*, Elsevier, vol. 85, no. 6, pp. 1073-1088, 2005.
- [184]. Etter W., and Moschytz G., "Noise Reduction by Noise Adaptive Spectral Magnitude Expansion", *Journal of Audio Eng. Soc.*, vol. 42, no 5, pp. 341-349, May 1994.
- [185]. Wolfe P. J. and Godsill S. J., "Perceptually Motivated Approaches to Music Restoration", *Journal of New Music Research*, vol. 30, no. 1, pp. 83-92, Jan 2001.
- [186]. Tsoukalas D. E., Mourjopoulos J., and Kokkinakis G., "Perceptual Filters for Audio Signal Enhancement", *Journal of Audio Eng. Soc.*, vol. 45, no 1/2, pp. 22-36, Jan/Feb 1997.
- [187]. Kalliris G., Dimoulas Ch., Papanikolaou G., *Broad-Band Acoustic Noise Reduction using a Novel Frequency Depended Parametric Wiener Filter: Implementations using Filter-bank, STFT and Wavelet Analysis/Synthesis Techniques*, Audio Engineering Society Preprint, Proceedings of the 110th AES Convention, Preprint 5382, Amsterdam, May 2001.
- [188]. Zwicker E. and Fastl H., "Psychoacoustics Facts and Models", *Springer Series in Information Sciences*, ISBN 3-540-52600-5, Springer-Verlag Berlin 1990.
- [189]. McAulay R.J. and Malpass M.L., "Speech enhancement using a soft-decision noise suppression filter", *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. ASSP-28, no. 2, pp. 137-145, Apr. 1980.
- [190]. Ramarapu P. K. and Maher R. C., "Methods for reducing audible artifacts in a wavelet based broad-band denoising system", *Journal of Audio Eng. Soc.*, vol. 46, no. 3, pp. 178-190, March 1998.
- [191]. Δημούλας Χ., "Μετασχηματισμός Κυρματιδίων: Θεωρία, αλγόριθμοι, εφαρμογές", μεταπτυχιακή εργασία στο μάθημα: *Μέθοδοι επεξεργασίας σήματος και εφαρμογές*, (Επιβλέπων Καθ.: Σ. Μ. Πανάς), Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών ΑΠΘ, Θεσσαλονίκη, 1999.
- [192]. Misiti M., Misiti Y., Oppenheim G. and Poggi J-M, "Wavelet TOOLBOOX, For Use with Matlab, v.3", The Math Works Inc., July 2004.
- [193]. National Instruments Corporation, *Signal Processing Toolset User Manual*, Austin, Texas, USA, Part Number 322142C-01, edition December 2002.

- [194]. Daubechies I., "Orthonormal bases of compactly supported wavelets", *Commun. Pure Appl. Math.*, vol. 41, pp. 909-996, 1988.
- [195]. Rioul Ol. and Vetterli M., "Wavelet and Signal Processing", *IEEE Signal Processing Magazine*, vol. 8, no. 4, pp. 14-38, Oct 1991.
- [196]. Qian, S., and D. Chen, "Joint Time-Frequency Analysis", Englewood, Cliffs, N.J.: Prentice-Hall, ISBN 0 13254384 2, 1996.
- [197]. Mallat S., "A Wavelet Tour of Signal Processing", 2nd edition, ISBN: 0-12-466606-X, Academic Press, Elsevier (USA), 1999.
- [198]. Unser M. and Blu T., "Wavelet Theory Demistified", *IEEE Transactions on Signal Processing*, vol. 51, no. 2, pp. 470 - 483, Feb 2003.
- [199]. National Instruments Corporation, Sound and Vibration Toolkit User Manual, Austin, Texas, USA, Part Number 322194C-01, edition April 2004.
- [200]. Ruiz N., Rosa M., López F., Jarabo P., "Adaptive wavelet-packet analysis for audio coding purposes", *Signal Processing*, Elsevier, vol. 83, no. 5, pp. 919 - 929, 2003.
- [201]. Johnston J., "Transform Coding of Audio Signals Using Perceptual Noise Criteria", *IEEE Journal on Selected Areas in Communications*, vol. 6, no. 2, pp. 314-323, Feb 1988.
- [202]. Painter T. and Spanias A., "Perceptual Coding of Digital Audio", *Proceeding of the IEEE*, vol.88, no.4, pp. 451 - 515, April 2000.
- [203]. Agerkvist F. T., "A Time - Frequency Auditory Model Using Wavelet Packets", *Journal of Audio Eng. Soc.*, vol. 44, no 1/2, pp. 37-50, Jan/Feb 1994.
- [204]. Yao J., Zhang Y-T, "Bionic wavelet transform: A new time-frequency method based on auditory model", *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 8, pp. 856-863, Aug 2001.
- [205]. Rosa M., López F., Jarabo P., Maldonado S., Ruiz N., "A new algorithm for translating psycho-acoustic information to the wavelet domain", *Signal Processing*, Elsevier, vol. 81, no. 3, pp. 519-531, 2001.
- [206]. Rioul Ol. and Duhamel P., "Fast Algorithms for Discrete and Continuous Wavelets Transforms, *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 569-586, March 1992.
- [207]. Pittner S. and Kamarthi S. V., "Feature extraction from wavelet coefficients for pattern recognition tasks", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 1, pp. 83-88, Jan 1999.
- [208]. Coifman R., Wickerhauser M. V., "Entropy Based Algorithms for Best Basis Selection", *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 713-718, March 1992.
- [209]. Coifman R. and Wickerhauser M. V., "Adapted waveform 'de-noising' for medical signals and images", *IEEE Eng. Med. Biol.*, vol. 14, no. 5, pp. 578-586, Sept.-Oct. 1995.
- [210]. Berger J., Coifman R. D., Goldberg M. J., "Removing Noise from Music using Local Trigonometric Bases and Wavelet Packets", *Journal of Audio Eng. Soc.*, vol. 42, no 10, pp. 808-818, Oct 1994.
- [211]. Donoho D. L., "De-Noising by Soft-Thresholding", *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613-627, May 1995.
- [212]. Ranta R., Louis-Dorr V., Heinrich Ch., and Wolf D., "Interpretation and improvement of an iterative wavelet-based denoising method", *IEEE Signal Processing Letters*, vol. 10, no. 8, pp. 239-241, 2003.
- [213]. Ruiz N., Rosa M., López F., Jarabo P., "Adaptive wavelet-packet analysis for audio coding purposes", *Signal Processing*, Elsevier, vol. 83, no. 5, pp. 919 - 929, 2003.
- [214]. Gupta A., Joshi S. D., Prasad S., "A new method for estimating wavelet with desired features from a given signal", *Signal Processing*, Elsevier, vol. 85, no. 1, pp. 147-161, 2005.
- [215]. Dimoulas C., Kalliris G., Papanikolaou G., Kalampakas A., "Abdominal sounds pattern classification using advanced signal processing and artificial intelligence", *Proceedings of the International Conference*

- on Computational Intelligence for Modelling Control and Automation (CIMCA 2003), Vienna, pp. 71-82, Feb 2003.
- [216]. Δημούλας Χ., “ Εκπαίδευση Τεχνητού Νευρωνικού Δικτύου για την Αναγνώριση Προτύπων σε Σήματα Εντερικών Ήχων ”, μεταπτυχιακή εργασία στο μάθημα: *Εκφώνη & αυτόνομα συστήματα, νευρωνικά δίκτυα*, (Επιβλέπων Καθ.: Β. Πετρίδης), Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών ΑΠΘ, Θεσσαλονίκη, 2003.
- [217]. Dimoulas C., Kalliris G., Papanikolaou G., Kalampakas A., “Long term signal detection, segmentation and summarization using wavelets and fractal dimension: a bioacoustics application in gastrointestinal motility monitoring”, *Computers in Biology and Medicine* (2006), doi: 10.1016/j.compbimed.2006.08.013, Elsevier, 2006 (in press).
- [218]. Dimoulas C., Kalliris G., Papanikolaou G., Petridis V., Kalampakas A., “Bowel-sound pattern analysis using wavelets and neural networks with application to long-term, unsupervised, gastrointestinal motility monitoring”, *Expert Systems with Applications* (2006), doi:10.1016/j.eswa.2006.08.014, Elsevier, 2006 (in press).
- [219]. Chan C. P., Ching P. C., and Lee T., “Noisy speech recognition using de-noised multiresolution analysis acoustic features”, *The Journal of the Acoustical Society of America*, vol. 110, no. 5, pp. 2567-2574, Nov 2001.
- [220]. Sablatash M. and Cooklev T., “Compression of high-quality audio signals, including recent methods using wavelet packets”, *Digital Signal Processing*, Elsevier, vol. 6, no. 2, pp. 96-107, 1996.
- [221]. Dron J. P., Bolaers F., Rasolofondraibe I., “Improvement of the sensitivity of the scalar indicators (crest factor, kurtosis) using a de-noising method by spectral subtraction: application to the detection of defects in ball bearings”, *Journal of Sound and Vibration*, Elsevier, doi:10.1016/S0022-460X(03)00483-8, vol. 270, no. 1-2, pp. 61-73, 2004.
- [222]. Ghael S. P., Sayeed A. M., and Baraniuk R.G., “Improved wavelet de-noising via empirical Wiener filtering,” *Proceedings SPIE* 3169, 1997.
- [223]. Cohen L., “Time-Frequency Distributions - A Review”, invited paper, *Proceedings of the IEEE*, vol. 77, no. 7, pp. 941-981, Jul 1989.
- [224]. Kazama M. and Tohyama M., “Estimation of speech components by acf analysis in a noisy environment”, *Journal of Sound and Vibration*, Elsevier, doi:10.1006/jsvi.2000.3275, vol. 241, no. 1, pp. 41-52, 2001.
- [225]. Lou H. W. and Hu G. R., “An approach based on simplified KLT and wavelet transform for enhancing speech degraded by non-stationary wideband noise”, *Journal of Sound and Vibration*, Elsevier, doi:10.1016/S0022-460X(02)01556-0, vol. 268, no. 4, pp. 717-729, 2003.
- [226]. Δημούλας Χ., Αβδελίδης Κ., Καλλίρης Γ., Παπανικολάου Γ., “Ανάλυση ακουστικών φαινομένων παρατεταμένου χρόνου, και διαχείριση περιεχομένου με το πρωτόκολλο MPEG-7”, πρακτικά 3ου Πανελληνίου Συνεδρίου Ακουστικής ΕΛΙΝΑ: «ΑΚΟΥΣΤΙΚΗ 2006», Ηράκλειο Κρήτης, 18/9 - 19/9, 2006.
- [227]. Linsday A. T., and Herre J., “MPEG-7 and MPEG-7 audio – an overview”, *Journal of Audio Eng. Soc.*, vol. 49, no 7/8, pp. 589-594, Jul/Aug 2001.
- [228]. Kalliris G.M., Dimoulas C.A., Papanikolaou G. V., Avdelidis K., Passias T. N., and Stoitsis J. S., “Phoneme recognition for 3d modeled digital character talking emulation”, *Audio Engineering Society Preprint*, *Proceedings of the 112th AES Convention*, Preprint 5626, Munich, May 2002.
- [229]. Kostek B. and Czyzewski A., “An Approach to Automatic Classification of Musical Sounds”, *Audio Engineering Society Preprint*, *Proceedings of the 108th AES Convention*, Preprint 5115, Paris, February 2000.
- [230]. Kostek B. and Czyzewski A., “Representing Musical Instruments for their Automatic Classification”, *Journal of Audio Eng. Soc.*, vol. 49, no 9, pp. 768-785, Sept. 2001.
- [231]. Li D., Sethi I. K., Dimitrova N., McGee T., “Classification of general audio data for content-based retrieval”, *Pattern Recognition Letters*, Elsevier, vol. 22 no. 5, pp. 533-544, May 2001.

- [232]. Zhang T., Jay Kuo C.-C., "Audio content analysis for online audiovisual data segmentation and classification", *IEEE Transactions on speech and audio processing*, vol. 9, no. 4, pp. 441-457, May 2001.
- [233]. Lu L., Zhang H.-J., Jiang H., "Content analysis for audio classification and segmentation", *IEEE Transactions on speech and audio processing*, vol. 10, no. 7, pp. 504 - 516, Oct. 2002.
- [234]. Li Y., Narayanan S., Jay Kuo C.-C., "Content-based movie analysis and indexing based on audio visual cues", *IEEE Transactions on circuits and systems for video technology*, vol. 14, no. 8, pp. 1073-1085, Aug. 2004.
- [235]. Giannakis G. B., Tsatsanis M. K., "Signal detection and classification using matched Filtering and Higher Order Statistics", *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, no. 7, pp. 1284 - 1296, July 1990.
- [236]. Zhang J. Q., "An eigenvalue residuum-based criterion for detection of the number of sinusoids in white Gaussian noise", *Digital Signal Processing*, Elsevier, vol. 13, no. 2, pp. 275-283, 2003.
- [237]. Οικονομίδου Β., "Ανάλυση ιδιαιτεροτήτων τιμών", μεταπτυχιακή εργασία στο μάθημα: *Μέθοδοι επεξεργασίας σημάτων και εφαρμογές*, (Επιβλέπων Καθ.: Σ. Μ. Πανάς), Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών ΑΠΘ, Θεσσαλονίκη, 1998.
- [238]. Dubuc B., Quiniou J. F., Roques-Carmes C., Tricot C. and Zucker S.W., Evaluating the fractal dimension of profiles, *Physical Review A*, vol.39, no. 3, pp. 1500-1512, 1989.
- [239]. Esteller R., Vachtsevanos G., Echaz J., and Litt B., "A comparison of waveform fractal dimension algorithms", *IEEE Transactions on Circuits and Systems – I: Fundamental Theory and Applications*, vol. 48, no. 2, pp. 177-183, Feb. 2001.
- [240]. Katz M. J., "Fractals and the Analysis of Waveforms", *Computers in Biology and Medicine*, Elsevier, vol. 18, no. 3. pp. 145- 156, 1988.
- [241]. Petrosian A., "Kolmogorov complexity of finite sequences and recognition of different preictal EEG patterns," in *Proc. IEEE Symp. Computer-Based Medical Syst.*, pp. 212-217, 1995.
- [242]. Guido R. C., Pereira J. C., Fonseca E., Vieira L. S., and Sanchez F. L., "Trying different wavelets on the search for voice disorders sorting", in: *Proceedings of the 37th IEEE SSST – International South-eastern Symposium on System Theory*, Tuskegee, AL, vol. 1, IEEE Press, pp. 459-499, 2005.
- [243]. Accardo P., and Mumolo E., "An algorithm for the automatic differentiation between the speech of normals and patients with Friedreich's ataxia based on the short-term fractal dimension", *Computers in Biology and Medicine*, Elsevier, vol. 28, no. 1, pp. 75-89, 1998.
- [244]. Al-Akaidi M., "Fractal Speech Processing", Cambridge University Press, ISBN: 0521814588, 2004.
- [245]. Fekkai S., Al-Akaidi M., and Blackledge J., "A New Speech Synthesis Based on Fractal", *XI European Signal Processing Conference EUSIPCO, 2002*, Vol. II, pp. 351-354, 2002.
- [246]. Badwal R. S. S., "The application of fractal dimension to temporomandibular joint sounds", *Computers in Biology and Medicine*, Elsevier, vol. 23, no. 1, pp. 1-14, 1993.
- [247]. Brüel & Kjær Sound and Vibration Measurement A/S, "Basic Concepts of Sound", www.bk.dk (available online), Lecture Note BA 7666-11, 1, 1998.
- [248]. Li D. and Hu Y-H, "Energy-based collaborative source localization using acoustic microsensor array", *EURASIP Journal on Applied Signal Processing*, vol. 2003, no. 4, pp. 321-337, 2003.
- [249]. Sheng X. and Hu Y-H, "Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks", *IEEE Transactions on Signal Processing*, vol. 53, no. 1, pp.44-53, Jan. 2005.
- [250]. Sheng X. and Hu Y-H, "Energy based acoustic source localization", in Zhao F. and Guibas L. (Eds.), *IPSN 2003*, Springer –Verlang Berlin Heidelberg, LNCS 2634, pp. 285-300, 2003.
- [251]. Torrence C. and Compo G. P., "A Practical Guide to Wavelet Analysis", *Bulleting of the American Meteorological Society*, vol. 79, no 1, pp. 61-78, Jan 1998.

- [252]. Turkoglu I., Arslan A. and Ilkay E., "An expert system for diagnosis of the heart valve diseases", *Expert Systems with Applications*, Elsevier, vol. 23, no. 3, pp. 229-236, 2002.
- [253]. Aruna P., Puviarasan N. and Palaniappan B., "An investigation of neuro-fuzzy systems in psychosomatic disorders", *Expert Systems with Applications*, Elsevier, vol. 28, no. 4, pp. 673-679, 2005.
- [254]. İçer S., Kara S. and Güven A., "Comparison of multilayer perceptron training algorithms for portal venous doppler signals in the cirrhosis disease", *Expert Systems with Applications*, Elsevier (article in press), 2005.
- [255]. Yan H., Jiang Y., Peng C. and Li Q., "A multilayer perceptron-based medical decision support system for heart disease diagnosis", *Expert Systems with Applications*, Elsevier, vol. 30, no. 2, pp. 272-281, 2006.
- [256]. Subasi A., "EEG signal classification using wavelet feature extraction and a mixture of expert model", *Expert Systems with Applications*, Elsevier (article in press), 2006.
- [257]. Jiang Z. and Choi S., "A cardiac sound characteristic waveform method for in-home heart disorder monitoring with electric stethoscope", *Expert Systems with Applications*, Elsevier (article in press), 2006.
- [258]. Kuo H-C, Chang H-K and Wang Y-Z, "Symbiotic evolution-based design of fuzzy-neural diagnostic system for common acute abdominal pain", *Expert Systems with Applications*, Elsevier, vol. 27, no. 3, pp. 391-401, 2004.
- [259]. Aruna P., Puviarasan N. and Palaniappan B., "Diagnosis of gastrointestinal disorders using DIAGNET", *Expert Systems with Applications*, Elsevier (article in press), 2006.
- [260]. Banks S., "Signal Processing, Image Processing and Pattern Recognition", ISBN 0-13-812579-1, Prentice Hall International (UK) Ltd, 1990.
- [261]. Στριντζης Γ. Μ., "Αναγνώριση Προτύπων", Σημειώσεις, Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, 1999.
- [262]. Jain A.K., Duin R.P.W., and Mao J., "Statistical Pattern Recognition: A Review", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4-37, 2000.
- [263]. Demuth H. and Beale M., "Neural Network Toolbox for Use with MATLAB", The Math Works Inc., online: http://www.mathworks.com/access/helpdesk/help/pdf_doc/nnet/nnet.pdf, 2004.
- [264]. Bishop, C.M., "Neural Networks for Pattern Recognition", ISBN: 0-19-853849-9, Oxford: Oxford University Press, 1995.
- [265]. Roberts S.J. and Hanka, R., "An interpretation of Mahalanobis distance in the dual space", *Pattern Recognition*, Elsevier, vol. 15, no. 4, pp. 325-333, 1982.
- [266]. Hush D.R. and Horne B.G., "Progress in Supervised Neural Network, *IEEE Signal Processing Magazine*, vol. 10, no. 1, pp. 8-39, 1993.
- [267]. Sarle W. S., "Artificial Intelligence Frequently Asked Questions / Neural - Networks", available online: <ftp://ftp.sas.com/pub/neural/FAQ.html>, Cary, NC, USA, 1997-2002.
- [268]. Moody, J.E., Hanson, S.J., and Lippmann, R.P., "The Effective Number of Parameters: An Analysis of Generalization and Regularization in Nonlinear Learning Systems", *Advances in Neural Information Processing Systems*, vol. 4, pp. 847-854, 1992.
- [269]. Mc Loone S. and Irwin G., "Improving neural network training solutions using regularisation", *Neurocomputing*, Elsevier, vol. 37, no. 1-4, pp. 71-90, 2001.
- [270]. Riedmiller M. and Braun H., "A direct adaptive method for faster backpropagation learning: the RPROP algorithm", in: *Proceedings of the IEEE International Conference on Neural Networks*, San Francisco, pp. 586- 591, 1993.
- [271]. Belue L.M., and Bauer K.W., "Methods of determining input features for multilayer perceptrons, *Neurocomputing* Elsevier, vol. 7, no. 2, pp. 111-121, 1995.
- [272]. Steppe M.J., and Bauer K.W., "Improved feature screening in feedforward neural networks", *Neurocomputing*, Elsevier, vol. 13, no. 1, pp. 47-58, 1996.

- [273]. Bauer K.W., Alsing S.G. and Greene K.A. "Feature screening using signal-to-noise ratios", *Neurocomputing*, Elsevier, vol. 31, no. 1, pp. 29-44, 2000.
- [274]. Güler İ. and Übeyli E.D., "Feature saliency using signal-to-noise ratios in automated diagnostic systems developed for ECG beats", *Expert Systems with Applications*, Elsevier, vol. 28, no. 2, pp. 295-304, 2005.
- [275]. Burred J.J., and Lerch A., "Hierarchical Automatic Audio Signal Classification", *Journal of Audio Eng. Soc.*, vol. 52, no. 7/8, pp. 724-739, Jul./Aug. 2004.
- [276]. Casey M., "MPEG-7 sound recognition tools", *IEEE Transactions on circuits and systems for video technology*, vol. 11, no. 6, pp. 747, Jun. 2001.
- [277]. Kaburlasos V.G. and Petridis V., "Fuzzy Lattice Neurocomputing (FLN) models", *Neural Networks*, Elsevier, vol. 13, no. 10, pp. 1145-1169, 2000.
- [278]. Petridis V. and Kaburlasos V.G., "FINkNN: A Fuzzy Interval Number k-Nearest Neighbor Classifier for Prediction of Sugar Production from Populations of Samples", *Journal of Machine Learning Research*, vol. 4, 17-37, 2003.
- [279]. Avci E., Turkoglu I., Poyraz M., "Intelligent target recognition based on wavelet packet neural network", *Expert Systems with Applications*, Elsevier, vol. 29, no. 1, pp. 175-182, 2005.
- [280]. Martinez J. M., "MPEG-7 Overview (version 9)", ISO/IEC JTC1/SC29/WG11/N5525, Pattaya, Mar. 2003.
- [281]. Chiariglione L., "Description of MPEG-7 Audio Low Level Descriptors", ISO/IEC JTC 1/SC29/WG 11/N7708, Nice, FR, Oct. 2005.
- [282]. Kim H.-G., Moreau N., Sikora T., "MPEG-7 Audio and Beyond: Audio Content Indexing and Retrieval", ISBN-13 978-0-470-09334-4, John Wiley and Sons LTD, England, Jan. 2006.
- [283]. Αβδελίδης Κ. Α., "Δημιουργία οντολογίας κεφαλιού στο πρότυπο MPEG-7", μεταπτυχιακή εργασία στο μάθημα: *Ειδικά κεφάλαια στα Συστήματα Διαχείρισης Βάσεων Δεδομένων*, (Επιβλέπων Καθ.: Π. Α. Μήτσας), Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών ΑΠΘ, Θεσσαλονίκη, 2005.
- [284]. Καλλίρης Γ., Αβδελίδης Κ., Βαγιόκας Ν., Γούσιος Χ., Δημούλας Χ., Παπανικολάου Γ., Πασσιάδης Κ., Σεβαστιάδης Χ., "Θόρυβος και Δονήσεις σε εργασιακούς χώρους: εκτίμηση - πρόβλεψη, μέτρα προστασίας", πρακτικά 2ου Συνεδρίου Ακουστικής, ΕΛΙΝΑ, σελ. 421-430, Θεσσαλονίκη, 2004.
- [285]. Joint Special issue on "Adaptive systems & time delay estimation", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-29, no. 3, June 1981.
- [286]. Σεβαστιάδης Χ., "Προσαρμοζόμενο κατευθυντικό σόστημα μικροφώνων", Διπλωματική εργασία, Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών ΔΠΘ, Ξάνθη, Ιούνιος 1997.
- [287]. Βαμβακάς Ε., " Ανάπτυξη Μεθόδων Εντοπισμού και Ενίσχυσης Χωρικά Ανιχνευμένων Ακουστικών Πηγών", Διπλωματική εργασία, Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών ΑΠΘ, Θεσσαλονίκη, 2002.
- [288]. Dumbacher S., Blough J., Hallman D., and Wang P., "Source Identification Using Acoustic Array Techniques", *Proceedings of the SAE Noise and Vibration Conference*, vol. 2, pp 1023-1035, May 1995.
- [289]. Kim Y-H, "Can we hear the shape of a noise source", Plenary Lecture, 18th International Congress on Acoustics (ICA 2004), 4-9 Apr. 2004.
- [290]. Ginn B., Christensen J.J., Hald J., Mørkholt J., Schuhmacher A. and Blaabjerg C., "A review of array techniques for noise source location", 10th International Congress on Sound and Vibration, pp. 1889-1896, 2003.
- [291]. Chen J. C., Yao K., Hudson R. E., "Acoustic Source Localization and Beamforming: Theory and Practice", *EURASIP Journal on Applied Signal Processing*, vol. 4, pp. 359-370, 2003.
- [292]. Christensen J.J., and Hald J., "Beamforming", *Brüel & Kjaer Technical Review*, No. 1, 2004, (Ed.) Zaveri H.K., available online: www.bk.dk, pp. 1 – 48, 2004.

- [293]. Charleston S., Azimi-Sadjadi M. R., and González-Camarena R., "Interference Cancellation in Respiratory Sounds via a Multiresolution Joint Time-Delay and Signal-Estimation Scheme", *IEEE Transactions on Biomedical Engineering*, vol. 44, no. 10, pp. 1006-1019, Oct. 1997.
- [294]. Kompis M., Pasterkamp H., Motai Y., Wodicka G.R., "Spatial representation of thoracic sounds", *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 3, pp. 1661-1664, Oct.1998.
- [295]. Aarabi P., and Mavandadi S., "Robust sound localization using conditional time-frequency histograms", *Information Fusion*, Elsevier, vol. 4, pp. 111-122, 2003.
- [296]. Brüel & Kjær, "Sound intensity", available online: www.bk.dk, Booklet, revision: Sept. 1993.
- [297]. Crocker M.J. and Jacobsen F., "Sound Intensity", in: Crocker M. (Ed.), *Encyclopedia of acoustics*, ISBN 0-471-80465-7, John Wiley & Sons Inc, vol. 4, ch. 156, pp. 1855-1867, 1997.
- [298]. Maynard J.D., Williams E.G. and Lee Y., "Nearfield acoustical holography: I. Theory of generalized holography and the development of NAH", *J. Acoust. Soc. Am.*, vol. 78, no. 4, pp. 1395 - 1413, 1985.
- [299]. Wang Z. and Wu S. F., "Helmholtz equation-least-squares method for reconstructing the acoustic pressure field", *J. Acoust. Soc. Am.*, vol. 102, pp. 2020-2032, 1997.
- [300]. Hald J., "STSF - a unique technique for scan-based Near-field Acoustic Holography without restrictions on coherence", *Brüel&Kjær Technical Review*, No. 1, 1989, available online: www.bk.dk, pp. 1 - 50, 1989.
- [301]. Hald J., "Combined NAH and Beamforming Using the Same Array", in *Brüel&Kjær Technical Review*, No. 1, 2005, (Ed.) Zaveri H.K., available online: www.bk.dk, pp. 11-39, 2005.
- [302]. Steiner R. and Hald J., "Near-field acoustical holography without the errors and limitations caused by the use of spatial DFT", *Int. J. Acoust. Vib.*, vol. 6, no. 2, pp. 83-89, 2001.
- [303]. Hald J., "Patch Near-field Acoustical Holography Using a New Statistically Optimal Method", in *Brüel&Kjær Technical Review*, No. 1, 2005, (Ed.) Zaveri H.K., available online: www.bk.dk, pp. 40-50, 2005.
- [304]. Lokki T., Savioja L., Väänänen R., Huopaniemi J., and Takala T., "Creating Interactive Virtual Auditory Environments", *IEEE Computer Graphics and Applications*, vol. July/Aug. 2002, pp. 49-57, 2002.
- [305]. Weng J., and Guentchev K.Y., "Three-dimensional sound localization from a compact non-coplanar array of microphones using tree-based learning", *J. Acoust. Soc. Am.*, vol. 110, no. 1, pp. 310-323, July 2001.
- [306]. Malham, D. G., "TUTORIAL ARTICLE: Approaches to spatialisation", *Organised Sound*, vol. 3, no. 2, pp. 167-177, 1998.
- [307]. Gerzon M., "Practical Periphony: The Reproduction of Full-Sphere Sound", 65th Convention of Audio Eng. Soc., preprint number 1571, London, Feb. 1980.
- [308]. Gerzon M., "Ambisonics in Multichannel Broadcasting and Video", *J. Audio Eng. Soc.*, vol. 33, no.11, pp. 859-871, Nov. 1985.
- [309]. Jagger D. S., "New Developments and Improvements in Soundfield Microphone Technology", 75th AES Convention, Preprint Number 2064, Feb. 1984.
- [310]. SoundField, "SoundField Technology", available online: www.soundfield.com, 2006.
- [311]. SoundField, "Soundfield SPS422B Studio Microphone System User Guide", www.soundfield.com, 2006.
- [312]. Bamber J.C., "Acoustical characteristics of biological media", in: Crocker M. (Ed.), *Encyclopedia of acoustics*, ISBN 0-471-80465-7, John Wiley & Sons Inc, vol. 4, ch. 141, pp. 1703-1725, 1997.
- [313]. Freitas R. Jr., "Nanomedicine", available online: www.nanomedicine.com, 1999.
- [314]. Bauer B.B., Allen R. G., Budelman G.A., Gravereaux D.W., "Quadraphonic Matrix Perspective-Advances in SQ Encoding and Decoding Technology", *Journal of Audio Engineering Society*, vol. 21, no. 5, pp. 342-350, June 1973.

- [315]. Quadraphonic Quad, available online: www.quadraphonicquad.com, 2004.
- [316]. Dolby Laboratories, "Dolby Headphone Technology", available online: www.dolby.com, 2005.
- [317]. Mansy H., Sandler R., Royston T., and Jones D., "Testing sensors for body surface vibration measurements", Proceedings of the First Joint BMES/EMBS IEEE Conference, vol. 2, p. 800, Oct. 1999.
- [318]. Kraman S.S., Pressler G.A., Pasterkamp H., and Wodicka G.R., "Design, construction, and evaluation of a bioacoustic transducer testing (BATT) system for respiratory sounds", IEEE Transactions on Biomedical Engineering, vol. 53, no.8, pp. 1711- 1715, Aug. 2006.
- [319]. Martinez J.M., "Standards - MPEG-7 overview of MPEG-7 description tools, part 2", IEEE Multimedia, vol. 9, no. 3, pp. 83-93, July-Sept. 2002.
- [320]. Salembier P., Smith J.R., "MPEG-7 multimedia description schemes", IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 6, pp. 748-759, June 2001.
- [321]. Sikora T., "The MPEG-7 Visual Standard for Content Description – An Overview", IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 6, pp. 696 – 702, June 2001.
- [322]. Jeannin S., Divakaran A., "MPEG-7 visual motion descriptors", IEEE Transactions on Circuits and Systems for Video Technology, vol. 11, no. 6, pp. 720 – 724, June 2001.
- [323]. Day N., "MPEG-7 Applications", ISO/IEC JTC1/SC29/WG11/N4676, Jeju, Korea, Mar. 2002.
- [324]. Coimbra M.T. and Cunha J.P.-S., "MPEG-7 Visual Descriptors – Contributions for Automated Feature Extraction in Capsule Endoscopy", IEEE Transactions on Circuits and Systems for Video Technology, vol. 16, no. 5, pp. 628-637, May 2006.

Audio-visual processing and content management techniques, for the study of (human) bioacoustics' phenomena

PhD Thesis

C. A. DIMOULAS

ABSTRACT

The present doctoral thesis aims towards the development of new long-term, multi-channel, audio-visual processing techniques for the analysis of bioacoustics phenomena. The effort is focused on the study of the physiology of the lower gastrointestinal system, aiming at the support of medical research for the discovery of gastrointestinal motility patterns and the diagnosis of functional disorders. The term "processing" in this case is quite broad, incorporating the procedures of signal processing, content description, manipulation and analysis, that are applied to all the recorded bioacoustics signals, the auxiliary audio-visual surveillance information (for the monitoring of experiments and the subjects' status), and the extracted audio-video sequences describing the abdominal sound-field alterations. The thesis outline is as follows.

The main objective of the thesis, which is the technological support of medical research, is presented in the first chapter. A quick problem definition is initially attempted, presenting the difficulties of gastrointestinal motility monitoring, as well as a thorough state of research of all the biomedical engineering approaches that have been utilized in this area, including bioacoustics methods. The technical specifications of the current medical research protocol are also presented, in combination with the proposed biomedical recording, processing and analysis system, aiming to resolve all the problems related with the prolonged monitoring period and the unsupervised nature of the experimental procedure.

The second chapter focuses on the presentation of the data acquisition system, for both the bioacoustics signals and the audio-visual surveillance sequences. A quick theoretical introduction, including a state of research with the solutions employed in similar works, is initially deployed. It precedes the analysis of the current application demands and the design of a system capable of delivering the required functionalities, considering also easy future adaptation and collaboration with other psycho-physiology monitoring equipment. The modules that were tested and finally selected are then presented in a detailed manner, while conclusions and performance evaluation comments are also discussed.

The third chapter concerns signal enhancement via audio de-noising techniques. This is a very important operation that affects both the medical auscultation of gastrointestinal sound signals as well as the various aspects of automated processing. In this context, the issues that cause the background noise generation and the difficulties emerged from the presence of additive broadband noise are discussed. De-noising theory is quickly presented, with emphasis on Wiener filtering and wavelet de-noising techniques that were combined in the development of a novel "Wavelet Domain Wiener Filtering" approach. The proposed methodology is then analyzed, in combination with the specific de-noising particularities

and prospects, performance issues and computational demands. The algorithms, finally implemented, are tested in natural recordings and test signals, compared with existing methods, and rated via carefully setup qualitative and quantitative evaluation procedures.

The fourth chapter presents long-term processing procedures, such as scanning techniques for the detection of significant bioacoustics events. Difficulties in prolonged recordings' analysis and content manipulation are firstly discussed, in combination with the existed strategies in short-term and long-term abdominal sound analysis approaches. A new methodology is then described, combining both long-term summarization envelope signals, and short-term detected events. Windowed-based processing techniques and long-term denoising requirements are then analyzed. A new, multi-resolution scanning method, using wavelet energy comparison rules and fractal dimension formulas, is introduced, offering effective event detection, advanced indexing-based content manipulation, abstracting capabilities and topographic interpretation of the multi-channeled detected episodes. The performance of the new procedure is evaluated via qualitative and quantitative terms.

Chapter five focuses on the process of determination of abdominal sound patterns. Many researchers consider that the lack of a standardised terminology and the absence of related classification dictionary, are some of the major obstacles towards the establishment of bowel-sound diagnostic methods. In the current thesis, a systematic experimental procedure was carried out to study potential classification schemes for abdominal sound patterns, with the help of the collaborative medical research team. The work was accomplished via the incorporation of all the existing medical knowledge and the standardisation of experimental observations, resulting to the definition of two classification dictionaries. Correspondingly, two pattern recognition expert systems were implemented, using neural networks and syntactic - hierarchic pattern classification approaches. Both tools were evaluated after their training and proved to exhibit remarkable performance and generalization capabilities. The importance of the pattern analysis task is twofold: it allows the isolation of the interfering noise patterns, and it facilitates more descriptive studying tools, such as the temporal and the spatio-temporal pattern distribution analysis approaches.

The sixth chapter is focused on bioacoustics sound-field visualization, via sound source estimation algorithms and sound mapping tools. In this context, a general spatial-audio processing introduction is quickly presented, followed by a state of research analysis with the corresponding abdominal sound mapping techniques. A new sound-source location hybrid method is proposed, taking advantage of a three-axial accelerometer (analogously to the soundfield microphones) and the energy based localization numerical approaches. The implemented method is experimentally evaluated via synthetic abdominal propagation media and software simulation, overcoming the difficulties of older sound mapping approaches. Based on the sound localization results, various sound-field visualization tools are introduced in combination with surround sound reproduction (auscultation) techniques, to facilitate both short-term and long-term bowel-sound studies.

The seventh chapter deals with the issue of audio-visual content manipulation and analysis. An MPEG-7 compatible ontology was setup for this purpose, aiming to provide advanced browsing and cross-search analysis capabilities, for both prolonged recording and isolated events. Thus, a quick MPEG-7 introduction is initially deployed, followed by a detailed presentation of the deigned ontology. The content management description schemes refer to all the experimental data, which are the multi-channeled abdominal sound recordings, the audio and video surveillance signals, and the bioacoustics' sound-field

audio-visual sequences. The proposed ontology is quite open, taking advantage of the MPEG-7 standard capabilities, and allowing easy future extension and adaptation to combined psycho-physiological monitoring protocols.

In the eighth and last chapter the general conclusions on the total research contribution and the novelty aspects of the thesis, are presented. In general, we may distinguish two different areas that have been promoted from the current thesis results. First of all, the biomedical engineering field and the gastrointestinal medical sector are promoted, since a number of bio-signal processing methods and related analysis tools have been developed. Secondly, many of the proposed algorithms can be implemented or extended to be utilized to other audio-visual engineering applications. A complete list with the bibliographic reports utilized in this thesis is presented at the end of the corpus.

The contribution of the work is focused on the following points:

- Design and implementation of a new multi-channeled bioacoustic monitoring system that comprises the contact piezoelectric transducers, the tri-axial accelerometer, the additive sensors' attaching and adaptation modules. The new system offers high spatial sensitivity, advanced topographic analysis capabilities and remarkable tolerance to ambient noise.
- Development of an assistant audio-video surveillance system that allows for monitoring of the subjects and the experimental conditions, and facilitates automation procedures, via video motion detection techniques.
- Implementation of novel wavelet-domain Wiener filtering de-noising techniques, for signal enhancement and removal of the unwanted background noise. The new algorithms feature remarkable performance to most of the known types of gastrointestinal sounds, in contrast to the preceding methods that were concentrated on explosive bowel sounds processing. Their reduced complexity and computational load makes them ideal for long-term processing demands.
- Implementation of a new multi-resolution scanning procedure for event detection and abstraction of the prolonged recordings. The developed, wavelet-based long-term signal detection, segmentation and summarization method has been tested in natural recordings and test signals. It proved to be very efficient with compromised computational demands.
- Systematical study of the gastrointestinal sound patterns, via interdisciplinary research collaboration, exhaustive screening, various clustering procedures and taking advantage of the existed knowledge. As a result, two hierarchical pattern classification dictionaries have been defined, in combination with the corresponding pattern recognition systems, offering new study methodologies for both short-term and long-term abdominal sound recordings.
- Implementation of a new sound source localization algorithm, taking advantage of the tri-axial accelerometer components and the sound-energies of the peripheral piezoelectric transducers. Novel sound-field visualization and auralization tools have been developed facilitating the analysis process of the isolated bioacoustic events, and providing spatio-temporal summarization of the long-term gastrointestinal motility mechanisms.
- Design and implementation of a MPEG-7 adapted ontology for data manipulation and analysis of all the involved audio-visual information. The proposed ontology introduces a novel description scheme, allowing for easy browsing, searching and retrieval. In this manner, it is helpful to all the medical approaches of bowel sound analysis, promising to deliver a standardized protocol for the study of gastrointestinal motility.