

Beszéd Információs Rendszerek Pótzh 2006

1, Magyarozza meg a következő fogalmakat:

a, A magánhangzók képzését milyen három alapvető artikulációs paraméterrel jellemezhetjük?

- Ajakállás, nyelvállás, nyelvkapocs.

b, A mássalhangzókat milyen három alapvető jellemzési paraméterrel definiálhatjuk? Definiáljon egy C-t ezekkel!

-Képzés módja: zárhang, réshang, zár-réshang, nazális, pergetett, közelítő

-Képzés helye: két ajak, ajak-fog, fog, fogmeder, előső szájpaddlás, hátulsó szájpadd, gége

-Gerjesztési forma: zöngés, turbulens áramlás, vegyes

- Példa: zárhang, zöngés, két ajak (más szóval: bilabiális) : **B**

c, Melyek az összetett szerkezetű beszédhangok, milyen szerkezeti elemekből állnak? Adj legalább két példát!

- összetett szerkezetű beszédhang, mely belső időszerkezettel is rendelkezik, ami leírja a hangon belüli akusztikai jelenségek időtartam értékeit, azok arányait, összetett hangok: b p t k g gy ty c cs dz r dzs ny
- Szerkezet: zár + felpattanás(10-20ms)/réselem(50-80ms)

2, Számítsa ki a suttogott beszéd (átlagos hangnyomás: 1000 mikroP) és a kiabálás (átlagos hangnyomás 1 000 000 mikroP) közti dinamikatarományt dB-ben. Mi a jellegzetes különbség a suttogó és a normál beszéd köz spektrális színeképe közt?

MEGOLDÁS

$$p_{suttogas} = 1000\mu Pa = 10^{-3} Pa$$

$$p_{kiabalas} = 1000000\mu Pa = 1 Pa$$

$$L = 20 \cdot \lg \frac{P}{20 \cdot 10^{-6} Pa}, \text{ tehát:}$$

$$L_{suttogas} = 20 \cdot \lg \frac{10^{-3} Pa}{20 \cdot 10^{-6} Pa} = 33,98 dB$$

$$L_{kiabalas} = 20 \cdot \lg \frac{1 Pa}{20 \cdot 10^{-6} Pa} = 93,98 dB$$

$$\text{vagyis a különbség: } L_{kiabalas} - L_{suttogas} = 60 dB$$

Különbség a spektrális színeképek között:

Rendes beszéd --> zöngés --> vonalas színekép

Suttogó beszéd --> nincs alapfrekvencia --> folytonos színekép

3, Egy 10 kHz mintavételi frekvenciával és az alábbi $H(f)$ karakterisztikájú visszaállítóval működő, bemeneti szűrővel nem rendelkező mintavételező rendszer bemenetére egyidejűleg 0 dB szintű emberi beszédet valamint -6 dB szintű, 6 kHz-es frekvenciájú szinuszjel 50%-os kitöltési tényezőjű kapuzott változatát adjuk. A kapu 500 ms-ig semmit sem enged át, majd 500 ms-ig mindent átenged.

$H(f) = 1$ ha $0,3 < \text{abs}(f) < 4$; $(4,5 - \text{abs}(f))/0,5$ ha $4 < \text{abs}(f) < 4,5$; 0 egyébként

a, Mit kapunk a rendszer kimenetén?

b, Mi változik, ha a rendszer bemenetére is egy $H(f)$ karakterisztikájú szűrő kerül?

c, Milyen mintavételi frekvenciát és szűrőrendszert javasolna külön beszéd és külön (hasznos) morze jel átviteléhez, ha a szempont az alacsony költségű megvalósítás?

MEGOLDÁS

4,

a, Rajzolj fel egy szabály alapú szövegfelolvasó szerkezeti felépítését, a blokkok bemeneti és kimeneti pontjain jellemezd a működés lényegét!

b, Rajzolj fel egy korpusz alapú beszéd szintetizátor szerkezeti felépítését, a blokkok bemeneti és kimeneti pontjain jellemezd a működés lényegét!

MEGOLDÁS

5, Egy telefonos információs rendszerben a következő típusú angol nyelvű bemondást kell magyar nyelvű megoldással helyettesíteni:

"Date: dd-mm-yy" (dd:1-31, mm: Jan, Feb stb.. , yy: 1900-2100)

Adjameg a fenti üzenet jó minőségű magyar bemondásához szükséges nyelvi szerkezetet és a reális erőforrások felhasználása mellett szükséges elemeket!

MEGOLDÁS

6, Minek a mértékegysége a PHON és a dB? Mikor egyezik meg a PHON skála a dB skálával?

MEGOLDÁS

Def: Egy szinuszos hang PHON-ban mért hangossága megegyezik a vele azonos hangosságú 1 kHz-es hang akusztikai dB-ben mért intenzitásszintjével.

Tehát a PHON a szinuszos hangok hangosságának mértékegysége, a dB pedig a hangerősség. Akkor egyeznek meg, ha 1 kHz-es szinuszhangról van szó.

7, Mi a pszichoakusztikus modellezés lényege és melyek alkalmazási területei?

MEGOLDÁS

[besz_q2000.pdf](#)

Pszicho-akusztikus modell

- Idő - frekvencia leképezés
 - Keretekre vágás : rövid idejű (15 - 50 ms), átlapolódó (50%) keretek
 - Ablakozás
 - Fourier transzformáció
- Pszicho-akusztikus érzeti modellezés
 - Az emberi hallás modellezésén alapul, célja a hallható különbségek kiemelése, és a nem észlelhetőek elnyomása
 - monoton legyen a kapcsolat a belső távolság és az MOS között

Pszicho-akusztikus modell elemei

- Transzformálás az érzeti tartományra
 - nemlineáris frekvencia skálák (mel, bark...)
- Frekvencia elfedés
 - közeli frekvenciák esetén az erősebbik elnyomja a gyengébbet
- Időbeli elfedés
 - egymás utáni rövid impulzusokat egynek hallunk
 - egy erős hang elnyomja a környező gyengébbeket
- Pszicho-akusztikus hangosság
 - jel energia és hangosság kapcsolata nemlineáris

-- [Csapszi](#) - 2006.05.29.

-- dON - 2009.11.24.