2012 International Symposium on Ubiquitous Virtual Reality

Physical-Virtual Humans: Challenges and Opportunities

Greg Welch

The University of Central Florida
Institute for Simulation & Training
Computer Science Division, EECS
Orlando, FL USA 32826-3281
Email: welch@ucf.edu

The University of North Carolina at Chapel Hill
Department of Computer Science
Chapel Hill, NC USA 27599-3175
Email: welch@unc.edu

***Abstract*—Physical-virtual humans combine work in physical animatronics and virtual humans to achieve virtual humans with physical manifestations. In this article I introduce the idea of a physical-virtual human along with potential applications, potential concerns, and some historically related work. I then discuss challenges and opportunities related to applications involving physical-virtual avatars, patients, and students.**

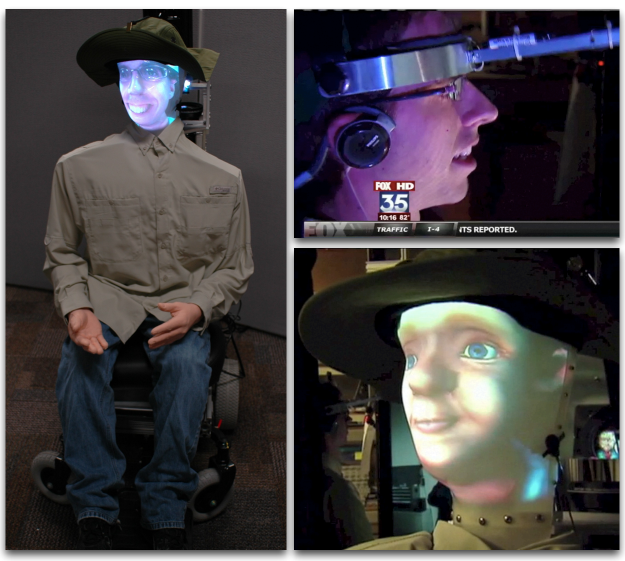

Fig. 1. Physical-Virtual Avatars (PVA) via Shader Lamps [1]. Left: Current PVA prototype at the University of Central Florida, with a real human agent and appearance. Top Right: Real human appearance obtained from a hear-worn camera. Bottom Right: Real human agency but synthetic human appearance from the TeachLive$^{\text{TM}}$ system [2].

I. Introduction

The notion of a *virtual human* has been around for many years, in computer games, Virtual Reality, and beyond. The inhabiters of virtual humans are generally classified as either *agents* or *avatars*. Agents have behavior that is purely synthetic, generally driven by some form of artificial intelligence or an expert system. Avatars on the other hand have behavior that derives from a real live human in a real-time, on-line fashion. As indicated by Blascovich and Bailenson [3] there is some debate about the etymological origins of the word "avatar," with likely sources including an online role-playing game *Habitat* developed by Lucasfilm Games, or Neal Stephenson's *Snow Crash* [4]. Today applications of virtual humans include telepresence, military and medical training, education, and healthcare. Indeed, the notion of mobile telepresence has been explored for years, for example by Tachi et al. [5], Paulos and Canny [6] and others [7], [8].

The notion of using a real, *physical* tele-robotic surrogate for viewing or manipulation in remote locations has been around for decades, and several studies have explored human interaction with a physical robot vs. a life-sized virtual robot [9], [10], [11], [12]. It turns out that physical robots more enjoyable, engaging, and effective. Recently in fact we have seen an increase in the use of robots in relatively complex human working environments such as homes, offices, and hospitals [13]. Advances include efforts to develop human-inspired [14] or human-like animatronic robots that in some ways resemble or imitate the appearance and behavior of specific humans or living creatures [15], [16], [17], [18]. Some robots are capable of changing shape in order to produce multiple expressions and identities [19], [20]. Others have included 2D cameras and displays to convey, at some level, the dynamic appearance of the "inhabiter" [5], [21], [22], [7].

One approach to developing physical avatars is to build a robotic replica of, or *surrogate* for, a particular person. This has been done for example at Walt Disney World to mimic U.S. presidents, by commercial companies such as Garner Holt who make animatronic humans for uses such as entertainment and training. Arguably the most extensive and effective human surrogates have been developed by Professor Hiroshi Ishiguro of the Intelligent Robotics Laboratory in the Graduate School of Engineering Science at Osaka University. One such example is the 50-actuator Geminoid HI-1 [23], and another is the 12-actuator Geminoid F.

Such animatronic (physical) humans can be very compelling, though sometimes considered uncanny as described in the next section. However such animatronic robotic humans will always and only look like one particular human:

978-0-7695-4766-4/12 $26.00 © 2012 IEEE
DOI 10.1109/ISUVR.2012.13

one skin tone, one gender, one shape—one person. This is in contrast to purely virtual humans, which have complete flexibility to change shape, skin tone, and gender for example. If one could combine *physical* presence and dynamic *virtual* appearance, one could convey important non-verbal information typically conveyed in human-human interactions such as between teachers and students, or physicians and patients. For example, a combination of eye contact, body proximity, forward leaning, smiling, and touch can convey intimacy and trust; while averted gaze, distance, and the absence of smiling and touch can indicate detachment [24]. For these and other reasons we are motivated to explore physical manifestations of virtual humans, i.e. *physical-virtual humans*. See Figure 1 for examples of recent work.

II. The Uncanny Valley

A phrase that sometimes occurs when interacting with such lifelike robots is the "uncanny valley," first articulated by Masahiro Mori in 1970 [25]. The basic idea is that as a robot's appearance becomes increasingly human-like, a human will have an increasingly positive response to that robot until a point is reached where the response changes quickly to one of repulsion. This is perhaps best illustrated by Fig. 1 from Mori's original paper [25], which is reproduced below in Figure 2 with added annotations "A" and "B" (explained below). While there have been a handful of rigorous empirical studies attempting to demonstrate the occurrence and boundaries of this effect, the results have been mixed [26]. For example, there is no clear consensus on whether the phenomenon exists in practice, and if it does then under what circumstances it does and to what degree there is a practical impact.

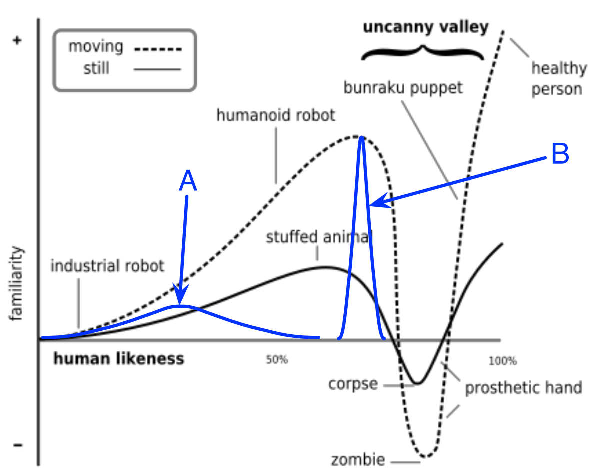

Fig. 2. A reproduction of Fig. 1 from Mori's original paper [25], with annotations A and B added to indicate regions of low and high confidence, respectively.

The related idea of uncanniness in humans actually reaches back to work by Sigmund Freud in 1919 [27], writing about a 1906 essay by Ernst Jentsch [28]. Freud defines uncanny as "doubts whether an apparently animate being is really alive; or conversely, whether a lifeless object might be, in fact, animate." This is possibly related to Pediophobia—historically a fear of dolls [29]. Freud describes how "In telling a story one of the most successful devices for easily creating uncanny effects is to leave the reader in uncertainty whether a particular figure in the story is a human being or an automaton..." and "...do it in such a way that his attention is not focused directly upon his uncertainty, so that he may not be led to go into the matter and clear it up immediately." In fact these concepts of confusion about what is really alive reaches back to "Der Sandmann" (1816)—a short story written (in German) by Ernst Theodor Wilhelm Hoffman, and "Tales of Hoffman" (1881)—an opera by Jacques Offenbach, based on stories from The Sandman. More recent theories blame a mismatch between appearance and behavior for giving rise to the uncanny valley, if it exists at all.

My own theory about the uncanny valley effect is that it is related to variations in the appearance or behavior that appear as "outliers" with respect to a decreasing uncertainty in the robot's (or avatar's) human nature, and a mental model for the human nature of the robot/avatar. Consider two cases. In the first case, if the appearance and behavior do not appear very human (e.g., to the left in Figure 2) then our uncertainty about "what it is" grows. In the face of a large uncertainty, as indicated by the notional probability density curve A in Figure 2, variations in appearance or behavior are not seen as startling because they are not outliers in the statistical sense—they are within the expected uncertainty. In the second case, if the appearance and behavior *do* appear very human like (e.g., to the right in Figure 2) then our uncertainty about "what it is" shrinks—we become more and more confident that it is human. In the face of such a small uncertainty (high confidence) in the human nature, as indicated by the notional probability density curve B in Figure 2, the same variation in appearance or behavior considered for the first case (above) might now be seen as startling because it *is* an outlier in the statistical sense. In other words, the degree to which something is uncanny is not *only* related to the degree to which we believe it is human, but to the magnitude of variations in appearance or behavior compared to our uncertainty about its human nature (our confidence that it is human).

Just as there is no clear consensus on whether the uncanny valley phenomenon exists for virtual avatars, there is no consensus about whether the phenomenon exists for physical-virtual avatars. This is something we would like to explore, once our prototypes are sufficiently smooth and accurate.

III. Physical-Virtual Humans

The idea of projecting video imagery onto a moving physical head goes back to at least 1980 and work by Michael Naimark [30] as shown in Figure 3. This system was designed for one person and recorded video only, but it did convey a physical presence with dynamic (pre-recorded) virtual appearance. In addition to the perceptual effects of having a physical presence in front of you, a physical head can

Fig. 3. Michael Naimark's film-based "Talking Head Projection" project [30]. Top Left: Adding head tracking components to the actor. Top Right: The talking head replay apparatus. Bottom: Playback of the film on a moving head, shown at two different points in time (different poses).

convey a significantly more accurate sense of avatar gaze direction than a conventional 2D virtual avatar can. This has been demonstrated for example in careful experiments by collaborator Prof. Amela Sadagic at the Naval Postgraduate School. (The work is undergoing submission for publication.)

Some details of the methods that we employ to achieve PVAs have previously been described and demonstrated with proof-of-concept prototype systems [1]. Yet many challenges remain. For example, there remain significant challenges in terms of projection onto moving objects, face capture/mapping, physical surface shape, mobility, and motion. We have developed methods to render onto moving surfaces with multiple proctors [31] but these do not account for static errors (e.g., calibration errors) or dynamic errors (e.g., latency-induced misregistration). We continue to work on these and more issues, for example developing methods employing closed-loop methods for ensuring physical-virtual registration, and methods for adapting the rendering to hide artifacts in cases where there is uncertainty. In the situation where there is human agency (e.g., see Figure 1) there is the issue of accurate dynamic tracking and mapping of the human head motion and facial appearance. In general this seems like a trivial problem as there would seem to be many options (e.g., FaceAPI by Seeing Machines, Inc.), however mapping the appearance onto a physical head makes it very challenging. Otherwise unnoticeable small errors become grotesque on a physical-virtual head. It's a challenge to "lock on" certain features (e.g., the nose) while allowing others to move (e.g., mouth or eyes). Head shape is also a challenge for various reasons, including the the desires to accommodate multiple "inhabiters" and synthetic animatronic motion.

A. Physical-Virtual Patients

2D virtual humans have been used to train and assess humans in interpersonal scenarios for applications including medical interviews and examinations1. However, some examinations require physical interactions with the patient and spatial awareness that are difficult to simulate using flat displays. Human Patient Simulators afford physical interactions, spatial awareness, and simulated physiological behaviors, but otherwise have static appearance-based interpersonal behavior.

As described in [32], [33], we used our PVA prototype to create an interactive training experience for medical students to conduct exams on a PVP. We performed a formative evaluation of the system (n=8) using medical educators and students previously trained in such exams. Each participant was introduced to the system, performed a patient interview and exam, responded to online questions addressing usability and co-presence, and participated in a guided discussion with the investigator(s). Exams lasted about 25 minutes, and discussions about 30 minutes. Most participants were positive about the paradigm. One reported "We dont have to move around a lot for this type of thing, but...I would think it will be a huge thing to learn how to move around an exam room with a patient...when you are for the first time seeing patients in real life, biggest thing that I thought about was I dont want to embarrass myself if I dont know how to move around."

B. Physical-Virtual Students

One education-related area of interest is telepresence for students and teachers. While computers can afford high-quality distance education experiences [34], a high percentage of students drop-out of these courses due to limited opportunities for interaction with other students and teachers [35]. For example, teachers who are present through televideo cannot take advantage of behavioral management techniques that involve movement throughout the classroom [36], nor can they write on chalkboards, annotate physical work products that students might create in the classroom, or even communicate with students collaborating in multiple small groups. Likewise, it can be difficult for teachers and *physically-present* students (real students) to monitor and engage in meaningful dialogue with students participating through televideo [37], [38]. Given that the quality of verbal and non-verbal interaction between teachers and students is a strong predictor of student involvement [39], the limitations of traditional distance education are serious barriers to learning [40]. While it is felt most acutely for the remote participants, the interaction limitations of televideo have negative effects upon learning for those in the traditional classrooms as well (e.g., teleconferencing with guest speakers).

Beyond telepresence for students and teachers, we are also exploring the use of synthetic physical-virtual appearance with human agency. See for example bottom-right of Figure 1. One example use is for therapy or training of students with special needs, e.g., students with autistic tendencies.

Acknowledgments

I would like to acknowledge collaborators at the University of Central Florida: Charlie Hughes, Arjun Nagendran, Remo Pillat, and Ryan Patrick; the University of North Carolina at Chapel Hill: Henry Fuchs, Adrian Ilie, Andrei State, Feng Zheng, Ryan Schubert, Peter Lincoln, and John Thomas; Nanyang Technological University: Gerald Seet and I-Ming Chen; and the University of Florida: Ben Lok and Diego Rivera-Gutierrez. My personal work in this area was inspired by Dr. Chris Macedonia, U.S. Army. Financial support comes from the U.S. Office of Naval Research and the U.S. National Science Foundation.

References

[1] P. Lincoln, G. Welch, A. Nashel, A. State, A. Ilie, and H. Fuchs, "Animatronic shader lamps avatars," Virtual Reality, pp. 1–14, 2010, 10.1007/s10055-010-0175-5. [Online]. Available: http://dx.doi.org/10.1007/s10055-010-0175-5

[2] D. P. Mapes, P. Tonner, and C. E. Hughes, "Geppetto: an environment for the efficient control and transmission of digital puppetry," in Proceedings of the 2011 international conference on Virtual and mixed reality: systems and applications - Volume Part II. Berlin, Heidelberg: Springer-Verlag, 2011, pp. 270–278. [Online]. Available: http://dl.acm.org/citation.cfm?id=2028716.2028749

[3] J. Blascovich and J. Bailenson, Infinite Reality: Avatars, Eternal Life, New Worlds, and the Dawn of the Virtual Revolution. HarperCollins, 2011. [Online]. Available: http://books.google.com/books?id=QcNDxzsbMPkC

[4] N. Stephenson, Snow Crash, ser. Bantam Spectra Book. Bantam Books, 2000. [Online]. Available: http://books.google.com/books?id=RMd3GpIFxcUC

[5] S. Tachi, "Telesar2 project page," http://projects.tachilab.org/telesar2/, May 2009.

[6] E. Paulos and J. Canny, "Prop 2 at the hcc symposium," http://www.prop.org/prop2/images_hcc.htm, March 2 2001.

[7] I. Health, "Intouch technologies, inc." http://www.intouchhealth.com/, July 2012.

[8] N. P. Jouppi, S. Iyer, S. Thomas, and A. Slayden, "Bireality: mutually-immersive telepresence," in MULTIMEDIA '04: Proceedings of the 12th annual ACM international conference on Multimedia. New York, NY, USA: ACM, 2004, pp. 860–867.

[9] W. Bainbridge, J. Hart, E. Kim, and B. Scassellati, "The effect of presence on human-robot interaction," in Robot and Human Interactive Communication, 2008. RO-MAN 2008. The 17th IEEE International Symposium on, Aug. 2008, pp. 701–706.

[10] C. Kidd and C. Breazeal, "Effect of a robot on user perceptions," in Intelligent Robots and Systems, 2004. (IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on, vol. 4, Sept.-2 Oct. 2004, pp. 3559–3564 vol.4.

[11] ——, "Robots at home: Understanding long-term human-robot interaction," in Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on, Sept. 2008, pp. 3230–3235.

[12] A. Powers, S. Kiesler, S. Fussell, and C. Torrey, "Comparing a computer agent with a humanoid robot," in HRI '07: Proceedings of the ACM/IEEE international conference on Human-robot interaction. New York, NY, USA: ACM, 2007, pp. 145–152.

[13] A. M. Bauer, D. Wollherr, and M. Buss, "Human-robot collaboration: a survey." I. J. Humanoid Robotics, vol. 5, no. 1, pp. 47–66, 2008.

[14] S. Coradeschi, H. Ishiguro, M. Asada, S. C. Shapiro, M. Thielscher, C. Breazeal, M. J. Mataric, and H. Ishida, "Human-inspired robots," IEEE Intelligent Systems, vol. 21, no. 4, pp. 74–85, 2006.

[15] C. Breazeal and B. Scassellati, "Robots that imitate humans," Trends in Cognitive Sciences, vol. 6, no. 11, pp. 481–487, November 2002. [Online]. Available: http://dx.doi.org/10.1016/S1364-6613(02)02016-8

[16] AIST, "Successful development of a robot with appearance and performance similar to humans," http://www.aist.go.jp/aist_e/latest_research/2009/20090513/20090513.html, May 2009.

[17] H. Ishiguro, "Intelligent Robotics Laboratory, Osaka University," http://www.is.sys.es.osaka-u.ac.jp/research/index.en.html, May 2009.

[18] R. Epstein, "My date with a robot," Scientific American Mind, vol. June/July, pp. 68–73, 2006.

[19] T. Laboratory, "Various face shape expression robot," http://www.takanishi.mech.waseda.ac.jp/top/research/docomo/index.htm, Takanishi Laboratory, WASEDA University, August 2009.

[20] T. Yotsukura, F. Nielsen, K. Binsted, S. Morishima, and C. S. Pinhanez, "Hypermask: Talking head projected onto real object," The Visual Computer, vol. 18, no. 2, pp. 111–120, April 2002. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.15.3623&rep=rep1&type=pdf

[21] J. Paradiso and J. Landay, "Guest editors' introduction: Cross-reality environments," Pervasive Computing, IEEE, vol. 8, no. 3, pp. 14–15, July-Sept. 2009.

[22] A.-M. Corley, "The reality of robot surrogates," http://spectrum.ieee.org/robotics/humanoids/the-reality-of-robot-surrogates/0, September 23 2009.

[23] E. Guizzo, "Hiroshi ishiguro: The man who made a copy of himself," IEEE Spectrum, 2010.

[24] R. Heath and J. Bryant, Human communication theory and research: concepts, contexts, and challenges, ser. LEA's communication series. L. Erlbaum, 2000. [Online]. Available: http://books.google.com/books?id=yZBmNgr8cMAC

[25] M. Mori, "The uncanny valley," Energy, vol. 7, no. 4, pp. 33–35, 1970.

[26] J. Seyama and R. S. Nagayama, "The uncanny valley: Effect of realism on the impression of artificial human faces," Presence: Teleoper. Virtual Environ., vol. 16, no. 4, pp. 337–351, 2007.

[27] S. Freud, "Freud, sigmund: Das unheimliche [1919]," Sigmund Freud Studienausgabe Bd, vol. 4, pp. 241–274, 1919.

[28] E. Jentsch, "On the psychology of the cncanny," Angelaki: Journal of Theoretical Humanities, vol. 2, no. 1, pp. 7–16, 1906. [Online]. Available: http://dx.doi.org/10.1080/09697259708571910

[29] M. Mufson, S. Slon, A. Dadloy, and K. Allison, Coping with Anxiety and Phobias. Harvard Health Publications, 2008. [Online]. Available: http://books.google.com/books?id=fBAVSQAACAAJ

[30] M. Naimark, "Talking head projection," http://www.naimark.net/projects/head.html, 1980.

[31] P. Lincoln, G. Welch, and H. Fuchs, "Continual surface-based multi-projector blending for moving objects," Proceedings of IEEE Virtual Reality Conference (VR), 2011, pp. 115–118, March 19–23 2011.

[32] D. Rivera-Gutierrez, G. Welch, P. Lincoln, M. Whitton, J. Cendan, D. A. Chesnutt, H. Fuchs, and B. Lok, "Shader lamps virtual patients: the physical representation of virtual patients," in To appear in Medicine Meets Virtual Reality 19 - NextMed, ser. Studies in Health Technology and Informatics. IOS Press, 2012.

[33] G. Welch, D. Rivera-Gutierrez, P. Lincoln, M. Whitton, J. Cendan, D. A. Chesnutt, H. Fuchs, B. Lok, and R. Skarbez, "Physical manifestations of virtual patients," Simulation in Healthcare, vol. 6, no. 6, p. 488, December 2011.

[34] R. M. Bernard, P. C. Abrami, Y. Lou, E. Borokhovski, and Et, "How does distance education compare with classroom instruction? a meta-analysis of the empirical literature," Review of Educational Research, vol. 74, no. 3, pp. 379–439, 2004.

[35] W. H. Hannum, M. J. Irvin, P.-W. Lei, and T. W. Farmer, "Effectiveness of using learner-centered principles on student retention in distance education courses in rural schools," Distance Education, vol. 29, no. 3, pp. 211–229, 2008. [Online]. Available: http://dx.doi.org/10.1080/01587910802395763

[36] J. Levin and J. F. Nolan, Principles of classroom management: A professional decision-making model, 4th ed. River, NJ: Allyn & Bacon, 2004.

[37] A. G. Chute, From Teletraining to e-Learning and Knowledge Management. Mahwah, NJ: Lawrence Erlbaum Associates, 2003, pp. 297–314.

[38] S. Naidu, Designing Instruction for e-Learning Environments. Mahwah, NJ: Lawrence Erlbaum Associates, 2003, pp. 349–366.

[39] M.-C. Opdenakker and J. V. Damme, "Teacher characteristics and teaching styles as effectiveness enhancing factors of classroom practice," Teaching and Teacher Education, vol. 22, no. 1, pp. 1–21, January 2006.

[40] J. J. G. van Merriënboer and J. Sweller, "Cognitive Load Theory and Complex Learning: Recent Developments and Future Directions," Educational Psychology Review, vol. 17, no. 2, pp. 147–177, June 2005. [Online]. Available: http://dx.doi.org/10.1007/s10648-005-3951-0