

Virtual Coaches over Mobile Video

David M. Krum

Sin-Hwa Kang

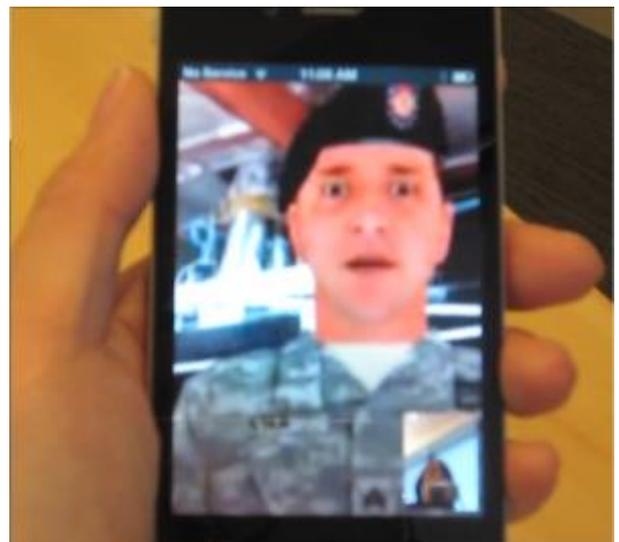
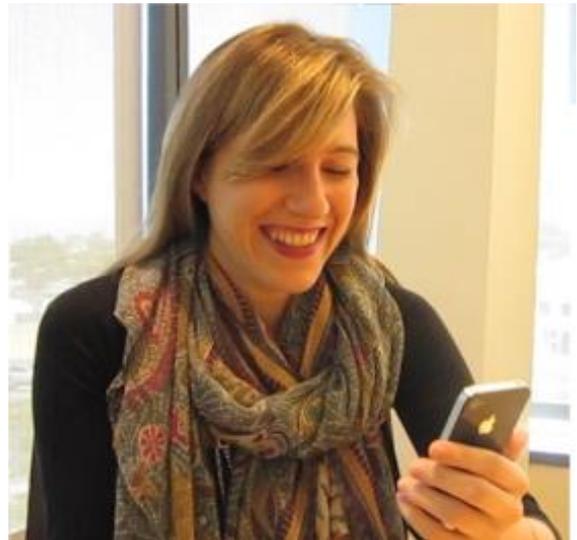
Mark Bolas

Institute for Creative Technologies
University of Southern California
{krum, kang, bolas}@ict.usc.edu
<http://ict.usc.edu>

Abstract

We hypothesize that the context of a smartphone, how a virtual human is presented within a smartphone app, and indeed, the nature of that app, can profoundly affect how the virtual human is perceived by a real human. We believe that virtual humans, presented over video chat services (such as Skype) and delivered using mobile phones, can be an effective way to deliver coaching applications. We propose to build a prototype system that allows virtual humans to initiate and receive Skype calls. This hardware will enable broadcast of the audio and video imagery of a character. Using this platform and a virtual human, we will conduct two user studies. The first study will examine factors involved in making a mobile video based character seem engaging and “real”. This study will examine how character appearance and the artifacts of the communication channel, such as video and audio quality, can affect rapport with a virtual human. The second study will examine ways to maintain a long-term relationship with a character, leveraging the character’s ability to call and interact with a real human over a longer period of time. These studies will help develop design guidelines for presenting virtual humans over mobile video.

Keywords: virtual humans; smartphones; communication channels; virtual coaches; video chat services; reciprocity; self-disclosure; social exchange theory



Virtual Coaches over Mobile Video

David M. Krum

Sin-Hwa Kang

Mark Bolas

Institute for Creative Technologies
University of Southern California
{krum, kang, bolas}@ict.usc.edu
<http://ict.usc.edu>

Abstract

We hypothesize that the context of a smartphone, how a virtual human is presented within a smartphone app, and indeed, the nature of that app, can profoundly affect how the virtual human is perceived by a real human. We believe that virtual humans, presented over video chat services (such as Skype) and delivered using mobile phones, can be an effective way to deliver coaching applications. We propose to build a prototype system that allows virtual humans to initiate and receive Skype calls. This hardware will enable broadcast of the audio and video imagery of a character. Using this platform and a virtual human, we will conduct two user studies. The first study will examine factors involved in making a mobile video based character seem engaging and “real”. This study will examine how character appearance and the artifacts of the communication channel, such as video and audio quality, can affect rapport with a virtual human. The second study will examine ways to maintain a long-term relationship with a character, leveraging the character’s ability to call and interact with a real human over a longer period of time. These studies will help develop design guidelines for presenting virtual humans over mobile video.

Keywords: virtual humans; smartphones; communication channels; virtual coaches; video chat services; reciprocity; self-disclosure; social exchange theory

1. Introduction

Virtual humans are beginning to show promise in applications such as entertainment, therapy, education, and training. In large part, this is due to their ability to evoke social responses in real humans. With the ubiquity of smartphones, the obvious next step for virtual humans is integration into various mobile apps. We believe that virtual humans, presented over video chat services (such as Skype) and delivered using mobile phones, can be an effective way to deliver counseling and coaching applications. However, we believe that this mobile video platform is very different than previously researched web-based and PC-based virtual human platforms. New research is needed to understand how to effectively present virtual characters on a mobile video platform. We hypothesize that factors such as the context of a smartphone, how the virtual human is presented within a smartphone app, and indeed, the nature of that app, can profoundly affect how a real human perceives the virtual human. We believe that virtual humans who communicate over videoconference services like Skype and Apple’s FaceTime have unique advantages over other forms of presentation, including characters in native smartphone apps. They may appear to be more realistic since they can communicate using video conferencing, much like real humans. We have also identified additional design questions involved in developing mobile virtual human experiences. Can video communication artifacts lower the realism requirement for virtual humans? What behaviors, visuals, and utterances might make

characters more relatable and increase rapport? Can reciprocity and other concepts from Social Exchange Theory encourage repeated interactions with the virtual character?

In this research, we propose to develop design guidelines for the deployment of virtual humans on smartphones, with a specific focus on mental wellbeing related coaching and counseling. To achieve this end, we will develop an apparatus that will allow virtual humans to interact over Skype. We will then perform two human subject studies. The first will examine if such characters can be made engaging and “real” by manipulating character appearance and taking advantage of audio and visual artifacts in the communications channel. The second will explore how these characters can establish and maintain long-term relationships, leveraging ideas from Social Exchange Theory. Our studies will focus on tasks and interactions pertaining to virtual humans acting as counseling coaches.

2. Related Work

Human-computer interaction studies suggest that people like to engage with an anthropomorphic agent in counseling interview interactions. This is also supported by Tourangeau et al. who noted that the presence of an interviewer could help retain users’ motivation [8]. Using immersive head mounted displays and large monitors, Van Vugt also found that people were highly engaged when interacting with virtual humans that presented an appearance similar to the user themselves and provided a helpful advice [9].

The study also discovered that male users presented negative reactions to virtual humans whose appearance was similar to their own faces when the character’s advice was unhelpful. Based on the results of the study, it is unknown if this general affinity for similar appearance would also apply to small non-immersive smartphone displays and if the negative male reaction to unhelpful, but similar characters might also manifest in some manner. Adaption of this study to a smartphone platform and the counseling domain would be of interest. Our notion that the context of the communications channel can influence a user’s reaction to a virtual character is supported in prior work by Takayama and Nass [7]. In their

study, an assistive agent was presented to users in a driving simulator. While the assistive content spoken by the agent did not differ, half of the drivers were told that the agent was on board the vehicle, while the other half were told that the agent was located elsewhere and remotely communicating with the vehicle. Drivers were more engaged with the “on-board” system and felt less discontentment; however, they drove faster than with the wireless system [7]. Our proposed work would likely uncover similar impacts of smartphone context, albeit with different specific reactions as smartphones and vehicles evoke different conceptual and contextual models in users.

3. Approach and Evaluation

At one level, the new idea of this research is to use a video conference-based virtual human prototype as a tool to determine design guidelines for a smartphone based virtual human coaches. At a deeper level, our proposal recognizes smartphones as a unique communications channel. A smartphone will color interactions with virtual humans due to the context, associations, limitations, and capabilities perceived by the user concerning the smartphone, the apps, and the virtual human. We believe the app matters and certain apps and presentations will make the virtual human appear more “real”. Thus the deeper idea behind this proposal is an attempt to build a model of how smartphone context can help or hinder the perceived realism, co-presence, and rapport between the virtual and real human. An understanding of this context will help designers of smartphone based virtual humans for counseling, but should also apply to a variety of other domains including education and training.

For this effort, we propose to build a prototype videoconference delivery platform for a virtual coach. The virtual coach software will leverage code from our previous virtual human characters, but will add behaviors and dialog appropriate for a video chat. The hardware will use a webcam to bring sound and imagery of the character into Skype. This platform will be used in user studies where metrics like co-presence, rapport, and self-disclosure can be used to gauge effects on interaction with the character. The research

effort will seek to demonstrate how a video chat-based virtual coach can effectively engage with users.

4. Preliminary Findings and Implications

We performed a compelling informal experiment, where we presented a puppeteered virtual human character to participants using Skype, a video chat application, on a smartphone (see the images in Figure 1). Three users were recruited at the authors' institution to interact with the system. The users voluntarily participated in the experiment without any compensation for their participation. They had never experienced the system prior to the experiment.

The prototype featured a virtual human that was displayed on a smartphone through Skype. The virtual human prototype was implemented using the Panda3D graphics library and the Microsoft Speech API using the Cepstral David voice. The virtual human interacted with our human users by means of a Wizard of Oz (WOZ) interface. Through this interface, the virtual human's speech was controlled by a human operator who chose an appropriate text to speech utterance in response to each of the human users' statements. Nonverbal responses were generated through puppeteering by a human operator with a Razer Hydra two-handed game controller. The human users were not aware that a human operator was controlling the virtual human. They were led to believe they were interacting with a fully autonomous artificially intelligent system. During the interaction, the virtual human asked users a couple of questions, such as "What role do you play at ICT?" We asked users' feelings about their interactions with the virtual human when they finished their conversations.

In general, they found this character to be quite engaging in spite of numerous audio/video artifacts and the small display size of the character. All three of the users also agreed to further contacted when asked by the virtual human. One of them even responded to the virtual human's question: "Can I get in touch with you later if I think of more questions?" by saying: "Do you want my cell number?" during the interaction. We conjectured that the use of Skype might have created a special

context for the user's interaction with the character. Since the virtual human "called" the user, the virtual human appeared to have greater agency. Users could also attribute video and audio defects present in the character to the losses and lags inherent to the video channel's network connection. The conception of Skype as a communications channel connecting one locale to another may also have reinforced the belief that the character had a real physical presence somewhere in the world.

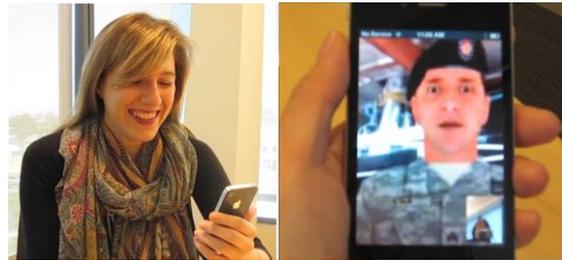


Figure 1: A user's interaction with a virtual coach on a smartphone using Skype.

We feel that a different scenario with a different context would evoke a very different response on the part of the user. A user who selected a particular smartphone app, waited for the app and character to initialize, and then began an interaction, would probably infer that the character, in some way, "lived" inside the smartphone. The character would be local and limited to the capabilities of that device. Defects in the character's presentation would not be attributable to distance or a lossy communications channel, but rather to limitations of the device's or the character's capabilities. From this scenario, we surmised that the context of the smartphone could play a large role in influencing the perceived realism of the character. We believe that it is important to understand how a smartphone's context will impact the perception of virtual characters.

5. Future Work and Contributions

We are currently designing and developing a character that could effectively communicate over mobile video and display the appropriate online social behaviors. Once we have the character, we will conduct two user studies.

In the first study, we will manipulate the appearance of virtual coaches, and examine effects on user interaction, in particular focusing on measures of social engagement

and rapport. We will also be observing the role of audio and visual channel artifacts on these measures. In this study, participants would be recruited from Craigslist and brought to our lab to interact with a virtual character on a prepared smartphone in a counseling related task. The results of this investigation will be compared with the outcomes of prior work that was conducted in an immersive head mounted display setting and also presented on full sized computer monitors [9]. The second proposed study will examine how a virtual coach can effectively utilize a longer-term relationship characterized by several interactions over a week or a month, initiated by the virtual coach. This study will compare reactions of users who are contacted once vs. users who are contacted several times by the video-based coach. We will be guided by Social Exchange Theory [2], exploring how promises, favors, gifts, or other forms of reciprocity might enhance relationships with repeated interactions. Participants would be specifically recruited if they had access to a smartphone, were willing to participate in several calls with a virtual character, and could fill out an online survey after each interaction. The survey will include measures that evaluate users' virtual rapport with the virtual character, partner perception, and so forth [4]. For the interview questions of a virtual coach, we will modify questions that were used in a previous study [5].

This research effort can be considered fruitful if new design ideas and guidelines for virtual human applications on mobile devices are generated, particularly with respect to video based virtual characters, long term relationships with virtual humans, virtual counseling coaches on mobile devices, and the effects of the smartphone context on virtual human interactions. Additional success can be measured if qualities of user interactions can be significantly influenced by design guidelines and implementation characteristics developed and employed in these studies. These developments will contribute to the design knowledge concerning the deployment of virtual humans.

References

- [1] Burns, M., Begale, M., Duffecy, J., Gergle, D., Karr, C., Giangrande, E., & Mohr, D. Harnessing context sensing to develop a mobile intervention for depression, *Journal of Medical Internet Research*, Vol. 13, 3: e55 (2011).
- [2] DeLamater, J.D. & Myers, D.J. *Social Psychology*, Cengage Learning (2010).
- [3] Gluhak, A., Presser, M., Zhu, L., Esfandiyari, S., & Kupschick, S. Towards mood based mobile services and applications, *EuroSSC (2007)*: 159-174.
- [4] Kang, S. & Gratch, J. Socially anxious people reveal more personal information with virtual counselors that talk about themselves using intimate human back stories, *The Annual Review of CyberTherapy and Telemedicine*, Vol. 181 (2012), 202-207.
- [5] Kang, S. & Gratch, J. Exploring users' social responses to computer counseling interviewers' behavior, Presented at the *International Communication Association Conference*, London, UK (2013).
- [6] Morris, M. & Guilak, F. Mobile heart health: Project highlight, *Journal of IEEE Pervasive Computing*, Vol. 8, 2 (2009): 57-61.
- [7] Takayama, L. & Nass, C. Driver safety and information from afar: An experimental driving simulator study of wireless vs. in-car information services. *International Journal of Human-Computer Studies*. 66, 3 (2008), 173-184.
- [8] Tourangeau, R. & Smith, T. W. Asking sensitive questions: The impact of data collection mode, question format, and question context. *The Public Opinion Quarterly*, Vol. 60, No. 2 (1996), 275-304.
- [9] Van Vugt, H., Bailenson, J., Hoorn, J., & Konijn, E. Effects of facial similarity on user responses to embodied agents. *ACM Transactions on Computer-Human Interaction*, 17, 2, Article 7 (2010), 7:1-7:27.