Computers in Human Behavior 55 (2016) 680–694

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Full length article

The design of virtual audiences: Noticeable and recognizable behavioral styles

Ni Kang [a, *], Willem-Paul Brinkman [a], M. Birna van Riemsdijk [a], Mark Neerincx [a, b]

[a] Department of Intelligent Systems, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands
[b] TNO Human Factors, Kampweg 5, 3769 DE Soesterberg, The Netherlands

ARTICLE INFO

Article history:
Received 9 February 2015
Received in revised form
6 October 2015
Accepted 11 October 2015
Available online 12 November 2015

Keywords:
Virtual audience
Expressive behavior
Bodily expression recognition
Simulated audience settings

ABSTRACT

Expressive virtual audiences are used in scientific research, psychotherapy, and training. To create an expressive virtual audience, developers need to know how specific audience behaviors are associated with certain characteristics of an audience, such as attitude, and how well people can recognize these characteristics. To examine this, four studies were conducted on a virtual audience and its behavioral models: (I) a perception study of a virtual audience showed that people ($n = 24$) could perceive changes in some of the mood, personality, and attitude parameters of the virtual audience; (II) a design experiment whereby individuals ($n = 24$) constructed 23 different audience scenarios indicated that the understanding of audience styles was consistent across individuals, and the clustering of similar settings of the virtual audience parameters revealed five distinct generic audience styles; (III) a perception validation study of these five audience styles showed that people ($n = 100$) could differentiate between some of the styles, and the audience's attentiveness was the most dominating audience characteristic that people perceived; (IV) the examination of the behavioral model of the virtual audience identified several typical audience behaviors for each style. We anticipate that future developers can use these findings to create distinct virtual audiences with recognizable behaviors.

1. Introduction

Virtual audiences can elicit responses in humans similar to those that are elicited by real human audiences (Slater, Pertaub, Barker, & Clark, 2006; Zanbaka, Ulinski, Goolkasian, & Hodges, 2007). This is used in scientific research (e.g., Kelly, Matheson, Martinez, Merali, & Anisman, 2007), psychotherapy (e.g., Powers & Emmelkamp, 2008), and training (e.g., Bissonnette, Dubé, Provencher, & Moreno Sala, 2015), because virtual environments are easier to configure and control than the real world. While some applications aim for a neutral audience (e.g., Wallergard, Jonsson, Osterberg, Johansson, & Karlson, 2011), others may benefit more from an expressive audience. For example, the treatment manuals of exposure therapy (Heimberg & Becker, 2002; Hofmann & Otto, 2008) suggest controlling the audience attitude as an effective means of controlling anxiety in a public speaking scenario; studies on stress responses explored variations of stress tests using supportive and non-supportive audiences (Kelly et al., 2007; Taylor et al., 2010). As virtual audiences in public speaking scenario are becoming more widely used, e.g., as part of the Trier Social Stress Test (TSST) (Kirschbaum, Pirke, & Hellhammer, 1993), and in exposure therapy for social anxiety disorder, an empirically validated expressive virtual audience appropriate for these applications is needed.

When individuals are exposed to a virtual environment and perform in front of a group of virtual humans, their belief, anxiety, and performance can be affected. For example, Wallergard et al. (2011) suggested that virtual audiences as part of a stress test can indeed, like human audiences, induce stress. Aymerich-Franch, Kizilcec, and Bailenson (2014) used a virtual audience to study the effects of self-representation on public speaking anxiety. When presenting in front of a virtual audience, the individuals could see in a virtual mirror their virtual reflection which was manipulated to be similar or dissimilar to themselves. Others (Anderson et al., 2013; Hartanto et al., 2015; Morina, Brinkman, Hartanto, Kampmann, & Emmelkamp, 2015) focused on giving people the experience of performing in front of an audience as part of exposure therapy for individuals with social anxiety disorder. This experience has also benefited non-clinical applications. For example,

* Corresponding author.
E-mail address: conniekdl@gmail.com (N. Kang).

http://dx.doi.org/10.1016/j.chb.2015.10.008
0747-5632/

Bautista and Boone (2015) let teachers be trained with virtual students to master their skills of content delivery and student management. Likewise, Bissonnette et al. (2015) trained performance arts students, in this case, young musicians to overcome their performance anxiety by performing in front of a virtual audience. The information expressed by virtual audiences can be used for various purposes. For example, the virtual audience in a public speaking training system manifested different attitudes as feedback for the speech performance (Chollet, Sratou, & Shapiro, 2014). Supportive and non-supportive audiences have been used to evoke different levels of anxiety (Kelly et al., 2007; Taylor et al., 2010). Thus, the expressiveness of a virtual audience, i.e., what information a virtual audience can express and whether people can recognize the information, becomes a key question when designing virtual audiences.

As virtual audiences are made up of individual virtual humans, the first step in the development is the generation of individual virtual humans with believable behavior. Extensive work has been done in simulating such behavior. This work ranges from facial expression of emotion (Broekens, Qu, & Brinkman, 2012), head movement (Wang, Lee, & Marsella, 2011), to full body posture simulation (Chollet, Ochs, & Pelachaud, 2014; Xu, Pelachaud, & Marsella, 2014). Besides emotions, Chollet, Ochs, et al. (2014) and Hu, Walker, Neff, and Tree (2015) demonstrated that attitude and even personality of an individual virtual human can effectively be expressed by body language. To make the virtual characters believable, dynamic behaviors, i.e., displaying sequences of behaviors instead of still images, are often required. These sequences can be pre-scripted (Hu et al., 2015), computed by crafted rules that specify which behavior should be generated in a certain context based on psychological knowledge and literature (Bevacqua, Sevin, Hyniewska, & Pelachaud, 2012), or generated by statistical models that predict body postures based on observation (Chollet, Ochs, et al., 2014).

Besides the behavior as individual virtual humans, audience members also respond to each other's behavior. Although work has been done on crowd behavior (Thalmann & Musse, 2013) such as path planning and interaction between individuals of pedestrians, Kang, Brinkman, Van Riemsdijk, and Neerincx (2013) specifically had looked at the interaction behavior in an audience. According to their audience model, when an individual audience member is looking at an audience member in the neighborhood, the member in the in the neighborhood responds by looking back.

Among various public situations, public speaking is a common scenario occurring in everybody's life, e.g., delivering a business proposal, teaching in class, or giving a speech at a wedding. In public speaking situations, body language is a main channel of expression for audiences. Knowledge about this is therefore essential for developers to develop audiences that can be tailored for the need of users at run time. Currently, studies on the effects of virtual audiences often used three audience styles, described as positive, neutral and negative (e.g., Pertaub, Slater, & Barker, 2002; Taylor et al., 2010). Their results showed the benefit and potential of varying audience styles. However, no explicit and unified descriptions or guidelines could be found for designing such virtual audiences. Therefore, it is still a challenge for future studies that needs either similar or different audience styles.

Limited research has been devoted to audience behavior in public speaking scenario. Poeschl and Doering (2012) and Tudor, Poeschl, and Doering (2013) provided some guidelines for behavioral design of realistic virtual audiences. They observed the behavior of a typical audience in a lecture and explored the behavioral patterns such as frequency, duration, and postural sequence of certain behavior category, e.g. paying attention. Kang et al. (2013) proposed a parameterized audience model to generate expressive audience behavior for public speaking scenarios. The generated behavior was controlled by model parameters that defined the audience members' moods, attitudes, and personalities. They showed that the simulated audience using this model could behave expressively with regard to the audience attitude, and that the behavioral styles can be controlled by modifying the model parameters. Still, it is currently unclear about how an audience behaves underlying an audience style, e.g., a positive audience or a bored audience, and let alone which mood, attitude, or personality trait is associated with a specific audience style.

To simulate audiences for a variety of public speaking scenarios, more understanding about audience style and the relation with individual audience member characteristics is needed. These could be scenarios such as business people listening to an investment proposal pitch, employees assembled to hear the management announcement of potential layoffs, or students attending a Friday afternoon lecture who are eager to leave. Audiences in these settings clearly behave differently. To simulate these audiences, a key question is how people differentiate between these audiences. Regardless of the narrative or the way people are dressed, are people able to recognize different audience styles in a similar way people are able to recognize different facial expression independent of the context, such as anger or sadness? And what are these audience styles?

To address these issues, the work presented in this paper uses an existing virtual audience environment (Kang et al., 2013) to address four questions: (1) what variations in audience characteristics, in particular, mood, personality, and attitude, result in perceivable variations in audience behavior? (2) What combination of individual audience members' characteristics do people use to design prescribed audience styles? (3) What audience styles do people recognize and (4) what are the typical audience postures and behaviors associated with specific audience styles? To answer these questions the paper first describes a paired comparison perception experiment, which is a classic psychophysical method that was used to determine peoples' sensitivity in noticing a specific quantitative difference in an audience characteristic, e.g., higher or lower arousal. After identifying which audience characteristic resulted in noticeable audience behavior differences, people were invited to use these characteristics to design audiences for a set of public speaking scenarios. Clustering the audience scenarios based on the similarity of the characteristic settings resulted in five audience styles. Videos of virtual audiences were made for each style, and people were invited to match audience style description to each video. The last step of the study was to examine the parameterized audience model and identify specific audience postures and behaviors that were characteristic for the behavioral styles.

2. Virtual audience model and simulation

The work in this paper revolves around a parameterized audience model (Fig. 1) (Kang et al., 2013) that underlies an audience of virtual humans in a virtual environment. This is a probabilistic model abstracted from observation of real human audiences. Behaviors of real audiences were recorded when they were listening to presentations on a topic they were interested in, were critical about, found boring, and were neutral about. The audience corpus (Kang, 2013) consists of 9600 coding units with a sampling interval of two seconds, specifying head, gaze, arm, hand, torso, and leg positions. To obtain a parameterized model, additional data was also collected about the audience members' personality (extroversion, agreeableness, openness, neuroticism, and conscientiousness), attitude towards the topic (interest, approval, eagerness for information, criticism, and impatience), mood (valence, arousal,