Internet-delivered multi-patient virtual reality exposure therapy system for the treatment of anxiety disorders

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

a. Delft University of Technology, Mekelweg 4, 2628 CD Delft, the Netherlands

b. TNO Human Factors, Dutch Institute of Applied Science, Soesterberg, The Netherlands

{N.Kang, W.P.Brinkman, M.B.vanRiemsdijk}@tudelft.nl Mark.Neerincx@tno.nl

ABSTRACT

Motivation –The project is to reduce the therapist's workload in virtual reality exposure therapy (VRET) for anxiety disorders and explore cognitive ergonomic factors in the design of an internet-delivered multi-patient treatment system.

Research questions – The aim can be achieved by system improvement on both the patient side and the therapist side, which leads to two main research questions. First, how should a virtual environment be designed to reduce the therapist's workload and give a relative high presence to patients? Second, in what way should a usable internet-delivered VRET system be designed which can treat multiple patients simultaneously?

Research approach – A case study of a VRET treatment system for social phobia is carried out to investigate the design of an internet-delivered VRET treatment system and its impact on the main actors, e.g. the patient and the therapist.

Keywords

Virtual reality, phobia treatment, tele-treatment, autonomous avatars, cognitive workload, multi-patient

INTRODUCTION

Anxiety disorder among mental disorders, is found in one-third of individuals with any mental disorders (World Health Organization International Consortium in Psychiatric Epidemiology, 2000). Patients that seek professional help are often treated with exposure in vivo (i.e. exposure to actual real-life situations), which is the gold standard for treatment of anxiety disorders. Although effective, this treatment has a number of serious drawbacks such as limitation of control on the exposure and inconvenient access to some exposures.

One of the newly studied treatments for anxiety disorders is Virtual Reality Exposure Therapy (VRET), especially for the treatment of specific phobia (e.g. fear of flying (Brinkman *et al*, 2010), social phobia (ter Heijden and Brinkman, 2011), PTSD and Panic disorder). Recent meta-studies (Gregg & Tarrier, 2007; Parsons & Rizzo, 2008; Powers & Emmelkamp, 2008) show that exposure in VR is as effective as exposure in vivo. Additionally it can overcome practical limitations of these treatments and increase the efficiency of therapist resources. Importantly patients are much more willing to undergo VRET treatment and are less likely to refuse the treatment compared to the often dreaded exposure in vivo (Garcia-Palacios *et al*, 2007). Furthermore, VRET may provide a substantial improvement in the control of the treatment, a substantial cost reduction, and a growth in treatment accessibility by delivering the treatment over the Internet.

Though VRET has many advantages over the traditional treatment in terms of accessibility, patients' preference and so on, the cognitive workload of therapist might be very high because the therapist needs to not only monitor the patient's mental and physical state, but also control the details of virtual world to evoke a certain level of anxiety. Taking the treatment of fear of flying for example, therapist needs to control the flight experience of a patient by controlling the behavior of the plane (taxiing, taking off, weather conditions, pilot calls, and turbulence). Moreover, high workload prevents VRET from further improvement in accessibility by developing an internet-delivered multi-patient treatment system.

To reduce therapist's workload, our research focuses mainly on two aspects: control of virtual environment and design of therapist interface. Our research will be demonstrated for the treatment of social phobia, using the environment of public speaking specifically.

Control of virtual environment is a more critical problem in the setting of social phobia. In the scenario of public speaking, which most experimental systems focus on (Harris, Kemmerling & North, 2002; Herbelin, 2005; Slater *et al* 2006), patients' anxiety is affected by the avatars' behavior. These avatars could be graphical characters, film-based avatars (e.g. Klinger, Bouchard *et al*, 2005), or predefined animations by programming. However, the avatars' behavior usually cannot be controlled or can only be changed by using another prepared video or animation. Another drawback of this is that when a treatment session is longer than the prepared material, the avatar behavior will be simply repeated on and on, which lowers the presence. On the other hand, if the animations are controllable, the therapist will allocate more time to the control of avatars to interact with the patient and hence the workload of the therapist increases.

To solve this problem, we refer to artificial intelligence to create a crowd of autonomous avatars in the VRET system so that the avatars can respond naturally without the direct need of therapist intervention and therefor reduced the workload. Moreover, easy access to the settings of virtual environment should be provided to therapists to adjust exposure level.

This project envisions an internet-delivered VRET system for the treatment of anxiety disorders. Before the treatment, therapists can set the hierarchy of gradually more difficult scenes for a specific patient so that therapists get rid of the complicated and frequent manipulation during treatment sessions and hence therapists might be less occupied during the treatment. Therefore, it becomes possible for them to treat multiple patients simultaneously.

To enhance the system support even further and thereby reduce the workload of therapists, the system will also provide a virtual therapist. It may not only automatically monitor the anxiety level of patients' giving therapists an insight into patients' state, but also take over some therapist's task when the therapist is in a high workload situation. For example, when the system detects a patient too scared to continue with the session, it will alert the therapist first and if the therapist cannot handle it in time, the system will adjust automatically the exposure to a prior agreed level.

To realize the vision, three sub-questions are proposed:

1) How should artificial intelligence be applied to create autonomous avatars with natural behaviors?
2) What kind of control of virtual environment should be provided to adjust the exposure level?
3) How should the automated treatment system be designed to collaborate with the therapist to treat multiple patients simultaneously?

RELATED WORK

Before describing our research, some related research will be introduced briefly.

Remote therapy

The idea of remote therapy may increase treatment accessibility greatly. A SWOT analysis (Rizzo & Kim, 2005) and an explorative study (Paping *et al*, 2010) on remote VRET show that remote therapy is very promising especially when it is suggested as homework.

Crowd simulation

An agents-based application is suggested to create autonomous avatars when complex individual behavior is concerned (Thalmann & Musse, 1999). To generate human-like behavior, agent model may include factors of personality, emotion, role and so on (e.g. Pelechano & O'brien, 2005), and most work focuses on generation of pedestrian and evacuation behavior of a crowd.

Cognitive Task Load and Adaptive Automation

The three-dimention model of cognitive task load shows that certain regions may lead to negtive effects on task performance (Neerincx, 2003). To keep the human within a bandwidth of workload for optimum performance, adaptive automation is suggested. The research on the naval command center (de Greef & Arciszewski, 2008) can be one example. Moreover, an out-of-loop performance problem may occur when an operator is under-load. The problem which may leads to serious failures can be directly linked with a low level of situation awareness (Endsley, 1996). Therefore, adequate situation awareness should be ensured when an automated system is designed.

RESEARCH PLAN

A four-year research plan is described as follows.

Firstly, a multi-agent system is applied to create autonomous avatars in the scenario of public speaking. To make the avatars' behavior more realistic, their behavior will be based on observations collected from a real audience. The observation will be analyzed to model the agent for the creation of autonomous avatars. Finally, the naturalness of avatars using multi-agent system will be evaluated by comparing avatars using different control methods. This part will be explained in details in the next section.

Secondly, an easy control of setting the exposure levels will be provided. At the start of a session the system will make working agreements with therapists to find out what kind of control will be delegate to the system under what kind of circumstances. The effect of anxiety-evoking elements in VRET might also be identified and assessed to establish a hierarchy of social scenes of different exposure levels if the therapists prefer a simple setting.

Thirdly, an internet-delivered VRET system will be designed. The design of therapist console for remote treatment will be investigated to ensure an adequate level of situation awareness and in the meanwhile, to reduce therapists' workload when treating multiple patients simultaneously. Therapists will be examined in series of simulated scenes on their capability of remotely controlling simultaneously multiple VRET sessions.

CURRENT RESEARCH

Figure 1: classroom with a audience of avatars

We currently focus on the first research question, which is how to apply a multi-agent system to create a crowd of autonomous avatars. The intelligent agents will be applied to the classroom environment for public speaking, shown in figure 1. The research is divided into two parts, behavioral modeling and an experiment.

Behavioral Modeling

To evoke anxiety of different levels, the avatars in the audience are designed to be able to show four kinds of attitude, which are interested, neutral, bored and critical. Each avatar varies in personality and physical state. Thus, each agent is supposed to decide its own behavior according to the mental and physical state, as well as perception of environment.

To build up such a model, we firstly observed audience behavior in different conditions. We conducted several sessions in which an audience of eight people sitting together was given four different presentations in a certain order and their behaviors were video recorded. After each presentation, lasting around seven minutes, the audience was asked to fill in a questionnaire on their emotional state using self-assessment manikin (Bradley and Lang, 1994), physical state and their attitude towards the presentation. The whole process was repeated with another audience of eight people and the presentations were given in a reverted order to control for the potential order effects. The audience was required to complete the IPIP-NEO personality inventory afterwards. All members of the audience were recruited from fellow PhD students studying computer science. Also nobody knew the speaker beforehand. The four presentations were designed to evoke the four attitudes and the settings are described below.

- To evoke interested attitude, the audience were told at the beginning that they may win a small prize if they listen carefully and get a high score in the quiz afterwards. The topic was a novel invention of robot gripper.
- The neutral topic was a software design method and there was no additional requirement for audience.
- The boring topic was a section from a book on ethics by Aristotle, which the speaker read aloud. The content was difficult to understand and the order of paragraphs had been disarranged so that it was impossible for the audience to follow.
- The critical presentation criticized all the PhD students in the audience, saying that they did worst compared to PhD students in other departments and in computer science at other universities. Also a very strict working discipline was proposed, such as working time from 9:00am to 6:00pm, with only half an hour lunch time.

We developed a coding scheme to classify over time the audience behavior in the video recorded. Two kinds of categorization of individual behavior were used, one for low level of behavior details, the other for higher level of attitude expression.. For the low level, behaviors were divided into four categories, which separately present the postures of arms, head, trunk and legs. This level was used for low level programming and animation modeling.

For the higher level, behaviors were grouped into four categories according to avatar attitude, which were interested, neutral, bored and critical. Behaviors were also categorized in terms of physical state, which were tired and energetic. Thus, the audience behavior can be described in an emotional and physical way.

We are now working on the agent modeling. According to the coding scheme, rules will be abstracted and summarized so that a human-like behavioral model for agent can be established and applied to the avatars. We propose that audience attitude, audience personality, energy level and the environmental perception should be included in the agent model. Besides, patients' speech is considered to be used as a part of environment perception for agents by speech detection.

Experiment Design

The aim of the experiment is to evaluate if the virtual audience controlled by intelligent agents can provide a relatively high presence compared with those controlled by other methods. We are also interested to investigate if different avatar attitude are also perceived as such.. Presence in this case focuses on the naturalness of avatar behavior and the ability to express the avatars' emotional and physical state. The conceptual model is shown as figure 2.

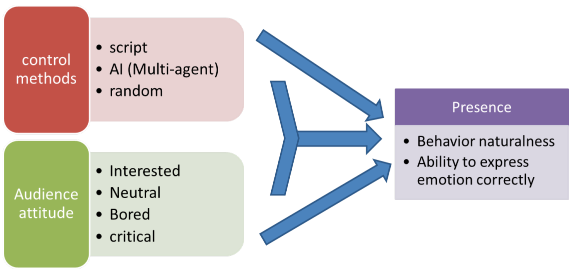

Figure 2: conceptual model of the experiment

In the experiment, three control methods of avatars will be compared. Besides the multi-agent system, one control method is called script method. Script control is to copy the audience's behaviors from the recorded video as direct input for the control of avatar behavior, and the events evoking certain behaviors are also copied into the virtual environment. Therefore, a sequence of behaviors or postures is defined beforehand for each avatar. This method is the most time-consuming one for a developer since the behavior sequence should be defined long enough to avoid repeats which would lower the presence experienced. The virtual environment using this method is regarded as the control condition for evaluating behavioral naturalness of the avatars. It is supposed to provide the highest level of presence because it is a copy from the real world.

The third method is random control where the avatars just pick randomly their behavior or posture regardless of the environment and the natural sequence of behaviors. The scenes using this method are expected to be at the lowest level of naturalness.

To carry out the experiment, animations will be programmed in 12 different conditions considering the combination of 3 control methods and 4 audience attitudes. For each condition 8 scenes will be created, each of which lasting 20 seconds. In the experiment, the scenes will be displayed in random order, and after each the participants will be asked to complete a questionnaire on behavior naturalness and perceived audience attitude. Additionally, participants will be asked to rate the presence of 12 scenes in different conditions, using the Igroup Presence Questionnaire (Schubert, Friedmann, & Regenbrecht, 2001).

Finally, the subjective data will be statistically analyzed to see whether and to what extent the intelligent agents help the avatars' behavior look natural. With this experiment we hope to obtain a better understanding how artificial intelligence can be applied to create autonomous avatars that exhibit natural behaviors, a first step towards the design of internet-delivered multi-patient VRET system.

REFERENCES

Brandley, M. M., and Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1), 49-59.

Brinkman, W.-P., van der Mast, C., Sandino, G., Gunawan, L.T., and Emmelkamp, P. (2010). The therapist user interface of a virtual reality exposure therapy system in the treatment of fear of flying. *Interacting with Computers*, 22(4), 299-310.

Endsley, M.R. (1996). Automation and situation awareness. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications*. Mahwah, NJ: Lawrence Erlbaum, 163-188

Garcia-Palacios, A., Botella, C., Hoffman, H., & Fabregat, S. (2007). Comparing acceptance and refusal rates of virtual reality exposure vs. in vivo exposure by patients with specific phobias. *Cyberpsychology & behavior*, 10(5), 722-724.

Greef, T. de, & Arciszewski, H. (2008). Combining adaptive automation and adaptive teams in a naval command centre. *Proceedings of the 15th European conference on Cognitive ergonomics the ergonomics of cool interaction - ECCE '08*.

Gregg, L., and Tarrier, N. (2007). Virtual reality in mental health : a review of the literature. *Soc Psychiatry Psychiatr Epidemiol*, 42(5), 343-354.

Harris, S. R., Kemmerling, R.L., and North, M.M. (2002). Brief virtual reality therapy for public speaking anxiety. *Cyberpsychology & behaviour*, 5(6), 543-550.

Herbelin, B. (2005). *Virtual reality exposure therapy for social phobia*. Doctoral dissertation, École polytechinque fédérale de Lausanne, Université Louis Pasteur, Strasbourg, France.

Klinger, E., Bouchard, S., Légeron P., Roy, S., Lauer, F., Chemin, I., and Nugues, P., (2005). Virtual reality therapy versus cognitive behavior therapy for social phobia: A preliminary controlled study. *Cyberpsychology & behavior*, 8(1), 76-87.

Neerincx, M.A. (2003). Cognitive task load design: model, methods and examples. In: E. Hollnagel (ed.), *Handbook of Cognitive Task Design*. Chapter 13. Mahwah, NJ: Lawrence Erlbaum Associates, 283-305.

Paping, C., Brinkman, W. P., & van der Mast, C. (2010). An explorative study into tele-delivered multi-patient virtual reality exposure therapy system. *Coping with Posttraumatic Stress Disorder in Returning Troops*, IOS press, 203-219.

Parsons, T. D., and Rizzo, A. A. (2008). Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: a meta-analysis. J *Behav Ther Exp Psychiatry*, 39(3), 250-261.

Pelechano, N., O'brien, K., Silverman, B., & Badler, N. (2005). Crowd simulation incorporating agent psychological models, roles and communication. *First International Workshop on Crowd Simulation*, 2, 21-30.

Powers, M. B., and Emmelkamp, P. M. (2008). Virtual reality exposure therapy for anxiety disorders: A meta-analysis. *J Anxiety Disord*, 22(3), 561-569.

Rizzo, A., & Kim, G. J. (2005). A SWOT analysis of the field of virtual reality rehabilitation and therapy. *Presence-Teleoperators and Virtual Environments*, 14(2), 119-146.

Scerbo, M. (1996). Theoretical perspectives on adaptive automation. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: theory and applications*, 37–63.

Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. *Presence-Teleoperators and Virtual Environments*, 10(3), 266-281.

Slater, M., Pertaub, D.-P., Barker, C., and Clark, D., (2006). An experimental study of fear of public speaking using a virtual environment. *Cyberpsychology and Behavior*, 9(5), 627-633.

ter Heijden, N., and Brinkman, W.-P. (2011). Design and evaluation of a VR exposure therapy system with automatic free speech interaction. *Journal of CyberTherapy and Rehabilitation*, 4(1), 41-55.

Thalmann, D., Musse, S. R., & Kallmann, M. (1999). Virtual Humans' Behaviour : Individuals, Groups , and Crowds. *Proceedings of Digital Media Futures*, 13–15.

World Health Organization International Consortium in Psychiatric Epidemiology (2000) Cross-national comparisons of the prevalences and correlates of mental disorders. *Bulletin of the World Health Organization*, 78, 413-426.