PAPER FOR VR FOR PUBLIC CONSUMPTION WORKSHOP

Virtual Reality and the Internal Experience

By

Alison McMahan and Rebecca Tortell

Abstract: Our research aims to develop a more transparent interface for virtual reality environments, an interface designed to read subconscious input from the user as well as player choices. Such an interface offers new possibilities for interactive fiction by enabling fictional "interior" experiences, something previously not thought possible with virtual reality. This paper is a report on a human subject study using heart rate and galvanic skin response to test the viability of adding biofeedback to the interface of a VR fiction experience for games. Using a test design based on video clips, we found that galvanic skin response measurement was suitably sensitive and that the system could learn over time.

1.Introduction

The overall goal of the prototype VRE we are working on, *Memesis*,[i] is to create a game that can be played in CAVEs or mini-CAVE installations in public places or at home on an expanded console box.

Our research aims to develop a more transparent interface for virtual reality environments, an interface designed to read subconscious input from the user as well as player choices. Using a test design based on video clips, we found that galvanic skin response measurement was suitably predictive of intensity felt by each subject. We also established that as the number of subjects increases, the accuracy of the regression will improve, making the system more accurate over time.

We are designing and producing an interactive virtual narrative that will make full use of the biofeedback interface. It is a short interactive experience for a 3D immersive environment using interface elements typical of such environments. The system collects biofeedback information from users as they experience the environment, and draws on that data to determine the outcome of the "game." There are three elements to this task: a new approach to interface design; a new way of thinking about and designing the relationship between the user and the interactive environment, i.e., how the user experiences a sense of presence; and an interactive narrative that takes full advantage of the other two design elements.

Our environment features a horror narrative that exposes users to a series of stimuli designed to evoke fear and anxiety. Initial interactive, narrative experiences on the first level of the game are designed to induce scares based on different types of horror experiences. Biofeedback data will be collected from users in the entry level and then used to direct him or her to one or maximum two linked, interactive, narrative experiences that (at least in theory) represent each user's greatest psychological fear.

2. The ultimate transparent interface

The goal of the project is to design a VR interactive fiction or game. The user will perceive a classic VR interface, and barely perceive the biofeedback interface. The classic interface is similar to that of a first person shooter game, with first person

[i] For more information about *Memesis*, including movies, see www.alisonmcmahan.com

perspective, a wand instead of a weapon to represent the player in the game, and an artificially intelligent environment that can respond and keep track of the player's movements. This traditional interface will be just enough to keep the user's mind engaged at the conscious level. The game will also incorporate the biofeedback data from the user's galvanic skin response and the environment will respond accordingly. The hypothesis is that as long as the response to the biofeedback data is well designed, the user should have a richer virtual experience with a highly increased sense of presence.

The preliminary database for the biofeedback interface was collected in a human subject study. Sixteen subjects over the age of 18 viewed five stimulating video clips and two non-stimulating control clips, as their skin conductance was recorded from electrodes placed on their palms and heart rate was recorded from electrodes on the wrists. The clips were chosen from recent Hollywood films made with a similar level of production value. They contained realistic depictions of sequences that might trigger a phobic reaction (sequences with a character teetering on the edge of a cliff or put at risk in an enclosed space, for example) or anxiety reaction (sequence depicting preparations for eye surgery). Some clips depicted events that could arouse psychological fears, such as the fear of humiliation or the fear of abandonment.

3. Results of Human Subject Study

The purpose of this experiment was to test how well the subjective arousal of stimuli could be measured by galvanic skin response (GSR) and heart rate. If successful, biofeedback could be used as a navigation interface for *Memesis*. We hypothesized that the GSR and heart rate data combined would provide the best indication of each subject's arousal while watching clips from Hollywood films. Films in this case were presumed functionally equivalent to a VR experience. We produced four different clip-orders, to which each subject was randomly assigned, in order to eliminate ordering effects. Subjects first filled out a short questionnaire, consisting of basic personal questions as well as ascertaining whether subjects had any pre-existing phobias. Each subject then viewed a baseline clip, followed by six experimental clips. Subjects were instructed after watching each clip to fill out a short questionnaire on which they rated its intensity (from 1 to 5) and provided a verbal description of their emotional state. The purpose of the introductory clip was to obtain a baseline against which to compare the arousal data gathered during the experiment.

For analysis, the data set was arranged to show, for each subject, average BPM (heart rate), minimum, mean, and maximum GSR level, number of GSR 'peaks', and the number of peaks adjusted by the length of each clip. A 'peak' was defined as a spike of GSR activity with a maximum point above 0.045V (For the purpose of comparison, baseline activity for each subject usually did not stray from between 0.040 and 0.041V). Using Friedman repeated-measures ANOVA on ranks (owing to a non-normal distribution of effects) for each of these factors, it was concluded that data from each movie clip differed significantly from that of each other ($P<0.05$), and had not occurred by random chance. Clip number five was the only one in which the data for each GSR-related factor differed significantly from the data of the baseline clip.

ANOVA, analysis of variance, compares the differences between sample means to examine whether the differences between them are statistically significant (i.e. have a less than 5% probability of occurring by random chance). The means in this case are the average data of all subjects (mean average BPM, mean adjusted number of peaks, etc.).

Regression analysis was used to determine what combination of biofeedback data was best able to predict the subject's intensity ratings of each clip. Linear regression procedures use several independent variables (in this case, each biofeedback data factor) to predict the value of a dependent variable (intensity). Linear regression attempts to fit a single line through several data points in order to make predictions, and describe how well the data points fit that line. Stepwise (backward) regression was used to determine the best predictor of the self-rated intensity from each clip (baseline intensity was assumed to be 1; data was not collected for this clip).

Backward stepwise regression begins by including all independent variables, and removing them one at a time, iteratively, to determine which produces the best prediction of the dependent variable. The intensity data was centered so as to have constant variance - mean intensity (2.6741) was subtracted from each value. With $P<0.001$, the adjusted GSR peaks variable was found to be the most predictive of intensity. The power of the test (alpha = 0.05) was 0.940.

Since the sample did not fall along a normal distribution, Spearman rank order correlation was used to measure the strength of correlation between the adjusted number of GSR peaks and the self-rated intensity. This is a non-parametric or non-directional test, which means that it makes no claim about which variable influences the other, and does not require normality or constant variance. The analysis showed a correlation coefficient of 0.281 between intensity and adjusted number of peaks ($P=0.003$). Thus, those variables tend to increase together 28.1% of the time.

One of the assumptions of regression procedures is a distribution of data point that approximates the normal distribution - a symmetric, bell-shaped curve. Our data had too many outlying data points to approximate this curve, and so failed the normality test. This is in large part due to a sample size ($n=16$) that may not have accurately reflected the full range of responses. Thus, while the regression model told us where to look for the best predictor, that model on its own provides only an imperfect approximation, and a non-parametric test (which does not require normality) was required. The regression test did point out where to look for the strongest correlation.

In this case, our two variables for correlation are adjusted number of GSR peaks and self-rated intensity. The test we performed could state that the two tended to increase together 28% of the time, but could not state whether a change in peaks would cause a change in intensity, or whether a change in intensity would cause a change in peaks. This is, however, only statistically speaking - we know factually that the intensity rating couldn't possibly have affected the number of peaks, since one occurred temporally after the other. The number of outlying data points means that the correlation (another procedure which involves drawing a line through data points) wasn't able to gather the minute differences required to determine which variable influenced which.

Most of the outlying data points originated from subject's responses to clip number five. As mentioned above, while all clips differed significantly from each other, only clip five differed from baseline arousal, indicating that the responses to that clip were much more dramatic than to any other. This may have been due to several factors - the narrative structure of the clip, the sound design, or dramatic visuals (further research forthcoming). Such a dramatic response is in itself very positive for this study, but in conjunction with the other clips, it generated problems for analysis.

We were surprised that heart rate was not included as a significant contributor to the variance of intensity - this may have been due to methodological difficulties. We suspect that, were maximum and minimum heart rate also included in the regression analysis, at least one of these would make a significant contribution.

The data from the human subject study will be incorporated into the initial database for the virtual game, so that the system's AI will know how to interpret a reading from a player. The AI will also continue to add player responses to its database, so that the more people play the game, the better the system will get at determining what the player is the most afraid of and rewarding them with the corresponding scenario.

4. Designing Narrative For Virtual Feedback

The hypothesis we wish to test in our environment is that by enabling the user to provide the computer with input at a subconscious level, the resulting interactive experience will be profoundly more effective at the moral and psychic level. This is why, of all the storytelling genres that would lend themselves to this prototype, the horror genre was chosen: because moral and psychic effect is precisely what the horror genre achieves best. Moral or emotional horror can be more effective in the VR environment because the altered status of the narratee (the receiver of the narration)

implicates the user morally and emotionally into the situation. (Future games are planned using other genres, such as fairy tales and romantic comedy.) The first step towards creating a modular horror experience was to analyze the horror genre itself, as it has manifested so far in film, television and prose fiction. Once an initial survey of the literature was complete, typical horror experiences were precisely defined and broken down into units or modules of experience that the game can recombine according to player input.

The biofeedback interface enables a shift in narrative strategies for interactive fiction. David Deutsch has pointed out the difference between "external experiences", an experience of something outside of one's own mind, and "internal experiences", an experience of something within one's own mind. He concludes that VR is suited for the former but not the latter, because though a VR can indirectly lead to an internal experience this cannot be replicated (crashing in a flight simulator is never as much fun the second time around) or even assured, because it would involve "replacing the user by a different person." (Deutsch, p. 104). We believe that the implementation of biofeedback technologies of various types could make the stimulation of internal experiences the norm for VR in the future. This requires a different approach to interactive narrative design, an approach that we are currently testing with *Memesis*..

5. Research Context

This project was originally inspired by projects that apply VR in the treatment of phobias, such as Larry Hodges' designs for virtual environments to treat patients with fear of heights and fear of spiders at the university of North Carolina, Chris Shaw's virtual reality environment for training fireman at Georgia Tech, and various others. The design of the game is partly inspired by the design of these virtual environments for clinical treatment and training. The difference is that while clinicians like Larry Hodges wish to help their patients overcome their phobias by exposing them to spiders or heights in a virtual environment, we wish to titillate our users with "shock horror" experiences based on their phobic reactions to events and situations in our environment. For the biofeedback interface to work it is critical that the game environment produce a heightened sense of presence in the player.

The term "presence" is a rather slippery one and is often used interchangeably with equally slippery terms such as "immersion" and "engagement." To elaborate a coherent set of aesthetic criteria that we could work from, we relied on previous research on presence by McMahan and the survey of research on presence by Lombard and Theresa Ditton.

Our initial search for an expanded interface was inspired by Bolter and Grusin's definition of immediacy and their concept of a transparent interface. [ii]

Many studies have measured normal and pathological reactions to emotional stimuli such as anxiety inducing situations (Litt 1999, Pury 1997, Tillfors et al 2001), provocative visual stimuli such as photographs (Paradiso et al 1999, Sheline et al. 2001, Thomas et al, 2001), computer games (MacDowell and Mandler 1989), and memories of previous emotionally-laden experiences (Dougherty et al 1999, Rosenberg et al. 2001). Most of these studies aimed to show a correlation between heightened emotional arousal and physiological effect generally, as reflected in MRIs or SCRs, and the methods used to evoke an arousal response were of like kind. Some studies have indicated a consistent concordance between response systems (Calvo, Miguel-Toabal, 1998, Alvarado 1997). The purpose of the current study is to ascertain if SCR or EKG measurements can be used effectively to determine, not

---

[ii] Jay David Bolter and Richard Grusin use *immediacy* to define a similar concept, in their book *Remediation* (Cambridge, MA: MIT Press, [1999] 2000):

Immediacy (or *transparent immediacy*): A style of visual representation whose goal is to make the viewer forget the presence of the medium (canvas, photographic film, cinema, and so on) and believe that he is in the presence of the object of representation. One of the two strategies of remediation; its opposite is *hypermediacy*, "A style of visual representation whose goal is to remind the viewer of the medium. One of the two strategies of remediation" (272-73).

whether certain stimuli evoke an arousal response at all, but which of various stimuli evokes the highest response in any given (normal) individual.

The design of the virtual reality environment was influenced by the work of Clive Fencott and Dave Pape and Josephine Anstey.[iii]

6. Conclusion

Our project started as an experiment in transparent interface design, but its biofeedback interface has led us to reconsider how interactive narrative works in virtual environments in general and with a biofeedback interface in particular. Our solution was to create a more episodic, or modularized narrative that emphasizes interactive experiences in more narrative based (and therefore more passive) experiences at the final levels. However the choice of second level is determined by the user's emotional state, and may aid users in gaining self-knowledge about their fears and phobias and lead to an internal experience.

REFERENCES

Alvarado, N. (1997). Arousal and Valence in the Direct Scaling of Emotional Response to Film Clips. Motivation and Emotions, Vol. 21, No. 4, 323-348.

Bechara, A., Damasio, H., Damasio, A.R. (2000). Emotion, decision making and the orbitofrontal cortex. Cerebral Cortex, 10(3): 295-307.

Calvo, M.G., Miguel-Tobal, J.J, (1998). The Anxiety Response: Concordance Among Components. Motivation and Emotion, Vol 22, No.3. 211-230.

Carlin, A.S., Hoffman, H. and Weghorst, S. (1997). Virtual Reality and Tactile Augmentation in the Treatment of Spider Phobia: A Case Study. *Behavior Research and Therapy*, *35*(2), 153-158.

Clader, J.J., Lawrence, A.D. & Young, A.W. (2001). Neuropsychology of fear and loathing. Nature Reviews Neuroscience 2(5) May 352-363.

Damasio A.R., Grabowksi, T.J., Bechara A., Damasio H. Ponto L.L.B., Parvizi, J. Hichwa, R.D. (2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. Nature Neuroscience 3 (10) October 1049-1056.

Deutsch, David, *The Fabric of Reality*, NY: Penguin Books, 1997.

Dougherty, D.D., Shin, L.M., Alpert, N.M., Pitman R.K., Orr, S.P., Lasko, M., Macklin, M.L, Fischman, A,J., & Rauch, S.L. (1999). Anger in healthy men: A PET study using script-driven imagery. Biological Psychiatry, 46(4) August 466-472.

Emmelkamp, P.M.G, Krijn, M. Hulsbosch, , L. de Vries, , S. , Schuemie , M.J. & van der Mast, C.A.P.G. (2002). Virtual Reality Treatment in Acrophobia: A Comparison with Exposure in Vivo. Behaviour Research & Therapy, May, Vol 40(5), 25-32

Garcia-Palacios, A., Hoffman, H.G., Kwong See, S., Tsai, A. and Botella-Arbona, C. (2001). Redefining therapeutic success with VR exposure therapy. *CyberPsychology and Behavior*, *4*(3), 341-348.

Hatada, T., Sakata, H., & Kusaka, H. (1980). Psychophysical analysis of the 'sensation of reality' induced by a visual wide-field display. SMPTE Journal, 89, 560-569.

Hodges, L., Rothbaum, B., Kooper, R, Opdyke, D., Meyer, T., North, M., de Graff, J. and Williford, J. Virtual environments for treating fear of heights. IEEE Computer 28, 7, (Jul. 1995),. 27-34.

Kesler, M.L., Anderson, A.H., Smith, C.D., Avison, M.J., Davis, C.E., Kryscio, R.J., & Blonder, L.X. (2001). Neural substrates of facial emotiona processing using fMRI. Congitive Brain Research 11(2) April 213-226.

Litt, M.D., Kalinowski, L., & Shafer, D. (1999) A Dental Fears Typology of Oral Surgery Patients: Matching Patients to Anxiety Interventions, Health Psychology, vol. 18(6), November, 614-624.

Lombard, M. (2000). Measuring presence: a literature-based approach to the development of a standardized paper-and-pencil instrument. Project Abstract Submitted for Presentation at Presence 2000: The Third International Workshop on Presence. http:www.presence-research.org/presence2000.html.

Lombard, M. (1995). Direct responses to people on the screen: Television and personal space. Communication Research, 22(3), 288-324.

Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. http://astro.temple.edu/~lombard/research.htm

[iii] www.resumbrae.com

Macdowell, K.A., Mandler, G. (1989). Constructions of Emotion: Discrepancy, Arousal, and Mood. Motivation and Emotion, Vol. 13, No.2, 105-124.

McMahan, Alison, (2003). Immersion, Engagement, and Presence: A Method for Analyzing 3-D Video Games. Routledge, pp. 67-86.

Mühlberger, A., Herrmann, M.J., Wiedemann, G., Ellgring, H., & Pauli, P. (2001). Repeated exposure of flight phobics to flights in virtual reality. Behaviour Research and Therapy 39 1033-1050.

Nijula, R. (1991). Psychological correlates of nonspecific skin conductance reponses. Pscyhophysiology, 28(1), 86-90.

North, M., North, S. and Coble, J. Effectiveness of virtual environment desensitization in the treatment of agoraphobia. Presence: Teleoperators and Virtual Environments 5, 3, (March 1996), 346-352

North, M.., North, S., and Coble J. Virtual Reality Therapy, An Innovative Paradigm. IPI Press, 1996.

Paradiso, s., Johnson, D., Andreasen, N.C., O'Leary, D.S. Watkins, G.L., Ponto, L.L.B., & Hichwa, R.D. (1999), Cerebral blood flow changes associated with attribution of emotional valence to pleasant, unpleasant, and neutral visual stimuli in a PET study of normal subjects. American Journal of Psychiatry 156 (10) October 1618-1629.

Prothero, J. D., Parker, D. E., Furness III, T. A., & Wells, M. J. (1995a). Towards a robust, quantitative measure for presence. In Proceedings of the Conference on Experimental Analysis and Measurement of Situation Awareness, 359-366. Available: http://www.hitl.washington.edu

Pury, C. L.S. & Mineka, S. (1997). Covariation bias for blood-injury stimuli and aversive outcomes. Behaviour Research and Therapy, Vol 35, Issue 1 January, 35-47.

Rothbaum, B., Hodges, L., Watson, B. Kassler, G., and Opdyke, D. Virtual reality exposure therapy in the treatment of fear of flying: A case report. Beh. Res. Ther. 34, 5-6 (May/Jun. 1996), 477-481.

M.J. Schuemie, C.A.P.G. van der Mast, M. Krijn, P.M.G. Emmelkamp (2002). Exploratory Design and Evaluation of a User Interface for Virtual Reality Exposure Therapy in: J.D. Westwood, H.M. Hoffman, R.A. Robb, D. Stredney (Eds.), Medicine Meets Virtual Reality 02/10, IOS Press, pp.468-474

M.J.Schuemie, C.A.P.G.van der Mast (2000). Design of Virtual Reality Exposure Therapy Systems - Task Analysis.Proceedings of the Revalidatie Anno 2000 Conference, Nederlands Instituut voor Psychologen, 17 en 18 november.

Shaw, Chris, http://www.gvu.gatech.edu/virtual/.

Slater, M., & Usoh, M. (1993). Representations systems, perceptual position, and presence in immersive virtual environments. Presence, 2(3), 221-233.

Strickland, D., Hodges, L. North, M. & Weghorst, S. (1997). Overcoming phobias by virtual exposure. Communications of the ACM 40, 8, August, 34-39.

Williams, L.M., Phillips, M.L., Brammer, M.J., Skerrett, D., Lagopoulos, J. , Rennie, C., Bahramali, H., Olivieri, G., David, A.S., Peduto, A., & Gordon, E. (2001). Arousal dissociated Amygdala and Hippocampal fear responses: Evidence from simultaneous fMRI and skin conductance recording. Neuroimage 14, 1070-1079.