Emotional personification of humanoids in Immersive Virtual Environments

Vinoba Vinayagamoorthy
Supervisors: *Mel Slater, Anthony Steed*
Department of Computer Science
University College London
Gower Street, London WC1E 6BT, UK

ABSTRACT
More physical realism demands an increase in behaviour realism in order to increase the levels of co-presence, realism and believability in virtual environment users.

This concept builds on the work done by others in the *Virtual Environment and Computer Graphics* group at University College London. In most cases a hypothesis is tackled by research into a particular behaviour. The behaviour in this research is *posture* and the control variables are various human *emotions* and *moods.*

There is an added factor that has to be taken into consideration. The purpose of this study is not only to examine the relationship between posture and emotions but also to deliver a virtual humanoid to incorporate into other project within *Equator*. Therefore there is an added importance in making the virtual humanoid completely generalised and manageable by a range of programs across various platforms.

Keywords
Virtual Reality, immersive virtual environments, avatars, photo-realism, behavioural realism, social presence, copresence, non-verbal behaviour, emotions and posture.

INTRODUCTION
Most virtual humanoids currently in use in virtual worlds have been designed to look human but do not include a convincing behaviour model within their structure. When users interact with these virtual humanoids, they lack the enrichment that non-verbal communication provides whilst communicating with each other in the real world [27].

Hypothesis
The comprehensive hypothesis behind this research in a nutshell is that a virtual humanoid endowed with realistic behaviour models can evoke better levels of presence, co-presence and believability in a user than an humanoid with no innate behaviour. In addition, it is postulated that an exaggerated behaviour model incorporated into a minimally rendered avatar will evoke a satisfactory state of realism hence compensating for the lack of photo-realism

RELATED RESEARCH IN VIRTUAL HUMANS
Psychological studies claim that the main channels for communication of the emotional state of mind are through the face, body and the tone of voice in that order. The face has evolved as a social signalling route while gestures and postures together with body movements are a secondary channel for emotional conveyance [9]. It has also been suggested and confirmed that the face conveys specific emotions while the body conveys the degree of intensity of the concerned emotion [12]. However more recently, research suggests that the body can display more specific emotional content than was previously thought [3]. Research is being carried out in embodying virtual humanoids and agents with behaviour models particularly controlled by emotions [5][8][13][15]. In terms of behavioural realism, and specifically eye gaze, Garau et al., [10] investigated the impact of avatar gaze on participants' perception of communication quality by comparing a random-gaze and inferred-gaze avatar. Lee et al. [16] presented a similar experiment comparing random, static and inferred eye animations. In terms of agent architecture, there have been a few good models [5][7] in the OZ project, [13] in USC, Improv [22], InViWo [23], Cathexis [26], BodyChat [28], Rea [8], and SimHuman [29].

NONVERBAL BEHAVIOUR: POSTURE
There are many models partitioning the different types of behaviour. The most widely suggests that up to 93% of communication is non-verbal (55% consists of body language, 38% is expressed through tone of voice and only 7% through words) [18]. Posture is just one of the channels used in non-verbal communication [1][4]. There are about a thousand stable human postures according to the anthropologist [14]. Incorporating all these postures into a computational model controlled by emotions is demanding. More on postures is available in [1][3][24].

Non-verbal signals can also be categorised by how far people are consciously aware of sending and receiving communication signals [3]. In the case of posture the sender is fully *unaware* of the conveyance of a non-verbal signal whereas the receiver is fully *aware* of the signal and at times trained in the interpretation of the signal.

The complexity of variations between postures is not known at this stage in the research hence it has been decided to concentrate on standing poses initially and move to other stances if time permits.

Each of the posture then has further variations corresponding to the different positions of the limbs and lean of the body.

- Lean *Forwards, Backwards, Sideways*
- Arms *Open, Closed, On hips*
- Head *Lowered, Raised, Tilted sideways*
- Legs *Stretched, Open, Crossed*

The posture portrayed by the body is thought to convey the degree of intensity of the emotions expressed by the face [3]. A forward lean of the body and head may be intended to convey a number of various emotions:

- Interest
- Pity
- Anger
- Curiosity
- Frustration

So is posture alone sufficient to decode the right message being sent by the owner or is posture only useful if accompanying a basic facial expression?

In the final virtual humanoid, all non-verbal messages (emotion or definition), being sent using a body pose will result from a combination of the aforementioned variables in varying intensities. In order to perceive the different effectual postures used by human beings in everyday life, it is necessary to decipher the messages that are sent using body posture as a major means of non-verbal communication.

Since this type of research is not widespread in the computer graphics community, it can only be obtained form psychological archives and case studies.

EMOTIONS

The second area relevant to this research is emotions. Expressing emotions is a form of body communication. The *emotion* of a person is the cumulative result of recent events and knowledge. Emotions are fluctuated and related to events. A prolonged, conscious emotional state leads to a *mood*. Since this research is related to the expression of emotions and not the generation/detection of emotions, it will not be dealing with the detection of emotions in goal-directed events. The virtual humanoid developed will be used to express different types of emotional states using body postural animation. There is some work into the subject area of using emotions as variables in animation: [2][6][13][15].

Models of Emotions

There are six different basic facial expressions defined in [9]. These are joy, surprise, disgust, fear, anger, and sadness. Yet in real life there are more than six basic emotions used in communication. There are some models categorising emotions. The OCC[1] model, categorises emotions into 22 groups [21]. It was used famously in the OZ project [6]. All these emotions in the OCC model can be matched up to one of the Ekman basic expression as shown below.

Joy	*Happy-for, Gloating, Joy, Pride, Admiration, Love, Hope, Satisfaction, Relief, Gratification, and Gratitude*
Sadness	*Resentment, Pity, Distress, Shame, Disappointment, and Remorse*
Anger	*Anger, Reproach, and Hate*
Fear	*Fear, and Fear-confirmed*
Surprise	
Disgust	

The concepts of surprise and disgust are not covered in the OCC model. There is a third expression, which may be categorised, has an emotional state: *interest*. The OCC model or Ekman has not covered *interest*. Is it presumptuous to assume that all the compound emotions described in the OCC model can be re-enacted using the basic emotions in different intensities? Such work has been researched into already to some extent [26].

The Concept of Moods

Eleven of the OCC emotions are positive and the remaining emotions are negative. Positive emotions lead to a good mood while negative emotions lead to a bad mood. Prolonged *joy* leads to a good mood whereas prolonged *anger*, *fear*, and *sadness* lead to a bad mood. Disgust, interest and surprise may lead to either mood or remain in a neutral mood.

A third mood (the neutral) is relevant in the ongoing research to encompass the control group. In a neutral mood the person will tend to hold back on displaying emotions readily and if expressed, it is done so with less intensity. Whereas in a bad mood they might react with more intensity to the emotion of *anger* than they would if they were in a good mood prior to the input of the stated emotion.

THE PLAN

Work has already been done into substantiating two hypotheses:

[1] A. Ortony, G. Clore and A. Collins

- Emotions can be conveyed through a shared immersive virtual environment [19].
- An inferred-gaze and random-gaze model were used to explore the combined impact on quality of communication of eye gaze model and visual appearance [11][16].

The results of the gaze experiments concluded that an increase in behaviour realism has to accompany an increase in photo-realism. This substantiates the arguments in both papers [11][25].

On the technical side, H-Anim [30] compliant avatars have been designed and built in order to represent a virtual human in an immersive setting.

Figure 1: Lower-realism avatar, higher-realism male avatar, higher-realism female avatar

A plugin was used to animate the humanoid's body in order to maintain a visually consistent virtual humanoid. This included inferring the position of the right elbow using a geometrical inverse kinematics when the user's tracked hand moved, and deducing the position of the avatar's knees when the user bent down. There were also some deductions involved in the rotation of the head and body. The body was not rotated to the same direction as the head unless there was significant translation (greater than 75 centimetres) associated with the user. This was to enable the user to nod, tilt and shake their head in the VE whilst in conversation.

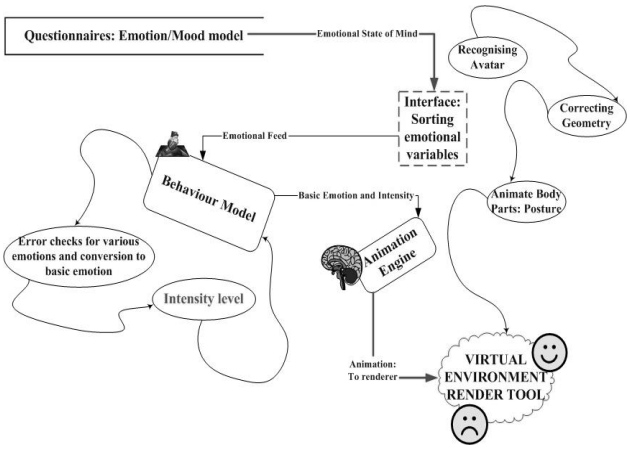

Figure 2: Envisaged architecture for humanoid animation

In the future, results got from psychological research will be used to build a posture-building functionality into the avatar's plugin thereby enabling the user to define the behaviour of the avatar in terms of emotional states. The resultant virtual humanoid should be able to take a number of input variables corresponding to the supposed emotional state of mind of the virtual humanoid and feed it to the behaviour model. The behaviour model then converts the compound emotion (OCC model) to a simpler set of emotions and conveys it to the animation engine. The animation layer will then produce the appropriate standing, sitting, squatting, kneeling or lying poses appropriate to the given emotional mood with the aid of either a script (autonomous agent) or the position of the user (avatar).

CONCLUSIONS

The most recent study on eye-gaze sought to investigate the impact of visual and behavioural realism in avatars on perceived quality of communication between participants meeting in a shared IVE. It was conclude that independent of head tracking; inferred eye animations can have a significant positive effect on participants' responses to an immersive interaction. The caveat is that they must have a certain degree of visual realism, since the lower-realism avatar did not appear to benefit from the inferred gaze model. This finding has implications for inexpensive ways of improving humanoid expressiveness using information readily available in the audio stream. The study also opens a path to explore in terms of other behaviours viz. posture.

ACKNOWLEDGMENTS

This research is supported by EQUATOR. I would also like to thank everyone, who has been conducive to this research so far, Celine Loscos (for *Pureform*), Andrea Brogni, Marco Gillies, David Swapp, Pip Lee Bull and fellow PhD students. I would also like to thank our collaborators at the University of North Carolina, Chapel Hill for their help in the *Internet2* project. This gratitude is extended to Alican Met for giving me the opportunity to take a glimpse into the world of virtual paranoia. Finally, I would like to acknowledge the volunteer work put in by the subjects during the running of the experiments.

REFERENCES

[1] Allbeck M. J. and Badler I. N., "Consistent Communication with Control", *Workshop on Non-Verbal and Verbal Communicative Acts to Achieve Contextual Embodied Agents at Autonomous Agents*, 2001

[2] Amaya K., Bruderlin A., and Calvert T., "Emotion from Motion", *Graphics Interface*, June 1996, pp. 222-229

[3] Argyle M., "Bodily Communication", Second Edition, 1998

[4] Badler I. N. and Allbeck M. J., "Towards Behavioural Consistency in Animated Agents", *Proceedings of Deformable Avatars*, 2000

[5] Bates J., "The Role of Emotion in Believeable Agents", *Communications of the ACM*, 37, July 1994, pp. 122-125

[6] Bates J., Loyall B. A., and Reilly S. W., "An Architecture for Action, Emotion and Social Behaviour", 1992,CMU-CS-92-144

[7] Bates J., Loyall B. A., and Reilly S. W., "Integrating Reactivity, Goals, and Emotion in a Broad Agent", 1992

[8] Cassell J., Bickmore W. T., Billinghurst M., Campbell L., Chang K., Vilhjálmsson H. H., and Yan H., "Embodiment in Conversational Interfaces: Rea", *SIGCHI*, 1999, pp. 520-527

[9] Ekman P., "Emotion in the Human Face", 1982

[10] Garau M., Slater M., Bee S., and Sasse M. A., "The impact of eye gaze on communication using humanoid avatars", *Conference on Human Factors and Computing Systems*, (Proceedings of the SIGCHI conference on Human factors in computing systems), March 2001, pp. 309-316

[11] Garau M., Slater M., Vinayagamoorthy V., Steed A., and Brogni A., "The Impact of Avatar Realism on the Perceived Quality of Communication in a Shared Immersive Virtual Environment", 2002

[12] Graham J. A., Ricci Bitti P., and Argyle M., "A cross-cultural study of the Communication of Emotion by Facial and Gestural cues", *Journal of Human Movement Studies*, 1975, pp. 178-182

[13] Gratch J. and Marsella S., "Tears and Fears: Modeling emotions and emotional behaviors in synthetic agents", *International Conference on Autonomous Agents*, 5, 2001

[14] Hewes G., "The Anthropology of Posture", *Scientific American*, 196, 1957, pp. 123-132

[15] Kshirsagar S. and Magnenat-Thalmann N., "A Multilayer Personality Model", 2002

[16] Lee S. P., Badler I. N., and Badler J., "Eyes Alive", *SIGGRAPH*, July 2002

[17] Loyall B. A. and Bates J., "Real-time control of Animated Broad Agents", 1993

[18] Mehrabian A., "Silent Messages", First Edition, Wandsworth, Belmont, California, 1971

[19] Mortensen J., Vinayagamoorthy V., Slater M., Steed A., Whitton M. C., and Lok B., "Collaboraton in Tele-Immersive Environments", *Eight Eurographics Workshop on Virtual Environments*, May 2002, pp. 93-101

[20] Niu X. and Vassileva J., "Agent Persona and Modelling Emotions", 26-2-2002

[21] Ortony A., Clore G., and Collins A., "Cognitive Structure of Emotions", 1988

[22] Perlin K. and Goldberg A., "Improv: A System for Scripting Interactive Actors in Virtual Worlds", *SIGGRAPH*, 1996

[23] Richard N., Codognet P., and Grumbach A., "The InViWo toolkit: Describing autonomous virtual agents and avatars", *Intelligent Virtual Agents*, 2001

[24] Rosenberg G. B. and Langer J., "A Study of Postural-Gestural communication", *Journal of Personality and Social Psychology*, 2, 1965, pp. 593-597

[25] Slater M. and Steed A., Chapter 9: Meeting people virtually: experiments in shared virtual environments, In "*The social life of avatars: presence and interaction in shared virtual environments*", New York, 2002

[26] Velásquez J. D., "Modeling Emotions and Other Motivations in Synthetic Agents", *American Association for Artificial Intelligence*, 1997

[27] Vilhjálmsson H. H., "Autonomous Communicative Behaviours in Avatars", 8-10-1996

[28] Vilhjálmsson H. H. and Cassell J., "BodyChat: Autonomous Communicative Behaviors in Avatars", 9-5-1998

[29] Vosinakis S. and Panayiotopoulos T., "SimHuman: A platform for real-time virtual agents with planning capabilities", *Intelligent Virtual Agents*, 2001

[30] Web3D Consortium, Humanoid Animation Working Group of the WEB3D Consortium, http://www.hanim.org/, Web