E-Tree: Emotionally Driven Augmented Reality Art

Stephen W. Gilroy[1], Marc Cavazza[1], Rémi Chaignon[1],
Satu-Marja Mäkelä[2], Markus Niranen[2], Elisabeth André[3], Thurid Vogt[3], Jérôme Urbain[4],
Mark Billinghurst[5], Hartmut Seichter[5] and Maurice Benayoun[6]

[1]School of Computing, University of Teesside, Middlesbrough TS1 3BA, UK;
s.w.gilroy@tees.ac.uk, m.o.cavazza@tees.ac.uk, r.chaignon@tees.ac.uk

[2]VTT Electronics, Finland, satu-marja.makela@vtt.fi, markus.niiranen@vtt.fi

[3] University of Augsburg, Germany, andre@informatik.uni-augsburg.de, thurid.vogt@informatik.uni-augsburg.de

[4] Faculté Polytechnique de Mons, Department of Electrical Engineering, Belgium; jerome.urbain@fpms.ac.be

[5] HITLabNZ, New Zealand; mark.billinghurst@hitlabnz.org, hartmut.seichter@hitlabnz.org

[6]Citu, Université Paris 1 Panthéon-Sorbonne, mb@benayoun.com

ABSTRACT

In this paper, we describe an Augmented Reality Art installation, which reacts to user behaviour using Multimodal analysis of affective signals. The installation features a virtual tree, whose growth is influenced by the perceived emotional response from spectators. The system implements a 'magic mirror' paradigm (using a large-screen display or projection system) and is based on the ARToolkit with extended representations for scene graphs. The system relies on a PAD dimensional model of affect to support the fusion of different affective modalities, while also supporting the representation of affective responses that relate to aesthetic impressions. The influence of affective input on the visual component is achieved by mapping affective data to an L-System governing virtual tree behaviour. We have performed an early evaluation of the system, both from the technical perspective and in terms of user experience. Post-hoc questionnaires were generally consistent with data from multimodal affective processing, and users rated the overall experience as positive and enjoyable, regardless of how proactive they were in their interaction with the installation.

Categories and Subject Descriptors

H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems – *augmented reality, evaluation.*

General Terms

Theory, Design, Experimentation.

Keywords

Affective Computing, Augmented Reality, Multimodal Interaction, Interactive Digital Arts.

1. INTRODUCTION

Interactive Digital Arts Installations rely on increasingly sophisticated input modalities, also taking advantage of the integration of sensors in Virtual Reality (VR) and Augmented Reality (AR) systems. In addition to supporting more natural interaction, this creates an opportunity for analyzing audience reactions using recent development in affective computing. The long-term perspective would be to address more explicitly the user experience targeted by artistic intentions. In previous work, we have investigated how the introduction of a semantic layer in the virtual world [1] could facilitate the implementation of artistic briefs in Virtual Reality Art by providing a more abstract level of description. However, we had not considered the semantic content of the interaction channels by which the users experienced the installation.

In this paper, we describe an Augmented Reality Art installation [2], which responds to user behaviour, using Multimodal analysis of affective signals. The underlying research objectives consist in being able in the long-term to explicitly support some aesthetic concepts related to user experience, mediated by affective input.

Recent Digital Artworks publicly exhibited by one of the authors (MB) have featured emotions as their main topic, such as "Emotional Traffic", "Frozen Feelings" or the "Emotion Vending Machine" [3]. This was an inspiration to explore further the possibility of artistic installations that could incorporate state-of-the-art work in affective computing, so that the artistic brief could explicitly address emotions and elements of the user experience.

This resulted in the production of an artistic brief supporting the investigation of affective interfaces in AR Art installations, the "Emotional Tree", or E-Tree. In this installation, a virtual tree is animated to display its growth, which is influenced by the perceived emotional response of the spectators, as interpreted by a dimensional model of affect. The tree exhibits sophisticated growth patterns in terms of branching, speed of growth, branch orientations and branch movement. Furthermore, growth cycles are not monotonic, and parts of the tree can fade as a consequence of negative responses then resume growing when negative influences cease. This creates the basis for a rich interactive experience, in which the installation induces a feedback loop, reactions to perceived user attitudes eliciting in turn new responses from users. In this context, emotions are the content of a metaphorical dialogue between the installation and the spectators, which provide substance for the interaction.

The tree's response is not intended as a representation of the affective state of spectators but rather an interpretation of users' affective response to the installation (closer to an aesthetic reaction in the context of interactive digital art).

2. AFFECTIVE AR ART: ARCHITECTURE AND OVERVIEW

An outline of the architecture of E-Tree installation is shown in Figure 1. Speech and video input of users is analysed by three components: Emotional Speech Recognition (ESR) for emotional features, Multi-Keyword Spotting (MKS) for recognition of pre-defined sequences of words and Video Feature Extraction (VFE) for detection and tracking of facial geometries and optical flow within a video stream. The features of these affective inputs are passed over UDP or TCP connections to Affective Interpretation modules, tailored to each component. The modules map input features to sets of values in our affective model. The combined output of this model is passed via UDP to a generational system that modifies the visual appearance of the E-Tree as a naturalistic tree structure, defined by an L-system [4]. Its growth is governed by rules that are modulated by the real-time interpretation of the spectators' emotional state as represented by the affective model. The user experience consists of observing tree growth, which follows specific patterns influenced by the perceived users' attitude. As the change in appearance and behaviour of the tree can induce further reactions in the user, the installation creates an affective feedback mechanism.

This model is based on Mehrabian's Pleasure-Arousal-Dominance (PAD) model [5]. The PAD model measures emotional tendencies and response along three dimensions: pleasure-displeasure, corresponding to cognitive evaluative judgements; arousal-nonarousal to levels of alertness and physical activity; and dominance-submissiveness to the feeling of control and influence over others and surroundings. These three dimensions provide a means to interpret diverse multimodal inputs as an affective representation of the user experience. Each input modality is interpreted in terms of a set of PAD dimensional values, conceptualised as a vector in the PAD model space. These vectors are combined by a weighted sum method to produce a representative emotional state, described by the calculated values of each dimension of the model. The input from affective modalities is integrated over the course of an interactive session to give a changing indication of the overall "mood" of the session. The model also integrates a measure of user interest, interpreted from user interaction, as a first step to incorporating the aesthetic components of Digital Art.

The emotional representation need not be transparent to the user, as the E-Tree is an artistic installation whose purpose is derived from artist intent, but it should be consistent, so that similar experiences cause similar tree reactions. In contrast, the final configuration of the tree will be unique, as the overall "experience", dynamically interpreted, will be different for each interaction with the installation. Russell [6] describes the interpretation of affect as a vector, characterising the placement of terms in the circumplex model as vectors from the neutral state, and the intensity of the emotion the length the vector. We apply a similar interpretation to the PAD model.

The choice of the PAD dimensional model had two main reasons. Firstly, the three dimensions have been shown to be optimal [7] for describing emotional state, being able to distinguish between states conflated in other models. Secondly, the dimensional approach allows combining and averaging of affective input from different sources, even if they are in a discrete category-like format (such as the "Ekmanian" emotions), provided a suitable mapping to PAD values can be devised. We can map two-factor Valence-Arousal models (such as that underlying the circumplex model [8]) to Pleasure and Arousal dimensions, This should allow us to integrate affective input that has been interpreted by components with different underlying models of emotion.

The graphical implementation of the tree makes use of a marker-driven AR system that displays the naturalistic tree situated in the environment of the participants. This is implemented using OSGART [9], which extends the OpenSceneGraph [10] 3D graphical framework to support AR applications, utilising ARToolkit [11] for marker detection and tracking. Users can directly interact with the installation by manipulating the markers and seeing the tree react appropriately on a "magic mirror" video display. We currently use a large (30 inch) monitor, but have also experimented with projection onto a wall to incorporate group participation.

For each affective input component, as well as the PAD interpretation of these features, a semantic interpretation can

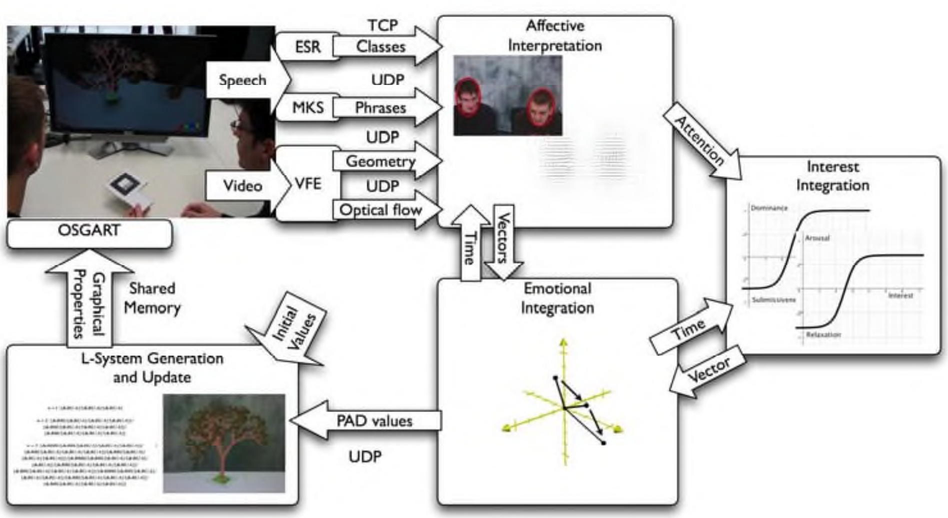

Figure 1. E-Tree System Architecture

result in interest "signals", where an indication of user interest is detected. Interest as an aesthetic property is interpreted as an integrated response that builds up with repeated signals. We use a model of signal reinforcement and decay described by Picard [12]. Positive interest changes build up if close together to give a larger and larger signal. Input is modelled as a separate component of affective input, builds up as components register interest increases, and naturally fades away if no interest-producing input is received. This unipolar interest representation is mapped to PAD dimensions using a sigmoid function, so a certain level of interest is needed before there is a noticeable affect response, while PAD values are kept in the correct range.

3. E-TREE BEHAVIOUR

In our implementation of the artistic brief, E-Tree grows and branches in a naturalistic manner, from an initial cluster of small shoots to a larger, many-branched tree with tapering boughs and coloured leaves. User interactions influence this growth so that a unique tree structure is create during each interactive session where the appearance of each branch reflects the integration of emotional state during the period of its growth and the overall look of the tree acts as a visual "record" or history of the installation experience. E-Tree also exhibits short-term behaviours and changes in appearance that are reflect the prevailing mood.

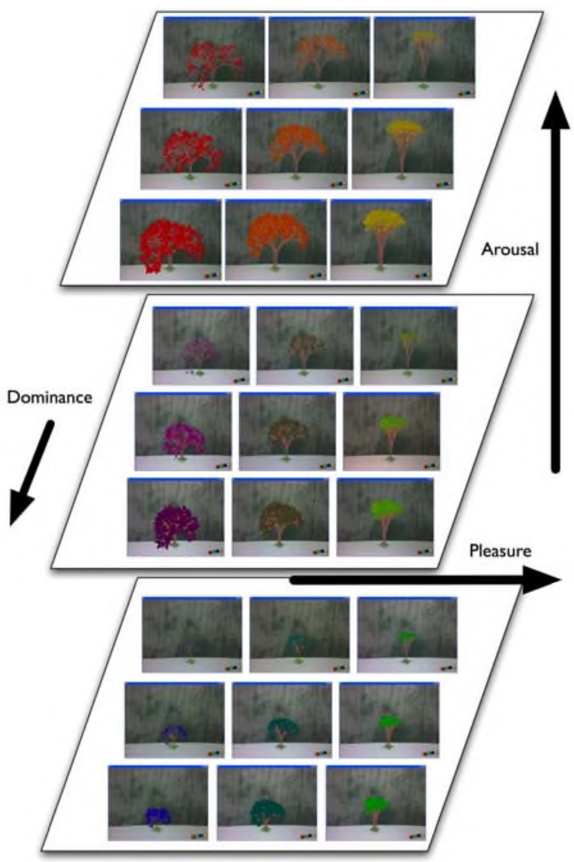

Figure 2. Relationship of PAD dimensions and E-Tree appearance.

The instantaneous reactive aspects of E-Tree are continually updated according the current affective state of the emotional model. These changes are neither permanent nor reflected in the final "historical" appearance of the tree, but serve to stimulate on-going reactions from participants, and provide some of the dynamics of the installation outlined in the brief. Figure 2 shows the effect of these changes related to the dimensions of affect used in our emotional model.

The perceived "pleasingness" of the installation, interpreted as the Pleasure PAD dimension, changes the overall scale of the tree. The more satisfied the participants, the bigger the tree will appear. The engagement of the user experience, manifested by the Arousal and Interest displayed by participants causes the tree to reach upwards, or droop downwards, a representation of user "alertness". Rapid changes in this droop give the twitching effects described in the artistic brief. Finally, expressiveness and control displayed in interactions is reflected in the overall thickness of branches and area of leaves. If a user acts in a determined, independent manner, controlling interaction, not just reacting, the tree will appear thicker than if they seem bored, anxious and docile. This is controlled by the Dominance dimension of affect in the PAD model. The hue of the leaves also changes depending on the combination of Pleasure and Arousal.

The growth and branching of the tree serve to record a history of the user experience. During periods of satisfaction and engagement, more branches will be produced and a denser system of branching will be created, while bored and uninterested periods will result in sparse foliage with shorter branches.

4. EVALUATION

For an initial evaluation of the E-Tree, 10 volunteers were observed interacting with the installation in pairs. Output was displayed on a large (30") LCD monitor, thus implementing a "magic mirror" paradigm for AR. Input devices included a microphone positioned with a stand to capture speech utterances and a USB camera set up so that both subjects could appear in frame for facial analysis. Our main goals when evaluating E-Tree were to verify that the concept was appealing to potential participants (in terms of engagement and encouraging interaction), and to validate the underlying hypotheses in terms of Multimodal affective processing. We recorded PAD values throughout interactive sessions, and users were asked to complete a set of semantic differential questions designed to assess their subjective evaluation of their engagement and interest in the Digital Art installation, together with how representative they found the reactions of the tree, and how realistic they found the tree behaviour (in terms of its branching and overall look).

The overall response was encouragingly positive. The average score for each question is given in Figure 3, where 1 indicates the most negative opinion and 5 the most positive. The E-Tree installation was perceived as being Interesting (4.0) as well as Responsive to user interaction (3.7). This indicates that the installation succeeds in engaging participants and communicating that they can affect the tree. The branching and growth was perceived as being fairly realistic (3.1) which hopefully indicates it was novel without being confusing.

Interestingly, participants in general did not feel that E-Tree represented their own emotions (2.2), although one participant pair did score a value of 5, indicating they perceived a high correlation with their own emotions. This may be explained by a terminological confusion between emotions displayed towards the installation and the actual emotional state of the user (there is no intention that the E-Tree should exhibit empathy). However, in some cases this can also be due to inaccurate processing of affective input.

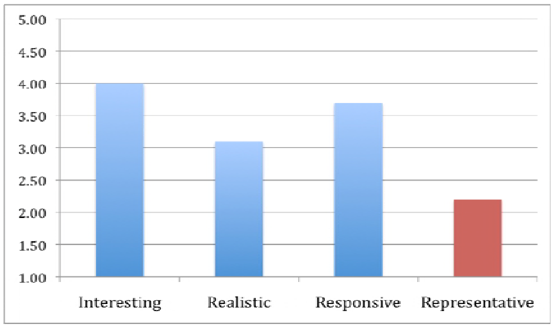

Figure 3. Mean subjective evaluation scores.

Participants engaged in exploratory behaviour when they realised that certain movements and vocal intonations had an affect on the appearance of the tree. No users appeared to move on to more "expressive" behaviour (that was not the aim of this version of the installation), though one user did start issuing vocal commands to the tree, but was eventually hindered by the limited vocabulary of the keyword component. A typical trace of PAD values for a "successful" interaction is shown in figure 4. There is interaction throughout the duration of the installation, resulting in substantial activity in each dimension, and all components were utilized (although Dominance was hindered by the lack of significant facial changes). Affective input appeared roughly correspondent with behaviour,

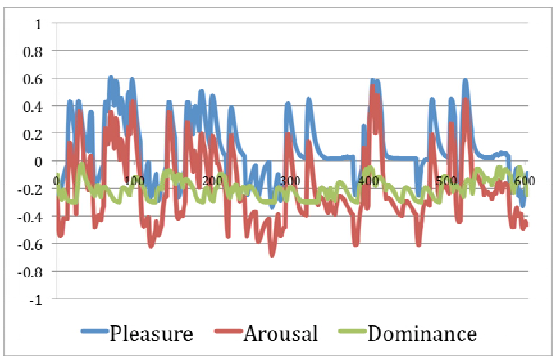

Figure 4. Typical PAD trace for an interaction

5. CONCLUSIONS

We have described one of the first examples of an AR Art Installation supporting affective interaction. While preliminary user feedback on the experience is certainly encouraging, additional work is required to further evaluate the affective processing component. One direction consists in monitoring real-time emotional states, for instance using physiological measurements, and comparing these to multimodal affective processing data. However, since the users' response is not necessarily a reflection of their own emotional state but an appreciation of the interactive experience, other directions for evaluation include entertainment theories [13] as well. The interactive experience itself can also be enriched, for instance by extending the input modalities and allowing direct "physical" interaction from the users.

6. ACKNOWLEDGMENTS

This work has been funded in part by the European Commission via the CALLAS Integrated Project. (ref. 034800, http://www.callas-newmedia.eu/)

7. REFERENCES

[1] Marc Cavazza , Jean-Luc Lugrin , Simon Hartley , Paolo Libardi , Matthew J. Barnes , Mikael Le Bras , Marc Le Renard , Louis Bec , Alok Nandi. 2004. New ways of worldmaking: the Alterne platform for VR art, Proceedings of the 12th annual ACM international conference on Multimedia, October 10-16, 2004, New York, NY, USA.

[2] Stephen W. Gilroy, Marc Cavazza, Rémi Chaignon, Satu-Marja Mäkelä ,Markus Niiranen, Elisabeth André, Thurid Vogt, Mark Billinghurst, Hartmut Seichter, and Maurice Benayoun. 2007. An Emotionally Responsive AR Art Installation. Proceedings of 6th International Symposium on Mixed and Augmented Reality (ISMAR 2007), Workshop on Mixed Reality, Entertainment and Art, pp.61–66.

[3] Maurice Benayoun. http://www.benayoun.com/works.php?id=01000

[4] Przemyslaw Prusinkiewicz, Aristid Lindenmayer. 1996. The Algorithmic Beauty of Plants. Springer-Verlag New York, Inc.

[5] Mehrabian, A., & Russell, J. A. 1974. An approach to environmental psychology. Cambridge, MA: MIT. J.A. Russell. 1989. Measures of Emotion. In *Emotion: Theory, Research and Experience*, ch.4, pp.83-111. Academic Press, Inc.

[6] J.A. Russell. 1989. Measures of Emotion. In *Emotion: Theory, Research and Experience*, ch.4, pp.83-111. Academic Press, Inc.

[7] Mehrabian, A. 1995. Framework for a comprehensive description and measurement of emotional states. Genetic, Social, and General Psychology Monographs, 121, 339-361.

[8] Russell. J. A. 1980. A circumplex model of affect. Journal of personality and social psychology, 39. 1161-1178

[9] J. Looser, R. Grasset, H. Seichter, M. Billinghurst. 2006. OSGART - A Pragmatic Approach to MR. In Industrial Workshop at ISMAR 2006, Santa Barbara, California, USA, October, 2006.

[10] H. Kato, M. Billinghurst. 1999. Marker Tracking and HMD Calibration for a Video-Based Augmented Reality Conferencing System. Proceedings of the Second IEEE and ACM International Workshop on Augmented Reality (IWAR 1999), San Francisco, California, USA, pp.85–95, October 1999.

[11] Don Burns, Robert Osfield. 2004. Open Scene Graph. Proceedings of the IEEE Virtual Reality 2004 (VR'04), p.265.

[12] Picard, R. 1997. Affective Computing, MIT Press.

[13] Bryant, J., Vorderer,. P. (Eds.) 2006. Psychology of Entertainment. Lawrence Erlbaum Associates, Inc.