Creative Expression of Emotions in Virtual Humans

Celso M. de Melo
Institute for Creative Technologies
University of Southern California
13274 Fiji Way, Marina Del Rey, CA 90292, USA
demelo@usc.edu

Jonathan Gratch
Institute for Creative Technologies
University of Southern California
13274 Fiji Way, Marina Del Rey, CA 90292, USA
gratch@ict.usc.edu

'Works of art (...) can be expressive of human qualities: one of the most characteristic and pervasive features of art is that percepts (lines, colors, progressions of musical tones) can be and are suffused with affect.'

This passage by John Hospers [1] emphasizes that it is in the nature of art to express emotions and that this expression can be accomplished through the formal elements of art. According to the author, the reason it is possible to perceive emotions in these elements is that people find analogies to the internal and external manifestations of their bodies when experiencing such emotions. Effectively, artists have been exploring for centuries the idea that it is possible to perceive emotions in line, space, mass, light, color, texture, pattern, sound and motion [2]. Yet, the virtual humans field, which tries to bring the richness of human multimodal expression into virtual worlds, has tended to limit itself to bodily forms of expression such as gesture, face and voice [3]. Inspiring in the arts, this PhD thesis aims at exploring new forms of expression of emotion for virtual humans which aren't necessarily limited to the body and explore the full breadth of the formal elements of visual art.

We began by creating a model for bodily expression in virtual humans [4][5]. The model reflects the bulk of the research in the area and, aside from providing graphical support for virtual humans, is crucial as a means of comparison to assess the impact of the new forms of expression we seek to explore. The model supports: animation of the body using keyframes and inverse kinematics; a psycholinguistics-based gesticulation model; skin expression including simulation of blushing, wrinkles and tears; facial expression using a pseudo-muscular model of the face; real-time integration with a text-to-speech synthesis including lipsync; and, a markup language to control multimodal expression.

Then, in our first attempt to go beyond the body to express emotions in virtual humans, we explored lights, shadows, filters and composition [6]. The lighting model is pixel-based and supports multiple light sources, three light types, shadows and several configuration parameters. Filters and composition allow manipulation of the virtual human pixels themselves by first rendering the virtual human into a texture, then manipulating it with vertex and pixel shaders and, then, rendering it to the screen. Several filters were explored (see Fig.1). A mechanism was then

created to support definition of rules to map emotions synthesized using the OCC emotion model [7] into these forms of expression.
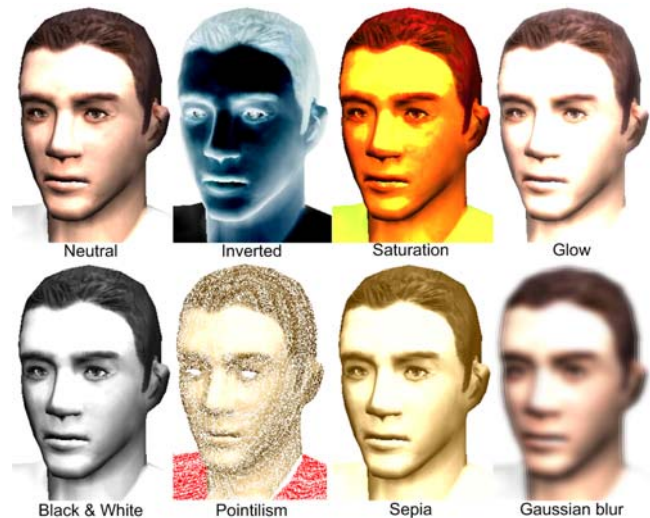

Figure 1. Filters to express emotions in virtual humans.

However, computer models which rely on predefined rules to achieve expressive goals are unlikely to ever be as flexible as the kind of expression we see in the arts. For instance, other models which have tried to map lighting parameters to affective states [8][9] only capture a small portion of the spectrum of manipulations of light and shadows we see in the visual arts [10][11]. Furthermore, artists are constantly breaking existing rules and creating new ways of expressing themselves. What is needed, then, is a computational cognitive model of the *creative process* the artist uses to express emotions. This is, therefore, a core subgoal for this PhD thesis.

Creativity is the process which results in products, real or abstract, which are both *novel* - i.e., original - and *appropriate* - i.e. valuable to a society or knowledge domain. Boden [12] distinguishes three types of computational models of creativity: *combinational*, where the focus is on unusual association of familiar ideas. Analogy-making, seen as sustained comparison, is a particular type of this kind of model; *exploratory*, where the idea is to generate creative outcomes from systematic exploration of a knowledge domain; and, *transformational*, where creative outcomes are generated by augmenting or breaking the rules of a knowledge domain. Genetic algorithms are described as belonging to this last class.

Thus, as a next step, we used genetic algorithms (GAs) to learn mappings between emotions and expression generated using lights and pixels. GAs are appropriate for several reasons. The GAs' clear separation between generation and evaluation of alternatives is convenient. Alternatives can be generated using biologically inspired operators – selection, mutation, crossover, etc. Evaluation, in turn, can rely on artificial critics, which define fitness functions from art theory, or on human critics. There are several advantages to bringing humans into the evaluation process: (a) humans can be used to fill in the gaps in the literature; (b) a mechanism is introduced to accommodate the individual, social and cultural values in artistic expression of emotion [2]. Furthermore, the artistic expression space is very large and GAs deal well with intractable search spaces.

In a preliminary study conducted to assess this approach, we asked fifteen subjects to evolve expression of sadness and joy through lights and filters using a special tool developed for that purpose. The results have shown that: (1) for both emotions, with each successive generation, the average fitness of the population of configurations of light and filters increases monotonically; (2) joy tends to be expressed with multiple colors, high saturation, high brightness and warm colors (see top row of Fig.2); (3) sadness tends to be expressed with total desaturation, low brightness, the sepia filter and high contrast (see bottom row of Fig.2). Overall, the results show that it is possible to express emotions using lights and pixels and, moreover, the evolutionary approach shows promising results with respect to learning mappings for, at least, sadness and joy. The results also suggest that there are idiosyncratic, cultural and general values at play. But further studies are still required to discern these factors. Another interesting thing is that these mappings were evolved by non-artists, which had no expertise in lighting or screen expression. This is in sharp contrast to approaches which rely on the expertise of artists to convey such effects.
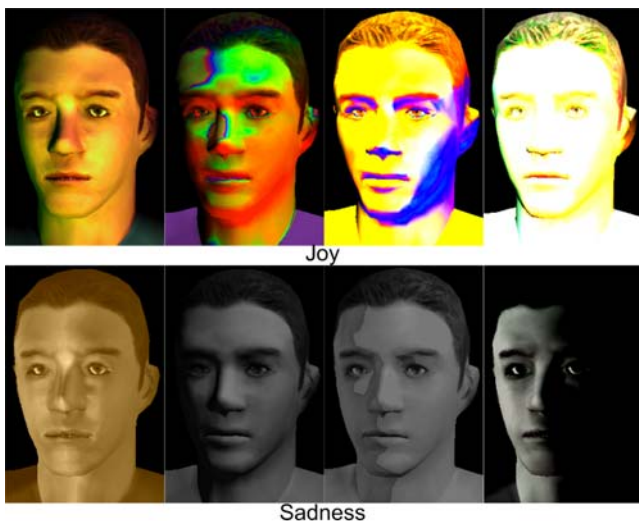

Figure 2. Evolved expression of joy and sadness through lights and filters using genetic algorithms.

However, despite the promising results using genetic algorithms, the creative process artists follow is more complex and diversified than the way a genetic algorithms works [2]. A *phase model for creativity* is more likely to be closer to the true creative process of the artist. This model defines a set of stages through which the creative process goes, in an iterative fashion, to reach the creative product. Building on the seminal model of Wallas [13], a seven-stage model is currently accepted [14]. These stages are: *information gathering*, where the creator learns about the knowledge domain; *preparation*, where the problem is identified; *incubation*, where the creator thinks about the problem, eventually as a background process resorting to unconscious thinking; *illumination*, where a promising solution emerges; *verification*, where the solution is thoroughly analyzed to verify its appropriateness; *communication*, where the creator communicates the solution and acts on feedback from its peers; and, *validation*, where the solution is validated and closure is achieved. Altogether, this process is argued to be iterative rather than sequential involving convergent thinking in some stages – information, preparation, verification, communication and validation – and divergent thinking in others – preparation, incubation, illumination and verification. Furthermore, the actual implementation of each one of these phases is likely to involve general principles as well as domain-specific principles [12]. Several computational models have already tried to simulate parts of this creative process when applied to the arts. Boden [12] surveys this area presenting examples for combinational, exploratory and transformational models. This thesis, thus, plans to draw on this knowledge to explore cognitive models of the creative process artists go through.

The systems described in the previous paragraph tend to focus on the simulation of the cognitive part of the creative process. This is in contrast, with the work described in the *non-photorealistic rendering* literature [15][16] which tends to focus on visual simulation of the techniques we see in the arts. Successful simulation of creative expression of emotions in virtual humans, as in our case, is likely to require integration of techniques from both fields. Therefore, we already began exploring some of the techniques from this field and applying them to virtual humans. Currently, we have implemented *mesh masks*, *silhouette* and *hatching* (see Fig.3). Mesh masks allow selective application of filters to subsets of the virtual human mesh in a way akin to collage or cubism [2]. Silhouette consists of drawing the virtual human edges with a configurable level-of-detail and hatching simulates a specific kind of pencil drawing. For this thesis, we seek to understand whether and how these (and other) techniques can be used to generate novel and appropriate forms of expression of emotions in virtual humans.

Finally, to evaluate the model, feedback shall be gathered from human subjects through interviews or questionnaires. Some of the issues we need to assess are: (a) How effectively is the model expressing emotions? (b) What is the importance of the new expression channels when compared with a model which uses only the usual bodily expression channels? Which one do subjects prefer? (c) Is the expression of emotions creative?

In summary, this PhD thesis seeks to develop a model for the expression of emotions in virtual humans which, going beyond the usual bodily expression channels, uses the formal elements of visual art to do so. To accomplish this, first, new expression channels shall be drawn from art theory and the non-photorealistic rendering literature. Then, several computational cognitive

models of the creative process of the artist shall be explored to define how are emotions expressed using these new expression channels. Finally, the model will be evaluated through questionnaires with human subjects. This thesis differs from previous work in the following: (1) it explores expression channels which go beyond the usual gesture, face and voice channels; (2) it applies (century-old) techniques from the visual arts to express emotions in virtual humans; (3) it seeks to integrate two fields which have had little interaction in the past, namely computational creativity and non-photorealistic rendering. Finally, this work is relevant to several different fields/industries: (a) *computational creativity/art theory*, as the model seeks to synthesize expression of emotions in a way akin to the arts; (b) *intelligent virtual agents/virtual humans*, as the model argues for the importance of going beyond bodily expression channels for the expression of emotions; (c) *computer graphics* and *non-photorealistic rendering*, as the model seeks to synthesize new forms of expression which cannot be automatically synthesized with the current state-of-the-art systems; (d) *psychology of creativity*, as there is a strong belief that the creative process artists use is very relevant to understanding how creativity works in other areas of life; (e) *gaming* and *movie industries*, which make extensive use of models of virtual humans.
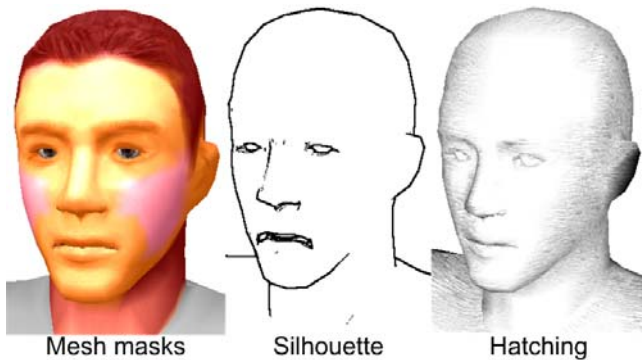

Figure 3. Mesh masks, silhouettes and hatching.

REFERENCES

[1] Hospers, J. 1967. Aesthetics, problems of. In The Encyclopedia of Philosophy, P. Edwards, Ed. Macmillan Publishing Co., New York, USA.

[2] Sayre, H. 2007. A World of Art, fifth ed. Prentice Hall, New Jersey, USA.

[3] Gratch, J., Rickel, J., Andre, E., Badler, N., Cassell, J. and Petajan, E. 2002. Creating interactive virtual humans: Some assembly required. ieee intelligent systems. IEEE Computer Society 17, 4, 54–63.

[4] de Melo, C. and Paiva, A. 2006. Multimodal expression in virtual humans. Computer Animation and Virtual Worlds Journal 17, 3-4, 239–248.

[5] de Melo, C. and Paiva, A. 2008. Modeling gesticulation expression in virtual humans. In New Advances in Virtual Humans: Artificial Intelligence, N. M.-T. et al., Ed. Springer-Verlag, Berlin Heidelberg, Germany.

[6] de Melo, C. and Paiva, A. 2007. The expression of emotions in virtual humans using lights, shadows, composition and filters. In Proceedings of Affective Computing Intelligent Interaction'07, 546–557.

[7] Ortony, A., Clore, G., and Collins, A. 1990. The Cognitive Structure of Emotions. Cambridge University Press, New York, USA.

[8] Tomlinson, B., Blumberg, B. and Nain, D. 2000. Expressive autonomous cinematography for interactive virtual environments. In Proceedings of the Fourth International Conference on Autonomous Agents.

[9] El-Nasr, M. and Horswill, I. 2004. Automatic lighting design for interactive entertainment. ACM Computers in Entertainment 2, 2.

[10] Birn, J. 2006. [digital] Lighting and Rendering, second ed. New Riders, California, USA.

[11] Millerson, G. 1999. Lighting for Television and Film, third ed. Focal Press, Oxford, USA.

[12] Boden, M. 2004. The Creative Mind - Myths and Mechanisms, second ed. Routledge, Routledge.

[13] Wallas G. (1926) The art of thought. New York, USA: Harcourt Brace

[14] Cropley A. (2006) In Praise of Convergent Thinking. Creativity Research Journal, 18(3), 391-404

[15] Gooch B and Gooch A. (2001) Non-Photorealistic Rendering. Massachusetts, USA: A K Peters

[16] Strothotte T. and Schlechtweg S. (2002). Non-Photorealistic Computer Graphics – Modeling, Rendering, and Animation. San Francisco, USA: Morgan Kaufmann Publishers