

The Art of Expressing Emotions in Virtual Humans

Celso de Melo and Ana Paiva

*University of Southern California and IST-Technical University of Lisbon
United States of America and Portugal*

1. Introduction

Artists express emotions through art. To accomplish this, they rely on lines, shapes, textures, color, light, sounds, music, words and the body (Sayre, 2007). It is the flexibility of artistic media which affords the expression of complex affective states (Bell, 1913; Oatley, 2003). In fact, many artists use art as a means to come to terms and understand the subtlety and peculiarity of their feelings (Collingwood, 1938). Furthermore, it is not uncommon for artists to claim they need these media, which are not limited to the body or language, for the expression of their feelings (Gardner, 1982). But art is not simply an outlet for the artist's emotions. From the perspective of the receiver, through its empathetic emotional response to a work of art, it is also seen as a means to learn about the human condition (Elliot, 1966; Oatley, 2003). Emotions are, therefore, intrinsically related to the value of art.

Affective computing has been neglecting the kind of expression we see in the arts. Instead, researchers have focused on expression signals, spanning face, voice and body, which accompany affective processes (Davidson, Scherer & Goldsmith, 2003). Four reasons could help explain why this has been so. First, it is natural to focus on overt expression signals as these are directly observable and, thus, can be measured objectively. In contrast, it is rather difficult to measure objectively what is being expressed in a work of art. In fact, whereas some argue for the importance of understanding the artist's production process and culture when interpreting an art artefact (Davies, 1987; Wollheim, 1980; Goodman, 1968), others argue that the interpretation and aesthetics of an artefact is independent of its creation process (Beardsley, 1958; Baxandall, 1987; Sontag, 1966). Second, art is certainly not limited to the expression of affective states and sometimes it is difficult to discern what constitutes such expression. In fact, there are several conceptions about what expression in the arts is: (a) it relates to beauty as the creative expression of beauty in nature (Batteux, 1969); (b) it relates to culture as the expression of the values of a given society (Geertz, 1976); (c) it relates to individuality as the expression of the artist's liberties and creativity (Kant, 1951); (d) finally, it relates to the expression of affective states. Third, the romantic perspective of art as the creative expression of affective states in all its complexity is problematic (Averill, Nunley & Tassinary, 1995; Oatley, 2003). In general, the subject of art is not the habitual emotions we experience in our daily lives but, subtle and peculiar affective states. Therefore, it is only natural that researchers choose to begin by addressing the simpler affective states. Finally, the affective sciences field is itself relatively recent and we are only now beginning

to move beyond all-encompassing grand theories of emotions to more specialized conceptions of affective phenomena (Davidson, Scherer & Goldsmith, 2003). It is, perhaps, time to start taking more seriously the kind of expression we see in the arts.

In this context, this work proposes to draw on accumulated knowledge from art theory to synthesize expression of emotions in virtual humans. Virtual humans are embodied characters which inhabit virtual worlds and look, think and act like humans (Gratch et al, 2002). The state-of-the-art in virtual human research is symptomatic of the narrow view of the affective computing field we described above. Indeed, researchers have, thus far, tended to focus on gesture (Cassell, 2000), face (Noh & Neumann, 1998) and voice (Schroder, 2004) for the expression of emotions. But, of course, in the digital medium we need not be limited to the body. In this sense, this work goes beyond embodiment and explores the expression of emotions in virtual humans using lights, shadows, composition and filters. Our approach, thus, focuses on two expression channels: lights and screen. In the first case, we inspire on the principles of lighting, regularly explored in theatre or film production (Millerson, 1999; Birn, 2006), to convey the virtual human's affective state through the environment's light sources. In the second case, we acknowledge that, at the meta level, virtual humans are no more than pixels in the screen which can be manipulated, in a way akin to the visual arts (Birn, 2006; Gross, 2007; Zettl, 2008), to convey affective states. Finally, because artistic expression is essentially a creative endeavor (Sayre, 2007), more than trying to find the right set of rules, we explore an evolutionary approach which relies on genetic algorithms to learn mappings from affective states to lighting and screen expression.

The remainder of the chapter is organized as follows. Section 2 provides background on virtual humans and describes the digital medium's potential for expression of emotions, focusing on lighting and screen expression. Section 3 describes our virtual human model, detailing the lighting and screen expression channels, and introduces our evolutionary model. Section 4 describes some of our results. Finally, section 5 draws some conclusions and discusses future work.

2. Background

2.1 Expression of emotions in the digital medium

Digital technology is a flexible medium for the expression of emotions. In virtual worlds, inhabited by virtual humans, besides embodiment, at least four expression channels can be identified: camera, lights, sound, and screen. The camera defines the view into the virtual world. Expressive control, which inspires on cinema and photography, is achieved through selection of shot, shot transitions, shot framing and manipulation of lens properties (Arijon, 1976; Block, 2001). Lights define which areas of the scene are illuminated and which are in shadow. Furthermore, lights define the color in the scene. Expressive control, which inspires in the visual arts, is achieved through manipulation of (Millerson, 1999; Birn, 2006): light type, placement and angle; shadow softness and falloff; color properties such as hue, brightness and saturation. Sound refers to literal sounds (e.g., dialogues), non-literal sounds (e.g., effects) and music. Expressive control, which inspires in drama and music, is achieved through selection of appropriate content for each kind of sound (Juslin & Sloboda, 2001; Zettl, 2008). Finally, the screen is a meta channel referring to the pixel-based screen itself. Expression control, which inspires on cinema and photography, is achieved through manipulation of pixel properties such as depth and color (Birn, 2006; Zettl, 2008). This work shall focus on the lighting and screen expression channels.

2.2 Expression of emotions in virtual humans

Virtual humans are embodied characters which inhabit virtual worlds (Gratch et al, 2002). First, virtual humans look like humans. Thus, research draws on computer graphics for models to control the body and face. Second, virtual humans think and act like humans. Thus, research draws on the social sciences for models to produce synchronized verbal and nonverbal communication as well as convey emotions and personality. Emotion synthesis usually resorts to cognitive appraisal theories of emotion, being the Ortony, Clore and Collins (OCC) theory (Ortony et al, 1988) one of the most commonly used. Emotion expression tends to focus on conveying emotions through synchronized and integrated gesture (Cassell, 2000), facial (Noh & Neumann, 1998) and vocal (Schroder, 2004) expression. In contrast, this work goes beyond the body using lights, shadows, composition and filters to express emotions.

A different line of research explores *motion modifiers* which add emotive qualities to neutral expression. Amaya (Amaya et al, 1996) uses signal processing techniques to capture the difference between neutral and emotional movement which would, then, be used to confer emotive properties to other motion data. Chi and colleagues (Chi et al., 2000) propose a system which adds expressiveness to existent motion data based on the effort and shape parameters of a dance movement observation technique called Laban Movement Analysis. Hartmann (Hartmann et al, 2005) draws from psychology six parameters for gesture modification: overall activation, spatial extent, temporal extent, fluidity, power and repetition. Finally, closer to this work, de Melo (de Melo & Paiva, 2005) proposes a model for expression of emotions using the camera, light and sound expression channels. However, this model did not focus on virtual humans, used a less sophisticated light channel than the one proposed here, did not explore screen expression and used simple rules instead of an evolutionary approach.

2.3 Expression of emotions using lights

This work explores *lighting* to express virtual humans' emotions. Lighting is the deliberate control of light to achieve expressive goals. Lighting can be used for the purpose of expression of emotions and aesthetics (Millerson, 1999; Birn, 2006). To achieve these goals, artists usually manipulate the following elements: (a) type, which defines whether the light is a point, directional or spotlight; (b) direction, which defines the angle; (c) color, which defines color properties. In Western culture, color is regularly manipulated to convey emotions (Fraser, 2004); (d) intensity, which defines exposure level; (e) softness, which defines how hard or soft the light is; (f) decay, which defines how light decays with distance; (g) throw pattern, which defines the shape of the light. Shadows occur in the absence of light. Though strictly related to lights, they tend to be independently controlled by artists. Shadows can also be used to express emotions and aesthetics (Millerson, 1999; Birn, 2006). In this case, this is achieved by manipulation of shadow softness and size. Lighting transitions change the elements of light and shadow in time. Transitions can be used to change the mood or atmosphere of the scene (Millerson 1999; Birn 2006).

2.4 Expression of emotions using pixels

At a meta level, virtual humans and virtual worlds can be seen as pixels in a screen. Thus, as in painting, photography or cinema, it is possible to manipulate the image itself for expressive reasons. In this view, this work explores composition and filtering for the

expression of emotions. Composition refers to the process of arranging different aspects of the objects in the scene into layers which are then manipulated and combined to form the final image (Birn, 2006). Here, aspects refer to the ambient, diffuse, specular, shadow, alpha or depth object components. Composition has two main advantages: increases efficiency as different aspects can be held fixed for several frames; and, increases expressiveness as each aspect can be controlled independently. Composition is a standard technique in film production. Filtering is a technique where the scene is rendered into a temporary texture which is then manipulated using *shaders* before being presented to the user (Zettl, 2008). Shaders replace parts of the traditional pipeline with programmable units (Moller & Haines, 2002). Vertex shaders modify vertex data such as position, color, normal and texture coordinates. Pixel shaders modify pixel data such as color and depth. Filtering has many advantages: it has constant performance independently of the scene complexity; it can be very expressive due to the variety of available filters (St-Laurent, 2004); and, it is scalable as several filters can be concatenated.

3. The model

3.1 Overview

The virtual human model is summarized in Fig.1. The virtual human itself is structured according to a three-layer architecture (Blumberg & Galyean, 1995; Perlin & Goldberg, 1996). The *geometry layer* defines a 54-bone human-based skeleton which is used to animate the skin. The skin is divided into body groups - head, torso, arms, hands and legs. The *animation layer* defines keyframe and procedural animation mechanisms. The *behaviour layer* defines speech and gesticulation expression. Finally, several expression modalities are built on top of this layered architecture. Bodily expression manipulates face, postures and gestures. Further details on bodily expression, which will not be addressed here, can be found in (de Melo & Paiva, 2006a; de Melo & Paiva, 2006b). Lighting expression explores the surrounding environment and manipulates lights and shadows. Screen expression manipulates the virtual human pixels themselves.

3.2 Lighting expression

Lighting expression relies on a local pixel-based lighting model. The model supports multiple sources, three light types and shadows using the shadow map technique (Moller & Haines, 2002). The detailed equations for the lighting model can be found in (de Melo & Paiva, 2007). Manipulation of light and shadow elements (subsection 2.3) is based on the following parameters: (a) *type*, which defines whether to use a directional, point or spotlight; (b) *direction* and *position*, which, according to type, control the light angle; (c) *ambient*, *diffuse* and *specular colors*, which define the color of each of the light's components in either RGB (red, green, blue) or HSB (hue, saturation and brightness) space; (d) *ambient*, *diffuse* and *specular intensity*, which define the intensity of each of the components' color. Setting intensity to 0 disables the component; (e) *attenuation*, *attnPower*, *attnMin*, *attnMax*, which simulate light falloff. Falloff is defined as $attenuation^{attnPower}$ and is 0 if the distance is less than *attnMin* and 1 beyond a distance of *attnMax*; (f) *throw pattern*, which constraints the light to a texture using component-wise multiplication; (g) *shadow color*, which defines the shadow color. If set to grays, shadows become transparent; if set to white, shadows are disabled; (h) *shadow softness*, which defines the falloff between light and shadow areas.

Finally, sophisticated lighting transitions, such as accelerations and decelerations, are based on parametric cubic curve interpolation (Moller & Haines, 2002) of parameters.

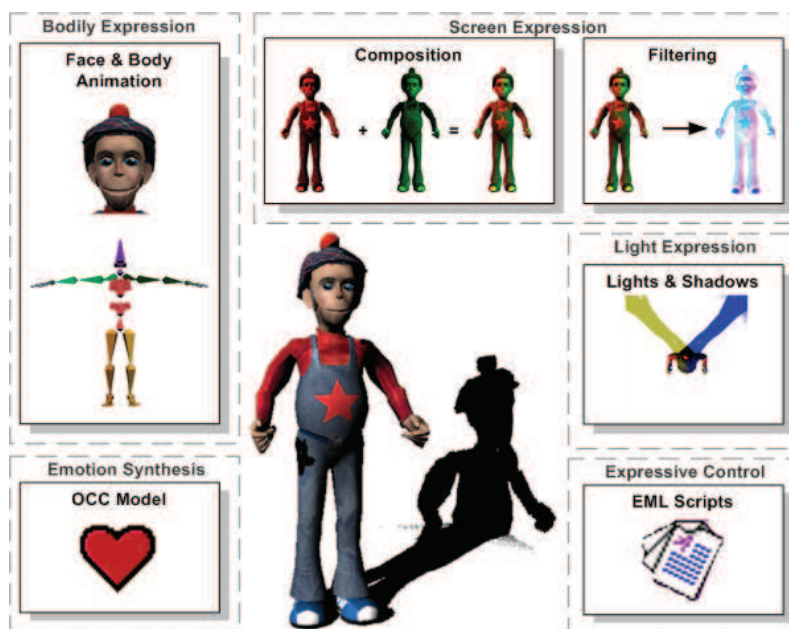


Fig. 1. The virtual human model

3.3 Screen expression

Screen expression explores composition and filtering. Filtering consists of rendering the scene to a temporary texture, modifying it using shaders and, then, presenting it to the user. Several filters have been explored in the literature (St-Laurent, 2004) and this work explores some of them: (a) the *contrast* filter, Fig.2-(b), which controls virtual human contrast and can be used to simulate exposure effects; (b) the *motion blur* filter, Fig.2-(c), which simulates motion blur and is usually used in film to convey nervousness; (c) the *HSB* filter, Fig.2-(c), which controls the virtual human hue, saturation and brightness; (d) the *style* filter, Fig.2-(d), which manipulates the virtual human's color properties to convey a stylized look. Filters can be concatenated to create compound effects and its parameters interpolated using parametric cubic curve interpolation (Moller & Haines, 2002).

Composition refers to the process of (Birn, 2006): arranging different aspects of the objects in the scene into layers; independently manipulating the layers for expressive reasons; combining the layers to form the final image. A layer is characterized as follows: (a) is associated with a subset of the objects which are rendered when the layer is rendered. These subsets need not be mutually exclusive; (b) can be rendered to a texture or the backbuffer. If rendered to a texture, filtering can be applied; (c) has an ordered list of filters which are

successively applied to the objects. Only applies if the layer is being rendered to a texture; (d) is associated with a subset of the lights in the scene. Objects in the layer are only affected by these lights; (e) defines a lighting mask, which defines which components of the associated lights apply to the objects; (f) can render only a subset of the virtual human's skin body groups. Finally, layer combination is defined by order and blending operation. The former defines the order in which layers are rendered into the backbuffer. The latter defines how are the pixels to be combined.

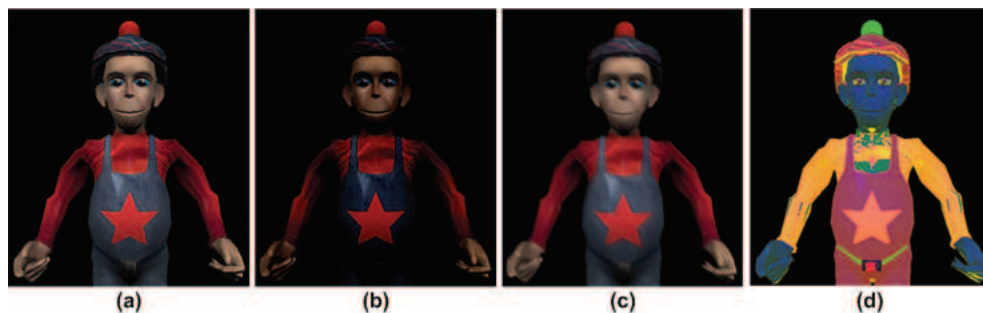


Fig. 2. Filtering manipulates the virtual human pixels. In (a) no filter is applied. In (b) the contrast filter is used to reduce contrast and create a more mysterious and harsh look. In (c) the motion blur is used to convey nervousness. In (d) the style filter, which is less concerned with photorealism, conveys an energetic look.

3.4 Synthesis of emotions

Emotion synthesis is based on the Ortony, Clore and Collins (OCC) model (Ortony, Clore & Collins, 1988). All 22 emotion types, local and global variables are implemented. Furthermore, emotion decay, reinforcement, arousal and mood are also considered. Emotion decay is, as suggested by Picard (1997), represented by an inverse exponential function. Emotion reinforcement is, so as to simulate the saturation effect (Picard, 1997), represented by a logarithmic function. Arousal, which relates to the physiological manifestation of emotions, is characterized as follows: is positive; decays linearly in time; reinforces with emotion eliciting; and, increases the elicited emotions' potential. Mood, which refers to the longer-term effects of emotions, is characterized as follows: can be negative or positive; converges to zero linearly in time; reinforces with emotion eliciting; if positive, increases the elicited emotions' potential, if negative, decreases it. Further details about this model can be found in (de Melo & Paiva, 2005).

3.5 Expression of emotions

Expression in the arts is a creative endeavor (Kant, 1951; Mill, 1965; Gombrich, 1960; Batteux, 1969; Sayre, 2007). Art is not a craft where artists can simply follow a set of rules to reach a result (Collingwood, 1938) and, according to the Romantic view, art is the creative expression of latent affective states (Oatley, 2003). Art can also be seen as the manifestation of the artist's individuality and liberties (Kant, 1951). In fact, Gombrich (1960) argues that this idiosyncrasy is inescapable as the artist's visual perceptions are necessarily confronted

with its mental schemas, including ideas and preconceptions. Thus, the simple rule-based approach is unlikely to capture the way artistic expression works. A better approach should, therefore, rely on machine learning, which would support adaptation to dynamic artistic values and automatic learning of new rules and sophisticated mappings between emotional states and bodily, environment and screen expression.

In this work, we propose an evolutionary approach for learning such mappings which is based on genetic algorithms. Genetic algorithms seem appropriate for several reasons. First, there are no available datasets exemplifying what correct expression of emotions using lights or screen is. Thus, standard supervised machine learning algorithms, which rely on a teacher, seem unsuitable. Furthermore, art varies according to time, individual, culture and what has been done before (Sayre 2007). Therefore, the artistic space should be explored in search of creative - i.e., new and aesthetic - expression. Genetic algorithms, defining a guided search, are, thus, appropriate. Second, the virtual humans field is new and novel forms of expression are available. Here, the genetic algorithm's clear separation between generation and evaluation of alternatives is appropriate. Alternatives, in this new artistic space, can be generated using biologically inspired operators - selection, mutation, crossover, etc. Evaluation, in turn, could rely on fitness functions drawn from art theory.

3.6 Evolutionary expression of emotions

The evolutionary model we propose revolves around two key entities: the *virtual human* and the *critic ensemble*. The virtual human tries to evolve the best way to express some affective state. For every possible state, it begins by generating a random set of *hypotheses*, which constitute a *population*. The population evolves resorting to a genetic algorithm under the influence of feedback from the critic ensemble. The ensemble is composed of human and artificial critics. The set of evolving populations (one per affective state) are kept on the *working memory*. The genetic algorithm only operates on populations in working memory. These can be saved persistently in the *long-term memory*. Furthermore, high fitness hypotheses (not necessarily from the same population) are saved in the long-term memory's *gallery*. Hypotheses from the gallery can, then, provide elements to the initial population thus, promoting high fitness hypotheses evolution. This model is summarized in Fig. 3.

At the core of the model lies a standard implementation of the genetic algorithm (Mitchell, 1999). The algorithm's inputs are: (a) the *critic ensemble* for ranking candidate hypotheses; (b) *stopping criteria* to end the algorithm; (c) the *size of the population*, p , to be maintained; (d) the *selection method*, SM , to select among the hypotheses in a population. In general, the higher the fitness, the higher the probability of being selected; (e) r , the *crossover rate*, defining the population percentage which is subjected to crossover; (f) m , the *mutation rate*, defining the population percentage subjected to mutation; (g) e , the *elitism rate*, defining the population percentage which propagates unchanged to the next generation.

The hypothesis space encodes the different expression modalities. Section 4 shows some possibilities for this encoding. Hypotheses are subjected to two genetic operators: crossover and mutation. Crossover takes two parent hypotheses from the current generation and creates two offspring by recombining portions of the parents. Mutation exists to provide a continuous source of variation in the population. This operator essentially randomizes the

values of a random number of the hypothesis' parameters. Selection among hypotheses is probabilistic and proportional to the fitness. Fitness is assigned by the critic ensemble.

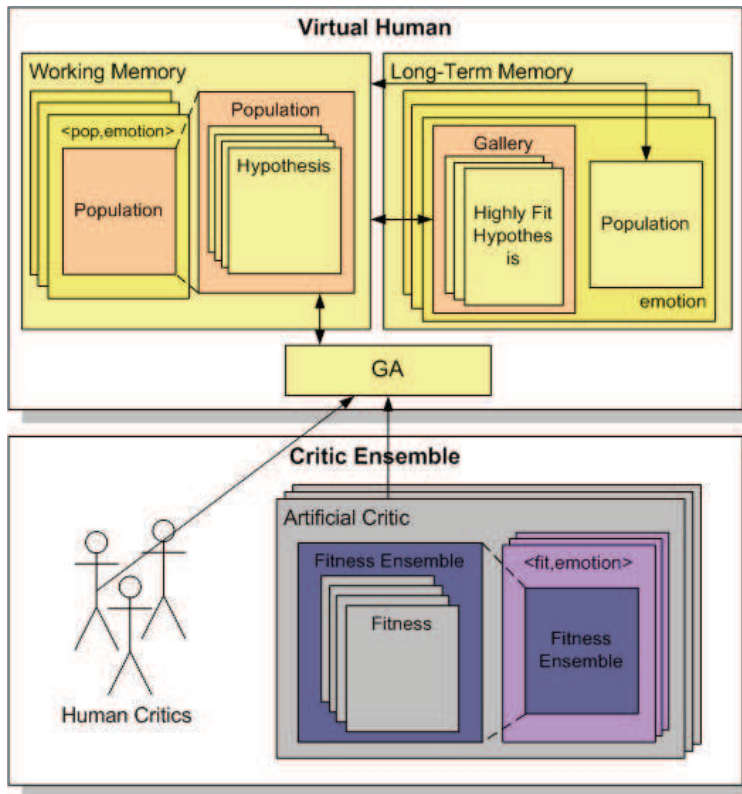


Fig. 3. The evolutionary model for expression of emotions in virtual humans

The critic ensemble defines several *critics* per affective state. Critics can be artificial, in which case fitness is inspired on art theory, or human, in which case fitness reflects the subjective opinion of the critic. An artificial critic consists of a set of *population fitness functions* and a set of *hypothesis fitness functions*. A population fitness function calculates a hypothesis' fitness with respect to the other hypotheses in the population. A hypothesis fitness function assigns an absolute fitness to each hypothesis, independently of the others. Both kinds of function are normalized to lie in the range [0.0;1.0] and the final fitness is calculated as the weighted average among all fitness functions. Regarding human critics, there are several advantages to bringing humans to the selection loop (Sayre, 2007): (a) art literature is far from being able to fully explain what is valued in the arts; (b) art is dynamic and values different things at different times. Furthermore, bringing humans into the evaluation process accommodates individual, social and cultural differences in expression of emotions (Keltner et al., 2003; Mesquita, 2003). Moreover, as discussed in future work, if the fitness functions are unknown, then the model might be made to learn from human experts. Two disadvantages are that humans may reduce variety in the population, causing convergence to a specific

style, and that the evolution process becomes much slower. For these reasons, the human fitness function (as well as any other) may be selectively deactivated.

Finally, notice the virtual human needs to keep track of several populations, one per affective state, even though only a single one is evolving at any instant of time. The working memory keeps the current state of evolving populations. In real life, creating an artistic product may take a long time (Sayre, 2007). Therefore, to accommodate this characteristic, the whole set of evolving populations can be saved, at any time, in long-term memory. Implementation-wise this corresponds to saving all information about the population in XML format. The interaction between working and long-term memory provides the foundations for life-long learning, where the virtual human can adapt to changing critics.

4. Results

The models we propose for lighting and screen expression are very expressive. For screen expression, consider the space which is defined by the application of one filter to each of the virtual human's body groups. Implementation-wise this is accomplished by having one composition layer per body group, where only that body group is active. Each layer is assigned its own filter. Rendering all layers will compose the whole virtual human. Examples of points in this exploration space are shown on the first eight images of Fig. 4. For lighting expression, let us restrict ourselves to configurations of the *three-point lighting technique* (Millerson, 1999; Birn, 2006). It is a configuration composed of the following light "roles": (a) *key light*, which is the main source of light focusing the character; (b) *fill light*, which is a low-intensity light that fills an area that is otherwise too dark; (c) *back light*, which is used to separate the character from the background. Moreover, we'll use only the key and fill lights, as these define the main illumination in the scene (Millerson 1999) and the back light can be computationally expensive (Birn 2006). Both lights are modeled as directional lights and, according to standard lighting practice (Birn 2006), only the key light is set to cast shadows. Examples of points in the lighting expression space are shown in the last seven images in Fig. 4.

Having shown that we are dealing with a very large expression space, what remains to be defined is how to define the mappings between affective states and lighting and screen expression. Here we'll exemplify our evolutionary model for the case of learning how to express 'anger' with lighting expression. The lighting hypothesis encoding is as described in the previous paragraph. Furthermore, we'll ignore all human critics and use a single artificial critic with the following fitness functions:

- The *red color* function, with weight 4.0, assigns higher fitness the smaller the Euclidean distance between the light's diffuse color to pure red. This function is applied both to the key and fill lights. Red is chosen because, in Western culture, red tends to be associated with excitation or nervousness (Fraser & Banks, 2004);
- The *low-angle illumination* function, with weight 1.5, assigns higher fitness the closer the Euclidean distance between the key light's angle about the x axis to 20° . The rationale is that illumination from below is unnatural, bizarre and frightening (Millerson, 1999);
- The *opaque shadow* function, with weight 1.0, assigns higher fitness the closer the key light's shadow opacity parameter is to 0. The rationale is that hard, crisp shadows convey a mysterious and harsh character (Millerson, 1999; Birn, 2006);

- The *low complexity* function, with weight 0.5, assigns higher fitness to less complex hypotheses. The rationale is that humans naturally value artefacts which can express many things in a simple manner (Machado, 2006). A lowest complexity hypothesis was defined to be one which has: diffuse color equal to a grayscale value (i.e. with equal R,G,B components); K_d equal to 2.0; and K_s equal to 0.0;
- The *key high-angle* function, with weight 0.5, assigns higher fitness the closer the Euclidean distance between the key light's angle about the x axis to $\pm 30^\circ$. This is a standard guideline for good illumination (Millerson, 1999). Notice this function contradicts the low-angle illumination function. This is acceptable, as guidelines in art can be contradictory (Sayre, 2007);
- The *key-fill symmetry* function, with weight 0.5, which assigns higher fitness if the fill light angles are symmetrical to the key light's. This is also a standard guideline for good illumination (Millerson, 1999);
- The *novelty* function, with weight 0.5, which assigns higher fitness the more novel the hypothesis is w.r.t. the rest of the population. This is a population fitness function. The idea is that novelty is usually appreciated in the arts (Sayre, 2007).

Having defined the artificial critic, the genetic algorithm was run for 50 iterations with the following parameters: $p = 50$, $r = 0.70$, $m = 0.00$, $e = 0.10$ and $SM = tournamentSelection$. The 44th generation was evaluated as having the best value of 2.3324, with its best hypothesis having a fitness of 0.9005. Five of the initial populations' hypotheses as well as five of the 44th generation are shown in Fig. 5. Finally, a graph showing population value and highest fitness evolution is shown in Fig. 6.

5. Conclusions and future work

This chapter argues for the importance of the kind of expression we find in the arts to the affective computing field. Emotions are intrinsically related to the value of art. For the artist, it is a means to come to terms with what he is experiencing affectively. For the audience, through empathy, it is a means to learn about the human condition. Furthermore, the kind of affective states which are reflected in works of art go beyond the kind of emotions we experience in our regular daily lives. The artist seeks to express an affective state in all its subtlety and peculiarity. This is why artists require flexible media and use lines, shapes, words, sounds and the body to express themselves. The affective computing field has been neglecting the kind of expression we find in the arts. Virtual humans are a case in point. Thus far, research has focused on the expression of emotions using body, face and voice. But why limit ourselves to the body?

Drawing on accumulated knowledge from the arts, we propose a virtual human model for the expression of emotions which goes beyond the body and uses lights, shadows, composition and filters. Regarding light expression, a pixel-based lighting model is defined which provides several control parameters. Parameter interpolation based on parametric cubic curves supports sophisticated lighting transitions. Regarding screen expression, filtering and composition are explored. Filtering consists of rendering the scene to a temporary texture, manipulating it using shaders and, then, presenting it to the user. Filters can be concatenated to generate a combined effect. In composition, aspects of the scene objects are separated into layers, which are subjected to independent lighting constraints and filters, before being combined to generate the final image. Regarding emotion synthesis, the OCC emotion model is integrated. To learn mappings from affective states to multimodal expression, we propose an evolutionary approach based on genetic algorithms.



Fig. 4. The lighting and screen expression space. The first eight pictures correspond to assigning different filters to the virtual human's body groups, without manipulating lights. The last seven pictures correspond to variations of the three-point lighting configuration, without manipulation of filters. Bodily expression is kept constant in all images.

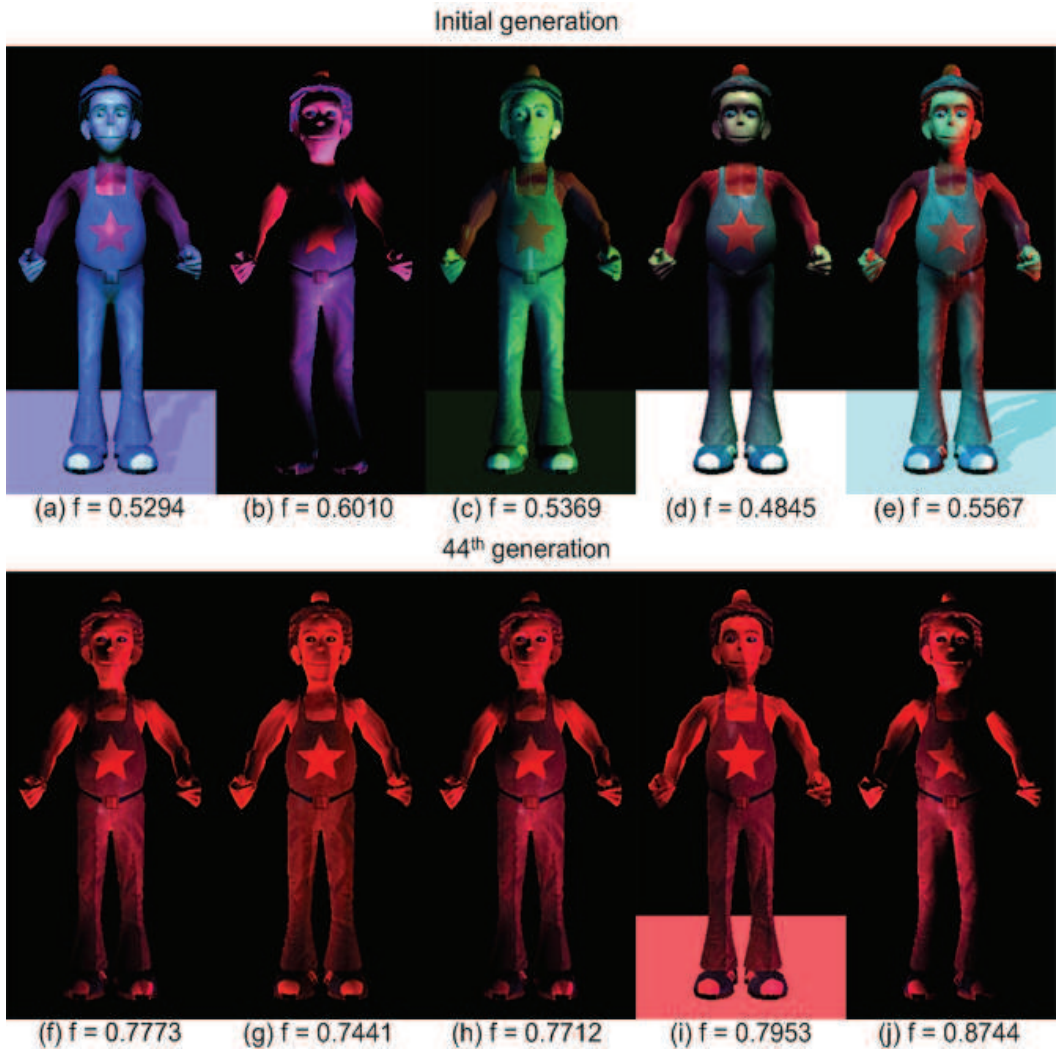


Fig. 5. The initial and 44th generation of a 50-iteration run of the evolutionary model

We've shown how broad our lighting and screen expression spaces are and demonstrated our evolutionary approach for the case of learning to express anger using lights. In the future, we'll explore fitness functions. This will require us to survey the literature on art theory and contact artists themselves. Artificial and human critics, respectively, accommodate, in our model, both of these forms of feedback. An interesting line of research would be to learn, through human critics, new fitness functions. In fact, this seems unavoidable, since it is clear for us that the literature is insufficient to fully comprehend the kind of expression we find in the arts. Furthermore, not only should the fitness functions be learned but, also their weights. The gallery could also be used to feed supervised learning

algorithms to generate models which explain highly fit hypotheses. These models could, then, feed a self-critic which would, in tandem with the usual artificial and human critics, influence the selection process. Finally, two obvious extensions to this work include exploring the camera and sound expression channels of which much knowledge already exists in the arts (Sayre, 2007).

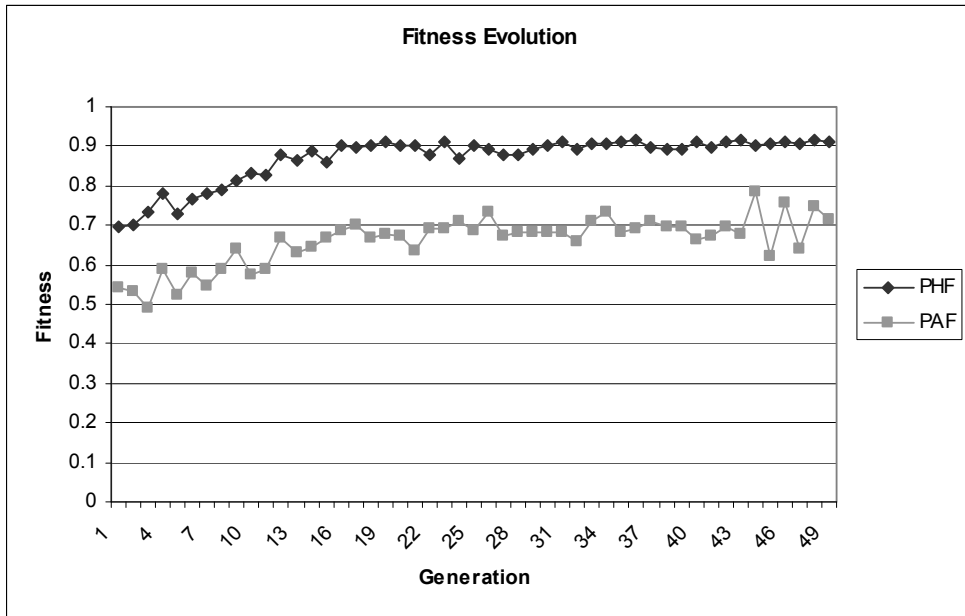


Fig. 6. Fitness evolution of a 50-iteration run of the evolutionary model

6. References

- Amaya, K., Bruderlin, A. & Calvert, T. (1996). Emotion from motion, *Proceedings of Graphics Interface'96*, pp.222-229
- Arijon, D. (1976). *Grammar of Film Language*, Hastings House, 0-80-382675-3, New York, USA
- Averill, J., Nunley, E. & Tassinary L. (1995). Voyages of the Heart: Living an Emotionally Creative Life, *Contemporary Psychology*, Vol. 40, No.6, pp.530, 0010-7549
- Batteux, C. (1969). *Les Beaux Arts Réduits à un meme Principe*, Slatkine Reprints, Genève, Switzerland
- Baxandall, M. (1987). *Patterns of Intention*, Yale University Press, 0-30-003763-5, New Haven, UK
- Bell, C. (1913). *Art*, Frederick A. Stokes, New York, USA
- Beardsley, M. (1958). *Aesthetics: Problem in the Philosophy of Criticism*, Harcourt, New York, USA

- Birn, J. (2006). *[digital] Lighting and Rendering – 2nd edn*, New Riders, 0-32-131631-2, London, UK
- Block, B. (2001). *The Visual Story: Seeing the Structure of Film, TV, and New Media*, Focal Press, 0-24-080467-8, Boston, USA
- Blumberg, B. & Galyean, T. (1995). Multi-Level Direction of Autonomous Creatures for Real-Time Virtual Environments, *Proceedings of SIGGRAPH'95*, Vol.30, No. 3, pp.173-182
- Cassell, J. (2000). Nudge, nudge, wink, wink: Elements of face-to-face conversation for embodied conversational agents, In: *Embodied Conversational Agents*, S. P. J. Cassell, J. Sullivan and E. Churchill, (Eds.), pp. 1-27, The MIT Press, 0-26-203278-3, Massachusetts, USA
- Chi, D., Costa, M., Zhao, L. & Badler, N. (2000). The EMOTE model for effort and shape, *Proceedings of SIGGRAPH 2000*, pp.173-182
- Collingwood, R. (1938). *The Principles of Art*, Clarendon Press, Oxford, UK
- Davidson, R., Scherer K. & Goldsmith J. (2003). *Handbook of Affective Sciences*, Oxford University Press, 0-19-512601-7, New York, USA
- Davies, S. (1987). Authenticity in Musical Performance, *The British Journal of Aesthetics*, Vol. 27, No.1, pp.39-50, doi:10.1093/bjaesthetics/27.1.39
- de Melo, C. & Paiva, A. (2005). Environment Expression: Expressing Emotions through Cameras, Lights and Music, *Proceedings of Affective Computing Intelligent Interaction (ACII'05)*, pp.715-722, Springer
- de Melo, C. & Paiva, A. (2006a). Multimodal Expression in Virtual Humans, *Computer Animation and Virtual Worlds Journal*, Vol.17, No.3, pp.215-220
- de Melo, C. & Paiva, A. (2006b). A Story about Gesticulation Expression, *Proceedings of Intelligent Virtual Agents (IVA'06)*, pp.270-281
- de Melo, C. & Paiva, A. (2007). The Expression of Emotions in Virtual Humans using Lights, Shadows, Composition and Filters, *Proceedings of Affective Computing Intelligent Interaction (ACII'07)*, pp.546-557
- Elliot, R. (1966). Aesthetic Theory and the Experience of Art paper, *Proceedings of the Aristotelian Society*, NS 67, III-26
- Fraser, T. & Banks, A. (2004). *Designer's Color Manual: the Complete Guide to Color Theory and Application*, Chronicle Books, San Francisco, USA
- Gardner, H. (1982). *Art, mind and brain: A cognitive approach to creativity*, Basic Books, 0-46-500445-8, New York, USA
- Geertz, C. (1976). Art as a Cultural System, *Modern Language Notes*, Vol. 91, No.6, pp.1473-1499
- Gombrich, E. (1960). *Art and Illusion; a Study in the Psychology of Pictorial Representation*, Pantheon Books, New York, USA
- Goodman, N. (1968). *Languages of Art; an Approach to a Theory of Symbols*, Bobbs-Merrill, Indianapolis, USA
- Gratch, J., Rickel, J., Andre, E., Badler, N., Cassell, J., Petajan, E. (2002). Creating Interactive Virtual Humans: Some Assembly Required, *IEEE Intelligent Systems*, Vol.17, No.4, pp.54-63

- Gross, L. & Ward, L. (2007). *Digital Moviemaking - 6th edn*, Thomson/Wadsworth, 0-49-505034-2, Belmont, USA
- Hartmann, B., Mancini, A. & Pelachaud, C. (2005). Implementing Expressive Gesture Synthesis for Embodied Conversational Agents, *Proceedings of Gesture Workshop*, pp.173-182, LNAI, Springer
- Juslin, P. & Sloboda, J. (2001). *Music and Emotion: Theory and Research*, Oxford University Press, 0-19-263188-8, New York, USA
- Kant, I. & Bernard, J. (1951). *Critique of judgment*, Hafner Pub. Co., New York, USA
- Keltner, D., Ekman, P., Gonzaga G. & Beer, J. (2003). Facial Expression of Emotion, In: *Handbook of Affective Sciences*, R. J. Davidson, K. R. Scherer and J. H. Goldsmith, (Eds.), 415-433, Oxford University Press, 0-19-512601-7, New York, USA
- Machado, F. (2006). *Inteligencia Artificial e Arte*, PhD thesis, Universidade de Coimbra
- Mesquita, B. (2003). Emotions as Dynamic Cultural Phenomena, In: *Handbook of Affective Sciences*, R. J. Davidson, K. R. Scherer and J. H. Goldsmith, (Eds.), 871-890, Oxford University Press, 0-19-512601-7, New York, USA
- Mill, J. (1965). What is Poetry?, In: *Mill's Essays on Literature and Society*, J. B. Schneewind, (Eds.), 103-117, Collier Books, New York, USA
- Millerson, G. (1999). *Lighting for Television and Film - 3rd edn*, Focal Press, 0-24-051582-X, Oxford, USA
- Mitchell, M. (1998). *An Introduction to Genetic Algorithms*, MIT Press, 0-26-263185-7, Massachusetts, USA
- Moller, T. & Haines, E. (2002). *Real-Time Rendering - 2nd edn*, AK Peters, 1-56-881182-9, Massachusetts, USA
- Noh, J. & Neumann, U. (1998). *A survey of facial modelling and animation techniques*, Technical report, USC Technical Report 99-705
- Oatley, K. (2003). Creative Expression and Communication of Emotions in the Visual and Narrative Arts, In: *Handbook of Affective Sciences*, R. J. Davidson, K. R. Scherer and J. H. Goldsmith, (Eds.), 481-502, Oxford University Press, 0-19-512601-7, New York, USA
- Ortony, A., Clore, G. & Collins, A. (1988). *The Cognitive Structure of Emotions*, Cambridge University Press, Cambridge, USA
- Perlin, K. & Goldberg, A. (1996). Improv: A System for Scripting Interactive Actors, *Virtual Worlds in Proceedings of SIGGRAPH'96*, pp.205-216
- Picard, R. (1997). *Affective Computing*, MIT Press, 0-58-500319-X, Massachusetts, USA
- Sayre, H. (2007). *A World of Art - 5th edn*, Prentice Hall, 978-0-132221-86-3, New Jersey, USA
- Schroder, M. (2004). *Speech and emotion research: an overview of research frameworks and a dimensional approach to emotional speech synthesis*, PhD thesis, Institute of Phonetics, Saarland University
- Sontag, S. (1966). *Against Interpretation and Other Essays*, Farrar, Straus & Giroux, New York, USA
- St-Laurent, S. (2004). *Shaders for Game Programmers and Artists*, Thomson/Course Technology, Massachusetts, USA

- Wollheim, R. (1980). *Art and its Objects – 2nd edn*, Cambridge University Press, 0-52-129706-0, Cambridge, UK
- Zettl, H. (2008). *Sight, Sound, Motion: Applied Media Aesthetics – 5th edn*, Thomson/Wadsworth, 0-49-509572-9, Belmont, USA



Affective Computing

Edited by Jimmy Or

ISBN 978-3-902613-23-3

Hard cover, 284 pages

Publisher I-Tech Education and Publishing

Published online 01, May, 2008

Published in print edition May, 2008

This book provides an overview of state of the art research in Affective Computing. It presents new ideas, original results and practical experiences in this increasingly important research field. The book consists of 23 chapters categorized into four sections. Since one of the most important means of human communication is facial expression, the first section of this book (Chapters 1 to 7) presents a research on synthesis and recognition of facial expressions. Given that we not only use the face but also body movements to express ourselves, in the second section (Chapters 8 to 11) we present a research on perception and generation of emotional expressions by using full-body motions. The third section of the book (Chapters 12 to 16) presents computational models on emotion, as well as findings from neuroscience research. In the last section of the book (Chapters 17 to 22) we present applications related to affective computing.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Celso de Melo and Ana Paiva (2008). The Art of Expressing Emotions in Virtual Humans, Affective Computing, Jimmy Or (Ed.), ISBN: 978-3-902613-23-3, InTech, Available from:

http://www.intechopen.com/books/affective_computing/the_art_of_expressing_emotions_in_virtual_humans

INTECH

open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2008 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](#), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.