Expression of affects in
Embodied Conversational Agents

S. Hyniewska[1], R. Niewiadomski[1], M. Mancini[2], C. Pelachaud[1,3]

LTCI, Télécom ParisTech
37/39 rue Dareau
75014, Paris - France

InfoMus Lab, DIST –
University of Genova
via Causa 13
I-16145, Genova - Italy

CNRS, France

{niewiado,hyniewska,pelachaud}
@telecom-paristech.fr

maurizio@infomus.org

Abstract:

Providing the virtual agents with expressive faculties is a contemporary challenge. The generated emotional behaviour for Embodied Conversational Agents (ECAs) particularly has to be credible and varied. Cowell and Stanney (2005) showed that well chosen behaviours and emotional expressions improve the credibility of an agent. Behaviour expressivity can be conveyed through: the choice of the nonverbal signals and their realisation.

Some researchers argue that creating an "internal state" for an ECA and giving the ECA the possibility to express emotions contributes to a richer human-machine interaction (Walker, Sproull and Subramani 1994). The internal states can be expressed on the behaviour level through the choice of the nonverbal signals and the modulation of their realisation.

In this Chapter several algorithms that enrich the behaviour repertoires of ECAs are presented. Although a great number of models use interpolations to create new expressions based on formerly defined basic emotions (e.g. Tsapatsoulis et al. 2002; Albrecht et al. 2005; Paradiso 2002), a particular attention is paid to the models of emotional expressions going beyond these so-called universal expressions. Some computational models of perceptively verified expressions have been created, mostly using fuzzy methods (Arya and DiPaola 2007; Arya et al. 2009). Niewiadomski and Pelachaud have also elaborated a model for the generation of complex facial displays like superposition or masking of expressions (Niewiadomski and Pelachaud 2007; see *Chapter by Niewiadomski et al.*).

Other researchers grounded their expression generation systems in a dimensional approach, with a link between points in the space defined mostly by the pleasure, arousal and dominance dimensions and the generated expressions (e.g. Zhang et al. 2007; Courgeon et al. 2008). Some models also integrate other modalities than the face into the expressions of emotions.

A different approach is to focus on the modelling of facial expressions as sequence of multimodal signals rather than expressions presented at their apex. Such a model of expression animation may increase the agent's communicative capabilities.

Another issue to be treated when modelling agents' displays of internal states is the expressive realisation of the behaviours. Different parameters have been defined as having an impact on the behaviour expressivity. This set of parameters is used in tasks such as animation synthesis, but also in the manual annotation of video corpora or automatic analysis.

Introduction

In this Chapter we present several algorithms that model expressive capabilities in embodied conversational agents. An embodied conversational agent (ECA) is a computer generated character with autonomous capacities in the verbal and non-verbal communication, that possesses a human-like appearance. Recent progress in the development of embodied conversational agents fosters expectations concerning their ability to express emotions in a credible way. Recent research in the field of emotional expressions in humans shows such expressions go beyond a simple facial expression presented at its apex. In fact, emotions recruit massive resources to cope with a relevant situation, leading to a high activation of the autonomous and somatic nervous system (Sander et al. 2005). This activation will have an impact on the behaviour of an individual: emotional cues can be observed in the speech parameters (Banse and Scherer 1996, Schuller et al. 2009), in posture (Coulson 2007) and so on (see *Chapter by Bänziger and Kaiser*). In general expressions should be seen as multimodal (see *Chapter by Scherer a*, *Chapter by Bänziger and Kaiser*; de Gelder et al. 2009) and composed of many signals (Levenson 2003, Shiota et al. 2003). Also the relation of each of these signals may contribute to the emotional interpretation of the overall expression (Scherer and Ellgring, 2007; Keltner 1995, Harrigan and O'Connell 1996). One may argue that a human-machine interaction should also include all these non-verbal signals emitted by the user and the embodied virtual interlocutor.

When working with virtual agents, it is important to keep in mind that the interaction of the humans with virtual characters is similar to the human-to-human interaction (Schilbach et al. 2006; Brave et al. 2005). Therefore, it seems plausible that for the emotional expression synthesis in ECAs it would be appropriate to apply a psychological model of human behaviour. Comprehending the processes implied in the emotional states is still a problem. Researchers involved in the animations of virtual characters rely on different models of emotional behaviour coming from distinct approaches. Classically, animators that seek emotions' definitions in the psychology domain see emotions in one of three major ways: in a discrete approach, that is as an automatic reaction to a situation, in a dimensional approach, i.e. as a state characterised by a position in a continuous multidimensional space, or in a componential approach as a dynamic cognitive evaluation of a situation. For a closer look at the mentioned emotional theories see *Chapter by Scherer a*. Different theories on emotions lead to different models of emotional behaviours in embodied conversational agents.

During a face-to-face interaction, whether between humans or in a human-machine context, an emotional message can be transmitted verbally or non-verbally. Since the first modelisation of a virtual face, work has been realised for the synthesis of emotional expressions (Parke 1972; Platt and Badler 1981). Studies have shown that agents that possess emotional expressivity faculties are engaging users more in the interactions than agents with non emotional behaviours (Walker, Sproull and Subramani 1994). In frustrating situations, agents that showed appropriate emotional reactions to users' situations also tended to diminish users' stress level, compared to an interaction with an agent showing no empathy (Mori, Prendinger and Ishizuka 2003; Prendinger, Mori and Ishizuka 2005). Similarly, when an agent has to interrupt the user in its task, the interruption is perceived as less frustrating when the agent's verbal messages are emotionally coloured (Picard and Liu 2007).
Although the relevance of full-body emotional activity has been stressed (e.g. Lazarus 1991, Argyle 1998; Bull 1987), some ECAs display no emotion expressions or only a limited number, mostly through the face. Usually facial expressions are defined using one of the two

standards of facial expression coding: the Facial Action Coding System (FACS; Ekman et al. 2002) and MPEG-4 (Ostermann 2002). FACS is an anatomically-based system that enables to describe the muscle movements that are perceived in the face. The minimum change that can be perceived is called an action unit (AU). The MPEG-4 is a graphic standard that defines the point displacements of the face and the joint rotation for the body.

In this Chapter different systems that enrich the expressive capabilities of an ECA are presented. First of all, several models of facial expressions have been proposed to extend the agents' facial behaviour. Most of them use some arithmetic operations like averaging facial parameters to generate new expressions. On the other hand, models of complex facial expressions like masking or superposition are usually modelled with fuzzy methods. Sometimes perceptive studies have been used to test the validity of the generated expressions and some models determined in that way have placed expressions in multidimensional spaces. Sequential expressions models have also been proposed by researchers. Some of these models were applied to the face signals, others to several modalities. Several models have also been developed in order to integrate several modalities of expressions into ECA animations. Finally, several agent systems exist that allow one to alter or modulate the way in which nonverbal behaviours are executed. These models are presented in the following Sections.

Facial expression models of emotion

Several models of facial expressions have been proposed to enrich the agents' facial behaviour. Most models use interpolations, mainly between pre-defined basic emotions' expressions. Other models are based on perceptive studies; mainly these have divided the facial area into different parts in order to generate the new expressions by reconstructing differentially from several pre-defined sub-expressions. A few models introduce some physiological changes, such as blushing or sweating into facial expression.

Whole face Interpolation models

The existing solutions usually compute new expressions by "averaging" the values of the parameters of the expressions of the "basic" emotions (see *Chapter by Scherer a*). The model called Emotion Disc (Ruttkay et al. 2003) uses a bi-linear interpolation between two basic expressions and the neutral one. In the *Emotion Disc* six expressions are spread evenly around the disc, while the neutral expression is represented by the centre of the circle. The distance from the centre of the circle represents the intensity of expression. The spatial relations are used to establish the expression corresponding to any point of the Emotion Disc.
Models of Tsapatsoulis et al. (Tsapatsoulis et al. 2002) and Albrecht et al. (Albrecht et al. 2005) can be used to compute expressions based on a similar approach. Both use the expressions of two "neighbouring" emotions to compute the facial expressions for non-basic emotions. For this purpose they use different multidimensional spaces, in which emotional labels are placed. In both approaches new expressions are constructed starting from the six Ekman's expressions: anger, disgust, fear, happiness, sadness, and surprise. To be more precise, in Tsapatsoulis *et al.* (2002) a new expression is generated by looking for the spatially closest two basic emotions as defined within the dimensional space proposed by Whissell (1989) and Plutchik (1980) and weighting the parameters of these expressions with their coordinates. Albrecht et al. (2005) proposed an extended approach. The authors use a three

dimensional space of emotional states defined by activation, evaluation, and power as proposed in (Cowie et al. 1999) and anatomical model of the face is used. As a consequence, they work with a numerical representation of muscle contradictions.

Paradiso (Paradiso 2002) introduced an algebraic structure for facial expression transformations with arbitrary chosen operators. More specifically, he used three operators defined over a set of MPEG-4 facial animation parameters. The sum operator is the weighted mean of two expressions, when the amplifier operator is a product of a real number and an expression. The probable intention of the first parameter was to simulate blends of emotions. Finally, the overlap operator combines two expressions by choosing some facial expression parameters of the first expression and others from the second one. Then by using these arithmetical operations different facial expressions can be generated.

Facial region compositional models

Several authors suggest that the intermediate expressions generated using interpolations may not be (perceptually) valid (e.g. Arya et al. 2009, Pelachaud and Poggi 2002). Compositional approaches combine separately different regions of facial expressions (mouth, eyebrows, etc.). Among others, Mäkäräinen and Takala (2009) rely on the blending of existing expressions to synthesise new ones. Their system enables the blending of facial actions defined partly based on the FACS and the MPEG-4 standard. New expressions can be computed by combining facial actions at the AUs level or at the expressions level. In the latest case, the generated blends can be of two or three basic emotion expressions. Rather than using additive operator, Bui (Bui 2004) uses a set of fuzzy rules to determine the blending expressions of six basic emotions based on Ekman's findings (Ekman and Friesen 1975). A subset of rules is attributed to each pair of emotions. The fuzzy inference determines the degree of muscle contractions of the final expression as a function of the input emotion intensities. Niewiadomski and Pelachaud (2007) have also used a partitioning approach based on fuzzy inferences, in which each facial expression was defined by a set of eight facial areas. Each part of the face displays an emotion. In complex facial expressions, different emotions can be expressed on different areas of the face (see *Chapter by Niewiadomski et al.*). The authors use an algorithm (Niewiadomski and Pelachaud 2007) based on fuzzy similarity to generate superposed, masked, fake or inhibited expressions.

A different type of facial expression differentiation was considered by Rehm and André (2005). In a study on deceptive agents, they showed that users were able to differentiate between the agent displaying an expression of a felt emotion versus an expression of a fake emotion (Rehm and André 2005). For this purpose they manually defined facial expressions according to Ekman's description of expressions for fake emotions. These expressions are more asymmetric and miss reliable features.

Arya and colleagues (Arya and DiPaola 2007; Arya et al. 2009) propose a perceptively valid model for expression blending. That is the authors aim to create new expressions which can be attached to a meaning (e.g. an emotional state). In a perceptive study human participants had to create facial expressions associated to mixed emotions onto a 3D face model. For this purpose they were asked to illustrate short stories with blending expressions. The expressions were evaluated by participants in terms of different dimensions. From the results of this study, a set of fuzzy rules that link specific facial actions with the 3D space of valence, arousal and agency has been developed. Rules are generated from the statistical analysis of the images created in the experiment by participants. Contrary to Bui whose fuzzy rules were activated depending on the intensity of emotions, in Arya et al. the fuzzy values in three emotional

dimensions are used to activate the virtual character's face. Interestingly the blending expression is a combination of the emotion expressions that are provided as input to the model.

Mao and colleagues (Mao et al. 2008) propose a layered model of facial behaviour. The authors stress that three processes contribute to any expressive behaviour: emotional, physiological and social. In Mao et al. the generation of the expression is realised in three layers, the first one being the physiological, the second emotional and the third social layer. At moment the authors consider: 14 physiological variables on physiological level (e.g. adrenaline, blood pressure or sneezing), 36 emotional expressions (e.g. fear, reproach or satisfaction) and six social expressions on social level (e.g. disagree or wink). The final facial behaviour is composed of the output of each layer processed separately, while taking in count the priorities given to each layer. In the case of the second layer (emotional expressions) the output is the result of the processing of the fuzzy relation matrix between expressions and emotions. This matrix contains the mapping from the fuzzy emotion vector to the fuzzy facial vector. Each value *(e, f)* in this matrix is a degree of membership expressing the probability that an emotion *e* is mapped to the expression *f*. Thus the mapping between emotions and expressions is many-to-many. Given an input of a vector of emotional states, the output is the fuzzy facial vector that is defuzzified. Working in parallel, the first layer may influence the way facial behaviour is realised while the last layer may facilitate or inhibit emotional expressions and/or use some social signals instead of the direct expression of an internal state. The layers define the hierarchical system. The output of each layer may be modified by the output of the layer which has a higher priority.

Modelling physiological changes

Other researchers (de Melo and Gratch 2009; Jung et al. 2009) have shifted their focus of attention on the visible expressions of vegetative functions controlled by the autonomic nervous system, such as blushing, sweating or weeping that accompany many emotional states (e.g. according to Levenson 2003). Thus de Melo and Gratch (2009) concentrated on the integration of blushing, sweating and wrinkles into the facial animations. The tearing and sweating animation relies on the modelling of water's properties and dynamics. The wrinkles' changes are synchronised with the muscular-based model of the face. An evaluation study shows that wrinkles and blushing add to the expressivity of anger, sweating to the expressivity of fear expressions, wrinkles to the surprised expression, wrinkles and tears to the sadness expression and blushing to the shame expression when coupled with the appropriate facial muscle expression (de Melo and Gratch 2009).

Multidimensional models based on perceptive studies

Several models of emotional behaviour are placed in the PAD model which is a three dimensional model defining emotions in terms of pleasure (P), arousal (A) and dominance (D) (Mehrabian and Russell 1980).
Among others, Zhang and colleagues proposed an approach for the synthesis of facial expressions from PAD values (Zhang et al. 2007). It allows for the generation of facial expressions of any emotional state that is described in a term three PAD variables. First, the authors proposed a new parameterisation of facial expressions: Partial Expression Parameters (PEPs). Similarly to MPEG-4, each PEP defines a facial movement in a specific area of the face. The main advantage is that it covers MPEG-4 space with a similar amount of details, while it relies on a more restricted number of parameters. A perceptive study evaluated how

their set of PEPs is linked to participants' attributions of P, A and D values. The validity of the expressions generated from PAD values was confirmed in an evaluation study, where participants had to attribute the PAD and emotional labels to the perceived animations (Zhang et al. 2007).

The same dimensional model was also used in a study where participants navigated in a PAD space with corresponding facial animations using a 3D control device (Courgeon et al. 2008). Eight expressions (fear, admiration, anger, joy, reproach, relief, distress, satisfaction) were attributed to the extreme points of the three dimensions (valence, activation and dominance) while an interpolation of FAPS values allowed for the generation of intermediate expressions. The movement in space was recorded through three dimensional joystick movements, which include its vertical rotation (Courgeon et al. 2008).

Another facial expression model was based on the Russell and Mehrabian three dimensional model which relies on reverse engineering (Boukricha et al. 2009). An empirical study enabled the authors to map a correspondence between randomly generated facial expressions composed of several action units as defined with FACS (Ekman, Friesen and Hager, 2002) and ratings in term of PAD values. These PAD ratings resulted from naive participants' evaluation of bipolar adjectives using a likert scale (Semantic Differential Measures of Emotional State or Characteristic (Trait) Emotions, as proposed in Mehrabian and Russell1974). The evaluated expressions were placed in the dimensional space, where Dominance takes one of two discrete values (high or low dominance) while Pleasure and Activation values are mapped into a continuous space. A facial expressions' control space is thus constructed with multivariate regressions, which enabled the authors to associate a facial expression to each point in the space.

A similar method was applied previously by Grammer and Oberzaucher (2006), whose work relies only on the two dimensions of pleasure and arousal. The authors also performed a perceptive study to place randomly generated facial expressions in the dimensional space and apply to the results a multiple multivariate regression, enabling the mapping between AUs and the two dimensions. Moreover the authors validated this model by checking the position of the six basic emotions in their 2D space. In their approach they partially tried to integrate different theories of facial expressions of emotions. It can be used for the creation of facial expressions relying on the action units defined in the FACS (Ekman et al. 2002) and situated in the dimensional space.

Sequenced expression models of emotions

While the previous described models deal with static facial expressions (i.e. expressions described at their apex), few models have been proposed for creating dynamic expressions. Some use the discrete approach but offer means to act on the temporal course of the expressions (Ruttkay 2001; Stoiber et al, 2009). Others are based on the appraisal approach (Paleari and Lisetti 2006; Malatesta et al. 2007). Some works consider the expressions of emotions as the result of temporal sequences of facial actions. They do not link explicitly these sequences to any appraisal checks.

Ruttkay (2001) proposed a system that allows the human designer to modify a facial expression animation defined par default by a trapezoid attack-hold-delay. The system permits, for any single facial parameter, to define manually the course of the animation. The plausibility of the final animation is assured by a set of constraints. The constraints are defined on the key-points of the animation and concern facial animation parameters. One can,

for example, force the facial expressions to be symmetric (i.e. all facial parameters have identical values for each key-point). Stoiber and colleagues (2009) propose another interface for the generation of facial expression of a virtual character. Using its 2D custom control space the user might deform both the geometry and the texture of a facial model. The approach is based on the principal component analysis of the images database showing a variety of facial expressions of one subject. It allows generating both realistic still images as fluent sequences of expressions but deprived of any psychological grounding.

Other researchers were inspired by the componential process model (CPM, see *Chapter by Scherer* for more information on this theory). Paleari and Lisetti (2006) and Malatesta et al. (2009) use manually defined sequential expressions inspired by the CPM (Scherer and Ellgring 2007). They consider a limited number of emotions and put the emphasis on the temporal relations between the different dynamic elements of an expression. The authors are also interested in the manner in which the elements are linked to the consecutive stages of cognitive evaluations predicted by the CPM. In Paleari and Lisetti's work (2006) the different facial action parameters are activated at different moments and the expression evolves through time. The final result is an animation consisting of a sequence of several facial movements expressing cognitive evaluations. In the work realised by Malatesta and colleagues (2007), the expressions of anger, disgust, fear, joy and sadness were generated manually according to Scherer's predictions and the focus was on the intensities and on the temporal constraints of facial signals. This work differs from Paleari and Lisetti's work (Paleari and Lisetti 2006) where each expression is derived from the addition of a new AU to the former ones. What is more, Malatesta and colleagues compared the additive approach with the sequential one. Results show a recognition rate well above chance level in the case of the additive approach, whereas the sequential approach gives recognition results only marginally above random choice (Malatesta and colleagues 2009).

Several studies show that emotional expressions are composed of signals arranged in a specific sequence (Scherer and Ellgring, 2007; Shiota et al. 2003; Keltner 1995, Harrigan et O'Connell 1996). Keltner, for example, showed that it is the temporal unfolding of the nonverbal behaviours that enables to differentiate the expressions of embarrassment and amusement, which in some studies (e.g. Edelman and Hampson, 1981, cited after Keltner 1995) tend to be confused by judges as they have a similar set of signals involving smiling, numerous sideway gaze and head shifts (Keltner 1995).

In behaviour animations in ECAs, Pan and colleagues (Pan et al. 2007) proposed an approach to display emotions that cannot be expressed by static facial expressions but that are expressed by certain sequences of signals (facial expressions and head movements). First of all, certain sequences of signals were extracted from a video-corpus. From this real data, Pan et al. built a directed graph (called a motion graph) in which the arcs are the observed sequences of signals and the nodes are possible transitions between them. Different paths in the graph correspond to different expressions of emotions. Thus, new animations can be generated by reordering the observed displays. Niewiadomski and colleagues have also proposed a system which allows an ECA to display multimodal expressions that respect sequentiality constraints defined in an algorithm. Their system enables the generation of expressions of any duration, while respecting the duration of individual signals composing it and their order of occurrence (Niewiadomski et al. 2009; see *Chapter by Niewiadomski et al.*).

Multimodal expression models of emotions

An emotion being a *dynamic episode* that produces a sequence of response patterns on the level of gestures, voice and face (Scherer and Ellgring, 2007), it is interesting to introduce more than one modality of emotional expression into agent animations. Especially that not only body movements tend to influence the interpretation of the facial expression (Meeren, van Heijnsbergen and de Gelder 2005) but also some of them seem to be specific to particular emotional states (e.g. Wallbott 1998; Pollick et al. 2001).

Very few multimodal behaviour models have been created so far. Clavel and colleagues (2009) studied the input of facial and posture in ECAs' emotional expressions. One study showed that the integration of the facial and postural changes into the ECAs' emotional behaviour affects users' overall perception of basic emotions, and have an impact on the attribution of the valence and activation values to the animations. The participants attributed the intended valence to the animations, whether these presented the face only, the posture only or the face and posture conditions, but not the intended activation. The activation was correctly attributed in the posture only condition, in the majority of the face and posture animations, but not in the face only. A second study shows an improvement of the emotion recognition when facial and postural changes are congruent. The authors observed that the judgments were mainly based on the information sent by the face, although adding congruent postures improves the interpretation of the facial expression (Clavel et al. 2009).

Michael Kipp proposed a system that automatically generates nonverbal behaviours that are synchronised with the verbal content in four modalities using a set of predefined rules (Kipp 2006). These rules determine the triggering conditions of each behaviour in function of the text. Thus a nonverbal behaviour can be triggered, for example, by a particular word, sequence of words, type of sentence (e.g. question) or when the agent starts a turn. The system offers also the possibility to discover new rules. Similarly Hofer and Shimodaira (2007) proposed an approach to generate head movements based on speech. Their system uses Hidden Markov Models to generate a sequence of behaviours. Data to train the model was manually annotated with four classes of behaviours: postural shifts, shakes and nods, pauses, and movement.

Lance and Marsella (2007) also explored head and body movements occurring in emotional displays and more particularly during gaze shifts. They placed their work in the theoretical context of the PAD dimensional model. Lance and Marsella defined a set of parameters enabling the differentiation of the multimodal emotional displays from the neutral ones based on extractions from the recordings of acted emotional displays. The gaze, head and body movements' data was captured through three motion sensors. Animations generated based on the motion captures were evaluated by human coders in terms of arousal and dominance. A set of proposed parameters contains temporal scaling and spatial transformations. Consequently, emotionally neutral displays of gaze, head and body movements can be transformed using this model into multimodal displays showing, for example, different levels of dominance and arousal.

Behaviour expressivity models

Given that the quality of behavior execution can be specific to or at least influenced by emotional states (Wallbott and Scherer 1986), several agent systems allow one to alter or modulate the way in which nonverbal behaviours are executed (Nayak 2005; Kipp 2006; Allbeck and Badler 2003; Neff and Fiume 2005). An agent could produce a nonverbal signal with straight and quick arm/hand movements and another signal with smooth, curved and slow movements. In some systems these movement characteristics are statically assigned to an agent. In other cases they depend on the agent's emotional state or personality. For

example, an excited agent could perform quick and rapid movements while a depressed agent - slow and heavy ones.

Allbeck and her colleagues created a system to select the most appropriate nonverbal behaviours (gestures and facial expressions) and control the movement quality of the Jack agent (Badler et al. 1993) depending on its personality and emotional state (Allbeck and Badler 2003). The user can give commands to the agent to make it conduct some actions. The system analyses the user input and selects the most appropriate behaviours that have to be performed to accomplish the given task. The way in which the agent performs its movements is influenced by a set of high level parameters, embedded in the Expressive Motion Engine (EMOTE). EMOTE is an implementation of the *Effort* and *Shape* movement components of the Laban Movement Analysis system (Laban and Lawrence 1974). The Effort component defines how a movement is modified in terms of: its *space* (relation with the surrounding space: indirect vs. direct), *weight* (impact of movement: light vs. strong), *time* (urgency of movement: sustained vs. sudden) and *flow* (control of movement: free vs. bound). The Shape component modifies the movement's coordination (how the mover's body parts change their relative position: for example, a contracted posture is obtained by bending the shoulders and torso toward the legs), direction (how movements are performed in relation to the environment) and shaping (how movements relate to the horizontal, vertical and sagittal planes: spreading vs. enclosing, rising vs. sinking, advancing vs. retreating). The EMOTE parameters are applied to the agent's movements depending on the agent's personality and emotional state. Allbeck *et al.* use two parametric models of emotion and personality, in which emotional states and personality traits are expressed as predefined distributions of the EMOTE parameters.

Neff et al. (2004; 2005) implemented some key movement properties by reviewing arts and literature, such as theatre and dance. They found that the body and the movement characteristics such as balance, body silhouette (contour of the body), position of torso and shoulder influence the way in which people perceive others. They have implemented three motion properties into animated characters: the pose of the character, the timing of movements and the transition from one pose to another. The shape is the character's body posture, which is expressed by some of its properties, like its tension, balance and extent. A posture can be relaxed or tensed, depending for example on the character's emotional state. By modifying the position of the character's centre of gravity relative to its feet, posture must be varied to keep balance. The extent is the quantity of space used by the character. To perform a movement, the arms for example can be fully stretched away from the body, or kept near to it. The timing is described as tempo and rhythm. The tempo is simply the speed with which we perform a movement. The rhythm corresponds to the movement pattern. The same rhythm can be performed with different tempos. The tempo is linked to the character's emotional states: for example sadness is communicated through slow movements, while joy is associated to small and quick ones. Movements' transitions can happen smoothly by keeping the same constant speed, or be interrupted and accelerated/decelerated. These different styles can reveal different character's emotional states and attitudes. For example interrupted movements can communicate hesitation and doubt, while accelerated movement can be used to give emphasis.

Conclusions

This Chapter presents different models of emotional expressions for ECAs. As ECAs are interactive per definition, the social and affective behaviour is crucial for their communication, which should not be limited to the verbal canal.

In the nonverbal domain of ECAs, some researchers use simple interpolations between a limited number of facial expressions, such as the basic emotion expressions, to generate an unlimited number of new ones. Some authors reached beyond the linearity between expressions in their expression generation by partitioning the face and applying mostly fuzzy methods or by grounding the facial generation in a multidimensional model. Work has also been realised on the integration of sequencing in the presentation of expressive behaviour and of its multimodality. Several systems have also been developed to modify the behaviour expressivity characteristics, such as quality of movements or frequency of an expressive feedback of an agent.

Recapitulating, the overall tendency in the research on emotional expressions of ECAs is to enrich a set of multimodal emotional signals. Temporal aspects of behaviours that contribute to emotional expressions should be analysed both in the framework of theoretical advances in the affective sciences and from real data studies, whether from observational or experimental settings. Emotional displays are defined by a set of signals; however these signals themselves have to be generated in synchrony across modalities, while respecting the constraints imposed by the sequence of appearance and by the modulation of their expressivity parameters in concordance with the specific states.

ACKNOWLEDGMENTS

References

Albrecht, I., Schröder, M., Haber, J., Seidel, H. (2005). Mixed feelings: expression of non-basic emotions in a muscle-based talking head. Virtual Reality 8 (4), 201-212.

Allbeck, J., Badler, N. (2003). Representing and parameterizing agent behaviours. In: H. Prendinger and M. Ishizuka, editors, Life-like Characters: Tools, Affective Functions and Applications. Springer, Germany.

Arya, A., DiPaola, S. (2007). Multi-Space Behavioural Model for Face-based Affective Social Agents," EURASIP Journal on Image and Video Processing, Special Issue on Facial Image Processing.

Arya, A., DiPaola, S. & Parush, A. (2009). Perceptually Valid Facial Expressions for Character-based Applications. To appear in: International Journal of Computer Game Technology.

Argyle, M. (1988). Bodily Communication. Methuen & Co., 2nd edition, London.

Badler, N., Phillips, C., Webber, B. (1993). Simulating Humans: Computer Graphics Animation and Control. Oxford University Press.

Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. Journal of Personality and Social Psychology, 70, 3, 614–636.

Boukricha, H., Wachsmuth,I., Hofstaetter, A., Grammer, K. (2009). Pleasure-Arousal-Dominance Driven Facial Expression Simulation. In: Proceedings of Third International Conference on Affective Computing and Intelligent Interaction (ACII 2009), 119-125. Amsterdam, The Netherlands.

Bull, P.E. (1987). Posture and Gesture. Oxford: Pergamon Books. Pergamon International Series in Experimental Social Psychology, vol. 16.

Becker, C., Wachsmuth, I., Prendinger, H., Ishizuka, M. (2005). Evaluating Affective Feedback of the 3D Agent Max in a Competitive Cards Game. In: Tao J., Tan T., Picard R. W. (Eds.), Proceedings of the International Conference on Affective Computing and Intelligent Interaction (ACII 2005), Springer, China.

Brave, S., Nass, C., Hutchinson, K. (2005). Computers that Care: Investigating the Effects of Orientation of Emotion Exhibited by an Embodied Computer Agent. International Journal of Human-Computer Studies, vol. 62, 161-178.

Bui, T.D. (2004). Creating emotions and facial expressions for embodied agents. Ph.D. thesis, University of Twente, Department of Computer Science, Twente, The Netherlands.

Clavel, C. Plessier J., Martin, J.-C. Ach L., Morel B. (2009). Combining Facial and Postural Expressions of Emotions in a Virtual Character. In: Proceedings of the 9th International Conference on Intelligent Virtual Agents (IVA 2009), Amsterdam, The Netherlands, 287-300.

Coulson, M. (2007). Attributing Emotion to Static Body Postures: Recognition Accuracy, Confusions, and Viewpoint Dependence. Journal of Nonverbal Behavior, 28, 2, 114-139.

Courgeon, M., Martin, J. C., Jacquemin, C. (2008). User's gestural exploration of different virtual agents' expressive profiles. AAMAS 2008, 1237-1240

Cowell, A. J., Stanney, K. M. (2005). Manipulation of non-verbal interaction style and demographic embodiment to increase anthropomorphic computer character credibility. International journal of human-computer studies, 62, 2, 281–306.

Cowie, R. Douglas-Cowie, E., Apolloni, B., Taylor, J., Romano, A., Fellenz, W. (1999). What a neural net needs to know about emotion words. In: N. Mastorakis, ed., Computational Intelligence and Applications. Word Scientific Engineering Society, pages 109–114, Society Press.

de Gelder, B., Van den Stock, J., Meeren, H.K.M., Sinke, C.B.A., Kret, M.E., Tamietto, M. (2009). Standing up for the body. Recent progress in uncovering the networks involved in processing bodies and bodily expressions. Neuroscience and Biobehavioral Reviews.

Ekman, P. (1972). Universals and Cultural Differences in Facial Expression of Emotion. In R. Cole (ed.), Nebraska Symposium on Motivation, University of Nebraska Press, Lincoln, Nebraska, 207-283.

Ekman, P. (2003a). Darwin, masking, and facial expression. Ann. N.Y. Acad. Sci. 1000, 205-221.

Ekman, P. (2003b). The Face Revealed. Weidenfeld & Nicolson, United Kingdom, London.

Ekman, P., Friesen, W. (1975). Unmasking the Face: A guide to recognizing emotions from facial clues. Prentice-Hall, Inc.

Ekman, P., Friesen, W.V., Hager, J.C. (2002). Facial Action Coding System (FACS). A Human Face, Research Nexus, Salt Lake City, USA.

Gallaher, P.E. (1992). Individual differences in nonverbal behaviour: Dimensions of style. Journal of Personality and Social Psychology, 63(1) 133-145.

Grammer, K., Oberzaucher, E. (2006). The Reconstruction of Facial Expressions in Embodied Systems: New Approaches to an Old Problem, ZIF Mitteilungen, 2, 14-31.

Harrigan, J.A., O'Connell, D. M (1996). How do you look when feeling anxious? Facial displays of anxiety. Personality and Individual Differences, volume 21, number 2, 205-212.

Hofer, G., Shimodaira, H. (2007). Automatic head motion prediction from speech data. In: Proc. Interspeech 2007, Antwerp, Belgium.

Jung, Y., Weber, C., Keil, J., Franke T. (2009). Real-Time Rendering of Skin Changes Caused by Emotions. In: Proceedings of 9th International Conference on Intelligent Virtual Agents (IVA 2009), Amsterdam, Holland.

Kaiser, S., Wehrle, T. (2001). Facial Expressions as Indicators of Appraisal Processes. In: K.R. Scherer, A. Schorr, T. Johnstone, (Eds.), Appraisal Processes in Emotions: Theory, Methods, Research, 285-300. New York: Oxford University Press.

Keltner, D. (1995). Signs of appeasement: Evidence for the distinct displays of embarrassment, amusement, and shame. In: P. Ekman & E.L. Rosenberg (Eds.), What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS). Series in affective science, 133-160. New York: Oxford University Press.

Keltner, D., Buswell, B.N. (1996). Evidence for the distinctness of embarrassment, shame, and guilt: A study of recalled antecedents and facial expressions of emotion. Cognition and Emotion, 10, 2, 155-172.

Keltner, D., Buswell, B.N. (1997). Embarrassment: its distinct form and appeasement functions, Psychological Bulletin, vol. 122, 250–270.

Kipp, M. (2006). Creativity Meets Automation: Combining Nonverbal Action Authoring with Rules and Machine Learning, In: Proceedings of 6th conference on Intelligent Virtual Agents, vol. 4133 of Lecture Notes in Computer Science, Springer, 230-242.

Laban, R., Lawrence, F.C. (1974). Effort: Economy in body movement. Plays, Inc, Boston.

Lance, B., Marsella, S. (2007). Emotionally expressive head and body movements during gaze shifts. In: Proceedings of the 7th International Conference on Intelligent Virtual Agents (IVA 2007), Springer 72–85.

Lazarus, R. S. (1991). Emotion and adaptation. New York: Oxford University Press.

Levenson, R. W. (2003). *Autonomic specificity and emotion*. In: Davidson, R. J., Scherer, K. R., Goldsmith, H. H. (Eds.), Handbook of affective sciences. New York.

Mäkäräinen, M., Takala, T. (2009). An Approach for Creating and Blending Synthetic Facial Expressions of Emotion. In: Proceedings of 9th International Conference on Intelligent Virtual Agents (IVA 2009), Amsterdam, Holland, 243-249

Malatesta, L., Raouzaiou, A., Karpouzis, K., Kollias, S. (2007). MPEG-4 Facial Expression Synthesis, Personal and Ubiquitous Computing, Special issue on Emerging Multimodal Interfaces, Springer.

Mancini, M., Pelachaud, C. (2008). Distinctiveness in multimodal behaviors. In: Proceedings of the 7th Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2008), Estoril, Portugal.

Mao, X., Xue, Y., Li, Z.,Bao, H. (2008). Layered Fuzzy Facial Expression Generation: Social, Emotional and Physiological. In: Affective Computing, Focus on Emotion Expression, Synthesis and Recognition, Or, J. (Eds.), I-Tech Education and Publishing, Vienna, Austria, 185-218.

Meeren, H. K. M., van Heijnsbergen, C. & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. Proceedings of the National Academy of Sciences of the USA, 102*, 45, 16518-16523.

Mehrabian, A. (1980). Basic Dimensions for a General Psychological Theory: Implications for Personality, Social, Environmental, and Developmental Studies, Oelgeschlager, Gunn & Hain, Cambridge, Mass.

Mehrabian, A. (1995). Framework for a comprehensive description and measurement of emotional states. Genetic, Social, and General Psychology Monographs, vol. 121, 3, 339-361.

Mehrabian, A., Russell, J. A. (1974). An approach to environmental psychology. MIT Press, Cambridge.

Mori, J., Prendinger, H., Ishizuka, M. (2003). Evaluation of an embodied conversational agent with affective behavior. In : Rosenschein, J. & Wooldridge, M. (Eds.) Proceedings of International joint Conference on Autonomous Agents and Multi-Agent Systems

(AAMAS 2003), Workshop on Embodied Conversational Characters as Individuals, 58–61, Melbourne, Australia.

Nayak, V. (2005). Emotional expressiveness through the body language of characters in interactive game environments. PhD thesis, Media Arts and Technology University of California, Santa Barbara.

Neff, M., Fiume, E. (2004). Artistically based computer generation of expressive motion. In: Proceedings of the AISB Symposium on Language, Speech and Gesture for Expressive Characters, 29-39.

Neff, M., Fiume, E. (2005). AER: Aesthetic Exploration and Refinement for Expressive Character Animation. In: Proceedings of the 2005 ACM SIGGRAPH/Eurographics symposium on Computer animation, 161-170, New York, NY, USA: ACM Press.

Niewiadomski, R., Pelachaud, C. (2007). Fuzzy Similarity of Facial Expressions of Embodied Agents. In: Proceedings of the 7th International Conference on Intelligent Virtual Agents, 86-98. Paris, France.

Niewiadomski, R., Hyniewska, S., Pelachaud, C. (2009). Evaluation of Multimodal Sequential Expressions of Emotions in ECA. In: Proceedings of Third International Conference on Affective Computing and Intelligent Interaction (ACII 2009), 635-641. Amsterdam, The Netherlands.

Ostermann, J. (2002). Face animation in MPEG-4. In: Pandzic, I.S. and R. Forchheimer (Eds.), MPEG-4 Facial Animation - The Standard Implementation and Applications, 17-55. Wiley, England.

Paleari, M., Lisetti, C.L. (2006). Psychologically Grounded Avatars Expressions. In: 1rst Workshop on Emotion And Computing at KI 2006, 29th Annual Conference on Artificial Intelligence, Bremen, Germany.

Pan, X. Gillies, M. Sezgin, T. M., Loscos, C. (2007). Expressing Complex Mental States through Facial Expressions. In: Proceedings of Second Affective Computing and Intelligent Interaction Conference (ACII 2007), Lecture Notes in Computer Science, Springer, 745-746.

Paradiso, A. (2002). An Algebra for combining MPEG-4 compliant Facial Animations, Proc. of Workshop Lifelike Animated Agents – Tools, Affective Functions, and Applications in conjunction with the Seventh Pacific Rim International Conference on Artificial Intelligence, Tokyo, Japan.

Parke, F. (1972). Computer generated animation of faces. In: ACM National Conference, vol. 1, 451–457, Salt Lake City, USA.

Pelachaud, C., Poggi, I. (2002). Subtleties of Facial Expressions in Embodied Agents, Journal of Visualization and Computer Animation, vol. 13, 301-312.

Picard, R., Liu, K. (2007). Relative Subjective Count and Assessment of Interruptive Technologies Applied to Mobile Monitoring of Stress. International Journal of Human-Computer Studies, 65, 396–375

Platt, S. M., Badler, N. I. (1981). Animating facial expressions, In: Proceedings of Conference on Computer graphics and interactive techniques (SIGGRAPH '81), ACM Press, 15, 245-252.

Pollick, F., Paterson, H.M., Bruderlin, A., Sanford, A.J. (2001). Perceiving Affect from Arm Movement, Cognition, vol. 82, 51-61.

Plutchik, R. (1980). Emotions: A psychoevolutionary Synthesis. Harper & Row, New York, NY, USA.

Prendinger, H., Mori, J. et Ishizuka, M. (2005). Using human physiology to evaluate subtle expressivity of a virtual quizmaster in a mathematical game. International Journal of Human-Computer Studies, 62, 231–245.

Rehm, M., André, E. (2005). Catch Me If You Can - Exploring Lying Agents in Social Settings, In: Dignum, F., Dignum, V., Koenig, S. Kraus, S., Singh, M. P., Wooldridge, M. (Eds.). In: Proceedings of International Joint Conference on Autonomous Agents And Multi-Agent Systems (AAMAS), ACM, Utrecht, The Netherlands, 937-944.

Russell J. (1980). A Circumplex Model of Affect. Journal of Personnality and Social Psychology, vol. 39, 1161-1178.

Ruttkay, Z. (2001). Constraint-based facial animation. Int. Journal of Constraints, vol. 6, 85-113.

Ruttkay, Z., Noot, H., ten Hagen, P. (2003). Emotion disc and emotion squares: Tools to explore the facial expression face. Computer Graphics Forum, 22, 1, 49-53.

Sander, D., Grandjean, D., Scherer, K. R. (2005). A systems approach to appraisal mechanisms in emotion. Neural Networks, 18, 317-352.

Scherer, K.R. (2001). Appraisal considered as a process of multilevel sequential checking. In: K.R. Scherer, A. Schorr, and T. Johnstone, editors, Appraisal Processes in Emotion: Theory, Methods, Research. Oxford University Press, 92–119.

Scherer, K.R., Ellgring, H. (2007). Multimodal Expression of Emotion: Affect Programs or Componential Appraisal Patterns? Emotion, vol. 7, 158-171.

Scherer, K.R, Wallbott, H.G. (1985). Analysis of nonverbal behaviour. Handbook of discourse analysis, 2, 199-230.

Schilbach, L., Wohlschläger, A., Krämer, N., Newen, N., Shah, A., Fink, G., Vogeley, K. (2006). Being with virtual others: neural correlates of social interaction. Neuropsycholgia, vol. 44, 718-730.

Schuller, B., Batliner, A., Steidl, S., Seppi, D. (2009) . Emotion Recognition from Speech: Putting ASR in the Loop. In: Proceedings o fthe Int. Conf. on Acoustics, Speech, and Signal Processing ICASSP 2009, Taipei, Taiwan, IEEE, 4585–4588.

Shiota, M. N., Campos, B., Keltner, D. (2003). The faces of positive emotion: prototype displays of awe, amusement, and pride. Annals of the New York Academy of Sciences, 1000, 296-299.

Stoiber, N., Séguier, R., Breton, G. (2009). Automatic design of a control interface for a synthetic face. IUI 2009, 207-216.

Tsapatsoulis, N., Raouzaiou, A., Kollias, S., Crowie, R., Douglas-Cowie, E. (2002). Emotion recognition and synthesis based on MPEG-4 FAPs. In: I. Pandzic, R. Forchheimer, R., (Eds.), MPEG-4 Facial Animation - The standard, implementations, applications. John Wiley & Sons, 141-168.

Whissell, C. (1989). The dictionary of affect in language. In: R. Plutchik, H. Kellerman, (Eds.), Emotion: Theory, Research, and Experience, vol. 4, The Measurement of Emotions, Academic Press. Inc., San Diego, 113-131.

Walker, J. H., Sproull, L., Subramani, R. (1994). Using a human face in an interface, Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence, Boston, Massachusetts, USA, 85-91.

Wallbott, H.G. (1998). Bodily expression of emotion. European Journal of Social Psychology, 28, 879-896.

Wallbott, H.G., Scherer, K. R. (1986). Cues and channels in emotion recognition. Journal of Personality and Social Psychology, 51, 4, 690-699.

Zhang, S., Wu, Z., Meng, H. M., Cai, L. (2007). Facial Expression Synthesis Using PAD Emotional Parameters for a Chinese Expressive Avatar. In: Proceedings of Second Affective Computing and Intelligent Interaction Conference (ACII 2007), Lecture Notes in Computer Science, Springer, 24-35.