Facial Expressions of Emotions for Virtual Characters

Magalie Ochs, Radoslaw Niewiadomski, and Catherine Pelachaud
CNRS-LTCI Télécom ParisTech
greta@telecom-paristech.fr

1. Introduction

Facial expressions convey information about emotional states, mood, intentions, stances, and so on. Even what seems a very simple signal such as head nod [18] or smile [37] can convey a large number of meanings. A slight change in their dynamism or morphology can be perceived by human observers and can be interpreted as transmitting different intentions and emotional states.

Embodied conversational agents (ECAs) are dialog partners to human users. As for human, they are endowed with human-like communicative capabilities. As such they ought to display a large repertoire of communicative and emotional behaviors. Thus when building a repertoire of nonverbal behaviors for virtual agents, one is faced with gathering an important variety of signals where subtle variation can alter their meaning (within a given discourse context). So one of the difficulties that needs to be addressed is to find the adequate level of signals description so it captures subtle variations while not over fitting the signals description. Another difficulty is to ensure the agent can convey multiple signals to convey any specific high level function. This is crucial for the agents not to appear to be too repetitive. So this issue concerns creating signals with variations without altering meaning. Another related difficulty is to ensure that facial expressions displayed by the agent correlate with events the agent is facing. This is particularly true when considering the appraisal theory that links the evaluation of an event and the unfolding of facial expressions. It is also important that human users interpret the created expressions as conveying specific messages. While it is necessary to be able to interpret the nonverbal behaviors of a virtual agent in a dialog context, it does not mean it should be highly readable. Indeed, highly recognizable expression may look too caricatural and may lose in naturalness.

The challenges above need to be addressed when building a repertoire of multimodal behaviors for ECAs. We review below various attempts, some relying on theoretical models, other on data analysis or even computational models.

2. Creation of lexicon of virtual character's facial expressions

To create a lexicon of virtual character's emotional facial expressions, two methods may be distinguished. A first method consists in exploiting the empirical and theoretical research in Human and Social Sciences on the characteristics of human's emotional faces (Section 2.1). A second method is based on the study of annotated corpus containing the expressions of emotions displayed by humans or virtual characters (Section 2.2).

2.1 Theoretical based lexicon of facial expressions

To create a repertoire of a virtual character's facial expressions, the method that is commonly used consists in exploiting the empirical and theoretical studies in Psychology that have highlighted the morphological and dynamic characteristics of human's facial expressions. Different theories lead to different approaches.

Categorical approach. Most of the computational models of virtual character's facial expressions are based on the *categorical approach* proposed by Ekman [15]. This approach is based on the hypothesis that humans categorize facial expressions of emotions into a number of categories similar across cultures: happy, fear, anger, surprise, disgust, and sadness (also known as the ``big six" basic emotions). Moreover, Ekman and his colleagues [16] have developed a system to describe human facial expressions, called *FACS* (Facial Action Coding System). This system is widely used in the domain of virtual characters to simulate emotional facial expressions. The Moving Pictures Experts Group MPEG-4 standards support facial animation by providing Facial Animation Parameters (FAPs) as well as a description of the expression of the six basic emotions [40].

Dimensional approach. To allow virtual characters to express a large number of emotional expressions, a *dimensional approach* was proposed by [REF]. In dimensional models, a new expression is often created by applying some arithmetical operations, such as linear interpolation, on numerical definitions of discrete emotions placed in the multi-dimensional space. For instance, the model called *Emotion Disc* [44] uses a bi-linear interpolation between two basic expressions and the neutral one. In this approach, six expressions are spread evenly around the disc, while the neutral expression is set at its center. The distance from the center of the circle and an expression represents its intensity. The spatial relations in 2D are used to establish the expression corresponding to any point of the Emotion Disc.

Two models by Tsapatsoulis and colleagues [50] and by Albrecht and colleagues [1} use a similar approach to compute new emotional displays. Both models use the expressions of two ``neighboring" basic emotions to compute a new facial expression. In Tsapatsoulis et al.'s [50] a new expression can be derived from a basic one by ``scaling" it, or by combining the spatially closest two basic emotions. In the latter case the parameters of these two expressions are weighted by their coordinates. Albrecht et al. [1] extend this approach by introducing a three dimensional space of emotional states defined by activation, evaluation, and power and an anatomical model of the face based on FACS [16].

Several other models of emotional behavior rely on a 3D space called PAD defining emotions in terms of pleasure (P), arousal (A) and dominance (D) [33]. The model proposed by Zhang and colleagues [53] is based on PAD and a new parameterization of facial expressions: Partial Expression Parameters (PEPs). Each PEP defines a facial movement in a specific area of the face. Compared to other existing parameterizations (e.g., MPEG-4 [40]), PEPs ensure a similar amount of details, while using less number of parameters. The authors linked PEPs with values of P, A and D by conducting an experimental study. The validity of the expressions generated from PAD values was further confirmed in an evaluation study, where participants had to attribute the PAD and emotional labels to several generated animations [53].

The same three dimensional model was also used in a study by Courgeon et al. [10] where participants navigated in a PAD space with corresponding facial animations using a 3D control device. 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 facial parameters defining an expression allowed for the generation of intermediate expressions [10].

The dimensional approach has the advantage allowing the generation of a large number of emotional facial expressions. However, the dynamic and the temporal characteristics of the expressions are generally not considered. Moreover, the large number of facial expressions poses the problem of the evaluation of all the generated emotional expressions.

Appraisal approach. Other models are based on an *appraisal approach* [45] such as Scherer's Componential Process Model [46]. This cognitive psychological approach considers that facial expressions of emotions reflect how an individual appraises and deals with his environment. In this approach, values of appraisal variables (e.g., novelty, intrinsic pleasantness, conduciveness and coping potential) are associated to the activation of action units (smallest units of perceptible facial activity defined in FACS).

Among others, Paleari and Lisetti [41] and Malatesta and colleagues [31] focus on the temporal relations between different facial actions predicted by the sequence of appraisal evaluations of the Scherer's model. In [41], the different facial actions are activated at different moments. The final animation that is generated on the virtual character's face is a sequence of several sub-expressions linked to the SECs cognitive evaluations.

In [31], the emotional expressions are created manually from sequences predicted by Scherer's theory. Differently from Paleari and Lisetti's work, each expression is derived by adding each new AU onto previous ones. What is more, Malatesta et al. [31] compare their additive approach with a sequential one. Results show an above chance level recognition in the case of the additive approach, and only marginally above random choice in the case of sequential approach [31].

Recently another partial implementation of the Scherer's model was proposed by Courgeon and colleagues [11]. In this model, the generation of facial expressions is directly driven by the evaluation of events appraised by the virtual character. For this purpose an appraisal module is implemented for a game-based scenario to associate to event the values of seven appraisal checks (expectedness, unpleasantness, goal hindrance, external causation, copying potential, immorality, and self-consistence). Four emotions are implemented and described by their appraisal profile: anger, sadness, guilt and joy. Expressions are generated at two levels. First, a temporary animation corresponding to the currently evaluated appraisal variable is displayed. When the evaluation of the event through all appraisal variables is finished, the system computes to which emotion corresponds the sequence of appraisal values and displays the corresponding full facial expression. The result of this evaluation can be one or more emotions. In the latter case the system displays a blend of emotions. The appraisal approach offers a detailed control on the single elements of facial expressions as well as on their dynamics (*i.e.* sequence). However, the appraisal theories are still incomplete regarding the facial action predictions. They are also complex and difficult to implement since the modeling of cognitive capabilities to infer the values of appraisal variables is required.

2.2 Corpus-based lexicon of facial expressions

To gather more subtle and natural expressions, other approaches are based on the analysis of annotated corpus of human or virtual faces.

2.2.1 Synthesis of emotional facial expressions from annotated human faces

To collect real data of persons expressing emotions, a first method consists in recording videos of actors having the instructions to express specific emotions. Another method consists in collecting spontaneous expressions by putting people in situations triggering various emotions. For instance, a common method to generate frustration is to simulate a bug in a computer program participants have to interact with. The second step is the annotation of the corpus to attribute labels to expressions, and to find out the morphological and dynamic characteristics of the emotion to create the lexicon of emotional facial expressions.

Based on an annotated corpus of humans expressing emotions, two approaches to synthesize virtual emotional faces have been explored. The facial expressions can be synthesized at a very low-level by retargeting the points tracked on a human face to a virtual mesh or, at the higher level, using copy-synthesis approach. In the later the virtual character's expressions are synthesized from the manual annotation of the human facial behavior.

The synthesized facial expressions are labeled and stored in the lexicon using low-level animation format such as MPEG4 [40] or FACS [16].

These two different approaches for building a lexicon of facial expressions were used in [35] to build a repository of laughs. The authors built a lexicon of facial expressions using high-level procedural animation synthesis from manual annotation and low-level data-driven animation synthesis based on optical motion capture system. The first approach consists of manually annotating the facial expressions using FACS coding [16]. Then the FACS based manual annotation of each episode is converted into Behavior Markup Language (BML) [51]. BML is an XML-like standard script language used to control the behavior of a virtual character, including the face. The other method uses machine learning algorithm and motion capture data. The 3D points of 27 markers are captured for each frame of the expression and then are retargeted to the virtual mesh using Temporal Restricted Boltzmann Machines [52]. In this approach the model was trained to find a mapping between the 84 dimensional space of input data and the 68 FAPs in MPEG-4. The model once trained, can be successfully applied to different data sources only with a minimum manual tuning [35].

These two approaches offer different degrees of flexibility and control over the expression and different levels of realism and precision of the movements. Motion capture based animation is usually richer in movements and consequently it may be perceived as more realistic. Also, the motion capture data permits maintaining the temporal and dynamic characteristics of the original expression. At the same time, optical motion capture system is invasive as markers need to be placed on the actors' face and may limit their spontaneous reactions. It is also resource and time consuming.

On the other hand, describing animation by sequences of action units allows one to control precisely an animation and its meaning (e.g., by adding or removing AU6, a marker of the Duchenne smile) but has all the weaknesses of procedural approaches to facial animation. The animation is poor in details and the dynamics of the movements is not very realistic.

2.2.2 User-perceptive approach for emotional facial expression synthesis

As highlighted in [17], most of these corpus-based studies on emotional facial expressions consider a *top-down approach*. An emotion label is attributed to each facial expression. This approach supposes that to each emotion corresponds a facial expression. However, each emotion may be represented by different facial expressions. Some researchers have explored other methods to investigate different facial expressions for each emotion type. For instance, in [47, 17], a *bottom-up approach* is proposed. In this approach, one to several action units selected randomly are activated on a virtual face. Observers rated the randomly generated facial expressions using two emotional dimensions (pleasure and arousal) and, in a second step, these two dimensions are mapped onto emotion types. This method has the advantage not to restrain the emotional facial expressions to a limit number of emotion types. Moreover, the lexicon of virtual character's emotional facial expressions is directly created based on human's perception.

More recently, Boukricha and colleagues [9] run a perceptive study to investigate the link between randomly generated facial expressions composed of several action units and their perception along the three dimensional space PAD [33]. 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 by Mehrabian and Russell (1974). The evaluated expressions were placed in the dimensional space, where Dominance takes one of two discrete values (high or low dominance) and where Pleasure and Activation values were mapped onto a continuous space. A facial expression control space was constructed from multivariate regressions. It consisted in mapping AUs and the dimensions allowing one to generate a facial expression to any point of the 3D space.

Following a user-perceptive approach, another method to create a repertoire of emotional facial expressions for virtual characters, consists in collecting a corpus of virtual character's expressions directly created by users. This method breaks with the traditional approach used to create repertoire of expressions: instead of asking people to label existing expressions, users are at the heart of the creation process of virtual character's expressions. This method has been first used to identify the morphological and dynamic characteristics of different types of smile (amused, polite and embarrassed smiles) [37]. A web application has been developed to enable a user to easily create different smiles on a virtual character's face (Figure 1).

Figure 1: Screenshot of the interface to create virtual character's facial expressions

Through radio buttons on an interface, the user could generate any smile by choosing a combination of seven parameters (amplitude of smile, duration of smile, mouth opening, symmetry of the lip corner, lip press, and velocity of the onset and offset of the smile). Two or three discrete values were considered for each of these parameters (for instance, small or large for the amplitude of the smile). When the user changes the value of one of the parameters, a virtual character shows automatically the corresponding animation. Considering all the possible combinations of the discrete values of the parameters, there are 192 different possible combinations, each one corresponding to a smile. Users were instructed to create one animation for each type of smile. The collected corpus contained 348 descriptions for each smile (amused, embarrassed, and polite). Based on this smile corpus and on a decision tree classification technique, an algorithm has been defined to determine the morphological and dynamic characteristics of the smile types [37]. As in the approach proposed in [17], the advantage of such a method is to consider, not only one single expression for each smile type but a variety of facial expressions. That enables one to increase the repertoire of the virtual character's expressions.

2.3 Facial expressions rendering

The past 10 years have seen many works on creating photorealistic skin rendering. Incredible results have been obtained that imitate extremely well digital photos of human faces. Several aspects of rendering need to be addressed. At first there is the skin rendering. Skin is a complex material which is translucent and partially reflects light. Wrinkles are also an important feature of realistic faces. Wrinkles can be due to muscular contraction due to expressions and can be static due to aging. Other communicative features of faces include tears, pallor and blushing.

The Emily project led by Debevec's group [2] aims to pass the Turing test in facial performance; i.e., it aims to reproduce with a very high degree of realism the rendering and the animation of faces. To this aim a sophisticated scan system called Light Stage was built; it is made of a dome of hundreds of LED that allows capturing the face of an actress with lighting coming from every direction. The LEDs can be modified to simulate various lighting conditions. From the captured images of the actress the subsurface and the specular reflections of the actress' face are separated. Moreover ambient occlusion corresponding to self-shadowing and inter-reflections along cavities of the nose, eyes and mouth corners is taken care of. The final rendering of the face is using a hybrid normal rendering algorithm [30]. Animation is obtained by capturing the actress doing various expressions and creating corresponding blend-shapes.

Stoiber et al. (2009) [48] developed an algorithm to reproduce facial rendering and animation with high resolution in real-time and with less heavy device than the previous model. A camera mounted on a helmet worn by the actor records her facial expressions. Facial expressions of a participant are recorded and tracked using contours in real-time. Using motion models the recorded facial expressions are reproduced onto synthetic models after some retargeting. The dynamism of the facial expressions is maintained: the motion models are learned on motion capture data and integrate how facial expressions dynamisms are non-linear and dependent on movement amplitude [48].

Jimenez and Guterriez [21] proposed a model that renders perceptually realistic human skin. To obtain real time, their idea was to translate the simulation of subsurface scattering effects from texture to screen space. Later on the authors added expressive wrinkles onto their model [20]. Wrinkles are designed as normal maps that are added to base normal maps and blend shapes in a weighted manner. To validate that their approximation did not introduce loss in realism, the authors conducted perceptual studies. Participants viewed images of faces rendered with different lighting conditions; they had to choose which faces were the closest to real human faces.

Other models of wrinkles have been proposed. They can be gathered into two main approaches: geometry and texture based methods. While geometric methods simulate dynamic wrinkles by deforming directly mesh geometry [10], texture based methods use bump mapping [8]. For instance, Niewiadomski et al. [34] use the texture based technique called Screen Space Bump Wrinkle to model wrinkles. In this approach the surface normal vector is modified before the lighting computation, thus using the new bent normal in the render process gives visually satisfying result without changing the surface geometry. Simulation of wrinkle effects is performed by computing the perturbed Normal vector in screen space with Pixel Shader. Thus the complexity of the computation only depends on the number of pixels, and not on the number of vertices of the facial model. Twelve groups of wrinkles related to different AUs are defined as texture. In runtime, when an action unit is activated on the face mesh, the GPU receives its intensity value and computes the corresponding wrinkles. The final result is the composition of all active wrinkles (see Figures 2 and 3).

Figure 2: Wrinkles for the raised eyebrows movement (AU1+AU2)

Figure 3: Wrinkles corresponding to action AU4 and AU6

de Melo and Gratch [13] integrate not only wrinkles but also blushing, tears and sweating effects into facial rendering. Wrinkles are modeled using bump mapping and, then, are synchronized with the muscular based facial model. Interestingly, the simulated wrinkles are copied from the pictures of a human displaying the respective wrinkle configuration. Photographed wrinkles are then converted to grayscale, blurred and applied onto the virtual human texture. The tearing and sweating animation are also modeled through bump mapping and relies on the modeling of water's properties and dynamics.
Works by [28, 10] are examples of geometry deformation approaches. Mesh editing tools are used to define control points which are perpendicular to the wrinkles, and to define the influence regions associated to each wrinkling curve. The wrinkling behavior is controlled through a set of parameters that specifies the way the mesh deforms.

3. The virtual character's emotional facial expressions in interaction

The virtual character's emotional facial expressions are generally constructed without considering the context of the interaction. To take advantage of a virtual character displaying emotions, the emotional facial expressions should be in accordance with the situation in the interaction.

3.1 Expressions of emotions in context

Several researches have shown that virtual character expressing emotions enhance human-machine interaction. One well-known effect of the expressions of emotions is the increasing of the virtual character's believability by creating an *illusion of life* [4, 49]. Concerning the effects of the virtual character's emotions on the users, it has been shown that emotions have generally positive effects, for instance on user's perception of the virtual character [32] on user's satisfaction [19], and to create a good relationship with the virtual character [7]. However, some studies [5, 39] have highlighted the importance of the *social context*. The social context includes the situation in which the user and the virtual characters are (place, actions, etc.), the social roles of the participants of the interaction, the cultural context, and the social norms [43]. For instance, as shown in [5], an emotional virtual character may have a different impact on users depending on its social role. For example, an emotional learning companion leads to better effects on users' perception than an emotional tutor [5]. The situation in which the emotions are expressed, i.e. when and which emotions are displayed during the interaction, plays an important role on users' behavior. Some emotional behaviors, such as the expression of empathy, seem to enhance the interaction whereas the expressions of self-emotions may have few impacts [5]. Moreover, an emotion expressed in an inappropriate situation may even have negative effects on the interaction, for instance by deteriorating users' perception of virtual characters [39].

To automatically compute the emotions that a virtual character should express, some existing tools can be used. For instance, the open-source computational model FATIMA (FearnotAffecTIve Mind Architecture) [22] computes the emotions of a virtual character elicited by events occurring in the environment, considering the influence of the character's personality, social relations, culture, and empathy. Others computational model of emotions have been developed (for more details, see Chapter 5 by Gratch and Marsella). Such a model of emotions may compute not only one emotion type but several ones. Indeed, an event may elicit a sad emotion in the virtual character, but also a little bit of surprise, and at the same time, the social norms indicate that in this situation the virtual character should express joy. This kind of situation occurs every day in human life. To create emotional reflexive agent, and not an impulsive one with an emotional behavior similar to a child, a virtual character should convey these emotional subtleties [38]. Consequently, the expressions of emotions might result in a combination of several emotion types. Several computational models have been proposed for the synthesis of blending of emotions (for instance [1, 38]).

3.2 Perceptive studies of emotional virtual characters

Virtual character's facial expressions of emotions may have significant impacts on the interaction, and more particularly on users' behavior. A same facial expression of emotion may lead to positive or negative effects depending on the social context in which the emotion is displayed. Consequently, an important step is the evaluation of the virtual character's expressing emotions.

Perceptive studies of emotional virtual characters may be considered at two levels:

- a *context-free level* of evaluation: only the perception of the emotional facial expressions is evaluated without considering any information about the context;
- a *in-context level* of evaluation: the perception of a virtual character expressing emotions is evaluated in a particular context of interaction.

Context-free level of evaluation. In a *context-free level* of evaluation, the objective is to validate that the emotional facial expressions are recognized with the expected intensities and types (or dimensions such as pleasure and dominance as proposed in [17]). The context of the interaction is not considered. The method generally used to perform such an evaluation is to present videos of virtual characters expressing the emotions and to ask users to indicate the recognized emotion types and intensity through forced choice questionnaire.

To capture the uncertainty of the users on the recognized emotions, Likert scales (with for instance different levels of agreement: Strongly disagree, Disagree, Slightly disagree, Neither agree nor disagree, Slightly agree, Agree, Strongly agree) could be used to collect the users' responses.

Several perceptive studies at a context-free level have highlighted the role of the dynamics, intensity and rendering of synthesized facial expressions of virtual characters on the user's perception. For instance, Katsyri and Sams [24] showed that synthetic dynamic expressions were identified better than static ones only for expressions whose static displays were not similar. In a similar study of Noëland et al. [36], the effect of the dynamics was however not observed. Bartneck and Reichenbach's work [3] shows that the higher intensity expressions were better recognized. In [10], they have shown that the application of wrinkles increases the agent's expressivity but does not improve the recognition. Also in de Melo [13], wrinkles, blushing and sweating add to the expressivity of some stereotypical (basic) expressions such as anger, fear or sadness.

Keltner [25] shows that expressions of emotion are rather a sequence of facial actions than full-blown single-shot displays. The later ones occur rarely in real-life interactions. Single facial actions are ambiguous as they can be the components of different facial expressions. The action unit identification on virtual faces can be more challenging than the identification of the stereotypical full-blown expressions. While the latter can be easily identified from a subset of features (i.e. some information is redundant) the identification of atomic facial actions is local and it requires that the attention is made on details. Nevertheless it can be supposed, following for example the appraisal theory, that the apparition of each action unit in the sequence is significant and meaningful, thus, if some of these are not properly identified the meaning of a facial expression can be altered. The work presented in [34] is an example of perceptive study at a context-free level that focuses on the perception of single facial actions rather that on stereotypical expressions. The results have shown that single facial actions are better identified when they are dynamic and with higher intensity. On the other hand, intense expressions of single facial actions are perceived less natural and less realistic. Finally, wrinkles did not improve significantly the identification of facial actions [34].

Even if the evaluation at a context-free level is done out of the context of an interaction, some elements of the social context may impact users' perception. The gender of the virtual character as well as the gender of the users rated the expression might lead to different perception. For instance, as shown in [27, 23], women are more sensitive to non-verbal signs and more able to decode facial expressions cues, even for virtual characters' faces. Moreover, women make more extreme judgment ratings than men when decoding facial expressions.

In-context level of evaluation. One main limit with the evaluation presented above is the lack of interactivity with the user. Indeed, the user remains passive since she is not involved in the conversation with the virtual character. The *in-context level* of evaluation consists in studying the effects of emotions expressions on the overall interaction. The objective, in this case, is to measure the benefits of the emotional virtual character on the user who is involved in an interaction with the virtual character. More than an evaluation of the facial expressions themselves, in the *in-context level*, it is the emotional behavior of the virtual character (i.e. when and which emotions are displayed during the interaction) that is analyzed. As highlighted in the previous section, this kind of evaluation is all the more important since the effects of emotion expressions may vary from positive to negative depending in the circumstances in which they are expressed.

To measure the effects of a virtual character's emotional behavior, users generally interact with a virtual character in, at least, two conditions: a condition in which the virtual character does not express any emotion (control condition), and a condition in which users interact with the virtual character expressing emotions. A questionnaire at the end of the interaction is generally used to collect the users' overall perception of the virtual character and/or of the interaction. Objective measures may also be collected to analyze the effects of emotions. For instance, physiological sensors may be used to study the emotional reaction of users [6]. The analysis of users' performances (i.e. score of test or ability to recall information) may enable one to study the benefits of an emotional character on task achievement [32]. As proposed in [26], the length users interact with a virtual character may be an indication of users' engagement.

Several studies evaluate the appropriateness of the facial behavior to the context of an interaction. In the experiment by Walker, Sproull, and Subramani (1994) people liked the facial interface that displayed a negative expression less than the one which showed a neutral expression. However, in a card game the agent that displayed only positive expressions, irrespectively of the event, was evaluated less ``human being'' than the one that also expressed negative emotions [6]. In Rehm and André [42], the agent expressing emotions was compared with the agent showing additionally subtle expressions of deception. The agent with deceptive facial expressions was perceived as less credible and less trustworthy. In Lim and Aylett [29] study on interactive guide using appropriate emotional displays was perceived to be more believable, natural, and interesting than the agent without emotional displays. These results suggest that the choice of emotional displays influences the perception of the agent. They also highlight the role of the context in the judgment.

4. Conclusion

In conclusion, the emotional facial expressions of virtual characters are generally created with the assumption that virtual characters should display emotions as humans do, i.e. with the same morphological and dynamic characteristics of the face. Most of the models of emotional facial expressions are often based on either empirical or theoretical research in Human and Social Sciences. However, computational models may enable us to go beyond these methodologies by analyzing automatically generated facial expressions that humans may not simulate well on demand (i.e. during corpus generation with actors), and then that may be difficult to study. It allows increasing the variability in expressions for the virtual characters.

During human-machine interaction, the same emotional facial expression of a virtual character may have different effect, from positive to negative, on the user's perception depending on the situation in which the emotion is expressed. The display of emotions should be appropriate or plausible in the situation of the interaction. As highlighted in [14], the emotional facial expressions are appropriate if they meet expectations of what one is supposed to feel in a given situation. However, an emotional expression may be inappropriate but plausible when the expression is displayed in a situation even if the expression is not the appropriate one. The recent work of de Melo et al. [12] shows that the user applies "reverse appraisal" to interpret the virtual character's emotional expression and then to deduce information from virtual character's facial expressions regarding for instance its goal conduciveness. An emotional expression may then be displayed depending on the values of the appraisal variables the virtual character wants to convey to the user.

Acknowledgment. This research has been supported by the European Community Seventh Framework Program (FP7/2007-2013), under grant agreement no. 231287 (SSPNet) and under grant agreement no. 270780 (HILAIRE).

References

[1] I. Albrecht, M. Schroder, J. Haber, and H. Seidel. Mixed feelings: expression of non-basic emotions in a muscle-based talking head. Special issue of Journal of Virtual Reality on Language, Speech and Gesture, 8(4):201 - 212, 2005.

[2] O. Alexander, M. Rogers, W. Lambeth, J.Y. Chiang, W.C. Ma, C.C.Wang, and P.E. Debevec. The digital emily project: Achieving a photorealistic digital actor. IEEE Computer Graphics and Applications, 30(4):20-31, 2010.

[3] C. Bartneck and J. Reichenbach. Subtle emotional expressions of synthetic characters. International Journal Human-Computer Studies, 62(3):179-192, 2005.

[4] J. Bates. The role of emotion in believable agents. Communication of ACM, 37(7):122-125, July 1994.

[5] R. Beale and C. Creed. Affective interaction: How emotional agents affect users. International Journal of Human-Computer Studies, 67(9):755-776, 2009.

[6] C. Becker, I.Wachsmuth, H. Prendinger, and M. Ishizuka. Evaluating affective feedback of the 3D agent max in a competitive cards game. In Jianhua Tao, Tieniu Tan, and Rosalind W. Picard, editors, Proceedings of the International Conference on Affective Computing and Intelligent Interaction (ACII). Springer, 2005.

[7] T.W. Bickmore and R.W. Picard. Establishing and maintaining long-term human-computer relationships. ACM Transaction on Computer-Human Interaction (TOCHI), 12(2):293-327, 2005.

[8] J.F. Blinn. Simulation of wrinkled surfaces. In Proceedings of SIGGRAPH '78, pages 286-292, NY, USA, 1978. ACM.

[9] Boukricha, Wachsmuth H., Hofstaetter I., and K. A., Grammer. Pleasure arousal dominance driven facial expression simulation. In Proceedings of Third International Conference on Affective Computing and Intelligent Interaction (ACII 2009), pages 119-125, Amsterdam, The Netherlands, 2009.

[10] M. Courgeon, S. Buisine, and J.-C. Martin. Impact of expressive wrinkles on perception of a virtual character's facial expressions of emotions. In The 9th International Conference on Intelligent Virtual Agents, pages 201-214, Amsterdam, Holland, 2009.

[11] M. Courgeon, C. Clavel, and J.-C. Martin. Appraising emotional events during a real-time interactive game. In G. Castellano, J-C. Martin, J. Murray, K. Karpouzis, and C. Peters, editors, In Proceedings of the International Workshop on Affective-Aware Virtual Agents and Social Robots (AFFINE '09), New York, NY, USA, 2009. ACM.

[12] C. de Melo, P. Carnevale, and J. Gratch. The effect of virtual agent's emotion displays and appraisals on people's decision making in negotiation. In the 12th International Conference on Intelligent Virtual Agents, pages 53-66, Santa Cruz, USA, 2012.

[13] C. de Melo and J. Gratch. Expression of emotions using wrinkles, blushing, sweating and tears. In The 9th International Conference on Intelligent Virtual Agents, pages 188-200, Amsterdam, Holland, 2009.

[14] V. Demeure, R. Niewiadomski, and C. Pelachaud. How believability of virtual agent is related to warmth, competence, personification and embodiment? MIT Presence, 20(5):431-448, 2011.

[15] P. Ekman and W. V. Friesen. Unmasking the Face. A guide to recognizing emotions from facial clues. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1975.

[16] P. Ekman, W. V. Friesen, and J. C. Hager. The facial action coding system. Weidenfeld and Nicolson, 2002.

[17] K. Grammer and E. Oberzaucher. The reconstruction of facial expressions in embodied systems. ZiF : Mitteilungen, 2, 2006.

[18] D. Heylen. Head gestures, gaze and the principles of conversational structure. I. J. Humanoid Robotics, 3(3):241-267, 2006.

[19] K. Hone. Empathic agents to reduce user frustration: The effects of varying agent characteristics. Interacting with Computer, 18(2):227-245, March 2006.

[20] J. Jimenez, J.I. Echevarria, C. Oat, and D. Gutierrez. GPU Pro 2, chapter Practical and Realistic Facial Wrinkles Animation, pages 15-27. AK Peters Ltd., 2011.

[21] J. Jimenez, D. Whelan, V. Sundstedt, and D. Gutierrez. Real-time realistic skin translucency. Computer Graphics and Applications, IEEE, 30(4):32-41, 2010.

[22] D. João and A. Paiva. Feeling and reasoning: A computational model for emotional characters. In Carlos Bento, Amílcar Cardoso, and Gaël Dias, editors, Progress in Artificial Intelligence, volume 3808 of Lecture Notes in Computer Science, pages 127-140. Springer Berlin / Heidelberg, 2005.

[23] M. Katsikitis, I. Pilowsky, and J. M. Innes. Encoding and decoding of facial expression. Journal of General Psychology, 124:357-370, 1997.

[24] J. Katsyri and M. Sams. The effect of dynamics on identifying basic emotions from synthetic and natural faces. International Journal Human-Computer Studies, 66(4):233-242, 2008.

[25] D. Keltner. Signs of appeasement: Evidence for the distinct displays of embarrassment, amusement, and shame. Journal of Personality and Social Psychology, 68:441-454, 1995.

[26] J. Klein, Y. Moon, and R.W. Picard. This computer responds to user frustration. In Proceedings of the Conference on Human Factors in Computing Systems, pages 242-243, Pittsburgh, USA, may 1999. ACM Press.

[27] E. Krumhuber, A. Manstead, and A. Kappas. Temporal aspects of facial displays in person and expression perception. The effects of smile dynamics, head tilt and gender. Journal of Nonverbal Behavior, 31:39-56, 2007.

[28] C. Larboulette and M.-P. Cani. Real-time dynamic wrinkles. In Proceedings of CGI'04. IEEE, 2004.

[29] M.Y Lim and R. Aylett. Feel the difference: A guide with attitude! In The 7th International Conference on Intelligent Virtual Agents, pages 17-19, Paris, France, 2007.

[30] W.C. Ma, T. Hawkins, P. Peers, C.F. Chabert, M. Weiss, and P. Debevec. Rapid acquisition of specular and diffuse normal maps from polarized spherical gradient illumination. In Proceedings of the 18th Eurographics conference on Rendering Techniques, EGSR'07, pages 183-194, 2007.

[31] L. Malatesta, A. Raouzaiou, K. Karpouzis, and S.D. Kollias. Towards modeling embodied conversational agent character profiles using appraisal theory predictions in expression synthesis. Applied Intelligence, 30(1):58-64, 2009.

[32] H. Maldonado, J.R. Lee, S. Brave, C. Nass, H. Nakajima, R. Yamada, K. Iwamura, and Y. Morishima. We learn better together: enhancing elearning with emotional characters. In T. Koschmann, D. Suthers, and T.W. Chan, editors, Computer Supported Collaborative Learning 2005: The Next 10 Years!, pages 408-417. Lawrence Erlbaum Associates, Mahwah, NJ, 2004.

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

[34] R. Niewiadomski, J. Huang, and C. Pelachaud. Effect of facial cues on identification. In Proceedings of the 25th Annual Conference on Computer Animation and Social Agents (CASA 2012), pages 37-44, Singapore, 2012.

[35] R. Niewiadomski and C. Pelachaud. Towards multimodal expression of laughter. In The 12th International Conference on Intelligent Virtual Agents, pages 231-244, Santa Cruz, USA, 2012.

[36] S. Noëland, Dumoulin, T. Whalen, and J. Stewart. Recognizing emotions on static and animated avatar faces. In HAVE 2006, pages 99-104, Ottawa, Canada, 2006.

[37] M. Ochs, R. Niewiadomski, P. Brunet, and C. Pelachaud. Smiling virtual agent in social context. Cognitive Processing, Special Issue on Social Agents, pages 1-14, 2011.

[38] M. Ochs, R. Niewiadomski, C. Pelachaud, and D. Sadek. Intelligent expressions of emotions. In Affective Computing and Intelligent Interaction, pages 707-714. Springer, 2005.

[39] M. Ochs, C. Pelachaud, and D. Sadek. An empathic virtual dialog agent to improve human-machine interaction. In Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems - pages 89-96, 2008.

[40] J. Ostermann. Face animation in mpeg-4. In I.S. Pandzic and R. Forchheimer, editors, MPEG-4 Facial Animation - The Standard Implementation and Applications, pages 17-55. Wiley, England, 2002.

[41] M. Paleari and C. Lisetti. Psychologically grounded avatars expressions. In 1rst Workshop on Emotion and Computing at KI 2006, 29th Annual Conference on Artificial Intelligence, pages 14-19, 2006, Bremen, Germany,

[42] M. Rehm and E. André. Informing the design of embodied conversational agents by analyzing multimodal politeness behaviors in human-human communication. In Workshop on Conversational Informatics for Supporting Social Intelligence and Interaction, 2005.

[43] L.D. Riek and P. Robinson. Challenges and opportunities in building socially intelligent machines. IEEE Signal Processing, 28(3):146-149, 2011.

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

[45] K. Scherer, A. Schorr, and T. Johnstone. Appraisal Processes in Emotion: Theory, Methods, Research. Oxford University Press, 2001.

[46] K.R. Scherer. 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, pages 92-119. 2001.

[47] J. Snodgrass. Judgment of feeling states from facial behavior: A bottom-up approach. PhD thesis, University of British Columbia, 1992.

[48] N. Stoiber, G. Breton, and R. Séguier. Modeling short-term dynamics and variability for realistic interactive facial animation. IEEE Computer Graphics and Applications, 30(4):51-61, 2010.

[49] F. Thomas and O. Johnston. Disney Animation: The illusion of life. Abbeville Press, 1981.

[50] N. Tsapatsoulis, A. Raouzaiou, S. Kollias, R. Crowie, and E. Douglas- Cowie. Emotion recognition and synthesis based on mpeg-4 faps. In I.S. Pandzic and R. Forchheimer, editors, MPEG-4 Facial Animation - The standard, implementations, applications, pages 141-168. John Wiley & Sons, 2002.

[51] H. Vilhjalmsson, N. Cantelmo, J. Cassell, N. Ech-Chafai, M. Kipp, S. Kopp, M. Mancini, S. Marsella, A.N. Marshall, C. Pelachaud, Z. Ruttkay, K.R. Thórisson, H. van Welbergen, and R.J. van der Werf. The behavior markup language: Recent developments and challenges. In Catherine Pelachaud, Jean-Claude Martin, Elisabeth André, Gérard Chollet, Kostas Karpouzis, and Danielle Pelé, editors, IVA, volume LNAI 4722 of Lecture Notes in Computer Science, pages 99-111. Springer, 2007.

[52] M.D. Zeiler, G.W. Taylor, L. Sigal, I. Matthews, and R. Fergus. Facial expression transfer with input-output temporal restricted boltzmann machines. In Proceedings of the conference on Neural Information Processing Systems Foundation (NIPS), 2011.

[53] S. Zhang, Z. Wu, H. M. Meng, and L. Cai. Facial expression synthesis using PAD emotional parameters for a chinese expressive avatar. In Proceedings of Second Affective Computing and Intelligent Interaction Conference (ACII 2007), pages 24-35. Springer, 2007.