Doi: 10.1109/ACII.2009.5349569.

Evaluation of Multimodal Sequential Expressions of Emotions in ECA

Radosław Niewiadomski
Telecom ParisTech
37/39, rue Dareau - 75014 Paris, France
niewiado@telecom-paristech.fr

Sylwia Hyniewska
Telecom ParisTech
37/39, rue Dareau - 75014 Paris, France
hyniewska@telecom-paristech.fr

Catherine Pelachaud
CNRS-LTCI
37/39, rue Dareau - 75014 Paris, France
pelachau@telecom-paristech.fr

Abstract

A model of multimodal sequential expressions of emotion for an Embodied Conversational Agent was developed. The model is based on video annotations and on descriptions found in the literature. A language has been derived to describe expressions of emotions as a sequence of facial and body movement signals. An evaluation study of our model is presented in this paper. Animations of 8 sequential expressions corresponding to the emotions - anger, anxiety, cheerfulness, embarrassment, panic fear, pride, relief, and tension - were realized with our model. The recognition rate of these expressions is higher than the chance level making us believe that our model is able to generate recognizable expressions of emotions, even for the emotional expressions not considered to be universally recognized.

1. Introduction

The importance of emotional displays in embodied conversational agents (ECAs) was often stressed. The agent displaying emotional states is perceived as more engaging [26]. The emotional expressions were found to be useful in understanding ambiguous text messages [8] and in increasing user's attention [26]. They also influence the credibility and trustworthiness of an ECA [20]. The research on emotional expressions was strongly influenced by the work of Paul Ekman and his colleagues [6] who postulate the existence of universal emotional expressions that can be described at the apex. They described six such expressions displayed on the face: anger, disgust, joy, fear, sadness, surprise. Recent studies show that several emotions are expressed by a set or a sequence of different nonverbal behaviors which are arranged in a certain interval of time (e.g. [11, 24]) rather than by a static facial expression. The expressions of emotional states are dynamic and can be displayed over different modalities like face, gaze and head movement [11], gestures [11], or posture [19, 27]. Interestingly, these signals do not have to occur simultaneously [11].

For the purpose of this work we called these expressions *multimodal sequential expressions of emotions*. They may be composed of nonverbal behaviors displayed over different modalities, of a sequence of behaviors, or of expressions that change dynamically within one modality. A system that generates multimodal sequential expressions has recently been proposed in [16]. It is composed of a language that describes such expressions (from annotating real data) and of the algorithm that uses such a description to generate automatically the expressions of emotions. In this paper we present a first evaluation of the system. We are particularly interested in showing that multimodal sequential expressions are useful in the communication of emotions, even of emotions others than the universally recognized ones e.g. pride or relief.

In this evaluation we considered eight different emotional states: anger, anxiety, cheerfulness, embarrassment, panic fear, pride, relief, and tension. To create believable multimodal expressions for each emotion to be displayed by the agent, data is needed on the sequence of appearance of different behaviors. We took two approaches to reach this aim. On one hand we gathered data from the literature and on the other hand we annotated manually visual data.

The remaining part of this paper is structured as follows. The next section is dedicated to an overview of the computational models on multimodal sequential expressions. Then section 3 describes research of emotional displays as well as lists the results of the emotional displays evaluations. Section 4 describes the annotation process that was

978-1-4244-4799-2/09/$25.00 ©2009 IEEE

realized for the purpose of this study, while section 5 is a brief introduction to how the computation of multimodal sequential expressions is realized. The results of our evaluation study of multimodal sequential expressions are presented in section 6 and section 7 concludes the paper.

2. State of art

Several systems exist that use more complex displays to communicate emotions than facial expressions at their apex. Xueni Pan et al. [17] proposed an approach to display emotions that cannot be expressed by static facial expressions but are expressed by certain sequences of signals (facial 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 in which the arcs are the observed sequences of signals and nodes are possible transitions between them. Different paths in the graph correspond to different expressions of emotions. New animations can be generated by reordering the observed displays.

Ruttkay [22] proposed a system that allows a human designer to modify a facial expression animation defined by default by a trapezoid onset-apex-offset. For any single facial signal (e.g. FAP), one can manually define the course of its 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 for any of the facial animation parameters. A different interface for the generation of facial expressions of an avatar is proposed in [25]. Using 2D custom control space the user might deform both the geometry and the texture of a facial model. The approach is based on principal component analysis of the images database showing a variety of facial expressions of one subject. It allows one to generate realistic still images as a fluent sequence of expressions.

On the other hand, Paleari et al. [15] and Malatesta et al. [14] use manually defined sequential expressions inspired by Scherer's appraisal theory [23]. They consider a limited number of emotions and placed the emphasis on the temporal relations between different dynamic elements of the expression and their link with the consecutive stages of cognitive evaluations. A facial expression is not activated at once and with unity, rather animation parameters are activated at different moments. The final result is an animation of a sequence of several micro-expressions of cognitive evaluations. Both approaches, the additive [14] and the sequential one [14, 15], were evaluated subsequently in [14]. In this experiment the synthetic expressions of fear and hot anger, which could be either an additive or a sequential presentation of facial expressions were rated by 20 participants. Results show an above chance level recognition of the correct emotions in the case of the additive approach, whereas the sequential approach gives recognition results marginally above random choice.

The evaluation of the expressions of a robot, Kismet, from still pictures and from videos presenting multimodal dynamic expressions is described in [3]. In a first test seven still images of anger, disgust, fear, happiness, sorrow, surprise and stern expressions were evaluated by 17 participants. The participants chose the best match from ten labels. In a second test seven people evaluated 7 videos of Kismet performing 7 different emotional expressions involving face and body movements using a forced choice 7-element questionnaire. The participants performed significantly above chance level, with overall stronger recognition performance in a second test than on the still images alone. The data from the video studies suggest that witnessing the movement of the robot's face and body strengthens the recognition of the expression.

3. Background

Some observational studies [9–12, 21, 24] have explored the complexity of emotional expressions in terms of their dynamics and/or multimodality. Keltner [11] studied the sequence of facial and gestural movements in embarrassment. She relied on the analysis of their appearance frequencies in audio-visual data. The typical expression of embarrassment starts from a downward gaze or gaze shifts which are followed by "controlled" smiles. These smiles are often accompanied by pressed lips. At the end of the expression the movement of the head to the left was often observed as well as some face touching gestures [11].

Shiota and colleagues on the other hand, studied three positive emotions: awe, amusement, pride [24]. They showed that these three could be expressed by a set of possible signals, sometimes with asynchronous onsets, offsets and apices. Not all signals have to be present at the same time, for such to be recognized as a display of a particular emotional state. In the expression of pride [24], for example, Shiota et al. observe a mild Duchenne smile (AUs 6 + 12) with compressed lips (AU 24) and some straightening of the back. They note that pride can also be often accompanied by some pulling back of the shoulders to expose the chest and by a slight head lift.

Another emotion considered is anxiety [10] which is displayed using partial facial expressions of fear (i.e. the expression of fear by the open mouth), mouth stretching movements, eyes blinking, and non-Duchenne smiles. Expressions of other emotional states were also analyzed, among others, amusement [9, 24], shame [9, 11], awe [12, 24], confusion [21], and worry [21].

We gathered the data about other emotional expressions from the annotation of a video-corpus. This process is described in the next section.

4. Annotation

To create an expressive multimodal agent we need data about the behaviors. Besides gathering data from observational studies (see section 3), a work of annotation has been done. Five emotional states (anger, cheerfulness, panic fear, relief, and tension) have been annotated by a FACS expert with at least two cases per state. The audiovisual clips were extracted from several live TV programs, from the EmoTV corpus [1], the Belfast Naturalistic Emotional Database [4] and the HUMAINE database [5]. The people observed in the chosen clip extracts were non actors, placed in a natural emotional situation.

The video annotations were realized with Anvil v4.7.6 [13], a software enabling the description of audio visual recordings. The clips were described in term of facial, body and head movement and gaze changes. One track has also been defined for the emotional label and its intensity. The face changes have been annotated with the Facial Action Coding System [7] (one track per Action Unit), while the body, head and gaze movements have been described in an ad hoc format. Body changes were described verbally, as no succinct system of body movement relevant to emotions is in use at the moment. Head and gaze movements were also described verbally, in order to indicate the contextual meaningful reference rather than simply locational references. For example, in panic fear, head is described as moving away from the threatening stimulus or in another case turning towards other people for social support rather than turning left or right. The attributed emotion was defined mostly from the situation, e.g. in a video of a person awaiting an extremely important event, the emotional state was annotated as tension. Although only a very short extract was annotated, limited strictly to the emotional expression, a longer part of the video clip was viewed to enable the comprehension of the context.

In the emotional states that have been annotated we found that the face is often not the only mode of expression of any internal state. For example in relief, it is the body's drop of tension that is most spectacular, with its efferent movements, such as jaw dropping, backward projection of the head or thrust of the hands into the air. Often the facial mimicry is minimal and mostly insufficient for the understanding of the affective changes, e.g. closing of the eyes and mouth opening per se do not lead to an attribution of the label "relief". What is more, some movements seem to be an expression of arousal: the general shaking of body parts (mostly of the head) is present in the strong emotional states of joy, anger and tension. In some other states, however, one can see some movements typical of that emotional state, e.g. in cheerfulness characteristic large upward arm movements and changes in the posture involving numerous head tilts and forward torso movements were observed. In relief, the movements may be also large, however not repeated but consisting rather of one movement (one throw of the arms in the air; one semi-circular throw of the head backwards). There is generally only a limited number of changes of posture from a tensed and immobile to a more relaxed one and a change in the focus (head and gaze shifting from the important stimulus to the surrounding). The latter change of focus is also observed in the panic fear states: after the first reaction and evaluation of the threatening stimulus, the person turns towards the surrounding environment, scanning it without paying attention to any particular elements. The arm movements are quick, jerky and numerous. Most of the time they are self-center, e.g. a hand touching the chest or the face, or a hand covering the mouth or the eyes. The hand displacements can be very limited, such as in anxiety, although the body is generally shifted backwards substantially, away from the threat by means of large movements. In tension, no shift of body is observed during the whole episode, the movements are scarce and in the large majority limited to the face. The head and the gaze are constantly focused on the important stimulus. In anger on the other hand the movements are numerous, with constant changes in all the modalities: in the face (mostly eyebrow actions and tensed actions of the mouth), arms (large beats, placement of the hands on the hips) and in the torso (forward movements).

To conclude, the manual annotation of the particular clips depicting emotional behavior was used to obtain relevant data for the creation of a repertoire of expressions for an ECA. With the individualized selection of clips, with among the criteria the feasibility element on the ECA platform, the annotations complete the theoretical and observational studies and procure additional information concerning emotional nonverbal behavior.

5. Computation of multimodal sequential expressions of emotions

A model of multimodal sequential expressions of emotions has been elaborated from two approaches, namely from the annotation of real data (see section 4) and from data reported from the literature (see section 3). To go beyond agents showing simply static facial expression of emotion, a representation scheme that encompasses dynamics of multimodal expressions is defined. Based on such a sequential multimodal behavior representation, an algorithm is drawn. The main task of this algorithm [16] is to generate emotional expressions that are composed of different signals partially ordered in time and belonging to different nonverbal communicative channels. These expressions can be of any duration while the duration of constitutive signals is limited (e.g. facial expressions of emotions usually are not longer than four seconds while the expression of surprise is much shorter [6]). We define for each emotional state a *behavior set* - a set of signals through which the emotion is displayed and a *constraint set* that defines the appearance and the temporal constraints between the signals of the behavior set.

The single signals like *frown*, *head nod* or *self-touch gesture* are described in the repositories of the agent's nonverbal behaviors. These signals are grouped in the *behavior sets*. Each behavior set associates one emotional state with a set of plausible signals that might by displayed by the agent. For each signal in a behavior set one may define the probability of occurrence, its minimum and maximum duration and its number of repetitions. All the signals that belong to one behavior set may occur in the displays of the emotion associated with it but their occurrence is not random (see section 3 and 4). We developed an XML-based language to describe a set of relations (i.e. a constraint set) between the signals of one behavior set. Two types of constraints are considered:

- temporal constraints specified by arithmetic operators defining the relation between the starting and ending time of a signal e.g. "signal s_i cannot start at the beginning of animation" or "signal s_i starts immediately after the signal s_j finishes". The temporal constraints are defined using arithmetical relations: $<$, $>$ and $=$;

- appearance constraints describing more general relations between the signals like inclusions or exclusions e.g. "signals s_i and s_j cannot co-occur" or "signal s_j cannot occur without signal s_i".

The constraints of both types are composed using the logical operators: *and, or, not*.

The algorithm works as follows. From the single label of an emotional state *e* (e.g. anger or embarrassment) the system generates sequences of multimodal expressions, i.e. the animation *A* of a given duration *t* composed of a sequence of signals $s_{j(i)}$ on different modalities. It does so by choosing an appropriate subset of signals from the behavior set BS_e, their durations, and order of display. The algorithm can generate several sequences of signals, each of them satisfying the constraints.

Example. In Figure 1 an animation for an expression of embarrassment is shown. The following images present the frames of an animation displaying respectively the signals: look right (Figure 1a), head down and gaze down (Figure 1b), gaze left (Figure 1c), head/gaze left and tensed smile (Figure 1d), gaze left (Figure 1f).

In Figure 2 an example of an animation for the expression of cheerfulness is shown. The following signals are displayed: open mouth smile (Figure 2a) which is accompanied (Figure 2b and 2c) by hands raising, closed mouth smile and head movements aside (Figure 2d and 2e), smile and raised hands (Figure 2f).

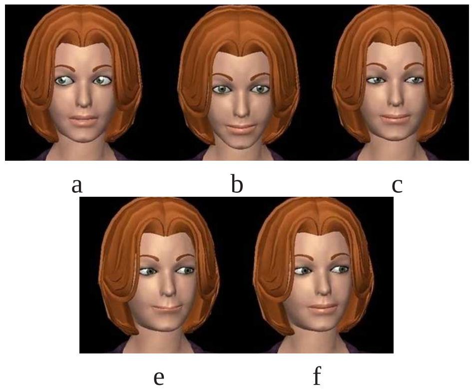

Figure 1. An example of a multimodal expression of embarrassment

6. Evaluation

We performed an evaluation study on multimodal sequential expressions of emotions. In this test we check if the emotions expressed by the agent are recognized by the participants. For this purpose we show the participants a set of animations generated by the Greta agent [2] displaying various emotional states. We ask them to attribute to each animation one emotional label.

6.1. Set-up scenario

For the purpose of this evaluation study we generated eight animations, one animation per emotional state, using the algorithm described in section 5. We consider the following emotions: *anger, anxiety, cheerfulness, embarrassment, panic fear, pride, relief, tension*. This arbitrary choice is motivated by the following:

C1) we want to differentiate between several positive emotional states. Usually all the positive emotions are described with the general label "joy" and associated

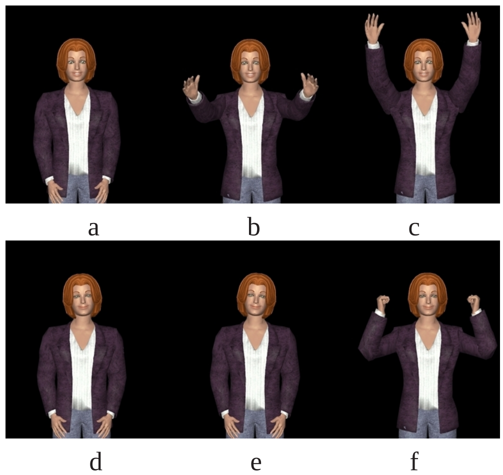

Figure 2. An example of a multimodal expression of cheerfulness

with the Duchenne smile. From section 3 we know that positive expressions can be differentiated. In this study we evaluate: cheerfulness, pride and relief.

C2) we want to differentiate expressions in which different types of smiles (Duchenne and non-Duchenne) might occur. Smiles are used to display positive emotions (e.g. in joy) but they also occur in negative expressions like embarrassment [11] or anxiety [10].

C3) finally we want also to differentiate different negative states to be used by the ECA (e.g. in the CALLAS EU Project[1]) like anxiety, tension, panic fear and we want to confront them with the expression of anger.

The behavior and constraint sets for the three emotional states of pride, embarrassment and anxiety were defined from the literature (see section 3). The sets of other five emotional states: anger, cheerfulness, panic fear, relief, tension were based on the annotation study described in section 4. Each video shows the agent displaying one emotional state. The agent is not speaking. The duration of each video is about 10 seconds.

6.2. Hypothesis

Our hypotheses are the following:

H1) each of the intended emotions is more often correctly recognized on the corresponding animation than chance level.

H2) for each animation the proper label will be attributed more often than any other label.

H3) if the hypothesis H2 is not confirmed, confusions are expected to occur between the emotional expressions of the three conditions, i.e. pride might by confused with relief and/or cheerfulness (C1: positive emotions ambiguity), embarrassment with anxiety or cheerfulness (C2: different smiles ambiguity), tension with anxiety or anger (C3: negative emotions ambiguity).

Each animation is shown twice. The habituation factor between the first and second viewing will be explored. We expect an increase in the numbers of correct answers.

6.3. Procedure

Participants accessed the evaluation study through a web browser. The explanation of the procedure to follow was described on a web page. Participants were asked to recognize the emotions displayed by the virtual agent. Eight animations presented in the previous section were used. Each study session consists of seeing twice the same set of eight videos presented in a random order. Each subject has to see all eight videos (turn 1) before seeing any of them for the second time (turn 2).

[1]www.callas-newmedia.eu

One study session is made of 16 web pages, each page presenting one animation a_i. After watching an animation the participants have to attribute one emotional label to the perceived emotional state from an 8-element list before they can pass to another page with a new animation. They cannot come back to the preceding animation a_{i-1} and they cannot jump to the animation a_{i+1} without providing answer to the current one. They cannot review the animation twice. No time constraint was put on the task. The animations were displayed in a random order, the emotional labels were ordered alphabetically. The participation in the study was anonymous.

Fifty three participants (25 women, 28 men) with a mean age of 28 years mainly from France (21%), Poland (21%) and Italy (15%) took part in the study. Nearly half of the participants comes with a computer science background (39%), the remaining being mainly from medicine (12%) and psychology (10%). The majority of the participants (73%) were graduates or post-graduates. None of them works in the domain of embodied conversational agents.

6.4. Results

The recognition level for each emotional expression in both turns is above chance level (which is 12,5%). In each turn, the most well recognized emotion was anger (93% both turns mean) while the least recognized was embarrassment (41% both turns mean). The number of correct vs. alternative answers in turn 1 and turn 2 was compared and the improvement was not significative (McNemar test, $p>.05$), therefore, although the analyses for each of the two turns are run, the means for both turns are stated for reference in the text when not otherwise specified.

In general, the proper label was attributed more often than any other label. For the animations of anger, cheerfulness, panic fear and relief, the correct labels were significantly more often attributed than any other ones in both turns (McNemar test, $p<.05$ in each turn). In the well recognized animations, the second mostly attributed label for anger (correctly attributed 93% both turns means) was tension (5%), while cheerfulness (70%) was labeled pride (25%), panic fear (61%) was labeled embarrassment (17%), relief (69%) was labeled cheerfulness (23%). For the remaining animations of anxiety, embarrassment, pride and tension the proper label was found but some confusions occurred. The strongest confusion occurs between anxiety and embarrassment. For the anxiety animation the number of attributions of the anxiety (43% both turns means) and of the embarrassment (36% both turns means) labels did not differ significantly (McNemar test, $p>.05$). In the embarrassment animation, embarrassment (41% both turns means) was confused with anxiety (36% both turns means)

	anger	anxiety	cheerful.	embarrass.	panic fear	pride	relief	tension
anger label	93	0	0	0	0	2	0	5
anxiety label	0	43	0	36	3	0	1	18
cheerfulness label	0	0	70	0	0	25	6	0
embarrassment label	0	36	0	41	1	0	0	23
panic fear label	2	11	0	17	61	0	7	2
pride label	0	4	14	6	0	45	26	5
relief label	0	0	23	1	0	8	69	0
tension label	4	21	1	24	3	2	0	46

Table 1. Matrix of confusions presented as percentages of participants' attributions of the eight emotional labels (means for both turns).

($p>.05$). In turn 2 embarrassment (40%) was also labeled tension (28%) ($p>.05$) (while in turn 1 it was labeled tension by 17%). Although on the limit of a significant difference ($p=.066$) some other confusions were found: pride (45% both turns means) was labeled relief (26% both turns means) in both turns and tension (49%) was labeled embarrassment (25%) in turn 2.

6.5. Discussion

The main aim of this evaluation study was to check if the multimodal sequential expressions generated according to the procedure described in sections 4 and 5 are recognized by the participants. The hypothesis H1 was verified: the recognition rate (41% - 93% both turns means) exceeds strongly chance level. The hypothesis H2 was only partially verified: although the number of attributions of correct labels was higher than that of alternatives, the difference was not significative for some emotions. The hypothesis H3 on the expected confusions to arise for similar emotions was partially confirmed. Similarity comes from cognitive representations and/or common behavior characteristics (see Conditions C1, C2 and C3). From results in Table 1, we can notice there are some confusions between the emotions of anxiety and embarrassment, embarrassment and tension, and finally pride and relief. Three out of four confusions occur due to the C3 explanation (difficulty in the negative states differentiation), while the fourth is due to the C2 (difficulty in the positive state differentiation). In general these pairs of emotions are placed in adjacent position to each other in dimensional spaces. For example in the Pluchik wheel [18], anxiety and embarrassment appear very close to each other: the angle for anxiety is 78.3 while for embarrassment it is 75.3. Those pairs of emotions also share signals from their behavioral sets (see section 5): anxiety and embarrassment show both a lot of gaze shift; anxiety and tension display lip and hand tensed actions; embarrassment and tension have both tensed lip shapes; and pride and relief are conveyed through smiles as well as up and backward body movements.

Our results show that even such subtly differentiated expressions like these of relief or of cheerfulness were recognized surprisingly well. These emotions probably would not have been recognized from still facial expressions in their apex. This claim needs however to be checked in future studies. The effect of habituation is small and the improvement as seen in correct labeling is not significative. Consequently multimodal sequential expressions may be used straight away, in short period interactions with the user.

While the recognition rate is quite high, we believe it could have been higher if behavior expressivity would have been considered. In the videos used for this perception study, emotions were conveyed through signals defined in the behavior set. Behavior execution did not vary, that is behaviors had the same expressive qualities in all the videos. However, body expressivity is an important cue to convey emotional states as claims Wallbott [27] and as we can infer from our corpus annotation (see section 4). The non adaptation of the behavior expressivity to the particular states might have influenced their recognition rate. Thus, we believe that the model of multimodal sequential expressions should be extended by a number of expressivity features in order to create an ECA with proper displays of emotions.

Last but not least, the emotions that received the highest recognition rate - anger, cheerfulness, panic fear and relief - are described by facial expressions as well as specific body and arm movements (e.g. anger with the hands on the hips and cheerfulness with raised arms). It seems that expressions of emotions that make used of the full body were better perceived compared to expressions of emotions conveyed mainly with the face (such as embarrassment and tension). However the use of multimodality in communicating emotions should be more carefully analyzed (for example by studying the added-value of each modality).

7. Conclusion

In this paper we presented the first evaluation of multimodal sequential expressions displayed by an ECA. These expressions go beyond Ekman's description of facial expressions of emotions described in their apex. In our evaluation study the ECA displayed emotions using a sequence of various nonverbal behaviors across modalities. The results show that multimodal sequential expressions allow for the differentiation of the emotional displays not considered to be universally recognized. In the case of all eight emotions the recognition rate significantly surpassed chance level. In particular certain positive emotional states like relief or cheerfulness were particularly well recognized.

Acknowledgement

Part of this research is supported by the EU FP6 Integrated Project CALLAS IP-CALLAS IST-034800.

References

[1] S. Abrilian, L. D. ans S. Buisine, and J.-C. Martin. EmoTV1: Annotation of real-life emotions for the specification of multimodal affective interfaces. In *11th International Conference on Human-Computer Interaction (HCII'2005)*, Las Vegas, Nevada, USA, 2005.

[2] E. Bevacqua, M. Mancini, R. Niewiadomski, and C. Pelachaud. An expressive ECA showing complex emotions. In *Proceedings of the AISB Annual Convention*, pages 208–216, Newcastle, UK, 2007.

[3] C. Breazeal. Emotion and sociable humanoid robots. *International Journal of Human-Computer Studies*, 59(1-2):119–155, July 2003.

[4] E. Douglas-Cowie, N. Campbell, and P. Roach. Emotional speech: Towards a new generation of databases. *Speech Communication*, 40(1–2):33–60, 2003.

[5] E. Douglas-Cowie, C. Cox, J.-C. Martin, L. Devillers, R. Cowie, I. Sneddon, M. McRorie, C. Pelachaud, C. Peters, O. Lowry, A. Batliner, and F. Honig. Humaine database. http://emotion-research.net/download/pilot-db/.

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

[7] P. Ekman and W. Friesen. *Facial Action Coding System*. Consulting Psychologists Press, 1978.

[8] C. Elliott. I picked up catapia and other stories: A multimodal approach to expressivity for 'emotionally intelligent' agents. In *First International Conference on Autonomous Agents*, pages 451–457, Marina del Rey, CA, 1997.

[9] J. Haidt and D. Keltner. Culture and facial expression: Open-ended methods find more expressions and a gradient of recognition. *Cognition and Emotion*, 13(3):225–266, 1999.

[10] J. A. Harrigan and D. M. O'Connell. How do you look when feeling anxious? facial displays of anxiety. *Personality and Individual Differences*, 21:205–212, 1996.

[11] 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.

[12] D. Keltner and J. Haidt. Approaching awe, a moral, spiritual, and aesthetic emotion. *Cognition and Emotion*, 17:297–314, 2003.

[13] M. Kipp. Anvil - a generic annotation tool for multimodal dialogue. In *Proceedings of the 7th European Conference on Speech Communication and Technology (Eurospeech)*, pages 1367–1370, Aalborg, September 2001.

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

[15] M.Paleari and C.L.Lisetti. Psychologically grounded avatars expressions. In *1rst Workshop on Emotion and Computing at KI 2006, 29th Annual Conference on Artificial Intelligence*, Bremen, Germany, 2006.

[16] R. Niewiadomski, S. Hyniewska, and C. Pelachaud. Modeling emotional expressions as sequences of behaviors. In *Proceedings of the 7th International Conference on Intelligent Virtual Agents (IVA)*, Amsterdam, Holland, 2009.

[17] X. Pan, M. Gillies, T. M. Sezgin, and C. Loscos. Expressing complex mental states through facial expressions. In *Second International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 745–746. Springer, 2007.

[18] R. Plutchik. *Emotion: Theory, research, and experience: Vol. 1. Theories of emotion*, chapter A general psychoevolutionary theory of emotion, pages 3–33. R. Plutchik & H. Kellerman, 1980.

[19] F. Pollick, H. Paterson, A. Bruderlin, and A. Sanford. Perceiving affect from arm movement. *Cognition*, 82:51 – 61, 2001.

[20] M. Rehm and E. André. Catch me if you can - exploring lying agents in social settings. In F. Dignum, V. Dignum, S. Koenig, S. Kraus, M. P. Singh, and M. Wooldridge, editors, *Proceedings of International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 937–944, Utrecht, The Netherlands, 2005. ACM.

[21] P. Rozin and A. Cohen. High frequency of facial expressions corresponding to confusion, concentration, and worry in an analysis of naturally occurring facial expressions of Americans. *Emotion*, 3(1):68–75, 2003.

[22] Z. Ruttkay. Constraint-based facial animation. *International Journal of Constraints*, 6:85–113, 2001.

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

[24] M. N. Shiota, B. Campos, and D. Keltner. The faces of positive emotion: Prototype displays of awe, amusement, and pride. *Annals of the New York Academy of Sciences*, 2003.

[25] N. Stoiber, R. Séguier, and G. Breton. Automatic design of a control interface for a synthetic face. In C. Conati, M. Bauer, N. Oliver, and D. S. Weld, editors, *IUI*, pages 207–216. ACM, 2009.

[26] J. Walker, L. Sproull, and R. Subramani. Using a human face in an interface. In *Proceedings of the SIGCHI conference on Human factors in computing systems: celebrating interdependence*, pages 85 – 91, Boston, Massachusetts, 1994.

[27] H. Wallbott. Bodily expression of emotion. *European Journal of Social Psychology*, 28:879–896, 1998.