

Online Gaming and Emotion Representation

A. Raouzaïou, K. Karpouzis, and Stefanos D. Kollias

Image, Video and Multimedia Systems Laboratory
National Technical University of Athens
157 80 Zographou, Athens, Greece
{araouz, kkar pou}@image.ntua.gr, stefanos@cs.ntua.gr

Abstract. The ability to simulate lifelike interactive characters has many applications in the gaming industry. A lifelike human face can enhance interactive applications by providing straightforward feedback to and from the users and stimulating emotional responses from them. Thus, the gaming and entertainment industries can benefit from employing believable, expressive characters since such features significantly enhance the atmosphere of a virtual world and communicate messages far more vividly than any textual or speech information. In this paper, we present an abstract means of description of facial expressions, by utilizing concepts included in the MPEG-4 standard. Furthermore, we exploit these concepts to synthesize a wide variety of expressions using a reduced representation, suitable for networked and lightweight applications.

1 Introduction

In the past five years online gaming has grown and reached a large and diverse market [2]. In addition to online-only games, available usually through downloads, the most popular PC games offer, in many cases, an online component, something that can also be found in some console games, hence online gaming is a developing field which needs more realistic agents. A few studies have already appeared in this direction [12]. Lifelike agents who could express their emotions would be more attractive as opponent players.

Research in facial expression analysis and synthesis has mainly concentrated on archetypal emotions. In particular, sadness, anger, joy, fear, disgust and surprise are categories of emotions that attracted most of the interest in human computer interaction environments. Very few studies [4] have appeared in the computer science literature, which explore non-archetypal emotions. This trend may be due to the great influence of the works of Ekman [9] and Friesen [4] who proposed that the archetypal emotions correspond to distinct facial expressions which are supposed to be universally recognizable across cultures. On the contrary psychological researchers have extensively investigated a broader variety of emotions. An extensive survey on emotion analysis can be found in [3].

Moreover, the MPEG-4 indicates an alternative way of modeling facial expressions and the underlying emotions, which is strongly influenced from neurophysiological and psychological studies (FAPs). The adoption of token-based animation in the MPEG-4 framework [7], [8] benefits the definition of emotional states, since the ex-

traction of simple, symbolic parameters is more appropriate to analyze, as well as synthesize facial expression.

In this paper we present a methodology for creating intermediate expressions based on archetypal ones and taking into account the results of Whissel's study [3]. Based on universal knowledge of the expression of emotions we form rules that utilize FAPs as linguistic variables. The actual partitioning in these rules is based on the analysis processing of actual video sequences showing variations of the six universal emotions; in order to come up with rules describing intermediate emotions, we utilize interpolating notions between "neighboring" emotions.

2 Emotion Representation

Psychologists have examined a broader set of emotions [4], but very few of the studies provide results which can be exploited in computer graphics and machine vision fields. One of these studies, carried out by Whissel [3], suggests that emotions are points in a space spanning a relatively small number of dimensions, which seem to occupy two axes: *activation* and *evaluation*. The expressions created with the present methodology taking into account the results of Whissel's study [3] and in particular the *activation* parameter, can be easily used for online gaming.

The activation-emotion space [3] is a representation that is both simple and capable of capturing a wide range of significant issues in emotion. It rests on a simplified treatment of two key themes (see Fig. 1. The Activation – emotion space):

- *Valence* (Evaluation level): the clearest common element of emotional states is that the person is materially influenced by feelings that are "valenced", i.e. they are centrally concerned with positive or negative evaluations of people or things or events. The link between emotion and valencing is widely agreed (horizontal axis).
- *Activation* level: research has recognized that emotional states involve dispositions to act in certain ways. A basic way of reflecting that theme turns out to be surprisingly useful. States are simply rated in terms of the associated activation level, i.e. the strength of the person's disposition to take some action rather than none (vertical axis).

A surprising amount of emotional discourse can be captured in terms of activation-emotion space. Perceived full-blown emotions are not evenly distributed in activation-emotion space; instead they tend to form a roughly circular pattern. From that and related evidence, Plutchic [11] shows that there is a circular structure inherent in emotionality. In this framework, identifying the center as a natural origin has several implications. Emotional strength can be measured as the distance from the origin to a given point in activation-evaluation space. An interesting implication is that strong emotions are more sharply distinct from each other than weaker emotions with the same emotional orientation. A related extension is to think of primary or basic emotions as cardinal points on the periphery of an emotion circle. Plutchik has offered a useful formulation of that idea, the "emotion wheel".

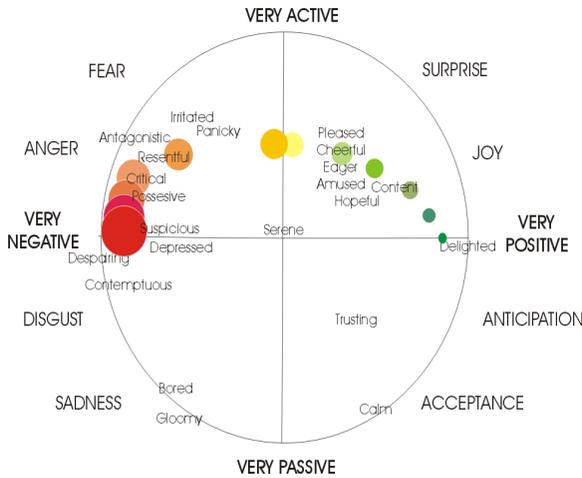


Fig. 1. The Activation – emotion space

3 Profiles Creation

In the following we define an *expression profile* to be a subset of the FAPs vocabulary, corresponding to a particular expression, accompanied with FAP intensities, i.e. the actual ranges of variation, which, if animated, creates the requested expression and can be used for expression synthesis in online gaming.

3.1 Creating Archetypal Expression Profiles

An archetypal *expression profile* is a set of FAPs accompanied by the corresponding range of variation, which, if animated, produces a visual representation of the corresponding emotion. Typically, a profile of an archetypal expression consists of a subset of the corresponding FAPs’ vocabulary coupled with the appropriate ranges of variation.

In order to define exact profiles for the archetypal expressions, we combined the following three steps:

- (a) we defined subsets of FAPs that are candidates to form an archetypal expression, by translating the proposed by psychological studies [5], [9] face formations to FAPs,
- (b) we used the corresponding ranges of variations obtained from statistics [1] and,
- (c) we animated the corresponding profiles to verify appropriateness of derived representations.

For our experiments on setting the archetypal expression profiles, we used the face model developed by the European Project *ACTS MoMuSys*, which is freely available at the website <http://www.iso.ch/ittf>. Table 1 shows examples of profiles of some archetypal expressions [1].

Table 1. Profiles for some archetypal emotions

<i>Profiles</i>	<i>FAPs and Range of Variation</i>
Anger ($P_A^{(0)}$)	$F_4 \in [22, 124], F_{31} \in [-131, -25], F_{37} \in [-136, -34], F_{33} \in [-189, -109], F_{34} \in [-183, -105], F_{35} \in [-101, -31], F_{36} \in [-108, -32], F_{37} \in [29, 85], F_{38} \in [27, 89]$
$P_A^{(1)}$	$F_{19} \in [-330, -200], F_{20} \in [-335, -205], F_{21} \in [200, 330], F_{22} \in [205, 335], F_{31} \in [-200, -80], F_{32} \in [-194, -74], F_{33} \in [-190, -70], F_{34} \in [-190, -70]$
Fear ($P_F^{(0)}$)	$F_5 \in [102, 480], F_5 \in [83, 353], F_{19} \in [118, 370], F_{20} \in [121, 377], F_{21} \in [118, 370], F_{22} \in [121, 377], F_{31} \in [35, 173], F_{37} \in [39, 183], F_{33} \in [14, 130], F_{34} \in [15, 135]$
$P_F^{(1)}$	$F_3 \in [400, 560], F_5 \in [333, 373], F_{19} \in [-400, -340], F_{20} \in [-407, -347], F_{21} \in [-400, -340], F_{22} \in [-407, -347]$
$P_F^{(2)}$	$F_3 \in [400, 560], F_5 \in [-240, -160], F_{19} \in [-630, -570], F_{20} \in [-630, -570], F_{21} \in [-630, -570], F_{22} \in [-630, -570], F_{31} \in [260, 340], F_{37} \in [260, 340], F_{33} \in [160, 240], F_{34} \in [160, 240], F_{35} \in [60, 140], F_{36} \in [60, 140]$

Figure 2 shows some examples of animated profiles. Fig. 2(a) shows a particular profile for the archetypal expression *anger*, while Fig. 2(b) and (c) show alternative profiles of the same expression. The difference between them is due to FAP intensities. Difference in FAP intensities is also shown in Figures 2(d) and (e), both illustrating the same profile of expression *surprise*. Finally Figure 2(f) shows an example of a profile of the expression *joy*.

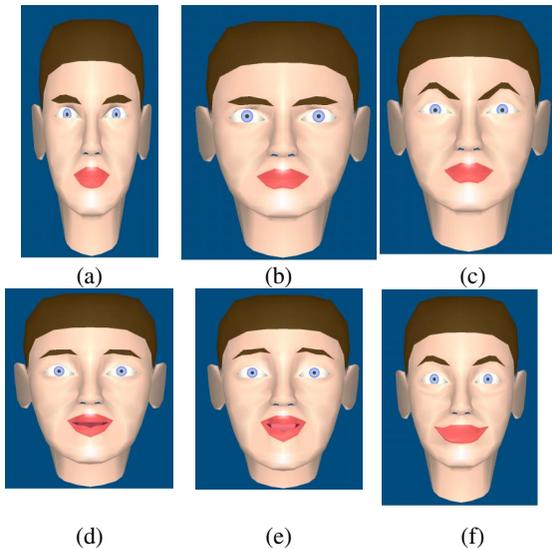


Fig. 2. Examples of animated profile: (a) – (c) Anger, (d) – (e) Surprise, (f) Joy

3.2 Creating Profiles for Intermediate Expressions

The limited number of studies, carried out by computer scientists and engineers [10], dealing with emotions other than the archetypal ones, lead us to search in other subject/discipline bibliographies. Psychologists examined a broader set of emotions [4], but very few of the corresponding studies provide exploitable results to computer graphics and machine vision fields, e.g. Whissel's study suggests that emotions are points in a space (Figure 1) spanning a relatively small number of dimensions, which in a first approximation, seem to occupy two axes: *activation* and *evaluation*, as shown in Table 2. *Activation* is the degree of arousal associated with the term, with terms like *patient* (at 3.3) representing a midpoint, *surprised* (over 6) representing high activation, and *bashful* (around 2) representing low *activation*. *Evaluation* is the degree of pleasantness associated with the term, with *guilty* (at 1.1) representing the negative extreme and *delighted* (at 6.4) representing the positive extreme. From the practical point of view, *evaluation* seems to express internal feelings of the subject and its estimation through face formations is intractable. On the other hand, *activation* is related to facial muscles' movement and can be easily estimated based on facial characteristics.

The third column in Table 2 represents the observation [11] that emotion terms are unevenly distributed through the space defined by dimensions like Whissel's. Instead, they form an approximately circular pattern called "*emotion wheel*". Shown values refer to an angular measure, which runs from *Acceptance* (0) to *Disgust* (180).

Table 2. Selected Words from Whissel's Study

	<i>Activation</i>	<i>Evaluation</i>	<i>Angle</i>
Terrified	6.3	3.4	75.7
Afraid	4.9	3.4	70.3
Worried	3.9	2.9	126
Angry	4.2	2.7	212
Patient	3.3	3.8	39.7
Sad	3.8	2.4	108.5
Delighted	4.2	6.4	318.6
Guilty	4	1.1	102.3
Bashful	2	2.7	74.7
Surprised	6.5	5.2	146.7

Same Universal Emotion Category. As a general rule, one can define six general categories, each characterized by an archetypal emotion; within each of these categories, intermediate expressions are described by different emotional intensities, as well as minor variation in expression details. From the synthetic point of view, emotions belonging to the same category can be rendered by animating the same FAPs using different intensities. In the case of expression profiles, this affects the range of variation of the corresponding FAPs which is appropriately translated; the fuzziness introduced by the varying scale of FAP intensities provides mildly differentiated output in similar situations. This ensures that the synthesis will not render "robot-like" animation, but drastically more realistic results.

For example, the emotion group *fear* also contains *worry* and *terror* [6] which can be synthesized by reducing or increasing the intensities of the employed FAPs, respectively.

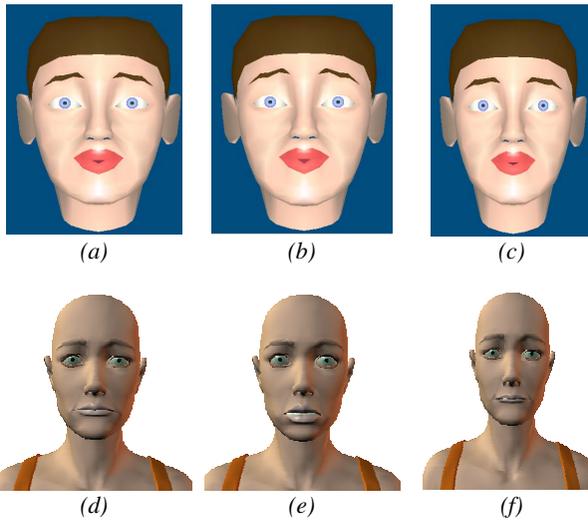


Fig. 3. Animated profiles for emotion terms (a, d) afraid, (b, e) terrified and (c, f) worried

Figures 3(a)-(c) show the resulting profiles for the terms *terrified* and *worried* emerged by the one of the profiles of *afraid* (in particular $P_F^{(2)}$). Figures 3(d)-(e) show the same profiles synthesized using the 3D model of the software package *Poser*, edition 4 of Curious Labs. The FAP values that we used are the median ones of the corresponding ranges of variation.

Emotions lying between archetypal ones. Creating profiles for emotions that do not clearly belong to a universal category is not straightforward. Apart from estimating the range of variations for FAPs, one should first define the vocabulary of FAPs for the particular emotion. In order to proceed we utilize both the “*emotion wheel*” of Plutchik [11] and especially the *angular* measure (shown also in Table 2), and the *activation* parameter.

One is able to synthesize intermediate emotions by combining the FAPs employed for the representation of universal ones. In our approach, FAPs that are common in both emotions are retained during synthesis, while emotions used in only one emotion are averaged with the respective neutral position. In the case of mutually exclusive FAPs, averaging of intensities usually favors the most exaggerated of the emotions that are combined, whereas FAPs with contradicting intensities are cancelled out.

Figure 4 shows the results of creating a profile for the emotion *guilt*, created using the face model of the European Project *ACTS MoMuSys*, as well as the 3D model of *Poser*, according to the calculated values of Table 3. Plutchik’s *angular* measure (see Table 2) shows that the emotion term *guilty* (angular measure 102.3 degrees) lies

between the archetypal emotion terms *afraid* (angular measure 70.3 degrees) and *sad* (angular measure 108.5 degrees), being closer to the latter.

Table 3. Activation and angular measures used to create the profile for the emotion *guilt*

<p>Afraid: (4.9, 70.3): $F_3 \in [400,560]$, $F_5 \in [-240,-160]$, $F_{19} \in [-630,-570]$, $F_{20} \in [-630,-570]$, $F_{21} \in [-630,-570]$, $F_{22} \in [-630,-570]$, $F_{31} \in [260,340]$, $F_{32} \in [260,340]$, $F_{33} \in [160,240]$, $F_{34} \in [160,240]$, $F_{35} \in [60,140]$, $F_{36} \in [60,140]$</p>
<p>Guilty: (4, 102.3): $F_5 \in [160,230]$, $F_5 \in [-100,-65]$, $F_{19} \in [-110,-310]$, $F_{20} \in [-120,-315]$, $F_{21} \in [-110,-310]$, $F_{22} \in [-120,-315]$, $F_{31} \in [61,167]$, $F_{32} \in [57,160]$, $F_{33} \in [65,100]$, $F_{34} \in [65,100]$, $F_{35} \in [25,60]$, $F_{36} \in [25,60]$</p>
<p>Sad: (3.9, 108.5): $F_{19} \in [-265,-41]$, $F_{20} \in [-270,-52]$, $F_{21} \in [-265,-41]$, $F_{22} \in [-270,-52]$, $F_{31} \in [30,140]$, $F_{32} \in [26,134]$</p>

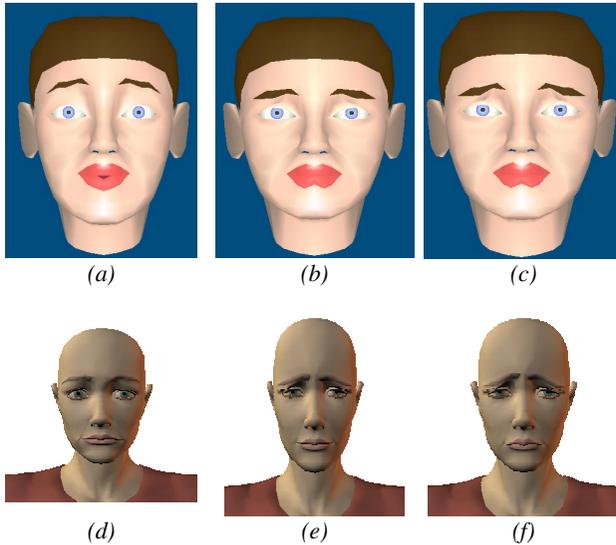


Fig. 4. Animated profiles for emotion terms: (a, d) afraid, (b, e) guilty and (c, f) sad

4 Conclusions

Online games have in many ways revolutionized the interactive entertainment industry creating an important new source of revenue and gamers. Hence, more lifelike agents are necessary for this new industry which will continue to grow. Facial animation is a great means of improving HCI applications and especially computer games, since it provides a powerful and universal means of expression and interaction.

In this paper we presented a parameterized approach of synthesizing realistic facial expressions using lightweight representations. This method employs concepts included in established standards, such as MPEG-4, which are widely supported in modern computers and standalone devices (e.g. PDAs, set-top boxes, etc.)

References

1. Raouzaiou, A., Tsapatsoulis, N., Karpouzis, K., Kollias, S.: Parameterized facial expression synthesis based on MPEG-4. *EURASIP Journal on Applied Signal Processing*, Vol. 2002, No. 10. Hindawi Publishing Corporation (2002) 1021–1038
2. IGDA Online Games Committee: *IGDA Online Games White Paper*. 2nd edition (2003)
3. Whissel, C.M.: The dictionary of affect in language. In: Plutchnik, R., Kellerman, H. (eds): *Emotion: Theory, research and experience: Vol 4, The measurement of emotions*. Academic Press, New York (1989)
4. EC TMR Project PHYSTA Report: *Review of Existing Techniques for Human Emotion Understanding and Applications in Human-Computer Interaction* (1998)
<http://www.image.ece.ntua.gr/physta/reports/emotionreview.htm>
5. Parke, F., Waters, K.: *Computer Facial Animation*. A K Peters (1996)
6. Faigin, G.: *The Artist's Complete Guide to Facial Expressions*. Watson-Guptill, New York (1990)
7. ISO/IEC 14496-1:2001 *Information technology – Coding of audio-visual objects – Part 1: Systems*, International Organization for Standardization. Swiss (2001)
8. Tekalp, M., Ostermann, J.: Face and 2-D mesh animation in MPEG-4. *Image Communication Journal*, Vol.15, Nos. 4-5 (2000) 387-421
9. Ekman, P.: Facial expression and Emotion. *Am. Psychologist*, Vol. 48 (1993) 384–392
10. Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., Taylor, J.: Emotion Recognition in Human-Computer Interaction. *IEEE Signal Processing Magazine* (2001) 32–80
11. Plutchik, R.: *Emotion: A psychoevolutionary synthesis*. Harper and Row New York (1980)
12. Morris, T.W.: *Conversational Agents for Game-Like Virtual Environments*. Artificial Intelligence and Interactive Entertainment. AAAI Press 82–86