Combining Facial and Postural Expressions of Emotions in a Virtual Character

Céline Clavel[1], Justine Plessier[2], Jean-Claude Martin[1], Laurent Ach[2], and Benoit Morel[2]

[1] LIMSI CNRS, BP 133
91403 Orsay cedex, France
[2] Cantoche, 68, rue d'Hauteville
75010 Paris, France
{celine.clavel,martin}@limsi.fr,
{jplessier,lach,bmorel}@cantoche.com

Abstract. Psychology suggests highly synchronized expressions of emotion across different modalities. Few experiments jointly studied the relative contribution of facial expression and body posture to the overall perception of emotion. Computational models for expressive virtual characters have to consider how such combinations will be perceived by users. This paper reports on two studies exploring how subjects perceived a virtual agent. The first study evaluates the contribution of the facial and postural expressions to the overall perception of basic emotion categories, as well as the valence and activation dimensions. The second study explores the impact of incongruent expressions on the perception of superposed emotions which are known to be frequent in everyday life. Our results suggest that the congruence of facial and bodily expression facilitates the recognition of emotion categories. Yet, judgments were mainly based on the emotion expressed in the face but were nevertheless affected by postures for the perception of the activation dimension.

Keywords: evaluation of virtual agents, affective interaction, conversational and non-verbal behavior, multimodal interaction with intelligent virtual agent.

1 Introduction

An emotion can be seen as "an episode of interrelated, synchronized changes in five components in response to an event of major significance to the organism" [33]. These five components are: the cognitive processing, the subjective feeling, the action tendencies, the physiological changes, and the motor expression. Thus the synchronization of expressions of emotion across different modalities is expected to be of major importance to our understanding of emotion perception. Yet, few studies explored jointly the contribution of the different nonverbal modalities such as posture and face to the overall perception of emotions [24] and did not use virtual characters as potential stimuli. Furthermore, several emotion theorists postulated the existence of emotion blends [8, 12, 13, 28, 31]. Klaus Scherer described three experiments in which the superposition of several emotions was observed.

Zs. Ruttkay et al. (Eds.): IVA 2009, LNAI 5773, pp. 287–300, 2009.

Virtual characters are now able to express basic emotions in individual modalities such as facial expressions and more recently bodily expressions. Computational models for generating nonverbal behaviors for expressive characters need to integrate the distribution of the expression of emotion across modalities based on how such combinations are perceived by users. Some models and related experimental studies considered the perception of superposition of emotions but mostly in facial expressions [26].

Our research aims to study how the face and the body posture jointly contribute to a multimodal congruent perception of a single emotion in terms of categories and dimensions. We also aim at studying the impact of incongruent facial and postural expressions on the perception of superposed emotions which are known to frequently occur in everyday life.

Section 2 summarizes related work in Psychology and Affective Computing. Section 3 describes the virtual character animation technology that we used for conducting experimental studies about the perception of combined facial and postural expressions. Sections 4 and 5 detail the two experiments that we conducted about the perception of congruent and incongruent expressions. After discussing the results, we provide possible directions for continuing to study nonverbal expression of superposed emotions, and assess their usefulness for computational models of nonverbal behaviors. A video presenting this research has been submitted along with the paper.[1]

2 Related Work

2.1 Emotions

Ekman proposed a set of characteristics that distinguish basic emotions from other affective phenomena (distinctive universal signals, distinctive physiology, automatic appraisal, distinctive universals in antecedent events, distinctive appearance developmentally, presence in other primates, quick onset, brief duration, unbidden occurrence, distinctive thoughts, distinctive subjective experience). Different sets of basic emotions have been proposed such as Joy, Surprise, Fear, Anger, Sadness, Disgust and Contempt [10]. Izard [21] also considers Interest, Shame and Guilt.

Independent and bipolar dimensions have also been proposed for representing emotional states. Russell proposed a 2D circumplex model of affects using a pleasure-displeasure dimension and an arousal-sleep dimension [29]. Other additional dimensions are also proposed such as dominance-submissiveness [30] or unpredictability [14].

In everyday life, several emotions often occur at the same time resulting in the expressions of blends of emotions [12, 32].

2.2 Facial and Bodily Expressions of Emotions

Ekman suggested that a given emotion can be expressed by a family of facial expressions. The number of expressions depends on the emotion [9, 10]. Joy and Contempt have a small set of expressions; Surprise and Disgust might have around 10 expressions; Fear, Sadness and Anger might have around 20 expressions. The recognition

[1] http://www.limsi.fr/Individu/martin/permanent/2009-iva/IVA09-Cantoche-comp.avi

rate was also observed to depend on the emotion. Gosselin and Kirouac [15] evaluated the recognition of 30 pictures of facial expressions collected by Ekman and Friesen. They observed that Joy, Anger, Surprise and Sadness were well recognized, whereas Fear and Disgust were less recognized.

A few studies considered how the body expresses emotion. Some researchers observed that body actions might provide information regarding the intensity of the felt emotion [7, 11]. Harrigan suggests that the body's positions and actions provide a backdrop for helping to interpret the meaning of more subtle facial and vocal affect [18]. Wallbott [36] nevertheless observed discriminative features of emotion categories in both posture and movement quality of acted behaviors collected in lab.

Digital corpora have been collected to get detailed information on the expression of emotion in the different modalities. The Geneva Multimodal Emotion Portrayal database is an audiovisual corpus containing more than 7,000 videos recordings of portraits of 18 emotions acted by 10 professional actors [3].

Postural expressions of emotions were collected by Berthouze et al. [4, 6, 25, 37] using a VICON motion-capture system. 12 subjects acted Anger, Joy and Sadness. Their movements were recorded at 34 points of the body. The VICON system did not record the position of the fingers, thus creating confusions between some emotions. Kleinsmith and Bianchi-Berthouze [22] tried to identify the postural features that enable us to recognize the emotional dimensions such as valence and arousal. 111 affective postures were presented to 5 observers who had to judge the postures according to emotional dimensions. The authors observed that some postural features inform the level of dimensions (e.g. the opening of the body is used to assess the activation).

Whereas the importance of synchronization of changes in the various components of emotion is acknowledged [33], few studies considered how the face and the body do combine to express emotions.

Gunes and Piccardi [16] observed that emotions were better recognized when subjects were able to rely on facial expressions and gestures rather than when limited to either facial expressions or gestures.

Scherer and Ellgring [35] compared the expression of emotion across acoustic, face and body modalities. Different emotions, including members of the same emotion family were selected: Hot Anger, Cold Anger, Panic Fear, Anxiety, Despair, Sadness, Elation, Happiness, Interest, Boredom, Shame, Pride, Disgust, and Contempt. Facial actions were coded using FACS [9]. The coding of body expression combined both functional and anatomical approaches. According to Scherer and Ellgring [34], the multimodal response pattern is driven by the presence or absence of behavioral urges produced by appraisal rather than triggered by emotion-specific affect programs. Indeed, the respective multimodal pattern occurs frequently for several emotions, including positive emotions.

Hietanen and Leppänen [20] studied the perception of combinations of static facial expressions and dynamic manual actions of the hands. The emotions (Happiness, Anger and “Neutrality”) expressed by the face and hands were the same (congruent expressions) or different (incongruent expressions). They observed that the judgments were mainly based on the facial expressions, but were also affected by the manual expressions. An effect of hand movement quality was observed for facial expressions of Happiness. When the happy face was combined with either happy or neutral hand movements, the happy response was the same. But the proportion of happy responses dropped when the happy face was combined with angry hand movements as compared to happy responses when the happy face was combined with either neutral or happy hand movements. Thus, the perception of facial expressions of emotions can be affected by the expressive qualities of hand movements.

Meeren et al. [24] studied subjects' perception of emotion displayed in static pictures combining facial expressions and body postures. The two modalities conveyed the same or different emotions (Fear, Anger). They observed that when the emotions conveyed by the two modalities differ, the recognition of the emotion conveyed by the facial expression is influenced by the emotion conveyed by the posture. Congruent emotional body language improves recognition of facial expression, and incongruent emotional postural expression biases facial judgment toward the emotion conveyed by the body.

2.3 Nonverbal Expression of Emotion in Virtual Characters

Virtual agents have been designed to display facial expressions of emotions including blends of several emotions using a decomposition of the face in several areas [1, 26]. Models of gesture expressivity have also been proposed focusing on movement quality represented by parameters such as overall activation, spatial extent, temporal extent, fluidity, power and repetition [27]. Postural expressivity was considered to enable virtual characters to display interpersonal attitudes [2, 27].

Few studies considered the combinations of facial expressions and bodily expression of emotion in virtual agents. Using a multimodal corpus exploratory approach, Buisine et al. [5] defined a method for replaying annotated gestures and facial expressions. The authors describe an experimental study exploring how subjects perceive different replays but do not compare how facial and postural expressions are perceived individually vs. jointly.

To summarize, expressions of emotions via the face and the posture have been studied individually in Psychology but rarely jointly. Current computational models for generating multimodal expressions of emotion in intelligent virtual agents are also limited with this respect.

3 A Virtual Agent Technology for Studying Facial and Postural Expressions of Emotions

The Living Actor™ technology was developed by Cantoche and launched in January 2002 for the Major Accounts market which includes e-marketing and e-learning. Our client's objectives depend on the mission of the Living Actor™ avatar (assistant, guide, tutor …) and all clients understand that emotional full body avatars are a unique personalized solution to communicate with the young generation as they are similar to the avatars they are used to interact with in video games. Today, more than 200 large companies in Europe and in the USA use Living Actor™ and the market is growing. According to Gartner (April 2007), by the end of 2011, 80 percent of active Internet users (and Fortune 500 enterprises) will have a "second life", but not necessarily *in* the virtual world called *Second Life*.

Living Actor™ is an innovative real time process for emotion-rich avatars driven solely by the user's voice that allows end users and companies to generate advanced personalized video calls, video content on mobile and chat-based services, or interactive content on web applications.

The system is based on speech-to-animation engine, which provides automatic animation of avatars by a human voice: Avatars' mouth, expressions, and gestures move in sync with the speech. After analyzing the audio signal (recorded or streamed) and detecting sound volume, pace of the speech, prosody, and others parameters, the engine associates and synchronizes it with the virtual character's animation data and the best behaviors and actions are automatically generated to realize the speech visualization by full-body avatars.

A multi-export engine enables the display of avatars in different formats (3D, Flash, AVI, 3GP, FLV). The avatars are compatible with various types of interfaces and devices depending on the technological specifications—ranging from PCs, and mobile phones to set-top boxes for interactive television, or large screen displays for mass audiences.

This industrial animation technology can benefit from experimental studies to evaluate how users will perceive the expressions of emotions displayed by the characters.

4 First Study: Facial and Postural Expressions of Congruent Emotion

This first study aimed at providing answers to the following question: how do different nonverbal modalities (face only, posture only, face and posture) compare in terms of how human subjects perceive the emotion expressed by a virtual character? As the study of combinations of facial and postural expressions of emotion is still exploratory, we decided to test a subset of basic emotions for which there is at least a minimum data in the literature to which we would be able to compare our results: Sadness, Fear, Anger, Surprise and Joy. For the same reason, we focused on a single expression for each emotion.

4.1 Methodology

We designed facial expressions of these emotions based on a selection of some of the distinctive clues described by Ekman [12]. We also designed postural expressions of emotions. We inspired from a database of acted postures [23]. This set of nonverbal expressions was implemented using the virtual character technology presented above. Some of the resulting set of stimuli used for the two experiments described in this paper is illustrated in a video submitted along with the current paper.

For this first study, we designed 15 canned animations of 3 seconds each (from a neutral pose to the expression of emotion and back to neutral pose). 5 animations were face-only expressions of the emotions that we selected (the rest of the body is hidden), 5 animations displayed the postural expression of the same emotions (the head is hidden), and 5 animations displayed combined facial and bodily expressions of emotions (Fig. 1). We decided to hide the irrelevant part of the body rather than involving a "neutral" expression to avoid any influence of this neutral expression on the expressive part.

59 students from an engineering school participated in this first study. They were divided randomly into two groups. Each group viewed only half of the animations.

Fig. 1. Expressions of anger in the face only (left), posture only (middle), and combination of face and posture (right)

A first animation of the character displaying a deictic gesture was used as a training animation so that subjects had seen the character once before evaluating its emotional expressions. After viewing an animation, subjects had to answer a questionnaire. They had to report how much they perceived each of the emotions in the animation using 5 points Likert scales. Subjects were able to select several emotion labels for each animation. They also had to rate the activation and the valence of the emotion.

4.2 Results

Emotion categories perceived in each animation. We computed repeated measures ANOVA in order to estimate the degree of recognition of the intended emotion in each modality (Table 1).

Anger was the most perceived emotion in our animations of Anger for the three modalities. *Sadness* was the most perceived emotion in our animations of Sadness for the three modalities. *Joy* was the most perceived emotion in the Face animation and Face+Posture designed for Joy. Nevertheless the Posture animation designed for Joy was equivalently perceived as Joy or Surprise. The Face animation designed for *Surprise* was equivalently perceived as Fear, Surprise, or Sadness. Surprise was the most perceived emotion in the Posture animation designed for Surprise. The Face+Posture animation designed for Surprise was equivalently perceived as Fear or Surprise. The Face animation designed for *Fear* was equivalently perceived as Sadness or Neutral. The Posture and Face+Posture animations designed for Fear were equivalently perceived as Surprise or Fear.

In summary, the animations designed for Anger, Sadness and Joy were well recognized. Animations designed for Surprise and Fear were both perceived as blends of Surprise and Fear.

These results are similar to Hall & Matsumoto where people perceived blends of emotions in static pictures intended to express a single emotion [17].

These results are also in line with those observed by Gosselin and Kirouac [15]. They observed that Ekman's pictures of facial expressions of Joy, Anger, Sadness and Surprise were better recognized than facial expressions of Fear and Disgust. Yet, in

Table 1. Average perception of emotion in the various modalities (min = 0, max = 4). Averages that are significantly different from others are displayed in bold.

Animations	Modalities	Joy	Anger	Fear	Surprise	Sad	Neutral	ANOVA *p*<,001
ANGER	Face	0,05	**3,21**	0,16	0,63	0,42	0,05	*F*(5, 90)=61,79
	Posture	0,05	**3,47**	0,28	0,5	0,28	0	*F*(4, 68)=100,82
	Face+Posture	0,05	**3,63**	0,21	0,05	0,31	0,05	*F*(5, 90)=176,44
JOY	Face	**2,31**	0,05	0,1	0,22	0,1	0,79	*F*(5, 85)=31,43
	Posture	**2,47**	0,84	0,26	**2,26**	0,16	0	*F*(4, 72)=14,90
	Face+Posture	**3,74**	0,05	0	1,58	0	0,5	*F*(3, 54)=124,44
SAD	Face	0,16	0,1	0,68	0,26	**3,16**	0,26	*F*(5, 90)=51,32
	Posture	0,05	0,28	0,22	0,44	**1,95**	0,89	*F*(5, 85)=7,90
	Face+Posture	0	0,05	0,58	0,37	**3,47**	0,37	*F*(4, 72)=67,86
SURPRISE	Face	0,12	0,28	**1,53**	**1,7**	1,69	0,39	*F*(5, 150)=18,95
	Posture	1,23	0,32	1,46	**2,89**	0,09	0,06	*F*(4, 132)=32,28
	Face+Posture	0,34	0,26	**2,38**	**3,02**	0,32	0,17	*F*(5, 165)=76,01
FEAR	Face	0,21	0,12	0,39	0,15	1,12	**1,8**	*F*(5, 150)=19,50
	Posture	0,64	0,17	**2,77**	**2,63**	0,67	0,08	*F*(5, 170)=48,17
	Face+Posture	0,26	0,41	**2,54**	**2,79**	0,6	0	*F*(4, 132)=59,59

our study, we observed that Surprise was confused with Fear for the three conditions. This means that the way we specified this animation has to be improved.

Emotional dimension perceived in each animation. With respect to the perception of dimensions of emotions, our results show that subjects perceived the intended valence for each condition, but not the intended activation. When rating face only animations of each emotion, we did not observe any significant difference between animations intended to display active versus passive emotions. Yet, subjects perceived correctly the activation dimension from our posture only animations (Table 2).

Table 2. Average perception of activation and valence for each emotion and for each modality (min = 0, max = 4)

Animations	Modalities	Active	Passive	Positive	Negative
ANGER	Face	1,42	1,37	0,31	**3,42**
	Posture	**3,16**	0,21	0,42	**3,31**
	Face+Posture	**3,21**	0,47	0,16	**3,63**
JOY	Face	0,78	**2,26**	**2,94**	0,37
	Posture	**3,21**	0,1	**2,36**	1,16
	Face+Posture	**3,37**	0,21	**3,63**	0,05
SAD	Face	1,47	1,31	0,53	**2,89**
	Posture	1,05	**1,84**	0,47	**2,37**
	Face+Posture	1,22	**2,05**	0,33	**3,1**
SURPRISE	Face	1,42	**2,11**	0,65	**2,64**
	Posture	**3**	0,32	1,85	1,21
	Face+Posture	**2,71**	0,62	0,82	**2,14**
FEAR	Face	0,77	**2,25**	0,71	**1,64**
	Posture	**2,78**	0,38	1	**2,15**
	Face+Posture	**2,82**	0,97	0,76	**2,26**

This effect of correct perception of activation via the posture and incorrect perception of activation via the face might be explained by the way we designed our animations. Yet, the fact that the effect was found for all our facial animations and not for any postural animation might also suggest that facial expressions might not be as appropriate as postural expressions for perceiving the activation of emotions.

Comparing the recognition of each emotion in the three modalities. Finally, we proceeded to a one-way ANOVA to compare the different modalities. Our goal was to study if the perceived emotion depends on the modality (e.g. does the level of perceived Anger depends on the modality? Do subjects perceive more Anger when this emotion is displayed both in the face and in the posture than when it is displayed in a single modality? Do face only and posture only are equivalent?).

The perception of *Anger* did not differ between the three conditions. The perception of *Joy* was higher in the Face+Posture condition than in the Face condition and in the Posture condition. The perception of *Sadness* was higher in the Face and in the Face+Posture conditions than in the Posture condition. The perceptions of *Surprise* and *Fear* were higher in the Posture and Face+Posture modalities than in the Face modality.

4.3 Discussion

This first study enabled us to validate some of our animations. The expressions of Joy, Anger and Sadness are well recognized by the subjects in the three modalities (Face only, Posture only and Face+Posture). The expressions of Surprise and Fear are more ambiguous and are less well recognized.

The results also show that the contribution of the different modalities depends on the emotion. For Anger, subjects recognize its expression regardless of the modality. For the other emotions, the mode of presentation influences recognition. However, it is still difficult to say if this is due to our animations or to a differentiated perceptive processing. Nevertheless, the results concerning the perception of the level of activation suggest that the perceptive processing is different for the two modalities.

5 Second Study: Facial and Postural Expressions of Incongruent Emotions

This second study aimed at providing answers to the following questions: how are perceived animations that blend two different emotions expressed in the face and in the posture? What is the influence of one modality over the other modality?

5.1 Methodology

We kept the emotions which were well recognized in the first study (Anger, Sadness and Joy) so as to avoid an effect of a bad recognition of a single emotion on the overall perception of a blend. Furthermore, these three emotions cover positive and negative valence, as well as low and high arousal.

We designed 6 canned animations of incongruent emotions (left column of Table 4). We also kept the 3 congruent animations combining the same emotion in both the Face and the Posture (Anger, Sadness and Joy) in order to compare them with the incongruent animations in the same experiment settings (Fig. 2 illustrates congruent and incongruent expressions). Each animation lasted 3 seconds (from a neutral pose to the expression of emotion and back to neutral pose).

Another group of 21 subjects (12 females, 9 males, average age 20 years) participated in the experiment. All of them were first or second year students at the University.

Fig. 2. Frame of congruent expressions of sadness in the face and in the posture (left) and incongruent facial expression of Joy and postural expression of Sadness (right)

As in the first study, a first animation of the character displaying a deictic gesture was used as a training animation so that subjects had seen the character once before evaluating its emotional expressions. After viewing an animation, the subjects had to answer a questionnaire. They had to report how much they perceived each of the emotions in the animation using 5 points Likert scales. Subjects were able to select several emotion labels for each animation. They also had to rate the level of activation and the valence of the emotion.

5.2 Results

Collected data was analyzed by repeated measures ANOVA and using t-test.

Emotion categories perceived in each animation. As observed in the first study, the intended emotion is well recognized in each congruent animation (Table 3).

Table 3. Average perception of emotion in congruent animations

	Joy	Anger	Fear	Surprise	Sadness	ANOVA $p<,001$
Face+Posture of Anger	0,00	**3,52**	0,05	0,19	0,10	$F(4, 76)=283,20$
Face+Posture of Joy	**3,67**	0,00	0,00	1,24	0,00	$F(1, 20)=85,00$
Face+Posture of Sadness	0,10	0,14	0,19	0,33	**2,19**	$F(5, 100)=34,75$

Regarding the 6 incongruent combinations, in all but one of the animations subjects reported perceiving only the emotion expressed via the face. The only exception is the animation combining the facial expression of Sadness with the postural expression of Anger. This animation was perceived as a blended expression of Sadness and Anger (Table 4).

Table 4. Average perception of emotion in incongruent animations

Face	Posture	Joy	Anger	Fear	Surprise	Sadness	Neutral	ANOVA *p*<,001
Joy	Sad	**2,05**	0,10	0,05	0,76	0,10	0,29	*F*(5, 100)=17,35
Joy	Anger	**2,43**	0,10	0,00	0,67	0,00	0,19	*F*(3, 60)=41,62
Anger	Joy	0,00	**3,00**	0,19	0,48	0,14	0,00	*F*(3, 60)=126,76
Anger	Sad	0,00	**2,81**	0,10	0,29	0,24	0,19	*F*(4, 80)=54,85
Sad	Anger	0,14	**1,52**	0,62	0,33	**1,29**	0,14	*F*(5, 100)=9,21
Sad	Joy	0,29	0,48	0,71	1,00	**1,95**	0,33	*F*(5, 100)=8,10

Table 5. Average perception of emotional dimension in incongruent animations

Face	Posture	Active	Passive	Positive	Negative
Joy	Sad	1,57	1,33	**2,62**	0,62
Joy	Anger	**2,52**	0,52	**2,95**	0,33
Anger	Joy	**2,95**	0,29	0,33	**2,95**
Anger	Sad	**2,05**	1,24	0,24	**3,14**
Sad	Anger	**2,38**	0,65	0,71	**2,52**
Sad	Joy	**2,67**	0,48	0,67	**2,52**

Perception of congruent and incongruent animations. Each congruent animation was compared to the four incongruent animations. For example, the animation of Anger was compared to 1) the animation combining the facial expression of Sadness with the postural expression of Anger, 2) the animation combining the facial expression of the Joy with the postural expression of Anger, 3) the animation combining the facial expression of Anger with the postural expression of Sadness, 4) and the animation combining the facial expression of Anger with the postural expression of Joy. The results show that the *Valence* and the *Activation* are better perceived when the congruent animations are displayed (Table 5). The only exception is the animation combining the facial expression of Anger with the postural expression of Joy. The perception of *Activation* for this animation is the same as for the congruent animations of Joy and of Anger. This result is not surprising since Anger and Joy are two emotions which have a high level of activation.

Comparison of incongruent animations displaying the same postural expression. Subjects perceive the emotions displayed in the facial expression. For example, the perception of *Sadness* was higher in the animation combining the facial expression of Sadness with the postural expression of Anger than in the animation combining the facial expression of Joy with the postural expression of Anger. Concerning the valence (Table 5), when the postural expression is negative, the perception of *Negative Valence* is higher in the animation combining the same negative valence for the facial and postural expression than in the animation combining different valence. Thus, the subjects perceive the emotion and the valence conveyed by the facial expressions.

Comparison of incongruent animations displaying the same facial expression. Subjects consider less the posture in order to elaborate their emotional judgment. Indeed, they do not report more the emotion displayed by the postural expression. For example, the perception of *Anger* is not higher in the animation combining the facial expression of Sadness with the postural expression of Anger than the animation combining the facial expression of Sadness with the postural expression of Joy. However, the perception of *Activation* is higher in the animation combining the same level of activation for the facial and postural expression than in the animations combining different levels of activation (Table 5).

Comparing incongruent animations. The results suggest that the categorical recognition is more influenced by the facial expression modality than by the postural modality. Furthermore, subjects preferentially used the postural modality to establish a judgment about the level of activation. For example, the perceptions of *Activation*, of *Negative Valence* and of *Sadness* were higher in the animation combining the facial expression of Sadness with the postural expression of Joy than in the animation combining the facial expression of Joy with the postural expression of Sadness. The perception of *Joy* was higher in the animation combining the facial expression of Joy with the postural expression of Sadness than in the animation combining the facial expression of Sadness with the postural expression of Joy. Thus, Joy and Sadness seem to be better identified when they are conveyed by facial expressions.

In summary, facial expressions are crucial to develop a perceptual judgment. The emotions are better recognized when they are conveyed by facial expressions than by the postural expressions.

6 Conclusions

Interest in body movement and gestures as components of emotional expression is relatively recent. Wallbott [36] observed some features of posture and movement quality that enabled him to discriminate between some emotion categories. On the contrary, Ekman and Friesen [11] suggested that bodily movements inform about the intensity of emotion, but not about emotion categories.

In this paper we described two experimental studies exploring how human subjects perceive combinations of facial and postural expressions of emotions. The first study aimed at assessing the relative contributions of the face and the posture in the perception of a single emotion. We observed that emotional congruence at the level of the body and the face facilitates the recognition of emotion categories and dimensions compared to face only or posture only presentations. Nevertheless, face only and posture only also enabled us to recognize emotion categories. The second study considered the influence of incongruent expressions. Judgments were mainly based on the facial expressions, but were nevertheless affected by postural expressions. In such incongruent case, posture revealed to be useful for assessing the perceived activation of the emotion.

Our results support the two theoretical positions mentioned above. Indeed, our first study shows that when postures are presented alone, subjects are able to perceive the emotion category. When postures are presented with an incongruent emotional combination, subjects rely on the posture to make judgments about the emotional activation. However in our studies, the dynamics of each emotional expression were not specified. Each expression had the same global duration, the same duration for attack, sustain and decay. These expressive properties of movement quality will be investigated jointly with postures since they also play a major role in the perception of emotion.

Coordination and synchronization of nonverbal modalities is an important issue for virtual characters to express emotions. Computational models of expressive characters need to consider how such combinations are perceived. Current studies seldom consider combinations of posture and face, and when they do, they consider static pictures [24] or stimuli what are not virtually integrated as a single body [19]. The originality of our work also lies in our integrated and animated stimuli. This methodology has the advantage of confronting the subjects to stimuli similar to virtual agent encountered on the web. It also helps us to understand how we perceive combinations of non-verbal expressions of emotion.

Acknowledgments

The research was partly supported by the project "ANR-07-TLOG-Affective Avatars". The authors thank the students who participated in the study.

References

1. Albrecht, I., Schröder, M., Haber, J., Seidel, H.-P.: Mixed feelings: Expression of non-basic emotions in a muscle-based talking head. Special issue of Journal of Virtual Reality on Language, Speech & Gesture 8(4) (2005)
2. Ballin, D., Gillies, M., Crabtree, B.: A Framework For Interpersonal Attitude And Non-Verbal Communication In Improvisational Visual Media Production. In: The First European Conference on Visual Media Production IEE, London (2004)
3. Bänziger, T., Scherer, K.: Using Actor Portrayals to Systematically Study Multimodal Emotion Expression: The GEMEP Corpus. In: Paiva, A., Prada, R., Picard, R.W. (eds.) ACII 2007. LNCS, vol. 4738, pp. 476–487. Springer, Heidelberg (2007)
4. Bianchi-Berthouze, N., Kleinsmith, A.: A categorical approach to affective gesture recognition. Connection Science 15(4), 259–269 (2003)
5. Buisine, S., Abrilian, S., Niewiadomski, R., Martin, J.C., Devillers, L., Pelachaud, C.: Perception of blended emotions: From video corpus to expressive agent. In: Gratch, J., Young, M., Aylett, R.S., Ballin, D., Olivier, P. (eds.) IVA 2006. LNCS (LNAI), vol. 4133, pp. 93–106. Springer, Heidelberg (2006)
6. De Silva, R., Bianchi-Berthouze, N.: Modelling human affective postures: an information theoretic characterization of posture features. Computer Animation and Virtual Worlds 15(3-4), 269–276 (2004)
7. Dittman, A.T.: The role of body movement in communication. In: Nonverbal behavior and communication. Lawrence Erlbaum, Mahwah (1987)
8. Ekman, P.: An argument for basic emotions. Cognition and Emotion 6, 169–200 (1992)
9. Ekman, P., Friesen, W.: Manual for the facial action coding system. Consulting Psychology Press, Palo Alto (1978)
10. Ekman, P., Friesen, W.: A new pan-cultural facial expression of emotion. Motivation and Emotion 10, 159–168 (1986)
11. Ekman, P., Friesen, W.V.: Nonverbal behavior and psychopathology. In: Friedman, R.J., Datz, M.M. (eds.) The psychology of depression: contemporary theory and research, pp. 203–232. Winston & Sons, Washington (1974)

12. Ekman, P., Friesen, W.V.: Unmasking the face. A guide to recognizing emotions from facial clues. Prentice-Hall Inc., Englewood Cliffs (1975)
13. Ekman, P., Friesen., W.: Unmasking the face: a guide to recognizing emotions from facial clues. Prentice-Hall, Englewood Cliffs (1975)
14. Fontaine, J.R., Scherer, K.R., Roesch, E.B., Ellsworth, P.: The world of emotion is not two-dimensional. Psychological Science 18(2), 1050–1057 (2007)
15. Gosselin, P., Kirouac, G.: Le décodage de prototypes emotionels faciaux. Canadian journal of experimental psychology 49(3), 313-329 (1995)
16. Gunes, H., Piccardi, M.: Bi-modal emotion recognition from expressive face and body gestures. Journal of Network & Computer Applications 30(4), 1334–1345 (2007)
17. Hall, J.A., Matsumoto, D.: Gender differences in judgments of multiple emotions from facial expressions. Emotion 4(2), 201–206 (2004)
18. Harrigan, J.A.: Proxemics, kinesics, and gaze. In: Harrigan, J.A., Rosenthal, R., Scherer, K. (eds.) The new handbook of methods in nonverbal behavior research. Series in Affective Science, pp. 137–198. Oxford University Press, Oxford (2005)
19. Hietanen, J.K., Leppänen, J.M.: Judgment of other people's facial expressions of emotions is influenced by their concurrent affective hand movements. Scandinavian Journal of Psychology 49, 221–230 (2008)
20. Hietanen, J.K., Leppänen, J.M.: Judgment of other people's facial expressions of emotions is influenced by their concurrent affective hand movements. Scandinavian Journal of Psychology 49, 221–230 (2008)
21. Izard, C.: Human Emotions. Plenum, New York (1977)
22. Kleinsmith, A., Bianchi-Berthouze, N.: Recognizing affective dimensions from body posture. In: Paiva, A.C.R., Prada, R., Picard, R.W. (eds.) ACII 2007. LNCS, vol. 4738, pp. 48–58. Springer, Heidelberg (2007)
23. Kleinsmith, A., De Silva, R., Bianchi-Berthouze, N.: Cross-Cultural Differences in Recognizing Affect from Body Posture. Interacting with Computers 18(6), 1371–1389 (2006)
24. Meeren, H.K.M., van Heijnsbergen, C., de Gelder, B.: Rapid perceptual integration of facial expression and emotional body language. Proceedings of the National Academy of Sciences of the USA, 16518–16523 (2005)
25. Mota, S., Picard, R.W.: Automated posture analysis for detecting learner's interest level. In: Workshop on Computer Vision and Pattern Recognition for Human-Computer Interaction (2003)
26. Niewiadomski, R.: A model of complex facial expressions in interpersonal relations for animated agents. PhD Thesis, University of Perugia (2007)
27. Pelachaud, C.: Multimodal expressive embodied conversational agents. In: Proceedings of the 13th annual ACM international conference on Multimedia. ACM, Hilton (2005)
28. Plutchik, R.: A general psychoevolutionary theory of emotion. In: Plutchik, R. (ed.) Emotion: Theory, research, and experience, Theories of emotion, Academic, New York. vol. 1, pp. 3–33 (1980)
29. Russell, J.A.: A circumplex model of affect. Journal of Personality and Social Psychology 39(6), 1161–1178 (1980)
30. Russell, J.A., Mehrabian, A.: Evidence for a three-factor theory of emotions. Journal of Research in Personality 11, 273–294 (1977)
31. Scherer, K.R.: The nonverbal dimension: A fad, a field, or a behavioral modality? The social dimension: European developments in social psychology, 160–183 (1984)
32. Scherer, K.R.: Analyzing Emotion Blends. In: Fischer, A. (ed.) Proceedings of the Xth Conference of the International Society for Research on Emotions, pp. 142–148 (1998)

33. Scherer, K.R.: Emotion. In: Stroebe, M.H.W. (ed.) Introduction to Social Psychology: A European perspective, pp. 151–191. Blackwell, Oxford (2000)
34. Scherer, K.R., Ellgring, H.: Are facial expressions of emotion produced by categorical affect programs or dynamically driven by appraisal? Emotion 7(1), 113–130 (2007)
35. Scherer, K.R., Ellgring, H.: Multimodal Expression of Emotion: Affect Programs or Componential Appraisal Patterns? Emotion 7(1) (2007)
36. Wallbott, H.G.: Bodily expression of emotion. European Journal of Social Psychology 28, 879–896 (1998)
37. Woo, W., Park, J., Iwadate, Y.: Emotion Analysis from Dance Performance Using Time-Delay Neural Networks. In: JCIS-CVPRIP 2000, pp. 374–377 (2000)