ISSN 2464-4617 (print)
ISSN 2464-4625 (CD)
Computer Science Research Notes
CSRN 2802
Short Papers Proceedings
http://www.WSCG.eu

Perception of basic emotion blends from facial expressions of virtual characters: pure, mixed, or complex?

Meeri Mäkäräinen
Aalto University
Finland
meeri.makarainen@aalto.fi

Jari Kätsyri
Maastricht University
The Netherlands
jari.katsyri@maastrichtuniversity.nl

Tapio Takala
Aalto University
Finland
tapio.takala@aalto.fi

ABSTRACT

As animated virtual characters in games, movies and other applications become more humanlike, it becomes more and more important to be able to imitate the complicated facial behaviour of a real human. So far, facial expression animation and research have been dominated by the basic emotions view, limited to the six universal expressions: anger, disgust, fear, joy, sadness and surprise. More complex facial expressions can be created by blending these basic emotions, but it is not clear how these blends are perceived. Are they still perceived as basic emotions or combinations of basic emotions, or are they perceived as expressions of more complex emotions? We used a series of online questionnaires to study the perception of all pairwise blends of basic emotions. The blends were produced as a sum of facial muscle activations in the two basic emotions, using a physically-based, animated face model. Our main finding is that several basic emotion blends with an opposite valence are perceived as complex emotions that are neither pure emotions nor their blends. Blends of basic emotions with a similar valence are typically perceived as pure basic emotions (e.g., a blend of anger and disgust is perceived as pure anger). Only one of the blends (joy+surprise) was perceived as a blend of two different basic emotions.

Keywords

virtual agent, basic emotions, facial expressions, mixed emotions, affective computing, perception.

1 INTRODUCTION

Animated characters are used widely in games, movies and virtual applications, and the recent advances in rendering and modeling techniques have enabled them to be highly human-like. This creates pressure to develop understanding of more fine-detailed facial expressions of animated characters, to enable the development of their facial behaviour in order to keep up with the development of appearance. Facial expressions of emotion are often conceptualized as discrete expressions of *basic emotions*. Anger, disgust, fear, joy, sadness, and surprise are considered as six basic emotions with universally recognizable characteristic facial expressions [6]. Although the basic emotion view remains debated, it still remains a useful basis for facial expression research. Dimensional emotion models, such as the pleasure-arousal-dominance model have been useful in the research on emotional states and reactions, but research on the perception of facial expressions has been largely based on the basic emotion approach. In particular, the facial expressions of most virtual agents have been based on the basic emotion view [11, 9].

Although the basic emotion approach may be a good basis for modeling facial behaviour in virtual agents, a set of six separate facial expressions is very limited for producing natural facial behaviour. For a wider variety of different facial expressions, a common solution is to create composite facial expressions that combine two basic emotions into one expression. Several techniques have been introduced, including interpolation in a two-dimensional emotion space [3], displaying different emotions in the upper and lower parts [14] or left and right halves of the face [1], and additive methods [16]. In some cases, a dimensional emotion model is used where each separate emotion is considered a point in an emotional space, and each emotion is combined only with its nearest neighbours [2].

Blending two basic emotions in one facial expression can be justified because the basic emotions are not mutually exclusive. For example, happiness and sadness [12], or amusement and disgust [10], can be experienced together. *Mixed emotions*, in which two or more basic emotions are present, can produce a powerful experience, such as the pleasure of listening to sad music. It is also possible for real humans to make facial expressions that combine two basic emotions. Sometimes the facial displays of old people are interpreted as mixed emotions, because their wrinkles suggest another emotion than their actual expression. Thus, mixed signals on the face are associated with mixed emotions. It is important for virtual agents to be able to express mixed emotions, as this increases the perceived empathy of the agent [15]. However, it is unclear whether the perception of blended expressions of virtual characters can be better understood as a perception of two simultaneous basic emotions or a perception of a third, more complex emotion.

So far, most studies on the perception of blends have used stimuli that are based on photographs [4]. However, from previous studies we know that there are some typical differences in how people perceive the basic emotions from natural and synthetic facial expressions, and therefore the perception of blends may also be different for natural and virtual faces. As for basic emotions, synthetic faces are typically poorer in communicating fear [9, 18] and disgust [9, 5], but on the other hand, they are often better than natural faces in communicating sadness [18, 5].

The main goal of this study is to investigate whether virtual agents can communicate complex emotions using blends of facial expressions of basic emotions. Due to differences in the perception of natural and synthetic faces, it is important to study the perception of blends specifically on synthetic faces. We use a physically-based facial model which is capable of producing all muscle actions required for the expressions of basic emotions, and combined muscle actions required for the blends.

We study all pairwise blends of basic expressions. The original basic expressions used to create a blend are called *parent expressions*, and the corresponding emotions are called *parent emotions*. The term *blend* refers to the combination of two parent expressions. Emotions that do not belong to the set of basic emotions are called *complex emotions*. The term *target emotion* is used in our analyses for an emotion that an expression is supposed to present (parents of a blend, or basic emotions in their original expressions).

This study consists of two experiments. The first experiment focuses on whether the used stimuli were perceived as pure basic emotions or combinations of two basic emotions. The second experiment focuses on whether blended facial expressions are perceived as expressing complex emotions.

2 VIRTUAL FACE MODEL

Various approaches can be adopted to model blends of facial expressions in virtual characters. Virtual faces themselves can be built in a variety of different ways, and several techniques can be used for constructing pure and blended facial expressions of basic emotions. The situation is made even more difficult by the fact that basic emotions and their blends can be expressed in several different ways even on a natural face.

To create the blends for this study, we used a facial animation model described in our previous work [13] (based on [20]). This facial model is physically based, and it has deformable skin and facial muscles. The facial tissue is implemented using a mass-string-damper model with two layers of cubical elements. The lower layer is attached to the bones which have been modeled beneath the tissue. Muscles are attached to the lower nodes in the top layer and to bone surface at the other end. The facial tissue has been modeled so that it is slightly asymmetric, similarly as in real human faces.

The animation model uses FACS (Facial Action Coding System) [7] as a control mechanism for facial expressions. FACS defines facial actions in terms of Action Units (AU). For example, AU12 is the action of the *Zygomaticus major* muscle pulling lip corners to a smile, and AU4 is the action of the *Corrugator supercilii* muscle making a frown. Our model includes facial examples of basic expressions that were created by selecting one prototypical AU combination defined in FACS for each emotion. The combinations used for the basic expressions are: joy 6+12 (with the addition of AU7), sadness 1+4+15, fear 1+2+4+5+20+26, anger 4+5+7+23, surprise 1+2+5+27 and disgust 10+17. The model also includes the possibility to blend any two facial expressions. This was implemented by adding the muscle activations of the parent expressions. This technique has advantages for creating blended expressions: muscle activations are anatomically correct, facial actions from both parent expressions remain present, and any pair of facial expressions can be blended easily.

3 EXPERIMENT 1: BASIC EMOTIONS IN BLENDS

The first experiment was designed to study the extent to which pure basic emotions are perceived in blends. Our secondary goal was to collect complex emotion words to be studied more thoroughly in Experiment 2. A further objective was to validate the basic expressions of our virtual face model.

29 volunteers were recruited via e-mails and social media. The sample consisted of 16 female and 13 male participants aged between 19 and 63 years ($M = 30.7, SD = 9.7$).

3.1 Methods

Using the model described above, we prepared 21 videos of expressions on a virtual face. They included the six basic emotions (anger, disgust, fear, joy, sadness and surprise) and all of their 15 pairwise blends. Neutral expression was not included, because dynamic but affectively neutral facial expressions would not have been analogous to our other stimuli. We recorded the transition from neutral to peak expression in a video clip with a duration of two seconds. The peak expression was reached in approximately one second.

As control stimuli, we used 21 corresponding videos produced by morphing basic emotion expressions posed by a human actor. These were selected from one actor (MO) in the Ekman and Friesen's Pictures of Facial Affect collection [8], which has become a standard database in this field. We used image morphing to blend basic expressions, which is a conventional method in facial expression studies (for example [4]). Photographs of each two basic expressions were blended with the ratio 50%–50% using the application MorphThing (http://www.MorphThing.com). Video sequences were created from the static expressions by morphing them with the neutral face using Sqirlz Morph software (http://www.xiberpix.com/SqirlzMorph.html). Similar approach has been used previously to create dynamic stimuli [17]. Only the face region was morphed, while the surrounding region was taken from the neutral face image.

Although forced-choice method is often used to study the recognition of basic emotions from facial expressions, this method is tied to a predefined list of emotion words and it makes the possibly incorrect assumption that a specific facial expression is only associated with one emotional state. In this study, we asked our participants to describe their perception of the facial expressions in more detail.

The evaluations were done using an online questionnaire. All 42 videos were evaluated one by one in random order. First, the participants rated each video on all six basic emotion dimensions using visual sliders on a scale ranging from no emotion (0) to extremely intense (100). This method enabled us to get detailed information not only about the recognized primary emotion, but also about the recognition of less intense secondary emotions. Second, to measure the recognition of complex emotions, the participants were asked to provide open responses to the question *"What other emotions do you see in the facial expression (if any)?"*

3.2 Analyses

To validate the modeled basic expressions, and to evaluate whether some of the blends are also perceived as *pure basic emotions*, we converted the six basic emotion ratings of each facial expression into a recognition score measuring whether one of them clearly dominates.

A facial expression can be thought to unambiguously display one *basic emotion* if people consistently give higher ratings to that emotion compared to all others. Thus, we defined the recognition score as the difference between the target emotion rating and the second highest rating (in case any other emotion received a higher rating than the target emotion, the score was negative). Formally, this recognition score can be defined as

$$RSpure_i(E) = R_i(E) - \max_{k \neq i}\{R_k(E)\}, \qquad (1)$$

where E is the facial expression, $R_k(E)$ is the rating for emotion k in the expression E, i is the targeted emotion, and k has six possible values: anger, disgust, fear, joy, sadness and surprise.

This recognition score is more accurate than a mean statistic for the targeted emotion, because it also takes into account how distinctive the target emotion was with respect to non-target emotions. A simple mean evaluation for a specific emotion can be high even though this emotion is considered a secondary emotion by most participants.

To measure whether a blend is recognized as a *mixed emotion*, we used a recognition score that has a positive value when both parent emotions receive higher ratings than any of the other emotions. This score is defined as

$$RSmix_{ij}(E) = \min\{R_i(E), R_j(E)\} - \max_{k \neq i; k \neq j}\{R_k(E)\}, \qquad (2)$$

where E is the facial expression, i and j are the targeted emotions with $R_i(E)$ and $R_j(E)$ their respective ratings, and $R_k(E)$ is the rating for a non-targeted emotion k. Again, the indices have six possible values: anger, disgust, fear, joy, sadness and surprise.

The score is positive only if the targeted emotions receive the highest and the second highest ratings among individual evaluations. Facial expressions meeting this strict requirement can be considered as unambiguous expressions of mixed emotions.

We used Wilcoxon signed-rank tests to determine whether the recognition scores were statistically

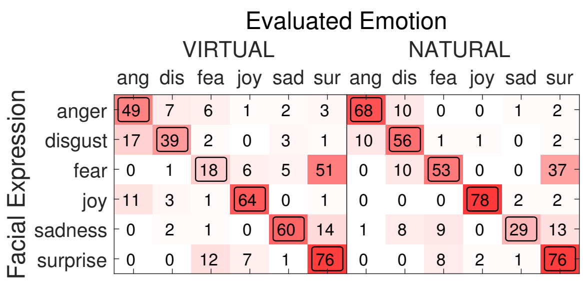
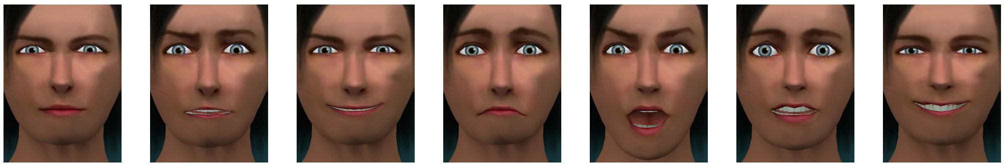


Figure 1: Perception of the expressions of basic emotions on virtual and natural faces presented as confusion matrices. Mean ratings of each emotion for each facial expression are presented as numbers and colour intensity (colour intensity is proportional to the number in the cell). Rectangles around numeric values indicate the targeted basic emotions.

ISBN 978-80-86943-41-1

emotion i	anger	disgust	fear	joy	sadness	surprise
$RSpure_i$ virtual	33.3***	18.5*	-33.3***	50.7***	45.6***	58.1***
$RSpure_i$ natural	56.7***	42.1***	11.4	74.7***	3.8	65.4***

Table 1: Virtual facial expressions of basic emotions at their emotional apex. Mean recognition scores $RSpure_i$ are listed for the basic expressions on virtual and natural faces. A positive score indicates that the expression was recognized correctly. Statistically significant scores are marked with asterisks.



emotion i emotion j	anger disgust	anger fear	anger joy	anger sadness	anger surprise	disgust fear	disgust joy
$RSmix_{ij}$ virtual	2.8	-38.4***	3.5	-7.3**	-16.5*	-42.1***	-19.0**
$RSpure_i$ virtual	**21.7****	-40.5***	1.7	-63.8***	-50.0***	-60.3***	-45.5***
$RSpure_j$ virtual	-31.4***	-19.7	-8.4	**54.7*****	17.1	-7.4	12.4
$RSmix_{ij}$ natural	9.7	-27.4***	-15.1**	-15.2**	-19.3**	-8.5	-3.3
$RSpure_i$ natural	-14.6	-5.7	**26.8****	14.9	-33.3***	**45.8*****	11.1
$RSpure_j$ natural	7.9	-37.9***	-47.9***	-37.7***	7.6	-59.0***	-22.8**

disgust sadness	disgust surprise	fear joy	fear sadness	fear surprise	joy sadness	joy surprise	sadness surprise
-0,1	-21,8**	-23,1**	-24,1**	8,1	-9,9*	**36,1*****	-25,3**
-63.8***	-45.9***	-40.6***	-26.5**	-45.8***	-0.4	-15.2*	-63.8***
61.5***	8.0	8.9	-11.4	**40.4*****	-22.2*	13.7	**23.1***
-5.5	-7.4*	-30.0***	-36.0***	**12.8***	-13.8**	2.7	-7.1
27.7***	**40.9*****	-29.1***	-13.3	-44.3***	0.9	-5.0	-31.8**
-35.1***	-49.9***	-17.1	-44.3***	**38.6****	-26.1***	-4.5	17.1

Table 2: The blends at their emotional apex displayed on the virtual face, and mean recognition scores $RSmix_{ij}$, $RSpure_i$ and $RSpure_j$ (statistically significant positive scores are in boldface).

different from zero. A nonparametric test was chosen because the recognition scores did not follow normal distribution. False-discovery rate correction at $\alpha = 0.05$ was applied to compensate for the multiple comparisons (102 comparisons: 6 for virtual basic expressions, 6 for natural basic expressions, 45 for virtual blends and 45 for natural blends). Statistical significance is indicated with the common asterisk notation * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ in Tables 1, 2 and 4.

3.3 Results

We first evaluated whether the basic emotion expressions were actually perceived as such. Expressions produced using the virtual face model were compared to expressions posed by a real human.

Figure 1 shows the mean emotion ratings for all *basic expressions* of virtual and natural faces. Visual inspection of this figure suggests that most targeted emotions are perceived as expected. The fearful virtual face was perceived as more surprised than fearful, and that also

ISBN 978-80-86943-41-1

the fearful natural face received high surprise ratings. With the natural face, the sad facial expression received remarkably low sadness ratings.

Recognition score results, visualized in Table 1, show that indeed all expressions except fear were recognized correctly from the virtual face, and all except fear and sadness were recognized correctly from the natural face. Thus we can be confident that our model reasonably well resembles a human face.

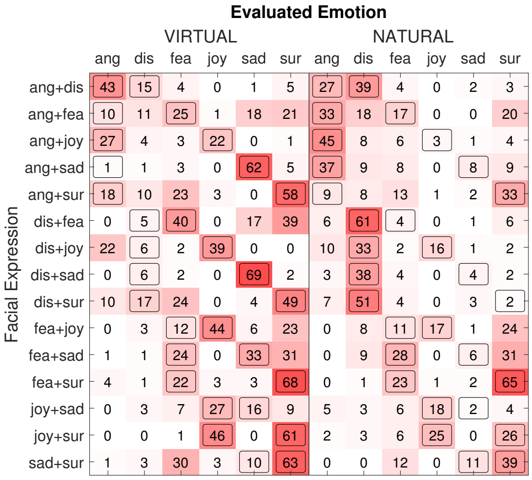


Figure 2: Perception of basic emotions in blends. The visualization is analogous to Figure 1.

The mean basic emotion ratings for all *virtual and natural blends* are presented in Figure 2. Visual inspection suggests that there are considerable differences between blends in how the ratings were distributed among basic emotions: Some of the expressions are recognized as one basic emotion, while others seem to express several emotions.

The recognition scores $RSmix_{ij}$ and $RSpure_i$ are collected in Table 2. They show that for the virtual face, only the blend of joy and surprise was recognized successfully as a mix of its parent emotions. Anger+disgust was perceived as anger, anger+sadness and disgust+sadness were both perceived as sadness, and fear+surprise and sadness+surprise were both perceived as surprise. As for the natural faces, only the blend of fear and surprise fulfilled the criteria of successful blend recognition, and even that blend was primarily perceived as surprise. Anger+joy was perceived as anger, whereas disgust+fear, disgust+sadness and disgust+surprise were all perceived as disgust.

As many of the blends were not perceived clearly as either mixtures of two basic emotions or instances of one basic emotion, the goal of our second experiment was to figure out whether some of the blends are perceived as expressions of more complex emotions. The open-ended part of our questionnaire offered a starting point: 17 participants out of 29 used the option of supplementing their answer using their own words. Altogether, additional descriptions were given 89 times. Experiment 2 was based on these answers.

4 EXPERIMENT 2: COMPLEX EMOTIONS IN BLENDS

To test whether some of the additional emotion words would describe perceptions of blends more accurately than the basic emotion words do, we paired these words with all facial expressions. The same videos of the virtual face that were used in Experiment 1 were used in this experiment also. The morphed natural faces were not used. Thus, there were 21 different facial expressions.

From all the additional words collected, we ignored references to non-genuine emotions (such as 'fake smile'). The remaining 17 words for complex emotions, together with the six basic emotions resulted in 23 emotion words altogether. When each facial expression was paired with each emotion word, we had $21 * 23 = 483$ expression-emotion pairs to study.

Using a three-stage procedure, we first narrowed down the list of expression-emotion pairs into those that appear more often than by chance, and then examined more carefully which expression-emotion pairs show evidence of complex emotion recognition.

This experiment was conducted using a crowdsourcing platform CrowdFlower (http://www.crowdflower.com/). The participants represented various backgrounds and were different in the different stages of the experiment.

4.1 Procedure and results

Stage 1 consisted of finding out which expression-emotion pairs are at least to some extent associated with each other. For each expression-emotion pair, the participants were asked *"Does this word describe the state of the person in this video?"* (yes or no). Ten evaluations were collected for each pair. All emotion words (with the corresponding expressions) that were mentioned six times or more were taken for further inspection in Stage 2. These words were: ambiguous, apologetic, disappointment, embarrassment, envy, malicious joy, revengeful, serious, shame, shock, and suspicious (words discarded at this stage were: cunning, concentration, despair, determination, interested and relief).

The purpose of Stage 2 was to further narrow down the list of expression-emotion pairs to only those with a strong association. All videos were shown to the participants with the question *"Which of these words best describes the emotion of the person in the video?"* The participants answered by making a forced-choice from one of the alternatives identified in Stage 1. It was also possible to select *"none of the above"*. 40 evaluations

ISBN 978-80-86943-41-1

Expression	Best descriptions of the expression	
	Targeted	Others
anger	anger	suspicious, revengeful, serious
disgust	disgust	envy
fear	fear	
joy	joy	malicious joy
sadness	sadness	
surprise	surprise	shock
ang+dis	anger	suspicious, revengeful
ang+fea	fear	disgust
ang+joy		malicious joy, revengeful
ang+sad	sadness	
ang+sur		shock
dis+fea	fear	disappointment
dis+joy	joy	malicious joy
dis+sad	sadness	
dis+sur		shock
fea+joy	joy	embarrassment
fea+sad	fear, sadness	disappointment
fea+sur		shock
joy+sad		ambiguous, apologetic, shame
joy+sur	surprise	
sad+sur		shock

Table 3: All emotion words that were selected as the best description for the corresponding expression more often than by chance. Targeted (basic) and other (complex) emotions in separate columns.

were collected for each facial expression. All emotion words that were mentioned more often than chance level as the best description for the facial expression were selected to the next list, presented in Table 3.

Stage 3 was conducted to identify blended facial expressions that are perceived as complex emotions. Candidates for this are the words in the bottom-right cell of Table 3, column "Others" for blended facial expressions. In total, 16 complex emotion terms for 12 blended facial expressions were considered at this stage (we also included as ’complex’ the basic emotion word disgust for the blend anger+fear, because it is neither of its parent emotions).

For each of the selected pairs between complex emotions and blended facial expressions, participants were asked to answer the question *“Which of these words best describes the emotion of the person in the video?”* with three possible choices: the complex emotion or either of its parent emotions. The same question was answered separately for three videos: the blended facial expression and both of the parent expressions. Each evaluation was conducted 120 times, resulting in nine frequency scores for each blend-emotion pair: three videos times three words.

Based on these scores we defined two indexes called *association* and *distinctiveness*. Association is positive if a complex emotion word describes the blend more accurately than either one of the parent emotion words. Distinctiveness, on the other hand, is positive if a complex emotion is associated specifically with the blend in contrast to the parent expressions. Formally we define them as

$$Ass(c,B) = S_c(B) - \max\{S_1(B), S_2(B)\}, \qquad (3)$$

$$Dis(c,B) = S_c(B) - \max\{S_c(P_1), S_c(P_2)\}, \qquad (4)$$

where c is a complex emotion, B a blend formed from parent expressions P_i (i=1..2), and $S_x(E)$ is the score of an emotion x (complex or parent) for the video of expression E.

Pearson’s chi-square test with Yates’ correction for continuity was used to determine whether association and distinctiveness values are different from zero with a statistical significance.

The results are collected in Table 4. The expression-emotion pairs are divided into three groups based on the strength of evidence they provide for the hypothesis that the complex emotion word unambiguously describes the blend (strong if both association and distinctiveness are positive, and weak if distinctiveness only is positive, and no evidence otherwise). For majority of the positive findings, parent emotions appear to be joy, sad, anger or surprise.

Expression-emotion pair	Association	Distinctiveness
joy+sad = apologetic	53***	73***
joy+sad = ambiguous	49***	71***
joy+sad = shame	43***	68***
joy+ang = malicious joy	58***	65***
joy+dis = malicious joy	59***	63***
joy+ang = revengeful	28***	54***
sur+sad = shock	28***	45***
sur+ang = shock	17*	44***
joy+fea = embarrassment	13	38***
sur+dis = shock	4	32***
sur+fea = shock	-8	28***
fea+ang = disgust	-45***	27***
ang+dis = revengeful	-70***	10
ang+dis = suspicious	-54***	7
fea+dis = disappointment	-77***	-4
sad+fea = disappointment	-49***	-8

Table 4: Association and distinctiveness values for all included expression-emotion pairs.

5 DISCUSSION

The present results show that most basic emotions were recognized very well from our virtual character. Although fear was incorrectly recognized as surprise, this confusion was also present in the natural face. Even though we did not explicitly compare recognition scores for the virtual and natural face, we note that the intensities tended to be lower for the virtual face. This observation is consistent with several previous facial animation studies [11, 9, 19, 18].

Only one virtual blend, joy+surprise was perceived as a mixture of its parent emotions. The natural blend fear+surprise was also recognized as a mixture of these two, but the perception of surprise was dominating. Also, it is noteworthy that the pure basic expression of fear was perceived as surprise both in our virtual and natural faces. The present scoring method was relatively strict, which may partly explain why no other blends were reliably recognized. However, the results show that it is not reasonable to assume that both parent emotions could be recognized from blends in general.

Our analysis revealed that five virtual blends and five natural blends were perceived as expressions of one of the parent emotions (bolded scores in Table 2). In the case of virtual fear+surprise blend, this result is trivial, since the facial expression of fear was also perceived as surprise. The other blends that were perceived as a parent emotion could be seen as partial evidence supporting the categorical emotion view. According to that view, when an expression gradually moves from anger to disgust (for example), it is perceived as pure anger until after a certain point it is perceived as pure disgust. Although congruent with our observations, only a minority of the blends were perceived as parent emotions, and thus we can't expect that this would generally happen.

The reason why some virtual blends were perceived as their parent emotion may be related to the blending method, which added together all muscle activities from both parent expressions. Some of the basic expressions include much greater and/or more visible movements than others, and thus in a blend the subtle movements may be overshadowed by the more prominent movements.

The blends that were strongly or weakly perceived as complex emotions (the two upper sections of Table 4) can be divided into two groups based on which parent expressions they consist of. The first group is joy blended with a negative emotion. These blends produce a variety of complex emotions which (with the exception of malicious joy) seem to be unique for each blend. However, a single blend can be described with several different emotion words. This is in accordance with the view that interpretation depends on context.

The second group is surprise blended with a negative emotion. All of these blends can, to some extent, be described with the word shock, which is believable considering the emotional content. This result may be somewhat questionable, however, because on Experiment 2 Stage 2, the basic expression of surprise was also often described as shock.

5.1 Limitations

This study was conducted using a single virtual face model and a single blending algorithm. The blends were created using only one facial expression of each basic emotion category. Although the used expressions of basic emotions were found relatively recognizable in comparison to natural facial expressions from a standard collection, they are not perfect, and specifically the expression of fear was poorly recognized. A wider variety of basic emotions could be used to create blends, and the faces could represent different individuals. Moreover, other animation methods for creating blends besides our additive method could be tested. Future studies might also consider whether different blending proportions of two expressions would produce different results.

Our facial model is crude compared to the highly photorealistic models used in movie industry. However, its visual fidelity is comparable to that of contemporary virtual agents used in interactive virtual reality and games. More advanced modeling and rendering may add details, such as wrinkles, that may cause new perceptual effects and different results.

Our stimuli were dynamic, but the brief motion from neutral to peak expression is still somewhat artificial. In real conversational situations facial expressions change continuously and follow each other. Some emotion blends may be expressed with two consecutive expressions, and blends are likely to momentarily occur when the emotional state changes. These kinds of temporal aspects are important in developing believable animated characters and virtual agents, and therefore future research should address also this issue.

6 CONCLUSION

As animated virtual characters become more humanlike, expectations towards their facial behavior increases. To be able to create believable facial expressions of emotions that imitate expressions of real humans, we need more understanding about how different facial expressions of virtual characters are perceived.

Our results demonstrate that people are often not able to correctly recognize the two basic emotions in a blend of facial expressions, but instead, some blends produce a perception of another, complex emotion. The blend of surprise with any negative emotion is often labeled as shock. On the other hand, blends with opposite valence (joy combined with a negative emotion) can be described with various complex emotion words. In real applications, the interpretation of these facial expressions would probably depend on context.

The results indicate that blended facial expressions of basic emotions can be used to increase the emotional expressiveness of virtual agents. To communicate more complex emotional states in addition to the basic emotions, it is important to blend not only facial expressions of emotions that are close to each other in a conceptual emotional space, such as anger and disgust, but also facial expressions that represent opposite emotional states, such as joy and sadness.

ISBN 978-80-86943-41-1

To our knowledge, this is the first study to systematically search for perceptions of complex emotions in pairwise blends of basic expressions. Its main contribution is to outline methodology and lay hypotheses for further research, while the detailed results and the scores used in the analysis may need revised studies with different facial models.

7 ACKNOWLEDGMENTS

This study was supported by the Academy of Finland, Doctoral Program in User-Centered Information Technology (UCIT), and by the H2020-MSCA-IF-2015 grant (no. 703493) to Jari Kätsyri.

8 REFERENCES

[1] Junghyun Ahn, Stephane Gobron, Daniel Thalmann, and Ronan Boulic. Asymmetric facial expressions: revealing richer emotions for embodied conversational agents. *Computer Animation and Virtual Worlds*, 24(6):539–551, 2013.

[2] Irene Albrecht, Marc Schröder, Jörg Haber, and Hans-Peter Seidel. Mixed feelings: expression of non-basic emotions in a muscle-based talking head. *Virtual Reality*, 8(4):201–212, 2005.

[3] Ali Arya, Steve DiPaola, and Avi Parush. Perceptually valid facial expressions for character-based application s. *International Journal of Computer Games Technology*, 2009, 2009.

[4] Andrew J Calder, Duncan Rowland, Andrew W Young, Ian Nimmo-Smith, Jill Keane, and David I Perrett. Caricaturing facial expressions. *Cognition*, 76(2):105–146, 2000.

[5] Miriam Dyck, Maren Winbeck, Susanne Leiberg, Yuhan Chen, Rurben C. Gur, and Klaus Mathiak. Recognition profile of emotions in natural and virtual faces. *PLoS ONE*, 3(11):e3628, 11 2008.

[6] P Ekman, WV Friesen, and P Ellsworth. What emotion categories or dimensions can observers judge from facial behaviour? in, p. ekman. *Emotion in the Human Face*, 1982.

[7] Paul Ekman, Wallace V. Friesen, and Joseph C. Hager. *Facial Action Coding System: Investigator's Guide*. A Human Face, 666 Malibu Drive, Salt Lake City UT 84107, USA, 2002.

[8] Paul Ekman, Wallace V Friesen, and Consulting Psychologists Press. *Pictures of facial affect*. consulting psychologists press, 1975.

[9] Marc Fabri, David Moore, and Dave Hobbs. Mediating the expression of emotion in educational collaborative virtual environments: an experimental study. *Virtual Reality*, 7(2):66–81, 2004.

[10] Scott H. Hemenover and Ulrich Schimmack. That's disgusting! ..., but very amusing: Mixed feelings of amusement and disgust. *Cognition & Emotion*, 21(5):1102–1113, 2007.

[11] Jari Kätsyri, Vasily Klucharev, Michael Frydrych, and Mikko Sams. Identification of synthetic and natural emotional facial expressions. In *AVSP 2003-International Conference on Audio-Visual Speech Processing*, 2003.

[12] Jeff T Larsen and A Peter McGraw. The case for mixed emotions. *Social and Personality Psychology Compass*, 8(6):263–274, 2014.

[13] Meeri Mäkäräinen and Tapio Takala. An approach for creating and blending synthetic facial expressions of emotion. In *IVA '09: Proceedings of the 9th International Conference on Intelligent Virtual Agents*, pages 243–249, Berlin, Heidelberg, 2009. Springer-Verlag.

[14] Jean-Claude Martin, Radoslaw Niewiadomski, Laurence Devillers, Stephanie Buisine, and Catherine Pelachaud. Multimodal complex emotions: Gesture expressivity and blended faci al expressions. *International Journal of Humanoid Robotics*, 3(3):269–291, 2006.

[15] Radoslaw Niewiadomski, Magalie Ochs, and Catherine Pelachaud. Expressions of empathy in ecas. In *Intelligent virtual agents*, pages 37–44. Springer, 2008.

[16] Catherine Pelachaud, Norman I. Badler, and Mark Steedman. Generating facial expressions for speech. *Cognitive Science*, 20(1):1–46, 1996.

[17] Wataru Sato, Takanori Kochiyama, Sakiko Yoshikawa, Eiichi Naito, and Michikazu Matsumura. Enhanced neural activity in response to dynamic facial expressions of emotion: an fmri study. *Cognitive Brain Research*, 20(1):81–91, 2004.

[18] Angela Tinwell, Mark Grimshaw, Debbie Abdel Nabi, and Andrew Williams. Facial expression of emotion and perception of the uncanny valley in virtual characters. *Computers in Human Behavior*, 27(2):741 – 749, 2011.

[19] T. Wehrle, S. Kaiser, S. Schmidt, and K. R. Scherer. Studying the dynamics of emotional expression using synthesized facial muscle movements. *Journal of Personality and Social Psychology*, 78:105–119, 2000.

[20] Yu Zhang, Edmond C. Prakash, and Eric Sung. Efficient modeling of an anatomy-based face and fast 3d facial expression synthesis. *Computer Graphics Forum*, 22(2):159–170, 2003.

 ISBN 978-80-86943-41-1