

The Impact of Combining Kinesthetic and Facial Expression Displays on Emotion Recognition by Users

Yoren Gaffary¹, Victoria Eyharabide², Jean-Claude Martin¹, and Mehdi Ammi¹

¹Human-Computer Communication, LIMSI-CNRS, Université Paris-Sud, Orsay, France

²Computer Sciences and Linguistics, Paris-Sorbonne University, Paris, France

Several studies have investigated the relevance of haptics to convey various types of emotions physically. This article investigates the improvement of the recognition rate of emotions using visuo-haptic feedback compared to facial and haptic expressions alone. Four experiments were conducted in which the recognition rates of emotions using facial, haptic and visuo-haptic expressions were tested. The first experiment evaluates the recognition rate of emotions using facial expressions. The second experiment collects a large corpus of 3D haptic expressions of certain emotions and subsequently identifies the relevant haptic expression for each emotion. The third experiment evaluates the selected haptic expressions through statistical and perceptive tests to retain the ones that result in the most accurate identification of the corresponding emotion. Finally, the fourth experiment studies the effect of visuo-haptic coupling on the recognition of the investigated emotions. Generally, emotions with high amplitudes of pleasure are better recognized in the visual modality. However, emotions with high activation are better recognized in the haptic modality. These results also highlighted the finding that participants are not equally aided by each modality when recognizing emotions efficiently. Beyond the recognition rate, multimodal expressions improved the sensation of presence and expressivity.

1. INTRODUCTION

Nonverbal communication comprises the primary method of interpersonal communication in everyday life (Mehrabian & Ferris, 1967). Nonverbal communication includes several channels, such as prosody, gestures, and postures. Affective communication exploits and combines these different channels to convey different emotions effectively (Scherer, 2000). Joy, surprise, and fear are examples of basic emotions that everyone can understand without specific learning (Ekman, 1992). Emotions play an essential role in interpersonal and affective communication with other people (Parkinson, Fischer, & Manstead, 2004; Scherer, 2005). Although the expression of emotions includes and exploits several modalities, existing studies mainly focus

on the study of facial expressions (Ahn, Bailenson, Fox, & Jabon, 2009; Courgeon, Clavel, & Martin, 2009; Ekman & Friesen, 1975) or gestural expressions (Buisine et al., 2014; Coulson, 2004; Dael, Mortillaro, & Scherer, 2011; Kubera & Wrighta, 2013; Wallbott, 1998). Affective Computing, which is the scientific field that studies the recognition and simulation of human affects, has proposed several solutions for mediated communication (Picard, 1997) with, for example, virtual characters (Courgeon, Martin, & Jacquemin, 2008). Unfortunately, these approaches are not yet able to reproduce the full potential of human-human affective communication and often remain limited to certain basic emotions, such as joy and anger.

Recently, some studies have investigated the potential of haptic channels to communicate emotions. Hertenstein, Keltner, App, Bulleit, and Jaskolka (2006) have shown how haptic expressions can effectively convey several emotions through direct contact between two humans. In the physiological field, Olausson et al. (2008) highlighted specific biological systems that were dedicated to the expression of emotions through this modality. Furthermore, several studies have highlighted the relationship between the stimulation of mechanoreceptors on the skin and the release of oxytocin, known as the “love hormone.” This hormone is involved in several social and love behaviors and thus in affective communication (Handlin, 2010). Other studies that have investigated human-human interactions have highlighted the role of warmth coupled with physical contact in effectively conveying different categories of positive emotions (Bargh & Shalev, 2012).

With the emergence of haptic interfaces, allowing tactile and kinesthetic stimulation (Chang, Hwang, & Ji, 2011), several studies have investigated haptic stimulation to communicate emotions, for example, between virtual characters and users. Basic research has exploited predefined cinematic and physical behaviors, which were identified in previous studies that investigated motor expressions for different emotions (Bonnet, Ammi, & Martin, 2011; Tsetserukou & Neviarouskaya, 2010), for example, swinging movements to express happiness and tapping movements to express disgust. However, these approaches were limited to a few basic emotions. Bailenson, Yee, Brave, Merget, and Koslow (2007) proposed a complete platform

Address correspondence to Yoren Gaffary, Human-Computer Communication, LIMSI-CNRS, Université Paris-Sud, B.P. 133, 91403, Orsay, Cedex, France. E-mail: yoren.gaffary@limsi.fr

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hihc.

for recording and rendering affective expressions with haptic channels through motor expressions. However, the results have shown low recognition rates for several emotions. Other research studies have used haptic feedback for long-distance interpersonal communication technologies. Multiple devices were developed to maintain physical or intimate contact between people in such long-distance interactions. For example, Park, Lim, and Nam (2010) presented the CheekTouch device. This device provided intimate tactile feedback via a mobile phone on the user's cheek, which corresponded to the partner's gestures. The HugMe platform developed by Cha, Eid, Rahal, and El Saddik (2008) is a synchronous haptic teleconferencing system that enables people to exchange physical stimuli to convey affection and intimacy.

Some researchers have attempted to combine haptic and facial expressions to improve the recognition of some emotions. Basori, Bade, Sunar, and Daman (2010) investigated the influence of a vibrotactile stimulation on the perception of two basic emotions (joy and anger), which were expressed by a 3D virtual character. They showed that the amplitude of the vibrotactile stimulus influenced the perception of the emotions: the higher the amplitude, the more the participants perceived the emotion of anger. Thereafter, Bickmore, Fernando, Ring, and Schulman (2010) studied the complementarity between the tactile feedback (with a pneumatic device) and a virtual character that presented three different facial expressions: positive, negative, and neutral. The results have shown nonsignificant tendencies of the influence of tactile feedback on the perception of the emotion's arousal.

More recently, Bonnet et al. (2011) combined parametric kinesthetic expressions, which were inspired by psychological studies and defined with basic parameters such as the amplitude and frequency of the movements, with facial expressions presenting a large set of emotions. The results have shown an improvement in the recognition of a limited number of emotions.

As previously shown, most studies investigated basic emotions and studied the influence of a haptic channel on the perception of the corresponding emotions. Our work proposes to investigate more complex emotions that present ambiguities in their facial expressions. The use of haptic feedback in combination with facial expressions is intended to improve the discrimination between those emotions. Most research in affective haptics exploited basic tactile and kinesthetic feedback based on haptic devices presenting 1 or 2 degrees of freedom or a limited workspace. We propose to investigate a 3D kinesthetic feedback to increase the expression possibilities. More particularly, we investigated the handshake configuration that corresponds to a natural physical interpersonal contact to support various emotions (Bailenson et al., 2007). Current work focused on the handshake gesture that can be easily supported by standard haptic devices.

Previous work exploited either synthetic haptic expressions (Yohanan, Chan, Hopkins, Sun, & Maclean, 2005)

or expressions recorded by a small number of participants (Bailenson et al., 2007). We propose to consider a larger corpus to better represent the natural haptic behaviors of people. The identification of the most representative haptic expressions of each studied emotion is based on a new statistical analysis approach that simultaneously considers several physical features. Once the haptic expressions were identified, we conducted an experiment to study the contribution of haptic feedback to the discrimination of ambiguous emotions expressed with facial expression of virtual characters. The study addresses static facial expressions to avoid issues related to the synchronicity and overlapping of visuo-haptic expressions to focus on basic visuo-haptic cues for the perception of emotions before investigating more complex coupling configurations (synchronous expressions, priming, etc.). We only provide congruent visuo-haptic expressions for an efficient integration of multimodal information while drawing the user's attention to the overall emotion expressed by both expressions. Finally, based on the results of previous studies, which have highlighted differences in the ways in which different people recognize some emotions expressed with facial expressions (Russell, 1994), we propose to investigate the relationship between the visuo-haptic coupling and interpersonal differences in the perception of emotions.

This article is organized as follows. Section 2 presents the selected set of emotions. Section 3 highlights the limitations of the facial expressions of virtual characters in conveying those emotions. Section 4 presents a collection of the corpus of affective haptic expressions and the selection of the most representative haptic expressions. Section 5 presents the evaluation of the selected haptic expressions to retain a single relevant haptic expression for each emotion. Section 6 presents the experimental study of the effect of the visuo-haptic coupling on the recognition of the investigated emotions and discusses the results. Finally, section 7 is a general discussion, and section 8 presents the conclusions and several perspectives.

2. THE EMOTIONS THAT WERE INVESTIGATED

The first step of this research comprised selecting a set of semantically close emotions that are difficult to differentiate with only visual cues. This work aimed to find discriminative haptic features to improve the differentiation of these ambiguous emotions. The dimensional approach for emotion representation suggests that emotions can be described using several uncorrelated and continuous dimensions, for example (Russell & Mehrabian, 1977),

- Pleasure (P): degree of well-being.
- Activation (A): degree of mental or physical activity.
- Dominance (D): degree of control of a situation.

Based on this 3D continuous representation, we have selected a set of eight emotions that have a positive activation. In fact, several researchers have highlighted the role of the

TABLE 1

Investigated Emotions and Their Estimated Location on the PAD Scales (Russell & Mehrabian, 1977)

Emotion	Pleasure	Arousal	Dominance
Joy	0.76	0.48	0.35
Elation	0.50	0.42	0.23
Disgust	-0.60	0.35	0.11
Contempt	-0.23	0.31	0.18
Anxiety	0.01	0.59	-0.15
Fear	-0.64	0.60	-0.43
Irritation	-0.58	0.40	0.01
Rage	-0.44	0.72	0.32

haptic modality, and more specifically of a kinesthetic channel, for the communication of active emotions (Bonnet et al., 2011).

In this half-space (i.e., Positive Activation), three basic emotions and three complex emotions (Golan, Baron-Cohen, & Hill, 2006) located in different quadrants have been selected:

- Basic emotions: Joy, disgust, and fear.
- Complex emotions: Elation, contempt, and anxiety.

Each complex emotion selected was closed in its meaning to a basic emotion. For instance, elation is closed to joy. In addition, except for anxiety and fear, each basic emotion and its corresponding complex emotion are in the same quadrant of the PAD space. For example, elation is in the same quadrant as joy. Two other complex emotions were investigated: Irritation and rage. These emotions are in the same quadrant as disgust and contempt. They will allow the study of the discrimination of semantically close emotions. Table 1 details the investigated emotions and their dimensions.

3. STUDY 1: RECOGNITION OF EMOTIONS FROM FACIAL EXPRESSIONS

3.1. Objective and Hypothesis

The objective of the first study was to investigate the recognition rate of the emotions that were conveyed by facial expressions. It aims at highlighting the limitations of facial expressions of virtual characters in conveying emotions. The hypothesis of this study is as follows:

H1: *Some emotions are difficult to discriminate using only the facial expressions of virtual characters.* Previous research studies (Bonnet et al., 2011; Courgeon, Clavel, Tan, & Martin, 2011) have shown low recognition rates for several emotions when using only facial expressions of virtual characters. This hypothesis is also inspired by the fact that we consider complex emotions in addition to basic emotions.

3.2. Facial Expressions

A series of eight static facial expressions, which corresponded to each selected emotion, were selected from the corpus of the MindReading database (Golan et al., 2006). Then, the MARC framework (Courgeon et al., 2008) was used to reproduce those expressions with a realistic 3D character. This software was successfully used in several projects (Clay et al., 2012; Courgeon & Martin, 2009).

3.3. Experimental Platform

The experiment ran on a computer that managed two main components: (a) the MARC virtual character displaying facial expressions (Courgeon et al., 2008) and (b) a graphical user interface (GUI) displaying instructions to participants and recording their answers using a form (see Figure 1). The platform manages the progress of the experiment without a human supervisor. This strategy is important because the presence of a person in the same room might influence the participant's answers.

3.4. Participants

Twenty-three people (four women, 19 men), with an average age of 33 ($SD = 11$), participated in the experiment. Twenty-one received a European education, one received an African education, and one received an Asian education. The influence of the gender or education on the recognition rate was not investigated because most of the participants were men with a European education.

3.5. Measures

The evaluation of this experiment was based on the measure of the emotion recognized for each presented facial expression. The participants were asked to select one emotion from the list of eight emotions.

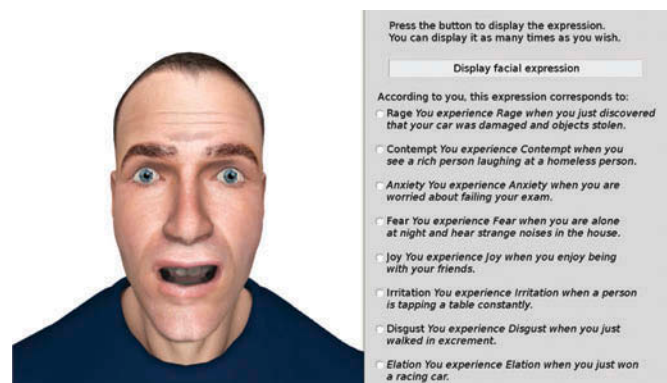


FIG. 1. The experimental platform. *Note.* The virtual character was displayed in the left part of the screen. The right part of the screen displayed the list of emotions (in French in this experiment). The participant had to select one emotion that corresponded to the emotion expressed by the virtual character.

3.6. Procedure

The participants were seated in front of a desk on which there was a screen, a keyboard, and a mouse. They were informed that they would see different facial expressions and would have to recognize the corresponding emotions. They began by filling out a form that indicated their age, gender, and cultural education. They next had a training phase wherein they had to evaluate the emotion surprise expressed by the facial expression of a virtual character to ensure that the task was clear. Then the participants were left alone in the room, and a series of eight facial expressions were presented to them in a random order. For each emotion, there was no visual display of the virtual character at the beginning. The participants had to press the display button to activate the visual display. Then the facial expression was displayed for 2 s. The participants could display it as many times as they wanted. Finally, they had to select one emotion from the list of eight emotions. The order of emotions in the list was presented in a random order. Each emotion from the list was followed by a sentence that presented a context in which this emotion appears. This approach was used to help subjects to understand the difference between close emotions (Wallbott, 1998).

3.7. Results

The results of emotions perceived for each presented facial expression were calculated. Table 2 displays the confusion matrix of the recognition rate. For the analysis of these results, the threshold of confusion is calculated; it expresses the level at which the emotion is considered to be perceived by the participant. This threshold corresponds to the chance level if the participants randomly select emotions. In this study, the chance level is equal to the total number of answers/number of possible answers (or in terms of a percentage, $100/8 = 12.5\%$). The eight emotions are discussed as follows:

- **Joy** was very well recognized (100%) and was not confused with other emotions.
- **Elation** presented a good recognition rate (78%) but was confused with joy (22%).
- **Disgust** was not well recognized but was still higher than the chance level (57%) and was confused with irritation (17%), contempt (13%) and anxiety (13%).
- **Contempt** was not well recognized but was still higher than the chance level (49%) and was confused with disgust (24%) and anxiety (23%).
- **Anxiety** was poorly recognized but was still higher than the chance level (32%) and was confused with contempt (32%) and disgust (28%).
- **Fear** was well recognized (74%) and was not confused with other emotions. All of the other emotions were recognized at a lower rate than the chance level (12.5%).
- **Irritation** was poorly recognized but was still higher than the chance level (26%) and was confused with rage (26%) and contempt (26%).
- **Rage** presented the lowest level of recognition (9%) and was confused with disgust (31%), irritation (31%) and fear (13%).

These results show that some expressions of emotions are very well recognized in facial expressions, whereas others were not well perceived and were confused with two or three emotions. Rage is a special case, in which the level of confusion with other emotions is very important. These results confirmed the difficulties in discriminating some emotions using facial expressions. The hypothesis H1 is, thus, validated.

3.8. Discussion

It should be noted that only two emotions presented positive pleasure (i.e., joy and elation). Previous research studies have

TABLE 2
The Recognition Rate of the Emotions (%)

	Joy	Elation	Disgust	Contempt	Anxiety	Fear	Irritation	Rage
Joy	100	22	0	0	0	4	0	0
Elation	0	78	0	0	0	0	0	0
Disgust	0	0	57	24	28	9	5	39
Contempt	0	0	13	49	32	0	26	4
Anxiety	0	0	13	23	32	9	12	4
Fear	0	0	0	0	4	74	5	13
Irritation	0	0	17	4	4	0	26	31
Rage	0	0	0	0	0	4	26	9
Total	100	100	100	100	100	100	100	100

Note. The columns correspond to facial expressions of an emotion. The rows correspond to the recognized emotions for this expression. The diagonal (black cells) corresponds to the percentage of participants who successfully recognized the target emotion. The gray cells correspond to false recognitions (but above chance level, i.e., $100/8 = 12.5\%$).

shown that facial expressions convey the positive dimension well (Bickmore et al., 2010; Castellano, 2008; Courgeon et al., 2011). This result explains in part the important recognition rate for these two emotions. Concerning the negative emotions, the low level of recognition can be explained by the imbalance between the number of positive and negative facial expressions. In fact, six emotions presented negative pleasure (disgust, contempt, anxiety, fear, irritation, and rage), which means that these emotions are in the same subspace and are semantically close (i.e., small PAD distances between them). Moreover, the results have shown that negative emotions are mainly confused with other negative emotions.

The objective of the following studies is to combine these facial expressions with suitable haptic feedback in an attempt to disambiguate the recognition of the expressed emotions.

4. STUDY 2: CREATION OF THE HAPTIC CORPUS AND SELECTION OF THE HAPTIC EXPRESSIONS OF EMOTIONS

4.1. Objectives

The objective of the second study was to collect a corpus of affective haptic expressions for the eight studied emotions. This corpus will be analyzed and filtered to identify the relevant haptic expressions (i.e., one or several expressions) for each emotion. Multiple protocols were defined in the literature to collect the expressions of the emotions in several modalities (Coan & Allen, 2007). Two main classes of expressions were used: acted and spontaneous (Audibert, Auberge, & Rilliard, 2008). This article exploits the acted expressions of emotions because they are easier to express and to collect in an experimental study. Moreover, this approach might provide haptic expressions that are exaggerated and that unambiguously convey a given emotion to the user. This approach comprises asking the participants to express the emotions as they are defined in a textual description. In this experiment, the participants express the emotions with a 6 DoF haptic device.

4.2. Experimental Platform

The experimental platform (Gaffary, Eyharabide, Martin, & Ammi, 2013b) used to collect the haptic expressions was based on a SensAble PHANTOM Desktop arm. This device enables the recording and subsequent rendering of haptic expressions. Moreover, it allows the generation of a wider variety of haptic expressions compared to devices used in previous research studies (Bailenson et al., 2007; Smith & Maclean, 2007). It enables the expression of 3D movements, which provides access to the depth component of the movement. Moreover, this device can exert forces of up to 7.9 N, which enable the generation of fast and jerky haptic expressions. All of these features should increase the number and the type of the haptic expressions compared to the standard 2D and 1D devices used in previous research.

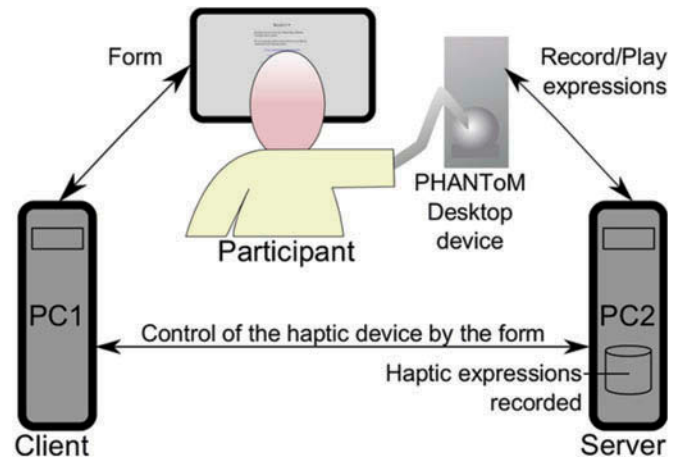


FIG. 2. The experimental platform used to collect haptic expressions. *Note.* The client node provides a form that displays the instructions and records the answers of the participants. The server node records and renders the haptic expressions made by the participants.

The platform was based on two computers (see Figure 2). The client node includes three components. First, a module manages the progress of the experiment and controls the overall software platform. Second, a module is used to record the haptic expressions, several measures, and information related to the participants. Third, a GUI is used to display the instructions to the participants and processes their keyboard and mouse inputs. Accordingly, the manager sends recording and displaying requests to the server, which manages the haptic device. The haptic device was connected to a dedicated computer to prevent haptic instabilities due to the calculation latency of the manager module. The haptic device is controlled with a low-level module (a haptic module) based on the OpenHaptics library.

4.3. Method

Participants

Forty subjects (eight women, 32 men), with an average age of 31 ($SD = 8$), participated in the experiment. Thirty-three of them were right-handed. Thirty-five received a European education. We did not analyze the influence of gender, handedness, or education on the collected haptic expressions, because the majority were European right-handed men.

Measures

Objective measures. The objective measures were based on the cinematic features of the haptic expressions. For each expression that was collected, we computed several measures that were used in different studies related to the affective communication with the haptic and gestural channels (Bailenson et al., 2007; Castellano, 2008). These measures were computed from the sequence of 3D-points $[(x, y, z)_1, (x, y, z)_2, \dots, (x, y, z)_n]$, which corresponded to the recorded movements with the haptic device. The sampling rate was 1 ms.

- **Distance:** the overall distance traveled by the device's end-effector (participant's hand).
- **Duration:** the overall duration of the haptic expression.
- **Amplitude:** the distance between the two farthest corners of the bounding box containing the haptic expression.
- **Mean speed:** the average speed of the end-effector.
- **Fluidity:** the degree of smoothness of the expression: $\sum_{t=0}^{n-1} |a(t+1) - a(t)| / n$, with $a(t)$ as the acceleration at time t .
- **Expansion Index:** the degree of contraction and expansion of the haptic expression (Castellano, 2008) $\sum_{t=0}^n d(p(t), isobar) / n$, with $d(p(t), isobar)$ as the Euclidian distance between the point taken at time t and the isobarycenter of the expression. A low value for the contraction index means that the movement is concentrated.
- **Major Axis:** the major axis of the gesture, computed with Singular Value Decomposition (SVD; Klement & Laub, 1980). We considered separately the X, Y, and Z coordinates of this vector.
- **Weight of the Major Axis:** the prevalence of the major axis on the movement (based on SVD).
- **Weight of the Second Major Axis:** the prevalence of the second major axis on the movement (based on SVD).
- **Repetitivity:** the estimation of the repetitive phases of the signal (Hartmann, Mancini, & Pelachaud, 2006). After computing the barycenter of the haptic expression and the major axis of the movement, all of the points that describe the expression on the major axis are projected. Then, one repetition is counted each time that the barycenter is crossed two times by the projection.

Subjective measures. Once the participant has finished expressing an emotion, the haptic device displays the recorded haptic expression for a self-evaluation by the participant. The participant has to rate, on a 7-point Likert scale, the level of confidence in the emotion that he or she has expressed. For the analysis, the expressions with a score less than or equal to 4 were removed; this score implies that the participant was not confident about the expressivity of his or her haptic expression.

Procedure

The participants were seated in front of a desk on which there was a screen, a keyboard, and a haptic device (Figure 3). The haptic device was positioned according to the user's dominant hand. The participants were instructed to keep the same body position during the entire experiment. The GUI displayed a text explaining the experiment and how the haptic device should be used to express the haptic expressions. The participants were asked to hold the device as if they were holding someone's wrist.

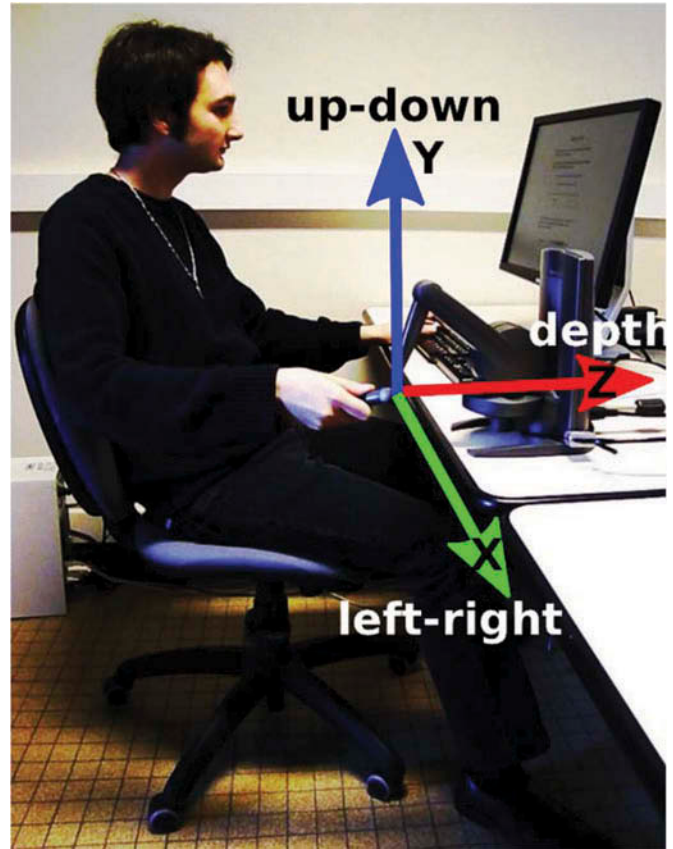


FIG. 3. The user manipulates the device as if he were holding the wrist of another person.

The participants began by filling out a questionnaire about their age and gender and whether they had previously used a haptic arm. The second step was the training. The participant explored the workspace of the haptic device and expressed a given emotion with the haptic device.

Once the training session was completed, the participants were asked to express each of the eight emotions. For each emotion, a textual description of a relevant emotional situation, selected from the MindReading database (Golan et al., 2006), was displayed near the emotional label. The description ensures a common understanding of the emotions' meanings (Wallbott, 1998). The order of presentation of the emotions was counterbalanced across subjects. The participants had 10 s to express the target emotion using the haptic device. Afterward, the recorded haptic expression was rendered to the participant through the same haptic device. Then, the participants had to report, via a 7-point Likert scale, their level of confidence about the expressed emotion. The users had only one trial to record each emotion, which was intended to decrease the duration of the whole experiment and to ensure a minimum spontaneity in the collected data. Forty haptic expressions were collected for each of the eight emotions. The haptic corpus, thus, comprises $40 \times 8 = 320$ haptic expressions of emotions.

Before the analysis, the corpus was filtered to retain only the haptic expressions that presented high levels of confidence. The applied threshold is 5/7 on the 7-point Likert scale. A total of 194 expressions fulfilled this criterion and were used in the following analysis. At this point it is important to emphasize that only 61% of the haptic expressions received a high level of confidence from participants, suggesting that the formation of such expressions is nonintuitive.

4.4. Selection of the Haptic Expressions

The objective of this step was to analyze and filter the haptic corpus (i.e., 194 expressions) to select the most representative haptic expressions (i.e., one, two, or three haptic expressions) for each emotion. The selected haptic expressions will be used in a perceptive study to identify one relevant haptic expression for each emotion. Thus, the filtering step played an important role because it reduced the number of tested haptic expressions from 194 expressions (which would require several hours for the perceptive test and could be painful for the participants) to a dozen expressions.

The approach adopted for the filtering used the Expectation–maximization clustering (EM-clustering) algorithm. This approach was successfully tested in our previous research (Gaffary, Eyharabide, Martin, & Ammi, 2013a, 2013b). The experimental results have highlighted several clusters (i.e., between one and four) of the haptic expressions for each of the emotions. The number of clusters corresponds to the different ways in which participants express this emotion with the haptic channel.

On the basis of these findings, we applied the EM-clustering algorithm on each emotion separately. Because the data set is small, this strategy was intended to highlight the main trends of the haptic expressions for each of the investigated emotions. The different objective measures (e.g., distance, fluidity, contraction index) were computed and used as descriptors for the haptic expressions. Then, a WEKA implementation of the EM-clustering was applied on all of the haptic expressions of each emotion. This algorithm enables estimation of the optimal number of clusters using the cross-validation method. In our case, this number is not known a priori. Furthermore, this algorithm captures the correlations and dependencies between the attributes. Figure 4 presents the different clusters identified. We observed that joy, anxiety, and fear presented only one cluster, whereas elation, disgust, content, irritation, and rage presented two clusters.

The different clusters include groups of haptic expressions that have several similar cinematic features but that present some differences (e.g., the 3D shape of the haptic expression of a given cinematic feature). To reduce the size of the haptic corpus for the perceptive study while accounting for these differences, we have proposed to retain four haptic expressions

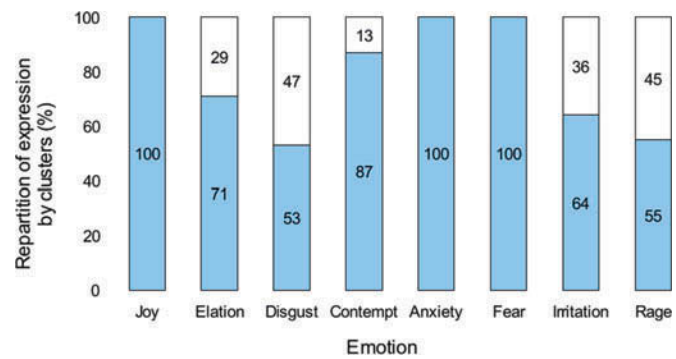


FIG. 4. Repartition of the expressions obtained for each emotion. *Note.* For joy, anxiety, and fear, there is a single cluster. For the other emotions, there are two clusters. This finding means that, for the latter emotions, the participants used at least two ways to express them.

for each emotion. As previously shown, the EM-clustering applied to the expressions of each emotion produced one or two clusters. For the emotions that presented one cluster, the four haptic expressions were selected in the same cluster. For the emotions that presented two clusters, two haptic expressions were selected in each cluster. Thus, four expressions were selected for each emotion. This selection approach provided 32 haptic expressions for the eight emotions, which is more reasonable than 194 expressions for the perceptive study. We retained the haptic expression that presents the shortest Euclidian distance to the barycenter of the cluster. The position of the barycenter and the Euclidian distances were calculated in the descriptor space. This criterion allowed the selection of haptic expressions that were close to the most often expressed haptic expressions for each cluster.

5. STUDY 3: PERCEPTIVE STUDY

5.1. Objectives

The objective of this third study was to evaluate the selected haptic expressions (i.e., 32 expressions) through a perceptive test to keep the best haptic expression for each emotion. In the next study, the haptic expressions selected in the perceptive study will be combined with the facial expressions previously investigated to study whether the combination improves the recognition and discrimination of the corresponding emotions.

5.2. Experimental Platform

The experimental platform is the same platform as was used in the previous experiment. However, the haptic device is used to render haptic expressions rather than to collect them. In other words, instead of asking the participants to express an emotion with the haptic device, the participants perceived an emotion through the movements made by the haptic device.

5.3. Method

Participants

Twenty-five people (six women, 19 men), with an average age of 29 ($SD = 9$), participated in this experiment. Twenty-three of them were right-handed. Twenty-four received a European education. One received a North African education.

Measures

This study was based on the recognition rate of the haptic expressions. Thus, for each presented haptic expression, the participants were required to select the emotion perceived from a list of eight emotions. They selected one or more emotions for a given haptic expression. This strategy was aimed at highlighting the haptic expressions that presented ambiguities between two or more emotions.

Procedure

The participants were informed that they would physically perceive movements with the hand that was holding the haptic device. For each haptic expression presented, they were required to select at least one emotion from the list. The order of presentation of the labels of emotions in the list was counter-balanced across participants. The participants received a paper document with a textual description of a relevant emotional situation for each of the eight studied emotions. These descriptions were based on the MindReading database (Golan et al., 2006). The descriptions ensure a common understanding of the emotions' meanings (Wallbott, 1998).

The participants received instructions regarding the ways to sit and to hold the haptic device (the same PHANTOM device as in the previous experiment). Then, a series of haptic expressions that were randomly extracted from the unfiltered haptic corpus (i.e., 194 expressions) were presented to the participants for training. Finally, the selected 32 haptic expressions were presented to the participants for evaluation. The presentation order of the haptic expressions was counterbalanced across the participants.

5.4. Definition of the Selection Scores

This section presents the criteria used to select the most relevant haptic expression for each emotion.

Let $reco_{exp}$ be the function that is defined by:

$$reco_{exp}(p, e) = \begin{cases} 1, & \text{if the participant } p \text{ recognized the} \\ & \text{emotion } e \text{ in the expression } exp \\ 0, & \text{otherwise} \end{cases}$$

This function expresses, for a given haptic expression exp , whether the participant p selected the emotion e . On the basis of this function, we first compute the function $score_1$:

$$score_1(exp) = \frac{\sum_p reco_{exp}(p, e_{target})}{\#p}$$

This function expresses the rate of participants who recognized the target emotion e_{target} from the haptic expression exp . However, this score function does not consider the participants who select several emotions for the same haptic expression. For example, if the participants perceived both joy and elation for a given haptic expression corresponding to elation, then this score cannot be used to discriminate those emotions. Therefore, a second score function $score_2$ was proposed considering the level of ambiguity for a given haptic expression.

$$score_2(exp) = \frac{\sum_p reco_{exp}(p, e_{target})}{\sum_p \sum_e reco_{exp}(p, e)}$$

This score function not only depends on the number of participants $\#p$ but also considers the total number of emotions $\sum_e reco_{exp}(p, e)$ recognized for each participant p for the haptic expression exp . This score function calculates the haptic expressions that can convey without ambiguity a target emotion e_{target} . Let $e1$ be a haptic expression that corresponds to joy and is recognized as joy and disgust. Let $e2$ be another haptic expression that corresponds to joy and is recognized as joy and elation. With this new score function, $e1$ and $e2$ have the same score. Therefore, it is necessary to distinguish between these two types of ambiguity. In fact, the objective of our research concerns the improvement of discrimination between close emotions with congruent haptic expressions using facial expressions. To solve that ambiguity, we use the PAD distance between the perceived emotions and the target emotion. Thus, we defined a third score function $score_3$.

$$score_3(exp)$$

$$= \frac{\sum_p reco_{exp}(p, e_{target})}{\underbrace{\sum_p \sum_{e \neq e_{target}} reco_{exp}(p, e) \times (d_{PADmax} - d_{PAD}(e, e_{target}))}_{\text{part I}} + \underbrace{\sum_p reco_{exp}(p, e_{target})}_{\text{part II}}}$$

where d_{PADmax} corresponds to the distance between the two furthest emotions of the studied set (precisely, between joy and fear; Russell & Mehrabian, 1977) and $d_{PADmax}(e, e_{target})$ is the Euclidian distance between an emotion e and the target emotion e_{target} on the PAD scales.

The denominator is based on the denominator of $score_2$, which is divided into two parts:

- **Part 1:** corresponds to the number of emotions that are different from the target emotion that is recognized by all of the participants $\left(\sum_p \sum_{e \neq e_{target}} reco_{exp}(p, e)\right)$. This part is weighted by the distance between each recognized emotion and the target emotion e_{target} . The closer the emotions, the higher the weight. This part decreases the global score of the haptic expressions that present ambiguities between the target emotion and the close emotions.

TABLE 3
The Principal Physical Features of the Selected Expression of Each Emotion

Measure	Joy	Elation	Disgust	Contempt	Anxiety	Fear	Irritation	Rage
Distance (m)	1.6	1.7	0.35	0.0015	0.98	1.2	0.36	3.1
Duration (s)	5.5	7.3	5.8	6.3	6.0	6.5	6.0	5.1
Amplitude (m)	0.28	0.19	0.16	$6.0e^{-4}$	0.24	0.20	0.060	0.28
M speed (m/s)	0.30	0.24	0.062	$2.5 e^{-4}$	0.17	0.20	0.061	0.62
Fluidity (m/s ²)	3.1	3.2	2.1	0.26	3.3	3.0	3.0	7.3
Contraction index (m)	0.075	0.045	0.050	0.0014	0.065	0.038	0.012	0.052
Major axis (X)	0.96	0.21	0.49	0.66	0.13	0.24	0.55	0.66
Major axis (Y)	0.23	0.85	0.73	0.73	0.75	0.96	0.24	0.11
Major axis (Z)	0.075	0.48	0.48	0.16	0.65	0.12	0.81	0.76
Weight of major axis	0.58	0.70	0.60	1.0	0.50	0.80	0.77	0.50
Weight of second major axis	0.30	0.17	0.37	0.0010	0.28	0.13	0.13	0.31
Repetitivity	2	7	2	2	2	0	3	6

- **Part 2:** corresponds to the number of participants who have recognized the target emotion e_{target} . This value corresponds to the denominator, to normalize the global score. Part 1 varies between 0 and $+\infty$.

We calculated the score for each haptic expression (i.e., 32 expressions) using the score function $score_3$. Then, for each emotion, we selected the haptic expression with the highest score among those that convey this emotion. Table 3 presents the selected haptic expressions and their principal physical features.

6. STUDY 4: VISUO-HAPTIC STUDY

6.1. Objectives and Hypotheses

The objective of this last experiment was to study the effect of the visuo-haptic coupling on the recognition of the investigated emotions. In addition to studying its effect on each of the emotions and on the whole set of emotions, we investigated whether some participants better exploit the visuo-haptic combination for the perception of some emotions. Other factors, such as the perceived expressivity, the presence of the virtual character, and the utility of using the haptic device are investigated to evaluate the relevance of the visuo-haptic strategy. The hypotheses of this experiment are as follows:

H2: *Recognition of all emotions is better in the visuo-haptic condition than in the visual and haptic conditions alone.* As participants have multiple clues to evaluate the emotion, it should be easier to discriminate emotions. In previous work (Bonnet et al., 2011), we found that the use of haptic feedback improves the recognition rate of facial expressions of virtual characters. Congruency of stimuli in different modalities has been observed to lead to a better recognition (de Gelder & van den Stock, 2011).

H3: *Recognition of an emotion is improved under the visuo-haptic condition compared to the visual and haptic conditions alone.* We expect not only the general recognition rate to improve but also the recognition rate of each emotion to improve.

H4: *Groups of participants improve their recognition of some emotions under the visuo-haptic condition compared to the visual and haptic conditions alone.* As the perception of emotions differs between people (Hamann & Canli, 2004), participants could have differing emotion recognition rates, depending on the modality of expression. Several studies have observed interindividual differences in the processing of facial expressions and its context (Lee, Choi, & Cho, 2012).

H5: *The visuo-haptic condition improves the sensation of presence compared to the visual and haptic conditions alone.* Previous studies observed an improvement of the sensation of presence in the case of multimodal interactions compared to mono-modal interactions (Tan, Courgeon, Bellik, & Martin, 2011).

H6: *The visuo-haptic condition provides a better expressivity for emotions compared to the visual and haptic conditions alone.* We expect the participants to perceive the virtual character to be more expressive in the visuo-haptic condition. Congruent combinations of facial expressions and body expressions displayed by a virtual character were perceived as being more intense than facial expressions and idle movements or no movements (Buisine et al. 2014).

The following two hypotheses aim to evaluate the acceptance and pertinence of the haptic feedback and device according to participants:

H7: *The participants prefer the visuo-haptic condition compared to the visual and haptic conditions alone.*

H8: The participants find the haptic device to be useful for the communication of emotions.

6.2. Experimental Platform

The experimental platform simultaneously provided visual and haptic feedback. It contained two computers (see Figure 5). The client node supports (a) the manager module that manages the progress of the experiment, (b) the MARC framework for the display of static facial expressions, and (c) the GUI that displayed the instructions to the participants and manipulated their inputs. As in the previous experiment, the haptic device (PHANTOM Desktop) was connected to a dedicated computer (a server node) to prevent haptic instabilities. The haptic device was controlled with a low-level module (a haptic module) that was based on the OpenHaptics library (<http://geomagic.com/en/products/open-haptics/overview/>). The client-server configuration exploited a UDP connection in a local network. The average delay time was approximately 32 ms.

When the platform displayed a visuo-haptic expression, the manager module sent requests to the MARC module, to generate the facial expression, and to the server node, to generate the haptic expression. After 2 s, the haptic and facial expressions were simultaneously rendered by the haptic device and on the screen, respectively. The picture of the facial expression was displayed for the whole duration of the haptic expression. In fact, the haptic expressions, which were generated by participants during the previous experiment, varied in duration. Once the haptic expression was completed, the display of the facial expression was removed.

6.3. Methods

Participants

Forty-one people (10 women, 31 men) participated in the experiment. The participants' average age was 27 ($SD = 9$).

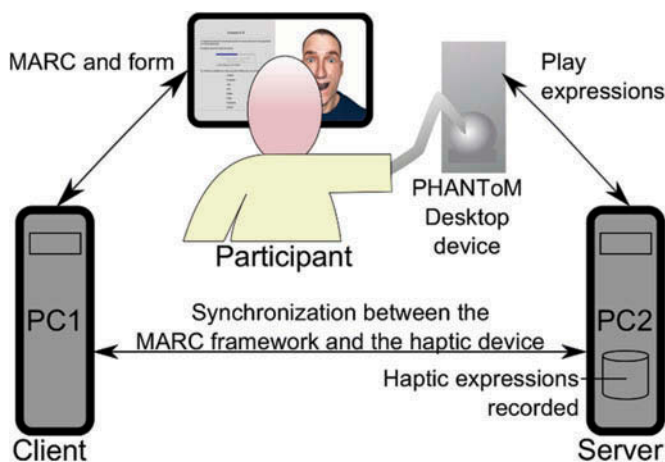


FIG. 5. The experimental platform that was used to study emotion recognition from facial and haptic expressions. *Note.* The client node provides a form that displays the instructions, displays the virtual agent's facial expression and records the answers of the participants. The server node records and renders the haptic expressions.

Thirty-six of the participants were right-handed. Regarding the participants' education, 23 of them received a European education, six received a North African education, two received an Asian education, and one received a South African education.

Conditions

To study the contribution of each modality and the role of the multimodal combination in the recognition of the emotions, three experimental conditions were investigated:

- **Visual condition:** Only facial expressions were displayed. These expressions are the same as those in Study 1.
- **Haptic condition:** Only haptic expressions were rendered. These expressions were selected in Studies 2 and 3.
- **Visuo-haptic condition:** Both congruent facial and haptic expressions were rendered.

The order of presentation of the three conditions was counter-balanced across the participant sample.

Measures

Several measures were collected for this experiment. First, we computed the recognition rate for the presented expressions. The participants were required to select the emotion perceived from the list of the eight emotions. They could select one or several emotions for a given expression. As explained previously, this evaluation aimed to highlight expressions that present ambiguity between two or more emotions.

Second, to evaluate the sensation of presence according to the three presented conditions, we presented the following questionnaire (QA; Bailenson, Blascovich, Beall, & Loomis, 2001) to the participants at the end of each condition:

- **Q1:** "I perceive that I am in the presence of another person in the room with me."
- **Q2:** "I feel that the person is watching me and is aware of my presence."
- **Q3:** "The thought that the person is not a real person crosses my mind often."
- **Q4:** "The person appears to be sentient (conscious and alive) to me."

The participants rated these questions using a 5-item Likert scale from 1 (*totally disagree*) to 5 (*totally agree*).

Third, after the previous questionnaire, we presented another questionnaire (QB) to evaluate the perceived expressivity for each condition. Questionnaire QB contains the following question:

- **Q5:** "I found the [Visual/Haptic/Visual+Haptic] rendering expressive." The participants rated this question using a 5-item Likert scale from 1 (*totally disagree*) to 5 (*totally agree*).

Finally, to evaluate the utility of the haptic device and the preference of the participants, we presented these two questions at the end of the experiment (Questionnaire QC):

- **Q6:** “What type of rendering did you prefer the most?” [Visual/Haptic/Visual+Haptic].
- **Q7:** “I found the haptic device to be useful for the perception or communication of emotions?” [Yes/No].

The participants selected one answer for each question.

Procedure

A participant started the experience by reading a document that describes the emotions that were considered. This document presented a typical context for each emotion that was described. Those contexts were issued from the MindReading database (Golan et al., 2006).

After reading the description, the participant was seated in front of a computer screen and a haptic device. Then the GUI interface displayed text that explained that a series of expressions corresponding to different emotions would be displayed with three types of stimuli (Visual, Haptic, and Visual+Haptic). The order of presentation of the expressions was random. Later, a second display of text informed the participant that, for each presented expression, he or she was to select the emotion perceived from a list of eight emotions. To motivate the participants, they were informed that their score of emotion recognition would be displayed at the end of the experiment. Finally, as in the previous studies, the participant was asked to hold the haptic device as if he or she were holding the wrist of another person.

For each condition, the experiment began with a short training session. Then a series of eight expressions were presented to the participants. The participants were required to click on a button to start the rendering of the expression. A progress bar indicated to them how long the expression was rendered. After each expression was presented, the participants completed a form to indicate the recognized emotion(s). At the end of each condition, questionnaires QA and QB were presented. At the end of the experiment, Questionnaire QC was presented. Finally, the score for each condition was displayed.

Results

For each condition, we computed the recognition score based on the emotion recognition of the participants for this condition. We used the following equation:

$$reco_{exp}(p, e) = \begin{cases} \frac{1}{\text{recognized}} & \text{if the participant } p \text{ recognized the emotion } e \text{ in the expression } exp, \text{ where recognized is the number of recognized emotions} \\ 0 & \text{otherwise} \end{cases}$$

No influence of the conditions was observed on the number of emotions recognized. On the basis of this score, we

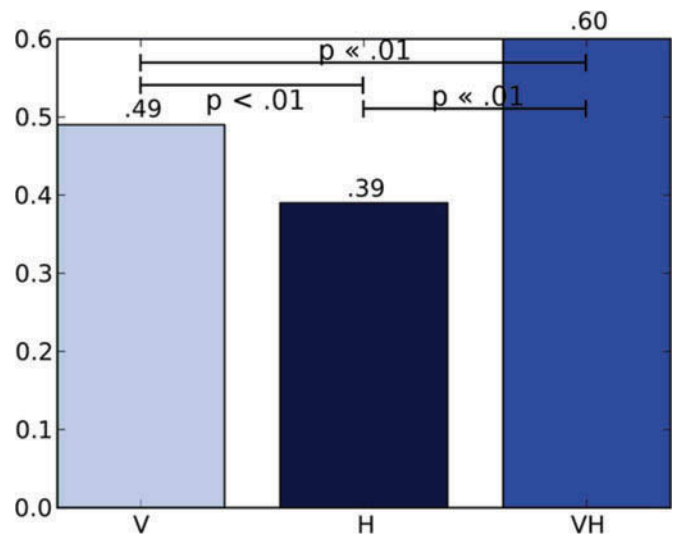


FIG. 6. The average recognition rates for all of the emotions, mixed according to the three conditions. *Note.* The recognition rate depends on the condition.

detail in the following sections the results that correspond to our hypotheses.

Recognition Rate According to the Conditions

The overall recognition rate of all of the emotions was combined to investigate the influence of each condition. Figure 6 presents the results of the mean recognition rate for the three conditions. The data did not follow a normal distribution. The Wilcoxon signed-rank test was utilized to highlight the influence of the conditions on the recognition rate.

The results showed that the visuo-haptic condition (VH) presented a significantly higher recognition rate (60%) than the visual condition (V; 49%, $p < .01$, $V = 96$) and the haptic condition (H; 39%, $p < .01$, $V = 84.5$) alone. Moreover, the haptic condition presented a significantly lower recognition rate than the visual condition ($p = .0085$, $V = 552$). On the basis of these results, H2 (recognition of all emotions is improved under the visuo-haptic condition compared to the visual and haptic conditions) is validated.

Recognition Rate According to the Emotions

The mean recognition rate of each emotion according to the three conditions was calculated (see Figure 7). A Friedman test showed an influence of the modality on the recognition rates, $\chi^2(2) = 7.1613$, $p = .028$. The following cases were observed:

- Similar recognition rates for the haptic condition and the visual condition ($V = H$): This is true for joy, contempt, anxiety, and irritation. We observed no significant influence of the three conditions ($p \geq .05$). This finding suggests that there is no predominance of haptic or visual cues for the recognition of these four emotions.

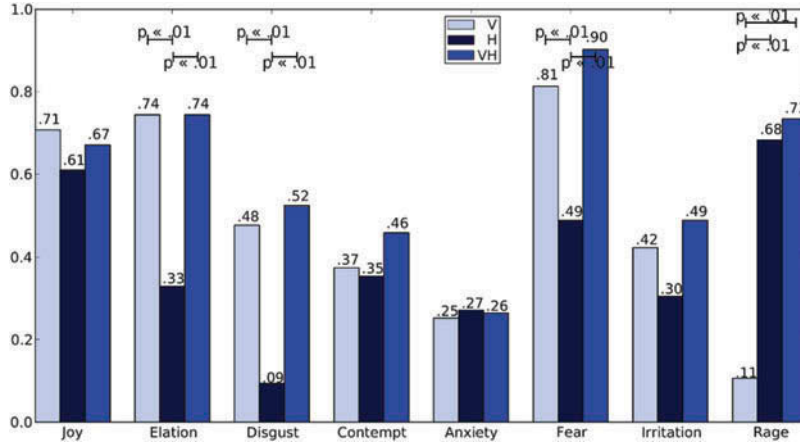


FIG. 7. The average recognition rates for each emotion for the three conditions. Elation, disgust, fear, and rage present significant differences depending on the condition.

- The visual condition presents higher recognition rates than the haptic condition ($V > H$, $p < .01$): This case concerns elation ($p < .01$, $V = 56$), disgust ($p < .01$, $V = 22$), and fear ($p < .01$, $V = 31.5$). We observed that the visual and visuo-haptic conditions display significantly higher recognition rates than the haptic condition. This finding suggests that the participants exploit more visual cues under the visual and visuo-haptic conditions for the recognition of these emotions.
- The haptic condition presents higher recognition rates than the visual condition ($V < H$): This case concerns rage ($p < .01$, $V = 364$). We observed that the visuo-haptic condition displayed a higher recognition rate than the visual condition ($p < .01$, $V = 480$). This finding suggests that the participants exploit the haptic cues for the recognition of this emotion.

On the basis of these results, H3 (recognition of an emotion is improved under the visuo-haptic condition compared to the visual and haptic conditions alone) is validated for rage but not for the other emotions. However, those results showed that the recognition rates for the visuo-haptic condition are close to those of the modality (V or H) presenting individually the higher recognition rates. The important standard deviation observed in recognition rates suggests also a great variability between participants. In the next section we investigate the individual differences in the participants regarding their recognition rates according to the channel of expression: visual or haptic.

Individual Differences

We also analyzed whether the haptic display was better perceived by some participants. More precisely, we study the sign of the difference between the recognition rates corresponding to two conditions: Visuo-haptic and visual alone ($\text{sgn}(VH - V)$; sgn is the sign function. This analysis aims to highlight the

participants for which haptics leads to an improvement of the recognition rate under the visuo-haptic approach. For each participant p and for each emotion e , we computed the following evaluation function:

$$\text{Eval}(p, e) = \begin{cases} 1, & \text{if } \text{sgn}(VH - V) > 0 \\ 0, & \text{otherwise} \end{cases}$$

The standard analysis approaches to study populations in data sets use clustering algorithms such as EM. However, the noncontinuous representation of the analyzed data ($\text{Eval}(p, e) = 1 \text{ or } 0$) could interfere with this type of clustering algorithm because the data does not present an order relation. To address this issue, we used a Multiple Factor Analysis (MFA; Thurstone, 1931) before applying the clustering. The MFA is suitable for this type of qualitative data and makes the classification more robust because the eigenvalues associated with the factorial axes give a clue as to the number of clusters required for clustering. Then, a K-means algorithm is applied to the data set by specifying the number of clusters identified.

The MFA algorithm was simultaneously applied on all of the participants' results $\{\text{Eval}(p, \{e\}_{e \in \text{emotions}})\}_{p \in \text{participants}}$. This step highlighted three different groups of participants. Thus, the K-means algorithm ($K = 3$) was applied to cluster the dataset. The results are detailed in Table 4. We observed that the improvement of the recognition rate under the visuo-haptic condition varied among the three groups. The population that corresponds to Cluster 1 presented an improvement for only one emotion: rage. The population that corresponds to Cluster 2 presented a global improvement for two emotions: disgust and rage. Finally, the population that corresponds to Cluster 3 presented a global improvement for several emotions: contempt, anxiety, fear, and irritation. This population (13% of the participants) more effectively exploits the haptic feedback to recognize the emotions.

TABLE 4
Results Obtained by the K-Means Algorithm

Cluster	Joy	Elation	Disgust	Contempt	Anxiety	Fear	Irritation	Rage
1 (46%)	21	5	0	37	21	26	0	100
2 (41%)	35	18	70	6	18	18	41	65
3 (13%)	0	40	20	100	100	60	80	40

Note. The numbers are the percentages of participants who perform better in the visuo-haptic condition than in the visual condition. Percentages that are greater than 50% are displayed in bold.

These different results support H4 (Groups of participants improve their recognition of some emotions under the visuo-haptic condition compared to the visual and haptic conditions alone).

To establish whether the improvement of the recognition rate under the visuo-haptic condition is due to haptics, we analyzed the sign of the difference between the recognition rate under the haptic and visual conditions $sgn(H - V)$. This approach aims to highlight the participants that recognize emotions with haptics equally well or better than they do with a visual display. We computed the following evaluation function:

$$Eval(p, e) = \begin{cases} 1, & \text{if } sgn(H - V) \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

The MFA algorithm identified four different groups. Then, the K-means algorithm ($K = 4$) was applied to cluster the data set. The results are detailed in Table 5. We observed that the population corresponding to Cluster A (49% of the whole population) recognizes all of the emotions with haptics equally well or better than they do with a visual display alone. Surprisingly, all of the participants who performed better under the visuo-haptic condition (Cluster 3, Table 4) were included in this cluster. This finding suggests that the participants who displayed the higher recognition rates in the haptic condition were also the ones who were helped the most by the addition of the haptic feedback in the visuo-haptic condition.

Subjective Measures

Figure 8 and Figure 9 present the subjective results.

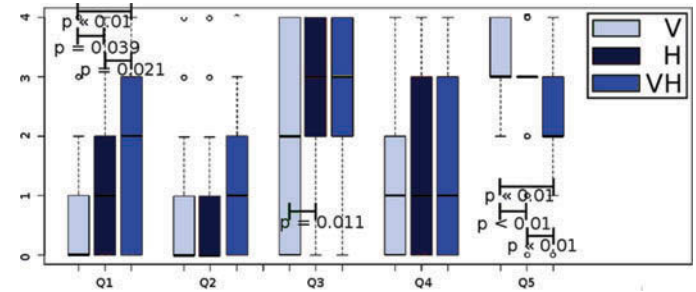


FIG. 8. The answers of the participants for the sensation of presence (Q1 to Q4) and perceived expressivity (Q5). Note. The black horizontal lines represent the medians. The error bars indicate the 25th and 75th percentiles. Circles represent outliers. All of these questions were asked to participants for each condition. Answers to Q1, Q3, and Q5 depend on the condition.

Sensation of presence. The statistical test (Wilcoxon) showed that there is no significant influence of the three conditions on Q2 and Q4. This finding is not surprising, because the experiment comprised a one-way interaction. The computer did not react to the participant's behavior. However, it remained interesting to investigate those two questions because we could not evaluate the perception of the avatar coupled to the haptic device before this experiment.

The visuo-haptic condition presented a significantly higher score than the two conditions alone (VH: 2.60 compared to H: 2.17, $p = .021$, $V = 57.5$, and V: 1.73, $p < .01$, $V = 10$) for Q1. Moreover, the haptic condition presented a significantly higher score than the visual condition (2.17 and 1.73, respectively, $p = .039$, $V = 50$). This finding suggested that the physical contact

TABLE 5
The Results Obtained by the K-Means Algorithm, Comparing the Haptic Versus Visual Condition

Cluster	Joy	Elation	Disgust	Contempt	Anxiety	Fear	Irritation	Rage
A (49%)	75	65	70	80	90	75	90	95
B (17%)	71	0	0	71	71	100	57	100
C (15%)	100	83	17	17	83	17	17	100
D (19%)	37	0	87	87	37	12	87	100

Note. The numbers are the percentage of bests in the haptic condition for each cluster. The numbers in bold are those greater than 50 (i.e., there are more participants who are better in the multimodal condition for this emotion than participants who are better in the visual condition in this cluster).

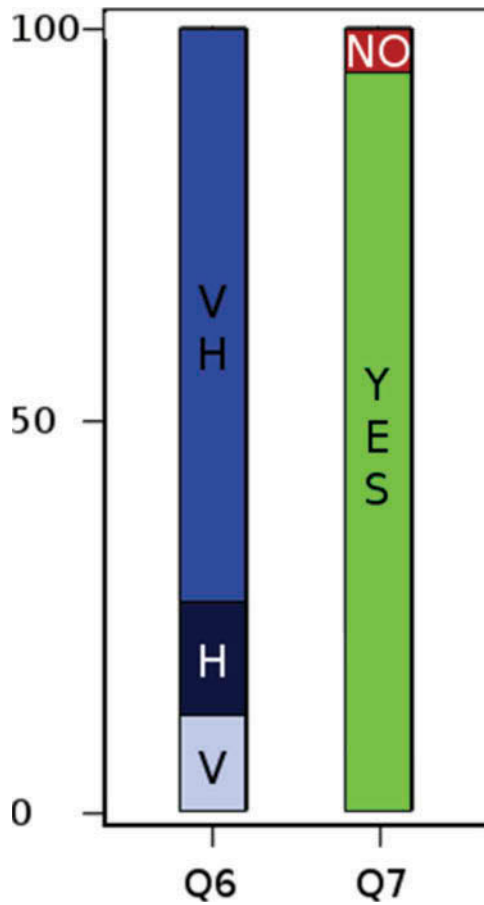


FIG. 9. The answers of the participants for preferred rendering (Q6) and usefulness of the haptic device for communicating emotions (Q7). Note. Those questions were asked after the participants completed the three conditions.

through the haptic interface and the dynamic behavior of the haptic expressions increased the sensation of presence.

The visual condition presented a significantly higher score than the haptic condition for question Q3 (2.63 and 1.95, respectively, $p = .011$, $V = 238.5$). Moreover, the visuo-haptic condition presented no significant differences when compared to the two other conditions. This finding suggests that the haptic channel better supported the realism of the virtual character, whereas the static visual expression limited the realism of the interaction.

All of these results validate H5 (The visuo-haptic condition improves the sensation of presence compared to the visual and haptic conditions).

Expressivity. The participants reported that they found the visuo-haptic condition to be the most expressive compared to the visual and haptic conditions alone (VH: 4.34 compared to V: 3.93, $p = .0093$, $V = 57$ and H: 2.19, $p < .01$, $V = 33$). In addition, the visual condition was considered to be more expressive than the haptic condition ($p < .01$, $V = 370$) in answers to Q5.

This result validated H6 (The visuo-haptic condition provides a better expressivity for emotions compared to the visual and haptic conditions alone).

Preferences of the participants. The participants reported that they preferred the visuo-haptic condition to the visual and haptic conditions alone (VH: 73%, V: 12%, H: 15%) in their answers to Q6. This result validated H7 (The participants prefer the visuo-haptic condition compared to the visual and haptic conditions).

Usefulness of the haptic device. Ninety-five percent of the participants found that the haptic device was useful and adapted to convey emotions (Q7). This result validated H8 (The participants find the haptic device useful for the communication of emotions).

7. GENERAL DISCUSSION

7.1. Haptic Expressions

Concerning the recognition of haptic expressions, the results observed by Bailenson et al. (2007), who exploited a similar protocol (with a different haptic device: an Immersion Impulse Engine 2000), differed from our results in several noticeable ways. This comparison mainly concerns three emotions that were common for the two studies: Joy, disgust, and fear. Table 6 details the recognition rates from both studies. The differences between the collected expressions and the investigated emotions are detailed in the following:

- **Joy.** We observed that the cinematic features extracted from the corresponding haptic expressions were similar in both studies (e.g., high distance and mean speed, low weight of the major axis). Two differences could explain this large difference in the recognition rate. First, we used the depth axis in our study, which was not available in the experiment of Bailenson et al.; however, our results have shown that this axis was used very little (0.075 in the selected expression). Second, the evaluation method in our case allowed the participant to check one or several emotions, whereas Bailenson et al. allowed checking only one emotion. In our study, the participants usually checked only one emotion (1.12) when this emotion was successfully

TABLE 6

The Comparison of Recognition Rates Observed for Three Basic Emotions (Joy, Disgust, and Fear) Between This Study and the Study of Bailenson et al. (2007)

Emotion	This Study	Bailenson et al. (2007)
Joy	63%	37.5%
Disgust	12%	31.3%
Fear	56%	37.5%

Note. The measures extracted from collected haptic expressions are quite similar in both studies. However, the sets of studied emotions are different. We assume that the differences observed are due to ambiguity between emotions. Numbers in bold correspond to the highest recognition rate for each emotion.

recognized, which means that there is no confusion with other emotions. One possible explanation for this large difference in the recognition rate concerns the confusion with other close emotions. In fact, Bailenson observed a similar recognition rate between joy and interest, which are close in the PAD space. Moreover, these emotions were also expressed with the same axis, in contrast to joy and elation in our study.

- **Disgust.** We observed that the cinematic features extracted from the corresponding haptic expressions were similar in both studies (e.g., a low distance and mean speed and a high weight of the major axis). The low recognition rate in our study can be explained by confusion with contempt. In fact, only 12% of the participants correctly recognized the haptic expression that corresponded to disgust, and 24% of the participants confused the haptic expressions of disgust and contempt. The absence of discriminative features of the haptic expressions of these two emotions might explain this confusion (see section 4).
- **Fear.** We observed that the cinematic features extracted from the corresponding haptic expressions were quite similar in both studies (e.g., a low distance and mean speed but an up-down axis for Bailenson and a left-right axis for us). However, in the study of Bailenson, the participants confused fear with sadness and disgust. The main differences between the two studies concerns the use of the depth axis (0.12; see Table 3). This finding might explain the results.

Beyond these local differences, the overall recognition rates of the two studies are similar: 39% in our study and 33% for the study of Bailenson et al. Moreover, both studies highlighted close discriminative features for common emotions. This finding suggests that there is a general way to express those three basic emotions using the haptic modality.

7.2. Combinations of Visual and Haptic Expressions

Concerning the combination of facial and haptic expressions, several studies highlighted the potential of facial expressions to efficiently convey the pleasure dimension and the potential of haptic expressions as gesture expressions to efficiently convey the activation dimension (Bickmore et al., 2010; Castellano, 2008; Courgeon et al., 2011). In our study, a Spearman's correlation test revealed that subjects are better at recognizing emotions that have a high amplitude of pleasure using the visual modality ($p = .046$, $\rho = 0.74$). Subjects also tend to be better at recognizing emotions that have a high activation when using the haptic modality ($p = .062$, $\rho = 0.48$). This finding suggests that the addition of haptic expressions should be effective for increasing the discrimination between facial expressions of emotions that are close in the pleasure dimension and far in the activation dimension.

Several studies regarding the visual or speech expression of emotions observed that different people do not rely on the same

clues to recognize emotions (Masuda et al., 2008). This finding explains why some emotions are rarely recognized by 100% of the people. The results of this study highlighted a similar behavior for the haptic channel. The participants did not recognize emotions in the same way using haptics. They do not rely on the same cues when both the visual and haptic channels are used simultaneously.

8. CONCLUSIONS

This article investigated the improvement of the recognition rate of emotions expressed by facial expressions by using a visuo-haptic combination. We observed that some emotions are difficult to discriminate when using facial expressions alone, especially when similar emotions are presented.

To improve the recognition rate, we proposed to combine congruent facial and haptic expressions. We collected a large corpus of affective haptic expressions. We explained how we extracted relevant expressions from this corpus. This approach combined (a) a statistical test to select the most recurrent expressions for each emotion and (b) a perceptive test to evaluate the recognition rate of those expressions. The results showed a small improvement in the recognition rate and highlighted some individual differences in the recognition of haptic expressions.

These conclusions suggest avenues for future research. In the last study of this work, facial and haptic expressions conveyed the same congruent emotions. However, those expressions were collected from different sources (facial expressions were modeled from videos of actors; 3D haptic expressions were collected in our own experiments). This approach might have limited the coherence of the presented expressions. Future studies will augment the size of the data set, investigate recognition times depending on the modality, and investigate improvements from using facial and haptic expressions for more spontaneous emotions. Moreover, to study more realistic affective interaction, dynamic facial expressions and issues related to synchronicity and overlapping of visuo-haptic expressions will be considered. Finally, we plan to consider other measures, such as brain activity (EEG) and facial muscles activity (EMG), to study and highlight whether emotions are well conveyed to subjects by the different investigated conditions (haptic, visual, and visuo-haptic).

REFERENCES

- Ahn, S., Bailenson, J., Fox, J., & Jabon, M. (2009). Using automated facial expression analysis for emotion and behavior prediction. In K. Doehring, C. von Scheve, & E. A. Konijn (Eds.), *The handbook of emotions and the mass media* (pp. 349–369). New York, NY: Routledge.
- Audibert, N., Auberge, V., & Rilliard, A. (2008). How we are not equally competent for discriminating acted from spontaneous expressive speech. *Proceedings of Speech Prosody*, 693–696.
- Bailenson, J., Blascovich, J., Beall, A., & Loomis, J. (2001). Equilibrium theory revisited: Mutual gaze and personal space in virtual environment. *Presence: Teleoperators*, 10, 583–598.
- Bailenson, J., Yee, N., Brave, S., Merget, D., & Koslow, D. (2007). Virtual interpersonal touch: Expressing and recognizing emotions through haptic devices. *Human-Computer Interaction*, 22, 325–353.

- Bargh, J., & Shalev, I. (2012). The substitutability of physical and social warmth in daily life. *Emotion*, 12, 154–162.
- Basori, A., Bade, A., Sunar, M., & Daman, D. (2010). Face-Touch: An emotional facial expression technique of avatar based on tactile vibration in virtual reality game. In *Virtual Reality* (pp. 107–127). Rijeka, Croatia: InTech.
- Bickmore, T., Fernando, R., Ring, L., & Schulman, D. (2010). Empathic Touch by Relational Agents. *IEEE Transactions of Affective Computing*, 1, 60–71.
- Bonnet, D., Ammi, M., & Martin, J.-C. (2011). Improvement of the recognition of facial expressions with haptic feedback. *IEEE Haptic Audio Visual Environments and Games International Workshop*, 81–87.
- Buisine, S., Courgeon, M., Charles, A., Clavel, C., Martin, J.-C., Tan, N., & Grynszpan, O. (2014). The role of body postures in the recognition of emotions in contextually-rich scenarios. *International Journal of Human-Computer Interaction*, 30, 52–62.
- Castellano, G. (2008). *Movement expressivity analysis in affective computers: from recognition to expression of emotion* (Unpublished doctoral dissertation). Department of Communication, Computer and System Sciences, University of Genoa, Italy.
- Cha, J., Eid, M., Rahal, L., & El Saddik, A. (2008). HugMe: An interpersonal haptic communication system. *International Workshop on Haptic Audio visual Environments and Games*, 99–102.
- Chang, W., Hwang, W., & Ji, Y. G. (2011). Haptic seat interfaces for driver information and warning systems. *International Journal of Human-Computer Interaction*, 27, 1119–1132.
- Clay, A., Couture, N., Nigay, L., De la Rivi re, J., Martin, J., Courgeon, M., . . . Domengero, G. (2012). Interactions and systems for augmenting a live dance performance. *International Symposium on Mixes and Augmented Reality*, 29–38.
- Coan, J. A., & Allen, J. J. B. (2007). *Handbook of emotion elicitation and assessment*. New York, NY: Oxford University Press.
- Coulson, M. (2004). Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence. *Journal of Nonverbal Behavior*, 28, 117–139.
- Courgeon, M., Clavel, C., & Martin, J.-C. (2009). Appraising emotional events during a real-time interactive game. *International Workshop on Affective-Aware Virtual Agents and Social Robots*, 1–5.
- Courgeon, M., Clavel, C., Tan, N., & Martin, J.-c. (2011). Front view vs. side view of facial and postural expressions of emotions in a virtual character. *Transactions on Edutainment VI*, 132–143.
- Courgeon, M., & Martin, J.-c. (2009). Impact of expressive wrinkles on perception of a virtual character's facial expressions of emotions. *International Conference on Intelligent Virtual Agents*, 201–214.
- Courgeon, M., Martin, J.-c., & Jacquemin, C. (2008). MARC: A Multimodal Affective and Reactive Character. *1st Workshop on Affective Interaction in Natural Environments*, 12–16.
- Dael, N., Mortillaro, M., & Scherer, K. (2011). Emotion expression in body action and posture. *Emotion*, 12, 1085–1101.
- de Gelder, B., & Van den Stock, J. (2011). Real faces, real emotions: Perceiving facial expressions in naturalistic contexts of voices, bodies and scenes. In A. J. Calder, G. Rhodes, J. V. Haxby, & M. H. Johnson (Eds.), *The handbook of face perception* (pp. 535–550). Oxford, UK: Oxford University Press.
- Ekman, P. (1992). Are there basic emotions? *Psychological Review*, 99, 550–553.
- Ekman, P., & Friesen, W. (1975). *Unmasking the face: A guide to recognizing emotions from facial clues*. Englewood Cliffs, NJ: Prentice Hall.
- Gaffary, Y., Eyharabide, V., Martin, J.-C., & Ammi, M. (2013a). Clustering approach to characterize haptic expressions of emotions. *ACM Transaction on Applied Perception*, 10(4).
- Gaffary, Y., Eyharabide, V., Martin, J.-C., & Ammi, M. (2013b). Comparison of statistical methods for analysis of affective haptic expressions. *International Workshop on Haptic and Audio Interaction Design*, 69–78.
- Golan, O., Baron-Cohen, S., & Hill, J. (2006). The Cambridge Mindreading Face-Voice Battery: Testing complex emotion recognition in adults with and without Asperger Syndrome. *Journal of Autism and Developmental Disorders*, 36, 169–183.
- Hamann, S., & Canli, T. (2004). Individual differences in emotion processing. *Current Opinion in Neurobiology*, 14, 233–238.
- Handlin, L. (2010). *Human-human and human-animal interaction: Some common physiological and psychological effects* (Unpublished doctoral dissertation). Faculty of Veterinary Medicine and Animal Science Department of Animal Environment and Health, Skara, Sweden.
- Hartmann, B., Mancini, M., & Pelachaud, C. (2006). Implementing expressive gesture synthesis for embodied conversational agents. *Gesture in Human-Computer Interaction and Simulation: 6th International Gesture Workshop*, 188–199.
- Hertenstein, M., Keltner, D., App, B., Bulleit, B., & Jaskolka, A. (2006). Touch communicates distinct emotions. *Emotion (Washington, D.C.)*, 6, 528–533.
- Klema, V., & Laub, A. (1980). The singular value decomposition: Its computation and some applications. *IEEE Transactions on Automatic Control*, 25, 164–176.
- Kubera, R., & Wrighta, F. P. (2013). Augmenting the instant messaging experience through the use of brain-computer interface and gestural technologies. *International Journal of Human-Computer Interaction*, 29, 178–191.
- Lee, T.-H., Choi, J.-S., & Cho, Y. S. (2012). Context modulation of facial emotion perception differed by individual difference. *PLoS ONE*, 7, e32987.
- Masuda, T., Ellsworth, P. C., Mesquita, B., Leu, J., Tanida, S., & Van de Veerdonk, E. (2008). Placing the face in context: Cultural differences in the perception of facial emotion. *Journal of Personality and Social Psychology*, 94, 365–381.
- Mehrabian, A., & Ferris, S. (1967). Inference of attitudes from nonverbal communication in two channels. *Journal of Consulting Psychology*, 3, 248–252.
- Olausson, H., Cole, J., Vallbo, A., McGlone, F., Elam, M., Kramer, H., & Bushnell, M. (2008). Unmyelinated tactile afferents have opposite effects on insular and somatosensory cortical processing. *Neuroscience Letters*, 436, 128–132.
- Park, Y.-w., Lim, C.-y., & Nam, T.-j. (2010). CheekTouch: An affective interaction technique while speaking on the mobile phone. *Extended Abstracts on Human Factors in Computing Systems*, 3241–3246.
- Parkinson, B., Fischer, A., & Manstead, A. (2004). *Emotion in social relations: Cultural, group, and interpersonal processes*. New York, NY: Psychology Press.
- Picard, R. (1997). *Affective computing* (pp. 85–89). Cambridge, MA: MIT Press.
- Russell, J. (1994). Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies. *Psychological Bulletin*, 115, 102–141.
- Russell, J., & Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. *Journal of Research in Personality*, 11, 273–294.
- Scherer, K. (2000). Emotion. In *Introduction to Social Psychology: A European perspective*, 3rd edition (pp. 151–191). Oxford, UK: Blackwell Publishing.
- Scherer, K. (2005). What are emotions? And how can they be measured? *Social Science Information*, 44, 695–729.
- Smith, J., & Maclean, K. (2007). Communicating emotion through a haptic link: Design space and methodology. *International Journal of Human-Computer Studies*, 65, 376–387.
- Tan, N., Courgeon, M., Bellik, Y., & Martin, J. (2011). A location-aware virtual character in a smart room: Effects on performance. *Presence and Adaptivity. Human Factors*, 399–402.
- Thurstone, L. (1931). Multiple factor analysis. *Psychological Review*, 38, 406–427.
- Tsetersukou, D., & Neviarouskaya, A. (2010). World's first wearable humanoid robot that augments our emotions. *1st Augmented Human International Conference*, 1–10.
- Wallbott, H. (1998). Bodily expression of emotion. *European Journal of Social Psychology*, 28, 879–896.
- Yohanan, S., Chan, M., Hopkins, J., Sun, H., & Maclean, K. (2005). Hapticat: Exploration of affective touch. *Proceedings of the International Conference on Multimodal Interfaces*, 222–229.

ABOUT THE AUTHORS

Yoren Gaffary is a Ph.D. student in Computer Science at the LIMSI-CNRS lab. He received a master's degree in Information, Learning and Cognition at Paris South University. His Ph.D. thesis concerns affective computing using mediated touch with robotic devices coupled with virtual humans.

Victoria Eyharabide is an associate professor at Paris-Sorbonne University. She received a Ph.D. in Computer Science in 2010 from UNICEN University. Her Ph.D. thesis proposes an ontology-based approach to personalize human-computer interactions using context-enriched user profiles. Her research interests are related to the study of emotions in multimodal human-computer interactions and computational models of

nonverbal expressions of affects for designing interactive virtual characters.

Jean-Claude Martin is a full professor of Computer Science at Paris-Sud University. He conducts his research at LIMSI-CNRS, where he is head of the Cognition Perception Use group. His research consists of designing interactive virtual agents that are inspired by psychological theories of emotions and social behaviors.

Mehdi Ammi is an associate professor at Paris-Sud University specializing in haptics for virtual reality and teleoperation. In particular, he is interested in all aspects of haptic processes ranging from physiological mechanisms to the search for operational methodologies designed to integrate the haptic modality in different type of applications.