Intensifying Emotional Reactions via Tactile Gestures in Immersive Films

GEORGIOS KARAFOTIAS, AKIKO TERANISHI, and GEORGIOS KORRES,
New York University Abu Dhabi
FRIEDERIKE EYSSEL, Bielefeld University
SCANDAR COPTI and MOHAMAD EID, New York University Abu Dhabi

The film industry continuously strives to make visitors' movie experience more immersive and thus, more captivating. This is realized through larger screens, sophisticated speaker systems, and high quality 2D and 3D content. Moreover, a recent trend in the film industry is to incorporate multiple interaction modalities, such as 4D film, to simulate rain, wind, vibration, and heat, in order to intensify viewers' emotional reactions. In this context, humans' sense of touch possesses significant potential for intensifying emotional reactions for the film experience beyond audio-visual sensory modalities. This article presents a framework for authoring tactile cues (tactile gestures as used in this article) and enabling automatic rendering of said gestures to intensify emotional reactions in an immersive film experience. To validate the proposed framework, we conducted an experimental study where tactile gestures are designed and evaluated for the ability to intensify four emotional reactions: high valence-high arousal, high valence-low arousal, low valence-high arousal, and low valence-low arousal. Using a haptic jacket, participants felt tactile gestures that are synchronized with the audio-visual contents of a film. Results demonstrated that (1) any tactile feedback generated a positive user experience; (2) the tactile feedback intensifies emotional reactions when the audio-visual stimuli elicit clear emotional responses, except for low arousal emotional response since tactile gestures seem to always generate excitement; (3) purposed tactile gestures do not seem to significantly outperform randomized tactile gesture for intensifying specific emotional reactions; and (4) using a haptic jacket is not distracting for the users.

CCS Concepts: • **Human-centered computing → Interaction paradigms**; *Virtual reality*; Haptic devices; User centered design;

Additional Key Words and Phrases: Immersive virtual reality, affective haptics, multimodal interaction, tactile gestures

ACM Reference Format:
Georgios Karafotias, Akiko Teranishi, Georgios Korres, Friederike Eyssel, Scandar Copti, and Mohamad Eid. 2017. Intensifying emotional reactions via tactile gestures in immersive films. ACM Trans. Multimedia Comput. Commun. Appl. 13, 3, Article 29 (June 2017), 17 pages.
DOI: http://dx.doi.org/10.1145/3092840

1. INTRODUCTION

Recent advancements in information technology allow the creation of more realistic and immersive multimedia systems, ranging from 3D authoring to sound spatialization,

---

Authors' addresses: G. Karafotias, A. Teranishi, G. Korres, S. Copti, and M. Eid, New York University Abu Dhabi, Saadiyat Island, Abu Dhabi, United Arab Emirates, P.O. Box 129188; emails: {georgios.karafotias, at132, george.korres, scandar.copti, mohamad.eid}@nyu.edu; F. Eyssel, Universität Bielefeld, Exzellenzcluster Cognitive Interaction Technology (CITEC), Forschungsbau, Inspiration 1 / Zehlendorfer Damm 199, 33619 Bielefeld; email: feyssel@cit-ec.uni-bielefeld.de.

DOI: http://dx.doi.org/10.1145/3092840

and higher quality 2D as well as 3D displays. However, these systems are still limited to the simulation of audio and visual senses. Studies in multimodal human computer interaction have shown that haptic feedback seems to enhance the quality of user experience [Reiner 2004].

Films have already been shown to reliably elicit discrete emotional reactions [Schaefer et al. 2010]. Existing research has demonstrated that haptic stimulation is successfully used to intensify emotional immersion during media consumption, and is particularly effective in communicating arousal and valence [Eid and Al Osman 2016]. The literature study showed that most of the research in haptic-based affect communication focuses on the use of tactile cues.

Various technologies have been proposed to allow users to feel tactile effects synchronized with audio-visual contents [Danieau et al. 2013], from vibration motors embedded in wearable devices to moving chairs in 4D films. Due to the complexity and perceptual challenges, designing tactile effects is mostly a manual task. This renders authoring tactile effects in real-time applications tedious and almost impossible. In order to investigate the capabilities of tactile cues to intensify emotional reactions, we propose a framework that elicits emotional reactions from an audio-visual system (such as a clip) in the form of arousal-valence, transforms these emotions to tactile cues, and renders the corresponding tactile cues in synchronization with the audio-video contents to the user. Note that we define a tactile gesture as a tactile stimulus that is derived from a hand gesture (including defined intensity, direction, body position, frequency, etc.).

The remainder of the article is organized as follows. Section 2 analyzes the related work in mapping emotions to haptic feedback and the impact of adding haptic feedback on top of an audio-visual experience. In Section 3, our proposed framework for authoring and rendering tactile stimuli is described. Section 4 presents an experimental study to showcase the authoring of tactile gestures and investigate if the tactile gestures intensify emotional reactions in an immersive film experience. Finally, Section 5 summarizes the findings of the study and provides future endeavors.

2. RELATED WORK

Today, haptic interfaces for displaying tactile cues are well studied. Wearable jacket-like interfaces with haptic features are available in the literature. An early work in this direction was the TapTap prototype, which records and plays back tactile patterns in order to experience the affective human touch [Bonanni et al. 2006]. A subsequent effort at Philips Research Europe demonstrated a haptic jacket (with 64 tactile actuators) that focuses on influencing the emotions of movie watchers [Lemmens et al. 2009]. Cha et al. [2009] implemented HugMe, a synchronous haptic teleconferencing system where a custom-made haptic jacket is utilized to display tactile feedback to a local user who is touched by a remote counterpart. Krishna et al. [2010] developed a haptic glove (named VibroGlove) to deliver facial expressions of an interacting partner to people who are blind or visually impaired. The partner's expressions, which can portray the six basic human emotions (love, joy, surprise, anger, sadness, and fear [Ekman and Friesen 1971]) and the neutral emotional state, are mapped through haptic "emoticons" to signals that actuate vibrotactile motors on the back of the glove.

Research has also demonstrated that tactile stimulation can affect or elicit emotional reactions. Kryssanov et al. [2009] proposed a computational model to first assess a sensation on scales of "pleasure-unpleasure" and "anxiety-boredom" and then render it through a haptic device. Results demonstrated correlation between pleasant emotion and soft tactile stimulation, boredom and smooth tactile perception, unpleasant emotion and rough-sticky-repulsion forces, relaxation and slick tactile sensation, and anxiety and attracting forces. A recent study [Arafsha et al. 2015] has utilized a haptic jacket with an array of vibrotactile motors, a temperature actuator, and a heartbeat sensor to facilitate the display of the six universal emotions. The overall quality of user immersion was enhanced when tactile cues were displayed while watching a movie. In this study, experts designed tactile gestures based on user survey and were not optimized to elicit emotional reactions. Other studies focused on mapping audio signal to tactile effects [Weddle and Yu 2013] or visual signals (particularly motion) into tactile effects [Rashid et al. 2013].

Meanwhile, authoring tactile gestures for various applications seems to be less mature. For instance, Immersion Corporation presented a programmable tactile display to control various attributes of the tactile gesture such as vibration, intensity, frequency, and pattern [Immersion Corporation 2016]. However, these parameters would have to be controlled individually and un-intuitively. Israr et al. [2014] strived to create a library of tactile gestures (called Feel Effect (FE)) where an FE comprises a haptic component and a semantic component [Israr et al. 2014]. The haptic component specifies how the sensation unfolds over time and location via parameter settings for stimulus onset asynchrony (SOA), duration, intensity, and ramp-up for each actuator in an array of vibration motors. The semantic component describes what experience the sensation feels like. Results confirmed that semantically similar FEs lie in close proximity in haptic parameter space. Similarly, tactile cues were authored manually by mindful experts by defining individual stimulation properties (intensity, frequency, body location, direction, etc.).

The ability to intuitively author tactile cues for intensifying emotional reactions would enable a wide range of applications in the field of affective haptics [Eid and Al Osman 2016]. This article proposes a framework for authoring tactile gestures in an intuitive and user-friendly manner by mapping hand motion attributes into tactile gesture attributes. The starting and ending points, velocity, and direction of the tactile gesture is derived from the corresponding parameters of the hand gesture. Note that the intensity of tactile stimulation is derived from the depth of the hand relative to the camera position (the closer the hand gesture to the camera, the stronger the intensity of tactile stimulation is). With the proposed framework, users can conveniently create tactile gestures with hand gestures and feel the corresponding tactile gesture, without the need for controlling the tactile gesture attributes.

The contributions of this article are to: (1) propose a framework for authoring and rendering tactile gestures that may intensify emotional reactions, (2) introduce an intuitive and user-friendly tactile gesture authoring technique using hand gestures, and (3) present a case study to demonstrate the developed framework and evaluate its effectiveness to intensify emotional reactions and enhance overall quality of user experience in an immersive virtual reality film environment.

3. PROPOSED FRAMEWORK

The proposed framework provides two major functionalities (see Figure 1): (1) authoring tactile gestures and storing them in a repository for later use, and (2) rendering tactile gestures in an immersive film environment. Rendering the tactile gestures is performed in three steps: (a) emotional cues are extracted from audio-visual media, (b) tactile gestures are elected based on the extracted emotional cues (valence/arousal), and (c) the elected tactile gesture is displayed using a haptic jacket. The platform works both manual or automatic authoring. In the automatic scenario, emotional reactions are detected, mapped into tactile gestures, and automatically rendered using a haptic jacket. In the manual scenario, the extraction of emotional reactions is performed offline and the audio-visual media is annotated with tactile gestures ahead of playing time.

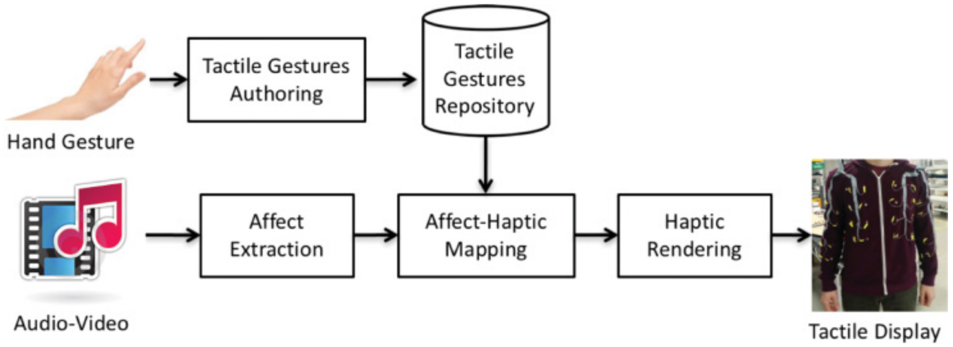

Fig. 1. Proposed framework for intensifying emotional reactions with tactile gestures.

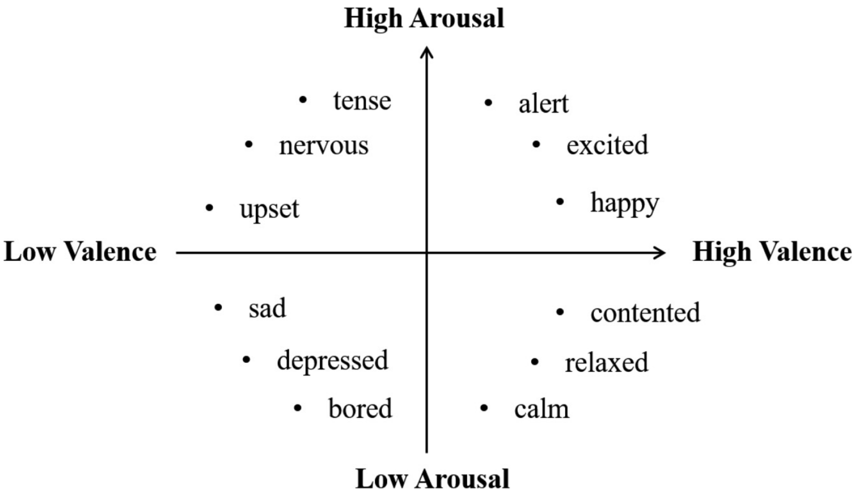

Fig. 2. Circumplex model of affect with arousal and valence dimensions.

3.1. Affect Extraction

The first step in the tactile gesture rendering process consists of analyzing the audio-visual contents to extract affective features that can then be mapped into haptic gestures. The abstract concept of emotion is modeled in the arousal-valence space in order to make it computationally easier (see Figure 2). Emotions are classified based on a continuous valued scale of two categories: arousal and valence [Posner et al. 2005]. Without loss of generality, they can be normalized to take values in the range of $[-1,+1]$ each. Estimating emotions on a continuous valued scale provides an essential framework for recognizing dynamics in emotions, tracking intensities over the course of time, and adapting to individual moods or personalities.

Previous studies extracted audio-visual features that are correlated with arousal and/or valence. Some of these features are correlated with arousal, some are correlated with valence, and some are correlated with both [Sun et al. 2009]. For instance, informative features for arousal estimation include loudness and energy of the audio signal, motion component, visual excitement, and shot duration. Features that correlate to valence include lighting, saturation, color energy, rhythm regularity, and pitch. As the aim of this study is to focus on authoring and rendering tactile gestures, the LIRIS-ACCEDE database [Baveye et al. 2015] of short annotated video-clips is utilized. The

Fig. 3. A. Exterior view of jacket; B. Front view of jacket turned inside-out; C. Back view of jacket turned inside-out.

clips are annotated with induced valence (negative and positive) and arousal (passive to active) values.

3.2. Affect Haptic Mapping

The second step in the tactile gesture rendering process comprises mapping the identified emotional responses (i.e., arousal-valence) into a corresponding tactile gesture that can be displayed using a haptic jacket. The concept is to generate a haptic cue that is capable of highlighting specific emotional reactions in a video. For instance, a scene showing an explosion implies high arousal and thus a tactile gesture with high arousal index will be elected (probably with abrupt, high intensity vibration). In such a case, video and audio analysis is used to detect explosions and then play the respective tactile gesture. Therefore, an arousal-valence state would be mapped to discrete values of these attributes that eventually shape the perception of the tactile gesture.

Affective haptics research has shown that the valence dimension of emotion is closely related to the frequency, continuity, position, and direction of perceived movement [Salminen et al. 2008]. Based on these findings, a mapping of the valence dimension of emotion to tactile gesture based on these parameters is implemented. A subsequent study suggested that a position-based mapping for valence is not the most effective to intensify emotional reactions [Nummenmaa et al. 2014]. The study concluded that the perceived direction of tactile gesture is more effective. Hence, valence is effectively conveyed through direction, rather than position. On the other hand, the arousal dimension of emotion is closely related to the intensity and frequency of tactile stimulation [Mazzoni and Bryan-Kinns 2015]. Therefore, the arousal dimension of emotion is best delivered using the intensity and frequency parameters. These findings are utilized to guide the user as they author tactile gestures to intensify particular emotional reactions.

3.3. Haptic Rendering

The last step in the tactile gesture rendering process is the display of the tactile gestures using a newly developed haptic jacket. The haptic jacket is a garment with embedded actuators that can simulate various tactile gestures. A network of 48 actuators is distributed over the flexible inner layer of a jacket as shown in Figure 3. A zipper allows for easy access to the actuators. The jacket, including all the hardware components detailed below, weighs 750 grams.

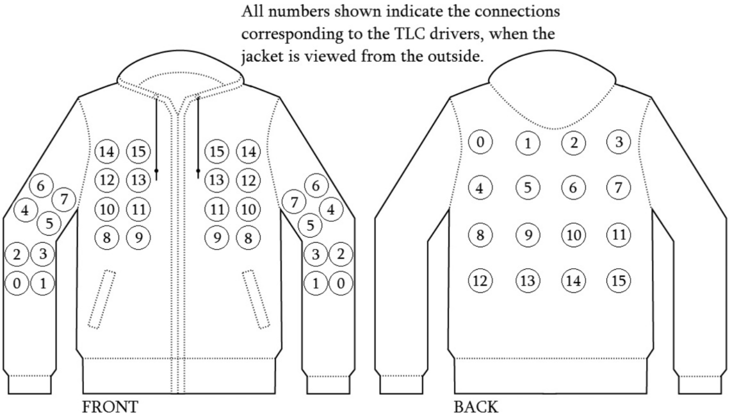

Fig. 4. Schematic showing the distribution of actuators on the front and arms (left) and back (right) of the jacket.

The main hardware components of the system are: 48 actuators distributed on the jacket, flexible cabling, and a custom-made printed circuit board (PCB) that contains the actuator control circuitry. The custom-made PCB contains a microcontroller, three 16-channel TLC drivers, a Bluetooth module, and a step-up regulator. An Arduino Nano microcontroller is used to control the entire system. It is programmed to receive serial commands from an external device to operate the actuators. The 16-channel TLC drivers are Pulse Width Modulation (PWM) units with 12-bit duty cycle control (0 - 4,095). The PWM technique is used to control the power supplied to the actuators. The three TLC drivers are daisy chained so that the number of PWM outputs is expanded to 48. In this way, the microcontroller can output to the 48 channels controlling the actuators at the same time.

A JY-MCU Bluetooth module is used for wireless communication between the microcontroller and a remote controller. Alternatively, the microcontroller can be connected to an external device through a hardware serial connection. On the jacket, the PCB is connected to the actuators using flexible cabling. Piezoceramic disk actuators were chosen because they are lightweight, thin, and inexpensive motors that can generate strong albeit non-audible vibration. The distribution and ordering of the actuators over the body is illustrated in Figure 4.

Note that the haptic jacket is designed to provide apparent tactile stimulation at the upper human body. There are eight motors placed at each of the arms - the motors are 4cm away from each other. Likewise, there are eight motors at each chest side (also distanced 4cm apart). Finally, the back of the jacket has 16 vibration motors that are placed 6cm apart (this is due to the fact that the back has lower sensitivity to tactile stimulation than the arms or chest areas). Funneling illusion is utilized to simulate apparent tactile motion. More details about the haptic jacket design can be found in our previous work [Lentini et al. 2016].

3.4. Tactile Gestures Authoring

The framework involves authoring tactile gestures that can be tested for correlation with specific emotional reactions (i.e., high arousal - high valence, high arousal - low valence, low arousal - high valence, and low arousal - low valence). The tactile gestures are authored via a software application developed in-house that allows for complete control over each of the 48 actuators by means of a video input interface. The application detects hand gestures and overlays these gestures on top of the 48 transducers map (as shown in Figure 5). Then, the transducers that collide with the hand gesture will

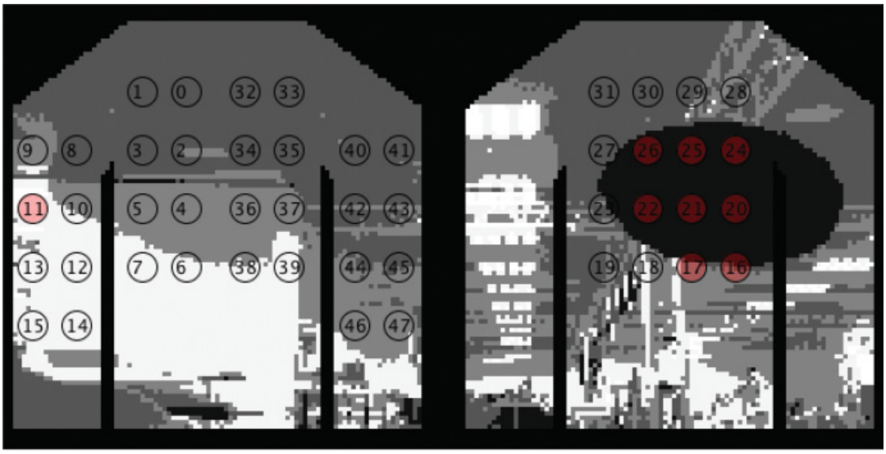

Fig. 5. Video-based gesture generation software. The front (left) and the back (right) vibration motors.

be activated over time. This will translate the hand gesture into tactile gesture. Hand gestures were chosen in order to provide an intuitive and natural means of generating tactile gestures that can be optimized without any technical knowledge about the tactile cue properties. To specify the intensity of actuation for each motor, a depth camera is used where a pixel depth is mapped into actuation intensity for the motor that overlaps with the corresponding pixel (the closer the pixel to the camera, the higher is the intensity). An outline of the front and back views of the jacket, with the positions of the actuators, is shown on a black background and the user is able to overlay a pattern onto the outline of the jacket, which actuates the motors that overlap with the regions selected by the user. Once the user is satisfied with the tactile gesture, the corresponding characteristics (intensity, frequency, direction of vibration, the mode of tactile motion (discrete versus continuous), and the position of tactile stimulation) are stored for later playback.

The depth stream is captured at a frame rate of 20Hz. Each frame is mapped into a 96-bit matrix that controls the intensity at which each motor is actuated. The indices of the elements refer to the number of the actuator, whereas each consecutive two-bits refer to the intensity of the vibration in a range of four steps. The motors are numbered 0 to 47 as shown in Figure 5.

The matrix is encoded into a 12-byte packet that is sent over to a microcontroller. The microcontroller is programmed to receive instructions from a remote controller in the form of bytes sent over a serial port. One byte at a time is read, decoded, and then saved into a matrix in the microcontrollers memory. After 12 bytes are received, the decoded matrix is used to set the output value for each channel. The PWM values are set as follows: “00” corresponds to zero duty cycle, or off; “01” corresponds to 1/4 of the maximum duty cycle, 1,023; “10” corresponds to 1/2 of the maximum duty cycle, 2,047; and “11” corresponds to 3/4 of the maximum duty cycle, 3,071. By updating the motors at 20Hz we can create the illusion of apparent motion [Israr and Poupyrev 2011]. The interface application also allows for a real-time control of the haptic jacket by using a camera. This can be used for quick testing of any gesture.

3.5. Tactile Gestures Repository

The tactile gestures repository is a database that stores previously created tactile gestures along with their respective emotional associations (such as arousal or valence index/range). This database is accessible by the Affect-Haptic Mapping component

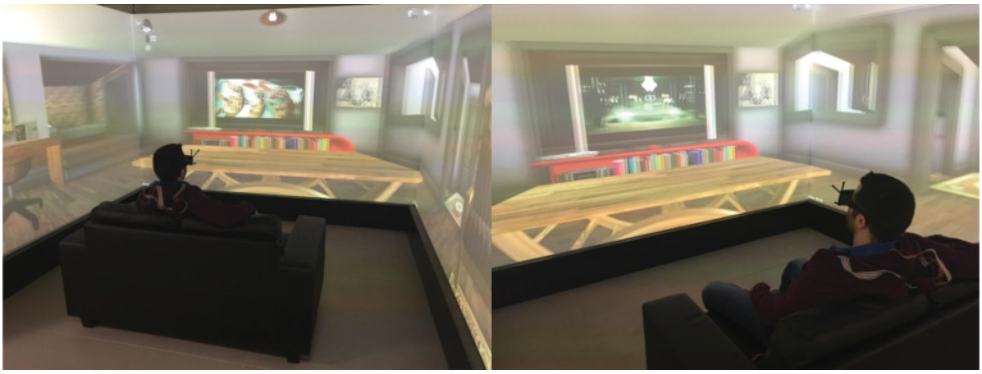

Fig. 6. Experimental setup.

to elect a tactile gesture that maximizes emotional reactions created by audio-visual media.

4. EXPERIMENTAL STUDY: TACTILE GESTURES TO INTENSIFY EMOTIONAL REACTIONS

By utilizing the developed framework, the aim of this study was to investigate whether there are distinct tactile gestures that are associated with intensifying specific emotional responses in an immersive virtual reality film experience. The study examines tactile gestures for emotional reactions in the arousal-valence space (high/low arousal/valence). Furthermore, this experimental study serves as an example realization of the proposed framework that elicits emotions from a film and renders corresponding tactile gestures.

4.1. Experimental Apparatus

The experimental testbed was composed of a three-sided CAVE system (walls size: 3mx2.25m, image size: 3mx1.875m, TechViz 3D driving software, 3D active glasses, and Advanced Realtime Tracking (ART) system), noise canceling headphones, and a custom-made haptic jacket that was developed at the Applied Interactive Multimedia research lab of New York University Abu Dhabi. The virtual environment comprises a living room with a large TV screen where the clip was exhibited (shown in Figure 6). A set of four short clips (with an average length of 111.75 seconds) was utilized, where each clip was associated with a specific emotional response and a respective tactile gesture.

4.2. Tactile Gestures Optimization

A pilot study was conducted with three participants to determine the best parameters to convey both arousal and valence in an intuitive, self-explanatory way. Both the average accuracy of recognition and the comments provided by the participants were used to alter parameters and optimize the gestures used in the next iteration of the optimization process. This process was repeated three times such that a total of six different gestures were tested for each parameter being evaluated (arousal and valence). One final round of testing involving the top five gestures for each parameter was conducted with a different set of three participants and the gestures were ranked. Results from these pilot tests, shown in Table I (arousal) and Table II (valence), are used to find the optimal frequency and intensity to map both valence and arousal effectively. An interesting finding of this pilot study is that the direction of apparent motion was more relevant than the stimuli position when attempting to convey valence.

Table I. The Configurations for the Five Candidate Tactile Gestures for Arousal

Rank	Low Arousal	High Arousal
1	Frequency: 0.5Hz Intensity: 40%	Frequency: 1.4Hz Intensity: 90%
2	Frequency: 0.7Hz Intensity: 40%	Frequency: 2Hz Intensity: 90%
3	Frequency: 0.5Hz Intensity: 30%	Frequency: 0.5Hz Intensity: 90%
4	Frequency: 0.45Hz Intensity: 50%	Frequency: 2Hz Intensity: 40%
5	Frequency: 1Hz Intensity: 50%	Frequency: 2Hz Intensity: 60%

Table II. The Configurations for the Five Candidate Tactile Gestures for Valence (Intensity: 50%)

Rank	Low Valence	High Valence
1	Frequency: 0.5Hz Mode: Discontinuous Direction: Downward and Out Position: Front Side and Arms	Frequency: 1.4Hz Mode: Continuous Direction: Upward and In Position: Front Side and Arms
2	Frequency: 0.3Hz Mode: Discontinuous Direction: Out Position: Front Side and Arms	Frequency: 2Hz Mode: Continuous Direction: Upward Position: Front Side and Arms
3	Frequency: 1Hz Mode: Discontinuous Direction: Out Position: Front Side and Arms	Frequency: 1Hz Mode: Continuous Direction: Out Position: Front Side and Arms
4	Frequency: 1Hz Mode: Discontinuous Direction: Downward Position: Front Side and Arms	Frequency: 2Hz Mode: Continuous Direction: Out Position: Front Side and Arms
5	Frequency: 2Hz Mode: Discontinuous Direction: Upward and In Position: Front Side and Arms	Frequency: 1Hz Mode: Continuous Direction: Downward and Out Position: Front Side and Arms

Consequently, four candidate tactile gestures (i.e., those that ranked first in each category) were considered for the experiment (Table III lists the configurations for these selected tactile gestures). As literature describes both arousal and valence as being functions of frequency, we established a linear relationship such that frequency increases with both arousal and valence.

4.3. Procedure

Sixty-three subjects participated in this study. They were divided into three groups (of 21 persons each): the “no haptics” group which did not experience any tactile gestures and served as a reference for the other two groups; the “randomized haptics” group which experienced a random (but always incorrect) mapping between the tactile gestures and the clip contents; and the “purposed haptics” group which experienced the correct-purposed mapping between the tactile gestures and the clip contents. Meaning that if, for example, a film includes a scene with high arousal and low valence values,

Table III. The Configurations for the Four Selected Tactile Gestures

	Low Valence	High Valence
Low Arousal	Frequency: 0.45Hz Intensity: 50% Mode: Discontinuous Direction: Out Position: Front Side and Arms	Frequency: 1Hz Intensity: 50% Mode: Continuous Direction: Upward Position: Front Side and Arms
High Arousal	Frequency: 0.5Hz Intensity: 90% Mode: Discontinuous Direction: Downward and Out Position: Front Side and Arms	Frequency: 1.4Hz Intensity: 90% Mode: Continuous Direction: Upward and In Position: Front Side and Arms

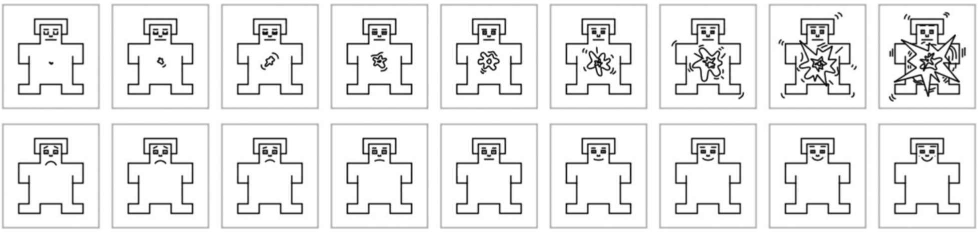

Fig. 7. SAM to assess arousal and valence. Arousal on top. Valence in bottom.

then the appropriate haptic gesture from Table III is displayed. The age range and the propensity to watch movies were similar within the groups, whereas the gender ratio was precisely the same (11 females and 10 males, per group).

First, we introduced the purpose of the experiment and explained the experimental setup. Each participant was then asked to wear the haptic jacket and the active 3D glasses in order to be fully immersed in the virtual space, and then sit comfortably in an armchair. Before the test session, participants were allowed to try out the system (specifically the vibrotactile feedback for the two haptics groups) so they would be well acquainted and more comfortable with the setup. During the session, the subjects watched a series of four clips supplemented with or without tactile gestures, based on their group. After each clip, they answered questions regarding the emotions depicted in that specific clip. At the end of the session, the participants completed a survey that assessed their overall quality of experience (QoE) and took part in a brief informal interview (see Appendix A).

4.4. Assessment Method

Participants rated the tactile gestures through the use of a modified version of the Self-Assessment Mannequin (SAM) [Bradley and Lang 1994]. This study used only two dimensions of SAM, the arousal and valence, as shown in Figure 7. A questionnaire was also designed to evaluate the quality of user experience, as shown in Appendix A. The questionnaire was built around the presence [Witmer and Singer 1998] and user satisfaction [Albert and Tullis 2013] concepts. Presence measures how much the user feels being physically connected to the clip contents. Witmer and Singer identified four factors to determine presence: control, realism, sensory, and distraction. We considered all factors except control - since the user is passively watching the clip. The realism factor determines how much the vibrotactile stimulation is natural and consistent with the users representation of the real world. The sensory factor characterizes how each modality is solicited during the interaction (audio, visual, and haptics). The distraction

Table IV. Evaluation Matrix (QoE and Distraction)

Factor	Questionnaire	Scale
Realism	1. Watching the movies in the virtual environment felt like a real-world experience.	1 (Strongly disagree) - 9 (Strongly agree)
Satisfaction	2. The overall experience of watching short clips is pleasant. 3. I would buy this technology if it becomes commercially available.	1 (Strongly disagree) - 9 (Strongly agree)
Sensory	4. Did the jacket enrich your experience while watching the movie?	1 (Strongly disagree) - 9 (Strongly agree)
Distraction	5. Did anything distract you from fully diving into the film experience? If so, please explain briefly.	Free writing and comments

Table V. Clip Name - Creator - Actual Valence/Arousal Values - Summary

		Valence	Arousal	Summary
Chatter	Leo Resnes	Low(-0.27)	High(0.76)	Girl uses social media to chat but witnesses something terrifying
On Time	Todd Wiseman	High(0.23)	Low(-0.37)	Preoccupations about time in modern human society
The Room of Franz Kafka	Fred. L'Epee	Low(-0.35)	Low(-0.48)	Kafka's paranoia, surrealism, and virtual perspective
Wanted	Ezel Domanic	High(0.73)	High(0.46)	Knowing your goals and striving to achieve them

factor identifies how much the user is disturbed by the apparatus used to create the immersive haptic experience (CAVE and haptic jacket). User satisfaction describes the acceptance and willingness to use immersive haptic systems.

Inspired by the previous work, we define, in this study, the quality of user experience in terms of three factors: realism, satisfaction, and sensory. To evaluate these factors, a total of four questions were administered using a rating on a nine-point Likert-scale (there were two questions about user satisfaction); then the quality of user experience is computed quantitatively by averaging these three factors. Moreover, we wanted to assess if (and how much) the haptic jacket was a distracting factor in the overall immersive experience, so we also included a question about distraction, which was evaluated non-quantitatively with written comments. All the factors are listed, along with their corresponding questions, in Table IV.

Table V lists the four different clips that were used in the experiment. These were selected from the LIRIS-ACCEDE database based on the combination of their recorded arousal/valence values. The values in this repository range from -1 to 1, so we considered a positive value as high and a negative value as low; for example the clip "Chatter" was selected because it contains a scene that induces low valence (-0.27) and at the same time high arousal (0.76).

4.5. Results and Discussion

Three groups are considered for analysis and comparison: "no haptics" (control group), "randomized haptics," and "purposed haptics." The independent variables are the realism, satisfaction, sensor, and distraction whereas the dependent variables are the QoE, the haptic audio visual correlation, and the recognition rate.

4.5.1. Quality of User Experience. The quality of user experience is defined as the aggregation of realism, satisfaction, and sensory. Satisfaction comprises two components: user satisfaction and commercial value. An outlier analysis was performed to determine if there were any values that should have been considered abnormal and thus eliminated. Using an interquartile range (IQR) value of 3.0, one value was found and

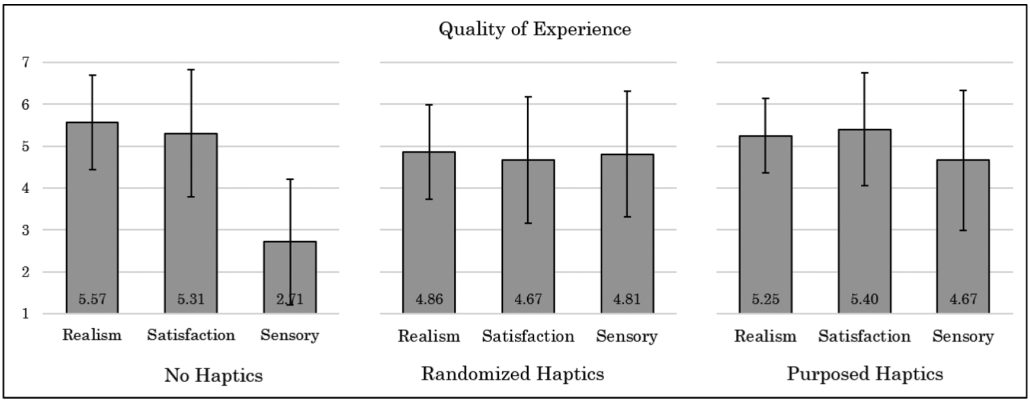

Fig. 8. QoE.

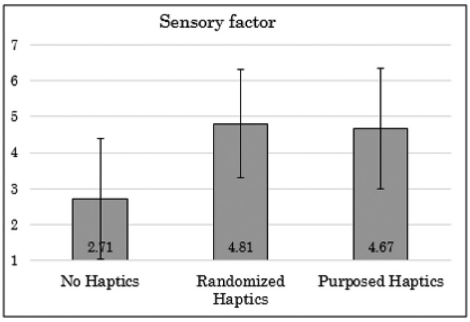

Fig. 9. Sensory factor of QoE.

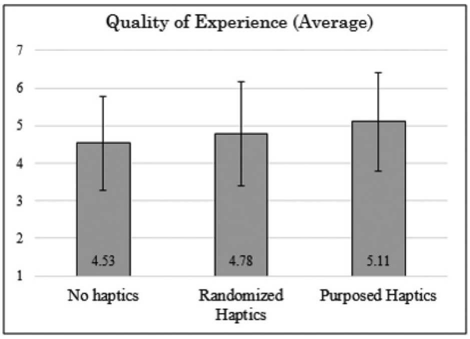

Fig. 10. Average QoE.

discarded in the realism factor of the QoE. The averages of each factor are shown in Figure 8.

ANOVA results showed that there are no significant differences between the three groups in the realism (F(2, 60) = 2.59, n.s.) and the satisfaction factors (F(2, 60) = 2.46, n.s.). However, there are significant differences in the sensory factor (F(2, 60) = 10.52, $p < .05$); see Figure 9. Post-hoc analysis (Tukey test; $p < .05$) revealed significant differences between the "no haptics" group and the "purposed haptics" group, and the "no haptics" group and the "randomized haptics" group; but there were no significant differences between the "randomized haptics" and "purposed haptics" groups. The difference in the sensory factor is an indication that having any kind of haptic feedback (either purposed or randomized) enriches the user experience (previous research presented a relationship between sensory factor and QoE [Hamam et al. 2013]). Note that the slight increase of rating for randomized haptics over purposed haptics in Figure 9 is not significant - this is also confirmed via statistical analysis.

Figure 10 shows the average value of the QoE for each group, along with standard deviations. "Purposed haptics" is rated with the highest QoE with 9.51% enhancement over the "no haptic" group. In terms of distraction, 98.41% of the participants did not mention the haptic jacket as a source of distraction. One subject noted that "The amount of wearable tech needed" was disconcerting. This probably includes the headphones, the 3D glasses, and the haptic jacket.

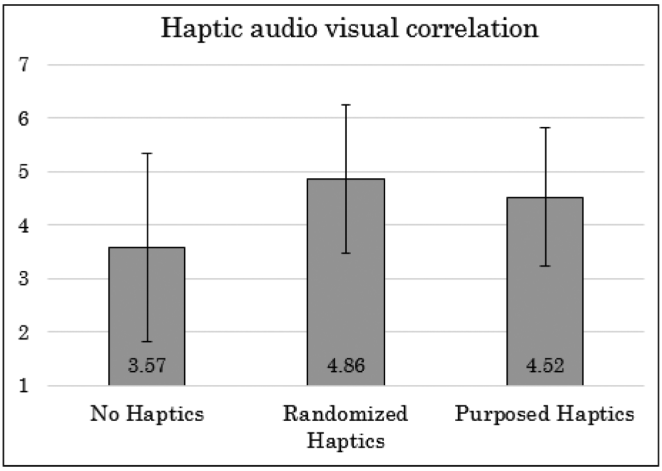

Fig. 11. Affective haptic audio visual correlation.

4.5.2. Affective Haptic-Audio-Video Correlation. To measure the effectiveness of the haptic feedback, we asked the subjects "Did that particular haptic-audio-visual feedback match the emotion depicted in the series of clips you watched?" Figure 11 shows that there are significant differences between the "no haptics" and "randomized haptics" groups, but there are no significant differences between the two haptics groups and between the "no haptics" and "purposed haptics" groups ($F(2, 60) = 3.98$, $p < .05$, Tukey post hoc). The "no haptics" group's low correlation score (with high standard deviation) can be justified by the fact that this group's members did not actually feel any haptic feedback. This implies that the audio-visual stimuli combination is the dominant one and the main source of emotion, while the haptic feedback can intensify and enhance this experience.

4.5.3. Recognition Rate by Emotion. The recognition rate was also measured amid the three groups. An accurate recognition is when the participant can correctly identify a clip as having high or low values in arousal and valence. A one-to-nine scale questionnaire was used to rate every clip after its viewing (see Appendix A for the questions and Figure 7 for the users' assessment method). The number of correct emotional recognitions for each subject was counted and then divided by the total to get the recognition rate for each emotional combination (high valence, low valence, high arousal, low arousal). For example, if the clip was of a high valence value and the participant rated the valence factor with a score higher than five, then this was considered as a correct recognition. If the score was lower than or equal to five, then this was an inaccurate recognition.

To statistically analyze the emotion recognition rate, we took into account that there are two clips for each emotion (for instance both "On Time" and "Wanted" clips include high valence values), so the total number of samples for each was 42. We then performed three one-way ANOVA analyses (high valence: $F(2, 123) = 0.18$, n.s.; low valence: $F(2, 123) = 0.05$, n.s.; and high arousal: $F(2, 123) = 0.24$, n.s.) which showed that there are no significant differences between the groups. Because the data for low arousal was not distributed normally, the KruskalWallis test was carried out ($H = 1.48$, $df = 2$, n.s.), which also confirmed that there are no significant differences between the groups in the low arousal case.

Nevertheless, there are some trends that we can discern (see Figure 12): high valence is well recognized in all groups (79%). Low valence is better perceived by the "purposed

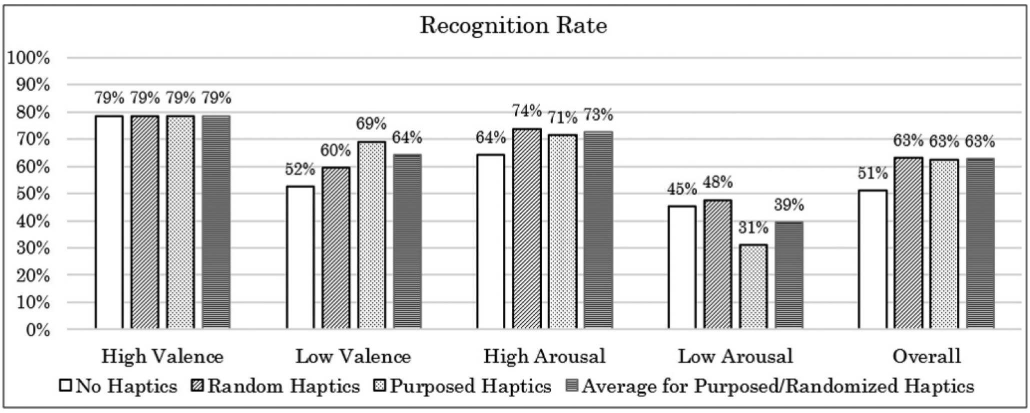

Fig. 12. Recognition rate by emotion.

haptics" group (17% difference with the "no haptics" group). High arousal is better identified by the two haptic groups over the "no haptics" group (the average of the two haptic groups is 9% higher). Low arousal is best discerned by the "no haptics" group since the tactile feedback elicited excitement that conflicted with the low arousal expectation (14% disparity with the "purposed haptics" group). Calculating the average recognition rate of the "purposed haptics" and the "randomized haptics" groups and comparing it to the recognition rate of the "no haptics" group for all four emotional combinations, we can see that the overall recognition rate is higher when there is some form of haptic feedback.

There is an increased recognition rate when the audio/video stimuli are accompanied by haptic feedback for arousal/valence reactions. The only exception is the low arousal sentiment, where the "no haptics" group had a higher recognition rate than the two haptics groups since the haptic feedback induced unwanted excitement. This is probably due to the fact that tactile gestures in general have induced excitement (active response) from the subjects and thus worked against intensifying low arousal. Therefore, eliciting low arousal must not consider the use of any tactile gestures.

4.5.4. Recognition Rate by Clip. Subjects watching the clip "Chatter" had better recognition rates for high arousal when "randomized" or "purposed" tactile gestures were applied, compared to the control scenario where there was no haptic feedback (arousal of 0.76). This was also the case with high valence with "Wanted" (valence of 0.73). This implies that when the audio-visual stimuli elicits clearly recognizable emotional reaction, the tactile gestures further intensify the corresponding emotional reaction. On the other hand, when emotions are less effectively elicited with audio-visual stimuli (like the clip "On Time" with arousal of 0.37 and the clip "The Room of Franz Kafka" with arousal of 0.48), the tactile gestures seem to reduce recognition rate. This implies that when the audio-visual stimuli are not clearly elicited, the tactile gestures might confuse the user and thus do not intensify emotional reactions.

Table VI shows the recognition rate for each clip among the three groups. From the two clips that contain high arousal values ("Chatter" and "Wanted"), we can conclude that haptic feedback increases the recognition rate in this case. On the contrary, when the clips enclose high valence values ("Wanted" and "On Time"), the recognition rate stays on the same level among the groups. In the clips with low arousal ("The Room of Franz Kafka" and "On Time"), we observe that the purposed group had the lowest recognition rate. This is probably because tactile feedback induced excitement and thus lowered the recognition rate with low arousal. And for the clips encompassing low

Table VI. Recognition Rate by Clip and Group

	No Haptics		Randomized		Purposed	
	Valence	Arousal	Valence	Arousal	Valence	Arousal
Chatter	57%	71%	57%	86%	67%	86%
On Time	71%	43%	67%	38%	67%	19%
The Room of Franz Kafka	48%	48%	62%	57%	71%	43%
Wanted	86%	57%	90%	62%	90%	57%

valence values (“Chatter” and “The Room of Franz Kafka”), the haptic stimuli always enhanced the recognition rate of emotions.

5. CONCLUSION AND FUTURE WORK

This article presented a framework to author and render tactile gestures to intensify emotional reactions (arousal and valence) in an immersive film environment. Results show that the overall user experience is enhanced with any kind of haptic feedback. Furthermore, 98.41% of the participants confirmed that the haptic jacket was not a source of distraction for them. The sensory factor increased the QoE when haptic feedback was used regardless of the particular tactile gesture (purposed or randomized). In addition, tactile gestures were not effective to intensify low arousal emotional reactions (around 30% recognition rate for low arousal for the groups with haptic feedback, which is in the vicinity of the chance level).

Future work will investigate further the affective haptic correlation, i.e., why the effectiveness of the purposed group was slightly lower than that of the random group (see Figure 11). The participants in the current experiment underwent only a short training with the haptic jacket before starting their session. We plan to compare the emotional experience between two groups: experienced haptic group and non-experienced haptic group. Future work would test if a group trained with the haptic gestures may result in an improved rating for purposed haptic feedback compared to random haptic feedback. Another future direction is implementing an automatic affective feature extraction (arousal, valence) from the film in real time. Currently, the affective video content from the LIRIS-ACCEDE database is used, where timestamps are recorded and hardcoded for each arousal and valence value. Finally, further analysis to find more effective tactile gestures for intensifying various emotional reactions will be considered. This is particularly important for low arousal since it was poorly recognized by the subjects.

APPENDIX A

In this appendix, we present the questionnaire we used in our experiment study.

For each clip:

(1) How was your emotional experience during the video clip? [Negative to Positive on SAM 9-point scale]
(2) How exciting was the video clip? [Not at all exciting to Very exciting on SAM 9-point scale]
(3) What was the film about? Please describe briefly (maximum 2 lines).

Overall questions:

(1) Watching the movies in the virtual environment felt like a real-world experience.
(2) The overall experience of watching short clips is pleasant.
(3) I would buy this technology if it becomes commercially available.
(4) Did anything distract you from fully diving into the film experience? If so, please explain briefly.

(5) How frequently do you watch movies?
(6) What do you think was the goal of this study? Please explain briefly.
(7) Did the jacket enrich your experience while watching the movie?
(8) Did the particular haptic feedback match the emotion depicted in the series of clips you watched?

REFERENCES

William Albert and Thomas Tullis. 2013. *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics*. Elsevier.

Faisal Arafsha, Kazi Masudul Alam, and Abdulmotaleb El Saddik. 2015. Design and development of a user centric affective haptic jacket. *Multimedia Tools and Applications* 74, 9 (2015), 3035–3052.

Yoann Baveye, Emmanuel Dellandrea, Christel Chamaret, and Liming Chen. 2015. LIRIS-ACCEDE: A video database for affective content analysis. *IEEE Transactions on Affective Computing* 6, 1 (2015), 43–55.

Leonardo Bonanni, Cati Vaucelle, Jeff Lieberman, and Orit Zuckerman. 2006. TapTap: A haptic wearable for asynchronous distributed touch therapy. In *CHI'06 Extended Abstracts on Human Factors in Computing Systems*. ACM, 580–585.

Margaret M. Bradley and Peter J. Lang. 1994. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry* 25, 1 (1994), 49–59.

Jongeun Cha, Mohamad Eid, Ahmad Barghout, A. S. M. Rahman, and Abdulmotaleb El Saddik. 2009. HugMe: Synchronous haptic teleconferencing. In *Proceedings of the 17th ACM International Conference on Multimedia*. ACM, 1135–1136.

Fabien Danieau, Anatole Lécuyer, Philippe Guillotel, Julien Fleureau, Nicolas Mollet, and Marc Christie. 2013. Enhancing audiovisual experience with haptic feedback: A survey on HAV. *IEEE Transactions on Haptics* 6, 2 (2013), 193–205.

Mohamad A. Eid and Hussein Al Osman. 2016. Affective haptics: Current research and future directions. *IEEE Access* 4 (2016), 26–40.

Paul Ekman and Wallace V. Friesen. 1971. Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology* 17, 2 (1971), 124.

Abdelwahab Hamam, Mohamad Eid, and Abdulmotaleb El Saddik. 2013. Effect of kinesthetic and tactile haptic feedback on the quality of experience of edutainment applications. *Multimedia Tools and Applications* 2, 67 (2013), 455–472.

Immersion Corporation. 2016. Immersion Corporation Website. Retrieved from http://ir.immersion.com.

Ali Israr and Ivan Poupyrev. 2011. Control space of apparent haptic motion. In *World Haptics Conference (WHC), 2011 IEEE*. IEEE, 457–462.

Ali Israr, Siyan Zhao, Kaitlyn Schwalje, Roberta Klatzky, and Jill Lehman. 2014. Feel effects: Enriching storytelling with haptic feedback. *ACM Transactions on Applied Perception (TAP)* 11, 3 (2014), 11.

Sreekar Krishna, Shantanu Bala, Troy McDaniel, Stephen McGuire, and Sethuraman Panchanathan. 2010. VibroGlove: An assistive technology aid for conveying facial expressions. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems, April* 10–15, 2010. 3637–3642.

Victor V. Kryssanov, Eric W. Cooper, Hitoshi Ogawa, and I. Kurose. 2009. A computational model to relay emotions with tactile stimuli. In *Proceedings of the 2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*. IEEE, 1–6.

Paul Lemmens, Floris Crompvoets, Dirk Brokken, Jack Van Den Eerenbeemd, and Gert-Jan de Vries. 2009. A body-conforming tactile jacket to enrich movie viewing. In *Proceedings of the EuroHaptics Conference, 2009 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics 2009. Third Joint*. IEEE, 7–12.

Rodrigo Lentini, Beatrice Ionascu, Friederike A. Eyssel, Scandar Copti, and Mohamad Eid. 2016. Authoring tactile gestures: Case study for emotion stimulation. *World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering* 10, 11 (2016), 1862–1867.

Antonella Mazzoni and Nick Bryan-Kinns. 2015. How does it feel like? An exploratory study of a prototype system to convey emotion through haptic wearable devices. In *Proceedings of the 2015 7th International Conference on Intelligent Technologies for Interactive Entertainment (INTETAIN'15)*. IEEE, 64–68.

Lauri Nummenmaa, Enrico Glerean, Riitta Hari, and Jari K. Hietanen. 2014. Bodily maps of emotions. *Proceedings of the National Academy of Sciences* 111, 2 (2014), 646–651.

Jonathan Posner, James A. Russell, and Bradley S. Peterson. 2005. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology* 17, 03 (2005), 715–734.

Munaf Rashid, S. A. R. Abu-Bakar, and Musa Mokji. 2013. Human emotion recognition from videos using spatio-temporal and audio features. *The Visual Computer* 29, 12 (2013), 1269–1275.

Miriam Reiner. 2004. The role of haptics in immersive telecommunication environments. *IEEE Transactions on Circuits and Systems for Video Technology* 14, 3 (2004), 392–401.

Katri Salminen, Veikko Surakka, Jani Lylykangas, Jukka Raisamo, Rami Saarinen, Roope Raisamo, Jussi Rantala, and Grigori Evreinov. 2008. Emotional and behavioral responses to haptic stimulation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1555–1562.

Alexandre Schaefer, Frédéric Nils, Xavier Sanchez, and Pierre Philippot. 2010. Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers. *Cognition and Emotion* 24, 7 (2010), 1153–1172.

Kai Sun, Junqing Yu, Yue Huang, and Xiaoqiang Hu. 2009. An improved valence-arousal emotion space for video affective content representation and recognition. In *Proceedings of the 2009 IEEE International Conference on Multimedia and Expo*. IEEE, 566–569.

Amaya Becvar Weddle and Hua Yu. 2013. How does audio-haptic enhancement influence emotional response to mobile media? In *Proceedings of the 2013 5th International Workshop on Quality of Multimedia Experience (QoMEX)*. IEEE, 158–163.

Bob G. Witmer and Michael J. Singer. 1998. Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments* 7, 3 (1998), 225–240.

Received October 2016; revised May 2017; accepted May 2017