Vibrotactile TV for Immersive Experience

Shafiq ur Réhman[*§], MSL Khan[*], Liu Li[††] and Haibo Li[†,††]
[*]i2lab, Department of Applied Physics and Electronics, Umeå University, SE-90187Umeå, Sweden.
[†]Media technology and interaction design, Royal Institute of Technology (KTH), Stockholm, Sweden.
[††] Nanjing University of Posts and Telecommunications, Nanjing 210003, China.

***Abstract*— Audio and video are two powerful media forms to shorten the distance between audience/viewer and actors or players in the TV and films. The recent research shows that today people are using more and more multimedia contents on mobile devices, such as tablets and smartphones. Therefore, an important question emerges - how can we render high-quality, personal immersive experiences to consumers on these systems? To give audience an immersive engagement that differs from 'watching a play', we have designed a study to render complete immersive media which include the 'emotional information' based on augmented vibrotactile-coding on the back of the user along with audio-video signal. The reported emotional responses to videos viewed with and without haptic enhancement, show that participants exhibited an increased emotional response to media with haptic enhancement. Overall, these studies suggest that the effectiveness of our approach and using a multisensory approach increase immersion and user satisfaction.**

I. INTRODUCTION

The recent trends in human computer interaction (HCI) strongly suggest that future computer interfaces and communications will no longer be bound to tradition input-output methods. Today innovative technologies have enabled the user to rethink the interaction scenario and consider more natural and intuitive methods for increased functionality and realism. Video allows great authenticity and realism, and it is becoming ubiquitous, in personal capturing and display devices, on the Internet and iTV. That is why researchers continue to improve the technologies and try to deliver more realistic audio and video (2D to 3D) experience. Immersion can be defined as the subjective experience of being fully involved in an environment and may be increased by a surround effect, sensory modalities and vividness through resolution. Today, consumer electronics have built-in technology for 2D-to- 3D image conversion and spatial sound rendering methods for introducing immersion. However, these systems are primarily limited to the stimulation of two senses, sight and hearing in order to simulate an immersive media experience. The way, we as humans interact with each other, is to engage almost all the senses, thus seeing, hearing, and touch. The viewer can be engaged at emotional level by rendering rich information and engaging several senses. The importance of tactile in communication is evident by recent ISO and MPEG-V [1, 2] standards. It is evident that haptics will be a prominent parts in various applications to provide immersiveness [3, 4, 5, 6, 10, 11].

[§]**Corresponding Author**: shafiq.urrehman@umu.se,
http://www.tfe.umu.se/forskning/i2lab/

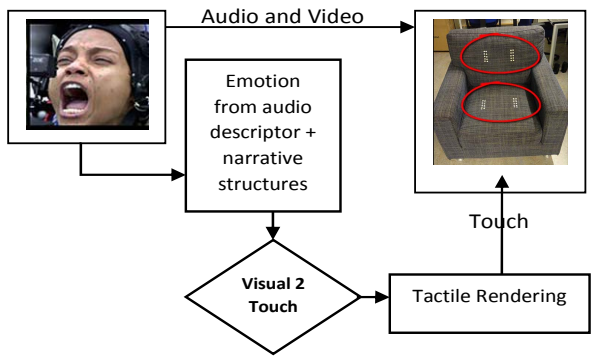


Fig. 1: Functional model for multimedia contents broadcasting with synchronized audio, video and haptic feelings.

In this work, we study the immersion potential of audio-video contents augmented with vibrotactile coding in handheld devices.

We consider an important question - *how can we render high quality, personal immersive experiences to consumers along with the traditional media*? The main goal of this investigation is to determine whether haptic (vibrotactile) augmented effects, rendered through vibration actuators attached to smart devices (phones and/or tablets) could provide an enhanced level of emotional impact and immersiveness to media contents.

To investigate the affective impression of vibrotactile rendering of multimedia contents, we organize this paper as following. Section 2 describes a pilot study that served as a preliminary step to categorize appropriate vibration patterns that are clearly distinguishable to elicit associations to functions that occur in the interaction with movie scenes. Based on our previous studies, a set of 12 vibration-tactors was used for this study described in Section 3. In this study the vibration patterns were rated by participants in order to examine their applicability and affective impression.

II. PILOT STUDY

In this work, we try to include both functionality and emotional impact of auditory, visual and tactile feedback. We have considered the effect of '*Affective Haptics*'; i.e., how multimedia systems affect a human's emotional condition by including the sense of touch [4]. To construct this research we manually extract 'emotions' from 'audio descriptor

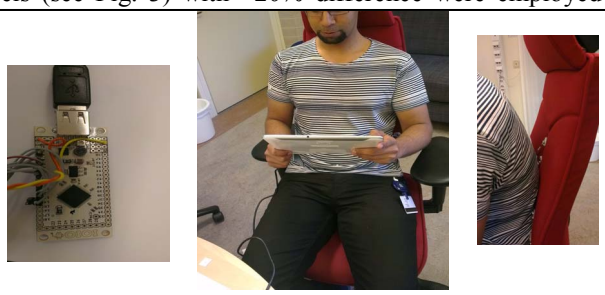
depicting emotional words' and narrative structure as in [7, 8]. Audio description (AD) refers to what is depicted on-screen at, or near to, the moment it is spoken and is provided for the visually impaired with films in some cinemas and on VHS/DVD releases; e.g. [10:35] Kim grows *alarmed*.

AD contains an additional insight into a character's emotional state which is given by audio description when the emotion is being depicted visually in the film. Similarly in film narrative relates to how audiences are engaged and how to provide emotion features as a connection between two physical events. For this work we manually extracted visibly manifested emotions as 'E-token' in time-coded audio description and classified them into six types (Joy, distress, like, dislike, hope, fear) as proposed by [9] and converted into three dimensions for this study.

A. Prototype Hardware Used

To carry out a user test, a vibrotactile mobile system was built. The prototype consisted of an IOIO device connected to the hand held device using a USB cable. This system was designed and was programmed to generate Pulse Width Modulation (PWM) at one of the pins and send signals to the vibration motor.

The extracted video emotion semantics (E-Tokens) from video Audio description contents are employed to generate vibrotactile signals and then sent them to the IOIO device (see Fig. 2). By controlling the IOIO's output to switch between on and off, the vibration motors (attached to the chair) could rotate at different frequencies and intensities. The vibrator was a low-cost coin motor used in pagers and cellular phone. The vibration frequencies had been selected in a range from 100 Hz to 1 kHz. To overcome the limitation such as, *stimuli masking* and *adaptation effect* and to enhance the perception of vibrating stimuli, 4 different inter-switched vibration signal levels (see Fig. 3) with ~20% difference were employed as proposed in [10]. Each vibration signal was used to present a specific 'emotion' based on E-Token.



a- Programmable IOIO device.

b- User with tablet-PC.

c- Vibration motor mounted on the chair.

Fig. 2: Application scenario: Rendering '*Affective Haptics'* from a video greatly engages the audiences into the contents and provides an immersive experience. For example, Audio/video contents are analyzed and audio/video/haptic contents are render to viewers through vibrations; a) Here '*Affective Haptics'* coding is programed in IOIO device and b) rendered on the back of the user.

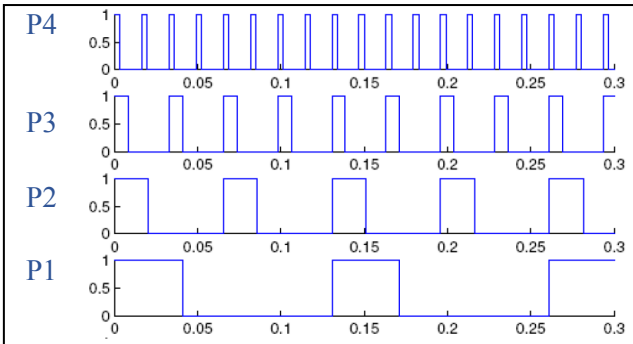


Fig. 3: Vibrotactile patterns used for this study.

For this study, we used a neutral laboratory setup in a sound-isolated test cabinet without any disturbing material. The test devices used in this study was GALAXY Tab 10.1 running Android's Ice Cream Sandwich OS and a custom build application which had E-Token enabled videos.

B. Participants and Stimuli

Total 10 participants aged between 24 and 37 (M = 29.40, SD = 4.15), three of them female, took part in this study. Participants were recruited from the campus of Umeå University Sweden, including both students and staff members. All the participants had mobile phones and had prior experience with mobile vibration. The test videos with E-Tokens were used as stimuli. Test videos contained action, thriller and horror movie scenes (filmed from a first-person perspective). The videos selected as test stimuli demonstrated a range of haptic effects that augmented audio-visual events. These audio-video contents augmented with '*Affective Haptics'* had all the required emotions which were tagged off-line. Each 'emotion' was tracked frame by frame. Time stamp and emotion stamp of scene were used for the vibrotactile stimuli. The motivation of our user test had been clearly explained to all the participants.

C. Procedure

First, we introduced the purpose of our experiments to the participants and explained how to use the system. Each participant held the tablet-PC and sat on chair mounted with the vibration motor (see Fig. 2). The study included two onsite sessions. Each video that had been viewed on the control device in the first session was subsequently viewed in the second session on a haptic enabled device, and vice versa. Before a test session, participants were first trained on how to use the devices and make subjective ratings using the survey instruments. When the participants finished the first round and started to perform the second one, they were considered to be more experienced. However, due to different stimuli, this did not affect the study of the effect of training process in the experiments.

The affective impression of all presented vibrotactile patterns was measured with the Self-Assessment Manikin (SAM) [12]. The affective impression of the stimuli is measured on the three affective dimensions of:

- *Valence:* ranging from a happily smiling face for a pleasant stimulus to an unhappily frowning face for an unpleasant stimulus.
- *Arousal*: from excited figure with eyes wide open (aroused) and a relaxed and sleepy figure (calm).
- *Dominance*: a manikin of increasing size is shown for high dominance values the manikin covers almost the complete frame and decreasing size for low dominance.

Additionally, to measure immersive levels from '*user satisfaction'* each participant was given a questionnaire in the end of experiment sessions. Each question was elaborated and discussed with the participants if required. The session had 30 minutes time duration approximately with 15-minutes time between each session for resting.

All the users were provided with a two-minute long training session before staring actual experiment. Before the experiment-session started, a short video of a train journey was shown to the participants, in order to adjust the mood of the participants to a fairly neutral state. The participant were presented with multimedia stimuli with and without '*Affective Haptics'* in random order; i.e. some were presented stimuli augmented with *vibrotactile coding* first session and then just simple audio-video stimuli in second session. SAM and functionality ratings for each session were collected directly thereafter. At the end of the second session, participants were engaged in informal interview by a moderator about their experience. Participants were asked to recall salient events from the study, and comment on their experience with and without 'Affective Haptics' enhanced multimedia contents.

TABLE I
MEAN RATINGS AND STANDARD DEVIATIONS (SD) OF THE THREE SAM SCALES

Affective Haptics	SAM	Mean	SD
With	Arousal	4.28	0.54
	Valence	2.88	0.84
	Dominance	3.68	0.62
Without	Arousal	2.03	1.05
	Valence	2.05	0.99
	Dominance	1.98	1.01

III. REUSLTS

This section describes the results on the collected ratings on the affective impression and user satisfaction. We used both direct measures, as well as self-reports from interview, in order to characterize participants' response.

A. *Affective Impression*

According to collected data, high mean valence and arousal ratings were given to pleasant stimuli with vibrotactile coding. There was significant increase in overall positive affect with augmented 'Affective Haptics' on multimedia experience on hand held devices.ie., with 'Affective Haptics' M=3.9, SD= 0.85 as compared to without M= 2.53, SD= 0.59. However very slight change was observed in negative affect with and without augmented 'Affective Haptics'; i.e., with M= 1.89, SD= 0.39 and without M= 2.07, SD= 0.49.

Table 1 shows mean ratings on the three SAM scales collected after each experimental session. It can be seen that the multimedia content is rated most pleasant and least arousing with the lowest effort. It can be concluded that vibrotactile feedback stimuli, are able to evoke an affective impression.

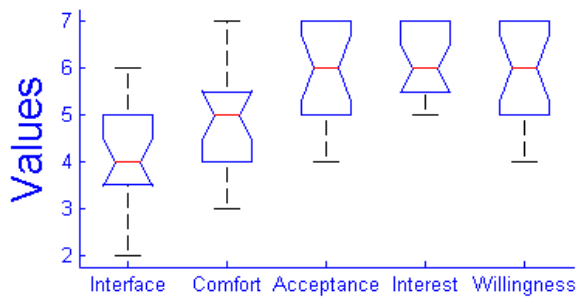


Fig. 4: Mean Questionnaire Score from User Satisfaction

B. *User Satisfaction*

Satisfaction refers to the comfort and acceptability of the system for its users and other people affected by its use [13]. A useful measure of user satisfaction can be made if the evaluation team's measure is based on observations of user attitudes towards the system. Thus, it is possible to measure user attitudes using a questionnaire, e.g., "is this application interesting? 1=very boring (negative) to 7=very interesting (positive)." The satisfaction levels of our experiments were measured by giving participants questionnaires after sessions. Its values represented strong disagreement (negative) and strong agreement (positive), respectively. Fig. 4 shows mean questionnaire responses to questionnaire questions. Each label on the x-axis in the figure represents a question where we used the following notation:

i. *Interface* - Is the human computer interface easy to use?
ii. *Comfort* - Is this application comfortable to use?
iii. *Acceptance* - Is this application acceptable?
iv. *Interest* - Is this application interesting?
v. *Willingness* - Are you willing to buy such a service at a cost of 199 SEK?

Experimental results indicated a high interest level for our application, mean score of 6.25. Participants gave an average score of 5.95 to acceptance and 4.25 was given by participants

to interface plus GUI. It show that more work is needed to be done on GUI and interface (hardware). An average rate of 5.758 on willingness to buy a service if available in market.

C. General results

According to the collected data, *Affective Haptics* increases the level of immersiveness and involvement participants experienced with the audio-video content. Participants in the study also tended to rate the same video content higher on a subjective quality of experience scale when it was accompanied by Affective Haptics enhancement. The main result of this study suggested by the presented analysis, is that participants exhibited an increased emotional response to media in presence of vibrotactile stimuli. The participants' affective responses to videos with and without vibrotactile enhancement are significantly different in both the valence and arousal dimensions. On average, participants reported significantly different responses to tactile-coded videos as compared to others. The user satisfaction data measures indicate that videos with vibrotactile enhancement generated more positive emotional response.

There are some limitation of this study as participants experienced video content in a lab situation and were motivated users; i.e., they were told in advance about experiencing enhanced media and making ratings. The user's response in a naturalistic context might be different. The participant have many suggestion and comments about the user interface as well.

IV. Concluding Remarks and Future Direction

This paper describes a preliminary studies investigating vibrotactile augmented audio-video contents on mobile devices. The presented study, conducted in controlled environment and evaluated by motivated participants with vibrotactile coding-enhanced mobile video experiences, provides evidence that haptic feedback can augment the experience of watching mobile videos, and intensify their emotional immersiveness. Based on the overall trends measured here, as well as participant's comments, we may conclude that more carefully designed haptic augmentation might achieve further emotional impacts.

Though this study, it is pointed out that a specific user experience benefit for tactile augmentation but there is need of automatic algorithm to generate haptic effects (E-tokens), instead of manual approaches. Current, ongoing work suggests that manually designed haptic content has a significantly stronger impact on user emotion and overall ratings. As a future work we will improve both hardware and software implementations as well as consider further studies in various context such as home environment, office environment and outdoor environment. We intend to conduct interaction experiments with focus on these aspects in the future. Usability measures will also be considered in order to complement the affective impression ratings.

References

[1] ISO 9241-920:2009, "Ergonomics of human-system interaction -- Part 920: Guidance on tactile and haptic interactions".

[2] ISO/IEC IS 23005-3 Sensory Information

[3] A. El Saddik, M. Orozco, M. Eid, J. Cha. Haptics Technologies, "Bringing Touch to Multimedia". Springer Series on Touch and Haptic Systems, 2011

[4] D. Tsetserukou, and A. Neviarouskaya, "iFeel_IM!: Augmenting Emotions during Online Communication," IEEE Computer Graphics and Applications, Vol. 30 Issue:5, pp. 72-80, Sept.-Oct. 2010.

[5] Paul M.C. Lemmens, Dirk Brokken, Floris M.H. Crompvoets, Jack van den Eerenbeemd, Gert-Jan de Vries. "Tactile Experiences". EuroHaptics 2010 (2010): 11.

[6] A. El Saddik, M. Orozco, M. Eid, and J. Cha, "Haptics Technologies: Bringing Touch to Multimedia," ISBN 978-3-642-22657-1, Springer-Verlag, 2011.

[7] TIWO – Television in Words, EPSRC research project, www.computing.surrey.ac.uk/ckm/tiwo_project/Index.html

[8] A. Salway and M Graham, "Extracting information about emotions in films". In ACM international conference on Multimedia (MULTIMEDIA '03), 2003.

[9] A. Ortony, G.L. Clore, and A. Collins, "The Cognitive Structure of Emotions". Cambridge University Press, 1988.

[10] S. ur Réhman and Li Liu, "iFeeling: Vibrotactile Rendering of Human Emotions on Mobile Phones", In Xiaoyi Jiang, Mathew Y. Ma, Chang Wen Chen (Eds) Mobile Multimedia Processing: Fundamentals, Methods, and Applications by Springer in Mobile Multimedia Processing : Fundamentals, Methods, and Applications (LNCS 5960: State-of-the-Art Surveys), pp. 1-20, 2010.

[11] S. ur Réhman, J. Sun, Li Liu and Haibo Li, "Turn your mobile into the football: rendering live football game by vibration", IEEE Transactions on Multimedia, vol. 10(6), pp.1022-1033, 2008.

[12] M. M. Bradley, and P. J. Lang, "Measuring emotion: The self-assessment manikin and the semantic differential", J. Behav. Ther. & Exp. Psychiar., 25(1), 49 -59,1994.

[13] Faulkner, X.: Usability Engineering. Palgrave Macmillan (2002)