

INVERSE AND TRANSITIVITY OF CROSS-MODAL CORRESPONDENCE IN MULSEMEDIA

Gebremariam Mesfin, Nadia Hussain, Alexandra Covaci, Gheorghita Ghinea

Department of Computer Science, Brunel University, London, UK
{gebremariam.assres; nadia.hussain; alexandra.covaci; george.ghinea}@brunel.ac.uk

ABSTRACT

The human brain has evolved to learn, adapt and operate optimally in multisensory settings. Exploring the multisensory-phenomena in a human computer interaction (HCI) context could help design interfaces and displays that tap into users' mental models. In this paper we explore the influence of crossmodal correspondences on the Quality of Experience (QoE) in a mulsemedia setup. A mulsemedia enhanced test-bed was developed to perform delivery of video enhanced with audio and haptic effects that were designed considering crossmodal congruencies principles. Accordingly, we performed an experiment on the correspondence from the visual features of selected videos into high/low pitch audio when they are experienced in conjunction with auto-generated haptic effects. Our results showed good insights into inverse and transitive cross-modal correspondences. In addition, the audio signal generated out of the angular shape (visual feature) and its corresponding haptic effect is found to have enhanced the viewers' quality of experience.

Index Terms – Mulsemedia, video, audio, haptic, crossmodal correspondence quality of experience

1. INTRODUCTION

We rarely experience senses in isolation. Our perceptual experience about the real world events is the result of different sensory signals. The synergy among different sensory modalities and the fusion of their information lead to multisensory integration. Multi-sensorial media (mulsemedia) represents media types that relate to human multiple sensory experience of the real world; and adds olfaction, haptic, gustatory, etc. into the audio/visual media in the digital world [1]. Mulsemedia is an active research area, which explores various issues related to the nature of the various senses; digital representation of the sensing, storage and emission; and on its impact on quality of experience [2-4].

Murray et al. [3] pointed out that the motivation for adding more media types into mulsemedia is to increase Quality of Experience (QoE). On the other hand, such additional media types can be incorporated into mulsemedia through crossmodal congruencies [5-7] that refer to correspondences between audio and haptic stimuli, and high-pitched tones are matched with yellow colour [8]. However, the inverse and transitivity characteristic in cross-modal association remains unexplored. Moreover, whilst

correspondences have been explored in the non-digital world, not the same can be said as far as the case of digital mulsemedia is concerned.

Accordingly, the study reported in this paper was focused on the cross-modal association between selected visual features of videos – colours, brightness and shape; and high/low audio pitches. In addition, the transitivity of visual stimuli (from the visual features of the videos) into haptic effects via the high/low-pitch audio mapping, and its impact on the overall QoE is given due emphasis. Thus, self-reported viewing experience of the participants' (divided into experimental and control) was captured for each video; and the cross-modal correspondence was analysed.

The paper is organized as follows. Related work is presented in section 2; methodology and results are provided in sections 3, and 4 respectively; finally, section 5 provides the conclusion.

2. RELATED WORK

In this section, we describe the related work on mulsemedia, quality of experience, and cross-modal correspondence.

2.1. Mulsemedia

Mulsemedia represents media types that relate to human-conscious multiple sensory experience of the real world objects and phenomena [1]. It incorporates additional media types for each of the senses of smell (olfaction), touch (haptic), taste (gustatory), heat (thermoception), and balance (equilibrioception) to the conventional audio/visual media.

As pointed out in [9-11], humans understand and assimilate the meaning of mulsemedia experiences through the process of capturing, interpreting, and combining information from numerous sensory organs in a bottom-up sensing. Thus, mulsemedia technologies targeting the rendering of media effects that stimulate one of these senses are introduced, and the sensory effects of each of these technologies need to be integrated with each other in order to simulate the multi sensorial nature of the world.

In this paper, we concentrate on mulsemedia experiences involving the senses of sight, hearing, and touch, which are briefly presented next.

2.1.1. Visual Sense

The sense of sight enables assimilation of textual, image, video, and animation information [1]. Such visual information is perceived when the light reflected from an object in the visual field enters the eye through the pupil and

passes through the lens, which projects an inverted image onto the retina at the back of the eye.

2.1.2. Auditory Sense

The sense of hearing is another human sensory system which enables the capture of audio signals such as sound, speech, and music [1]. Such audio signals are produced by a sequence of wave compressions in the air surrounding a vibrating source.

2.1.3. Haptic

The sense of touch (haptic) is among the most powerful means to communicate emotions [2]. Nowadays, haptics is one of the state-of-the-art media types added into the conventional audio/visual media, which can be used together to convey mulsemmedia experiences [1].

2.2. Quality of Experience

Le Callet et al. [12] describe QoE as the degree of delight or annoyance of the user of an application or service. Murray et al. [3] also affirmed that the motivation for pursuing more mulsemmedia components is to increase the level of user immersion and/or QoE. Thus, QoE is a fundamental measure on the impact of mulsemmedia on the fulfilment of a user's expectations on the utility and or enjoyment of the application or service. For example, studies [13, 14] pointed out that multimedia sequences integrated with olfactory content can partly mask a decreased movie quality and enhance the user's perceived QoE.

In light of the above, although it is widely accepted that the emergence of new technologies has a positive influence, gaining a higher level of QoE has never been straightforward. However, studies in various application domains such as in [4, 14] confirmed that significant QoE improvements can be made by integrating more mulsemmedia components in digital content.

Overall, engaging more mulsemmedia components enhances QoE by facilitating memory [15]. For example, audio/tactile cueing improves the speed and accuracy of tasks and reduces the amount of mental workload [16].

2.3. Cross-modal correspondence

Crossmodal correspondence describes the interactions between two or more different sensory modalities. It is a process that underlies synaesthesia [6, 7], and sensory substitution [5].

Haptic effects can be automatically generated from the audio content by transducing an audible signal into a signal suitable for vibration motors. Haptic devices (e.g., gaming vest) use a band-pass filter to isolate frequencies compatible with a targeted vibration motor and then amplify the output signal (haptic effects) on the device [17]. Most research applies this relatively straightforward technique to convert audio into vibrations. However, audio analysis techniques would be useful to extract specific features (e.g., frequency) to generate the desired haptic effect.

In the realm of crossmodal correspondence, positive correlations were identified between audio and visual features (e.g. colour, brightness, and shape) [18-20]. For example, high-pitched tones are matched with yellow colour, brighter surfaces [21], and angular shapes [19], while low-pitched tones match blue [8], dark surfaces [21], and rounded shapes [19].

In general, the existing literature has provided significant insights into crossmodal correspondences, especially in a non-digital context. However, some of the issues requiring further research are presented next.

Synaesthetic experiences such as the perception of colours from hearing audios are unidirectional and non-transitive [22]. In this respect, the inverse and transitive characteristic in cross-modal association remain unexplored. Thus, for each modality - audio pitch, visual feature, and haptic effect, we need to explore that:

- If there exists an association from audio pitch to visual features, and if the association from visual features to audio pitch also holds true (inverse relationship).
- If there exists an association from visual features to audio pitch and from audio pitch to haptic effect, whether then the association from visual feature to haptic effect also holds true (transitive relationship).

In addition, existing crossmodal correspondence research such as in [19, 23] are conducted based on image and text samples. Thus, the impact of using video samples (instead of images) and understanding the cross-modal correspondence as QoE is unexplored – and is what we undertook in the study which is presented next.

3. METHODOLOGY

3.1. Participants

We recruited 15 male and 9 female (24 in total) participants aged between 18-41 years and from a range of diverse backgrounds, nationalities, and education. All participants spoke English and were computer literate. They watched the

Table 1. The experimental videos

					
V1: Beach Scene, blue waves lapping on the shore	V2: Yellow sulphur springs, Danakil Desert, Ethiopia, smooth sound	V3: Solar Eclipse, sky turns dark and the moon appears, birds chirping in background	V4: Bright sun shining upon the Arctic with bright snow, smooth sound	V5: Angular Skyscrapers, smooth sound	V6: Bouncing Balls, smooth sound

six videos in two equal-sized groups - experimental and control – to which participants were randomly allocated.

3.2. Experimental Material

3.2.1. Video clips

Sample videos are selected based on the dominant features – blue, yellow, dark dominated, mostly bright, angular shaped objects, and rounded objects (see Table 1). Each of the 1366X768 pixel resolution videos were edited to 120 seconds long excerpts. The playback rate was 30 frames per second. These six sample videos, their original audio, and audio altered for high pitch (328Hz) and low pitch (41Hz) are employed in the experiment. The altered audio crossmodally matched the video content in terms of either colour (blue/yellow) or content (rounded/angular) [8], [19], [21].

3.2.2. Devices

We made use of the KOR-FX¹ gaming vest as a haptic device. Participants listened to the audio soundtrack through a pair of BOSE noise cancelling headphones.

3.2.3. Questionnaire

A set of five-point Likert scale questions were programmatically presented at the end of viewing each video as shown in Table 2.

Table 2. Questionnaire

Ref.	Question detail
Q1	I enjoyed watching the video clip whilst wearing a haptic vest.
Q2	The haptic vest effects were relevant to the video clip I was watching.
Q3	The vibration was annoying.
Q4	The haptic vest effects enhanced the sense of reality whilst watching the video clip.
Q5	The haptic vest effects enhanced my viewing experience.

3.3. Experimental setting

Our current experiment was focused on the crossmodal correspondence between visual features in video clips, audio pitch, haptic effect, and their impact on user QoE. It took place at Brunel University in a quiet room; where the actual time of the experiment lasted between 30-40 minutes per participant.

The factors, levels and possible responses corresponding to our experiment are as follows. Colour, brightness and shape of objects in videos are the factors as shown in the Fishbone² chart in Fig. 1. Levels, which describe value settings of each factor are - blue and yellow for colour, dark and bright for brightness, and angular and round for shape.

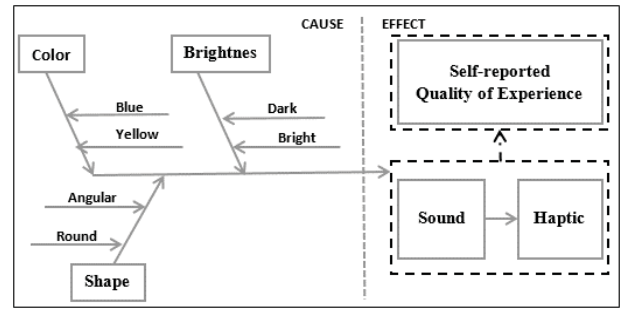


Fig. 1. Fishbone representation of factors

The self-reported responses, on the other hand, focused on identifying the association of the factors on audio and haptic effects and their impact on participants' QoE.

The study was conducted in experimental and control groups. The experimental and control groups watched videos with plain (altered) audio and original audio, respectively. In the former case, the videos are combined with high pitch (328Hz) or low pitch (41Hz) audios for each of the visual feature as shown in Table 3; which in turn generates the haptic stimulus. Thus, the senses of hearing and touch are generated from the visual features so as to render multisensory output. Table 3 shows our experimental setting which was performed by mapping each of the features (factors) of the videos and audio effects.

Table 3. Experimental setting

Factor	Video	Audio effects	Haptic effect
Color			
Blue	V1	Low pitch	Yes
Yellow	V2	High pitch	Yes
Brightness			
Dark	V3	Low pitch	Yes
Bright	V4	High pitch	Yes
Shape			
Angular	V5	High pitch	Yes
Round	V6	Low pitch	Yes

3.4. Experimental Process

Before proceeding with the experiments, a pilot study was carried out with two participants. We wanted to know whether we need to add or make changes in relation to the experimental setting. Hence, the participants reflected that they felt the high pitch audio was distracting. Consequently, we made a slight change by lowering the volume for the high pitch to make the participant feel more comfortable.

At the outset of the experiment, each participant was asked to put on the haptic vest and confirm s/he felt comfortable wearing it. Once this was accomplished, each participant viewed the videos in a random order so that order effects were minimised. We assisted participants with any queries they might have had whilst playing the video clips. At the end of the presentation of each video clip, the

¹ <http://www.korfx.com>

² <https://www.moresteam.com/toolbox/fishbone-igram.cfm>

participants' experience was captured using the set of questions given in Table 1.

4. RESULTS

4.1. Data analysis

Our analysis of the participants' reflections about their experience when viewing each video is performed. The analysis was conducted from the perspective of the high versus low audio pitch, original versus altered audio, as well as the overall QoE corresponding to each visual feature (in videos). The score of each negatively-phrased questions was converted by calculating six minus the recorded score (6 - score). Thus, the result is presented as follows.

4.1.1. High-pitch versus low-pitch audios

The result described below correspond to the groups of responses defined by the videos with high-pitch and low-pitch altered audio. Here, the group statistics, and the independent samples t-test result is shown in Table 4 and Table 5, respectively.

Table 4. Group Statistics for high-pitch and low-pitch audios

Grouping_2	N	Mean	Std. Deviation	Std. Error Mean
Score_2 High-pitch	180	3.24	1.279	.095
Low-pitch	180	3.38	1.140	.085

In

Table 4, there are 180 responses (5 questions by 6 participants for 6 videos) for each of the high-pitch and low-pitch groups. The mean response (out of five scale) is 3.24, and 3.38 for high-pitch and low-pitch, respectively.

Table 5. Independent Samples Test for high-pitch and low-pitch audios

	Levene's Test for Equality of Variances		t-test for Equality of Means		
	F	Sig.	t	df	Sig. (2-tailed)
Equal variances assumed	4.303	.039	-1.131	358	.259
Equal variances not assumed			-1.131	353.362	.259

	t-test for Equality of Means			
	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
			Lower	Upper
Equal variances assumed	-.144	.128	-.396	.107
Equal variances not assumed	-.144	.128	-.396	.107

As can be seen in Table 5, we undertook an independent samples t-test and results showed $t=-1.131$; $p=0.039$ highlighting statistically significant differences between the two groups at the 5% significance level.

4.1.2. Original versus altered audios

Similarly, the result of the data corresponding to the groups of responses defined by the videos with original and altered audios is described next.

Table 6 shows the mean and standard deviation of the groups who viewed with original and altered audios. Here, there are 360 responses (5 questions by 12 participants for 6 videos) for each of the original-audio and altered-audio groups; and the mean response is 3.44, and 3.31, respectively.

Table 6. Group Statistics for original and altered audios

Grouping_1	N	Mean	Std. Deviation	Std. Error Mean
Score_1 Original audio	360	3.44	.865	.046
Altered audio	360	3.31	1.212	.064

In Table 7, we consider the equal variances not assumed row because the standard deviation of the groups in Table 6 shows different values ($0.865 \neq 1.212$). Thus, the p-value in the t-test result in Table 7 is 0.104 (sig. 2-tailed $\geq .05$). This shows that at the 5% significance level, the difference in mean scores between the groups is statistically insignificant. In addition, the mean difference value (-0.128) is within the confidence interval of the difference indicating that the upper and lower value alternates between the groups.

Table 7. Independent Samples Test for original and altered audios

	Levene's Test for Equality of Variances		t-test for Equality of Means		
	F	Sig.	t	df	Sig. (2-tailed)
Equal variances assumed	70.526	.000	1.628	718	.104
Equal variances not assumed			1.628	649.473	.104

	t-test for Equality of Means			
	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
			Lower	Upper
Equal variances assumed	.128	.078	-.026	.282
Equal variances not assumed	.128	.078	-.026	.282

4.1.3. Analysis of responses of each question and video
 Analysis of the responses from respondents in the control and experimental groups corresponding to the specific questions and videos is also presented. Accordingly, the mean, standard deviation, and mean difference (t and p values) of responses to each question (and to each video) of experimental and control groups is described as follows.

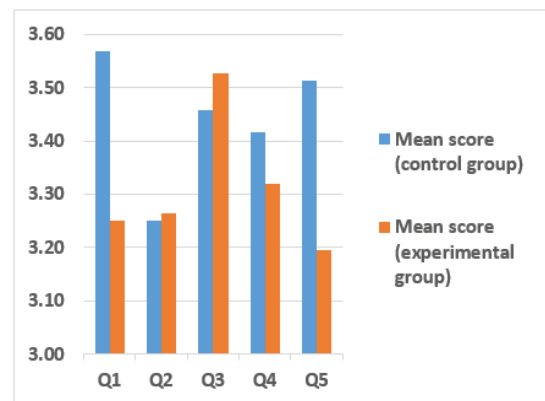


Fig. 2. Mean scores of the groups for each question

The graph in Fig. 2 depicts the mean scores of the control and experimental groups for each question. Here, it shows that the participants' average responses (in the experimental and control groups) to each question (for all of the videos) is generally above the neutral score (3.0).

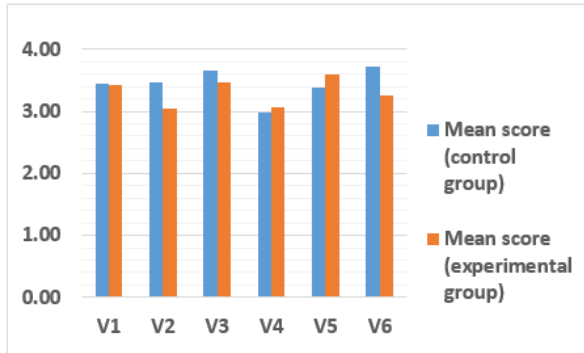


Fig. 3. Mean scores of the groups for each video

Positive responses (average score ≥ 3.5) are also registered to most of the questions for all videos except V4 by the control group as shown in Fig. 3. In addition, Fig. 3 shows that the experimental group's average score of the responses to most of the questions for V1, V3 and V5 is greater or equal to 3.5.

Table 8. T and P values of the mean difference of the groups' responses to each question

Questions	Videos											
	V1		V2		V3		V4		V5		V6	
	T	P	T	P	T	P	T	P	T	P	T	P
Q1	0.895	0.382	0.938	0.358	1.269	0.218	0.411	0.686	0.474	0.642	0.550	0.588
Q2	-1.211	0.239	0.175	0.863	0.364	0.720	-0.185	0.855	-0.519	0.609	1.165	0.257
Q3	-0.176	0.862	1.400	0.175	-1.000	0.328	-0.875	0.391	-0.684	0.501	0.364	0.719
Q4	0.000	1.000	1.149	0.267	0.205	0.840	0.000	1.000	-1.268	0.218	1.818	0.083
Q5	0.498	0.623	1.268	0.223	1.545	0.141	-0.178	0.861	-0.394	0.698	2.419	0.024

Table 8 shows the t-value (T) and p-value (P) of the mean difference in scores of the responses from the control and experimental groups. The results show that the difference in mean scores between the groups is statistically insignificant.

Accordingly, the cross-modal correspondence between the visual features, audio, and haptic effects; and their impact on the overall QoE is investigated. Further discussion about these results is provided next.

4.2. Discussion

In our sample dataset, participants reflected their viewing experience on each video. Here, our objective is to explore the QoE difference in viewing the videos with original versus altered and high-pitch versus low-pitch audio. In addition, we wanted to know the level of cross-modal correspondence shown by each feature in the videos. Accordingly, analysis of the data matching to each of the above mentioned objectives has shown substantial result on the cross-modal correspondence as described next.

4.2.1. Quality of experience

Our statistical test in
978-1-5386-1737-3/18/\$31.00 ©2018 IEEE

Table 4 and Table 5 shows a comparison of the impact of high-pitch versus low-pitch audios (and the auto-generated haptic effects) on QoE. The results show that the change in audio pitch has certain effect on the participants' QoE and the low-pitch audio seem to have better result. However, because the t-test for equality of means did not confirm this difference, there is insufficient evidence to confirm that the "merely" varying audio pitch and the haptic effects generated out of it has a direct association with QoE.

In our second statistical test, we compared the impact of the original and altered audios (and auto-generated haptic effects) on QoE and the original showed better result than the altered (see Table 6, and Table 7). However, this result is not substantiated by the difference in mean scores between the groups which is statistically insignificant. Thus, further discussion on the impact of the association between the visual features of each videos and the (high versus low) audio pitch on the QoE is required as presented next.

4.2.2. Cross-modal correspondence

Our analysis result on participants' reflections in Fig. 2 shows that the average level of agreement of the experimental group (with original audio) of participants to all of the questions in Table 2 (for all videos) is generally lower than that of the control group. This implies that the participants have enjoyed viewing the videos with original audios more than viewing them with the high/low pitched audio. Thus, the reverse cross-modal correspondence between the visual features and audio is generally low.

The p-values in Table 8 also show that the difference between the groups for all videos and for all questions is insignificant. This shows that the impact of the altered audio (and its haptic effect) on the perceived QoE by the experimental groups of participants is about the same as the original audio in the control group.

On the other hand, the average response from the experimental and control groups for all videos shown in Fig. 3 is above neutral which indicated that the transitively generated haptic effects have enhanced the QoE of both groups. In addition, a higher average response of the experimental group is observed on V5 showing existence of certain level of correspondence.

From the facts in the above paragraphs, we can deduce that the inverse cross-modal association from visual features to audio pitch is generally weak. However, certain level of association is observed from angular shape to high pitch audio.

Finally, the relevance of the transitively generated haptic effect for achieving better enjoyment, sense of reality, and enhanced QoE was generally positive. In addition, the haptic effect which was transitively generated out of the angular visual features have shown better enhancement on the participants' QoE of viewing the sample videos.

5. CONCLUSION

In this research, our review of related work shows that adding more media types into mulsemmedia enhances the

user's quality of experience and the cross-modal correspondence among the media types can contribute to this end. Because not all cross-modal correspondences are bi-directional and transitive, we conducted an experiment to identify the association from the visual features into varying audio pitches; and then haptic effects. Accordingly, our work revealed weak inverse cross-modal correspondence except that the certain level of association is observed from angular shape to high pitch audio. On the other hand, the transitively generated haptic effects were found to have enhanced the participants' quality of experience. In general, our findings showed good insights into inverse and transitive cross-modal correspondences using video samples. Thus, similar work will be repeated in the future by incorporating additional media types (e.g. olfaction) focusing on semantic congruence between elements of a video scene and mulsemmedia effects.

Acknowledgment - This work has been performed in the framework of the Horizon 2020 project NEWTON (ICT-688503) receiving funds from the European Union.

6. REFERENCES

- [1] Ghinea, G., Timmerer, C., Lin, W., & Gulliver, S. R. (2014). Mulsemmedia: State of the art, perspectives, and challenges. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 11(1s), 17.
- [2] Cingel, D. (2017, February). How Parents Engage Children in Tablet-Based Reading Experiences: An Exploration of Haptic Feedback. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (pp. 505-510). ACM.
- [3] Murray, N., Ademoye, O. A., Ghinea, G., & Muntean, G. M. (2017). A Tutorial for Olfaction-based Multisensorial Media Application Design and Evaluation. *ACM Computing Surveys (CSUR)*, 50(5), 67.
- [4] Zou, L., Tal, I., Covaci, A., Ibarrola, E., Ghinea, G., & Muntean, G. M. (2017, June). Can Multisensorial Media Improve Learner Experience? In *Proceedings of the 8th ACM on Multimedia Systems Conference* (pp. 315-320). ACM.
- [5] Lenay, C., Gapenne, O., Hanne-ton, S., Marque, C., & Genouëlle, C. (2003). Sensory substitution: limits and perspectives. *Touching for knowing*, 275-292.
- [6] Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., et al (2010). Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. *Psychological Science*, 21, 21–25.
- [7] Harvey, J. P. (2013). Sensory perception: lessons from synesthesia: using synesthesia to inform the understanding of sensory perception. *The Yale journal of biology and medicine*, 86(2), 203. Chicago
- [8] Simpson, R. H., Quinn, M., & Ausubel, D. P. (1956). Synaesthesia in children: Association of colors with pure tone frequencies. *The Journal of Genetic Psychology*, 89, 95–103.
- [9] Ademoye, O. A., & Ghinea, G. (2013). Information recall task impact in olfaction-enhanced multimedia. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 9(3), 17.
- [10] Ghinea, G., & Chen, S. Y. (2006). Perceived quality of multimedia educational content: A cognitive style approach. *Multimedia systems*, 11(3), 271-279.
- [11] Goldstein, E.B. (2013). *Sensation and Perception*. Cengage Learning
- [12] Le Callet, P., Möller, S., & Perki, A. (2012). Qualinet white paper on definitions of quality of experience. *European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003)*, 3.
- [13] Yuan, Z., Chen, S., Ghinea, G., & Muntean, G. M. (2014). User quality of experience of mulsemmedia applications. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 11(1s), 15.
- [14] Yuan, Z., Ghinea, G., & Muntean, G. M. (2015). Beyond multimedia adaptation: Quality of experience-aware multi-sensorial media delivery. *IEEE Transactions on Multimedia*, 17(1), 104-117.
- [15] Murray, M. M., Michel, C. M., de Peralta, R. G., Ortigue, S., Brunet, D., Andino, S. G., & Schneider, A. (2004). Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging. *Neuroimage*, 21(1), 125-135.
- [16] Hancock, P. A., Mercado, J. E., Merlo, J., & Van Erp, J. B. (2013). Improving target detection in visual search through the augmenting multi-sensory cues. *Ergonomics*, 56(5), 729-738.
- [17] Danieau, F., Lécuyer, A., Guillotel, P., Fleureau, J., Mollet, N., & Christie, M. (2013). Enhancing audiovisual experience with haptic feedback: a survey on HAV. *IEEE Transactions on Haptics*, 6(2), 193-205.
- [18] Marks, L. E. (1989). On cross-modal similarity: the perceptual structure of pitch, loudness, and brightness. *Journal of Experimental Psychology: Human Perception and Performance*, 15(3), 586.
- [19] Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, 73(4), 971-995.
- [20] Kanaya, S., Kariya, K., & Fujisaki, W. (2016). Cross-Modal Correspondence Among Vision, Audition, and Touch in Natural Objects: An Investigation of the Perceptual Properties of Wood. *Perception*, 45(10), 1099-1114.
- [21] Wicker, F. W. (1968). Mapping the intersensory regions of perceptual space. *The American Journal of Psychology*, 81, 178–188.
- [22] Ward, J., Huckstep, B., & Tsakanikos, E. (2006). Sound-colour synaesthesia: To what extent does it use cross-modal mechanisms common to us all? *Cortex*, 42(2), 264-280.
- [23] Iordanescu, L., Guzman-Martinez, E., Grabowecky, M., & Suzuki, S. (2008). Characteristic sounds facilitate visual search. *Psychonomic Bulletin & Review*, 15(3), 548-554.