AGE AND GENDER INFLUENCE ON PERCEIVED OLFACTORY & VISUAL MEDIA SYNCHRONIZATION

Niall Murray[1], Yuansong Qiao[1], Brian Lee[1], Gabriel-Miro Muntean[2], Karunakar A. K. [3]

[1]Software Research Institute, Athlone Institute of Technology, Ireland
[2]School of Electronic Engineering, Dublin City University, Ireland
[3]Manipal Institute of Technology, Manipal University, India
{nmurray, ysqiao}@research.ait.ie, blee@ait.ie, munteang@eeng.dcu.ie, karunakar.ak@manipal.edu

ABSTRACT

Lately, significant efforts have being put into proposing various solutions for increasing multimedia viewers' perceived quality levels. One innovative avenue is to enhance users' quality of experience (QoE) by extending the classic audio-visual multimedia content to stimulate also other human senses such as olfaction, tactile, etc. In this context, this paper focuses on olfaction-enhanced multimedia content and presents the results of an experimental study which looked at user perception of inter-stream synchronization between olfactory data and video, whereby the audio used provides no contextual information. The study investigates how age and gender influence users' perception of the temporal boundaries within which they perceive olfactory data and video to be synchronized. The impact on user QoE levels (considering sense of enjoyment, relevance and reality) during synchronous and asynchronous presentations of olfactory and video media is also analyzed and discussed. The results show that there are significant differences in terms of how users of various gender and age groups perceive the skew between olfaction and video content and in their QoE levels.

Index Terms— Olfactory Media, Multimedia Inter-stream Synchronization, Perceived Multimedia Quality, Quality of Experience

1. INTRODUCTION

Classic multimedia systems involve the combination and presentation of independent, discrete or continuous media components such as: text, animation, graphics, images, audio and video. As this list is being extended with so-called new media like e-touch [1], e-taste [2] and e-smell [3], there are expectations that multisensory-enhanced multimedia will result in significant increases in users' Quality of Experience (QoE) levels [4].

This paper studies the effect olfaction-enhanced multimedia content has on viewers' QoE levels and analyses the results of a study which looked at users' perception of synchronization between olfactory data and video (audio used was sound of a blowing fan). When compared with [5], the results of this work show that the removal of contextual audio, inter-media skew has a more significant impact on user detection of skew and scale of acceptable skew. Understanding the relationship between olfaction and visual media is crucial because, cross-modal effects, i.e. the interaction of the senses, can have a major influence on how environments are perceived, even to the extent that large amounts of detail perceived by one sense may be ignored when in the presence of other more dominant sensory inputs [6]. In many olfaction-enhanced applications, the audio may provide no contextual information on the scent or visual – hence understanding the relationship between olfaction and visual alone is an important question to address.

It has been shown that humans perceive smell differently based on a number of factors, including age, culture, mood, gender and life experiences [7]. Interestingly, little research has been carried out in terms of perception of olfactory data with other media [3][5][8][10][11] and no works document such results considering age and gender. This work, investigates how age and gender affect users' perception of inter-stream synchronization of olfactory and visual media.

The remainder of this paper is organized as follows. Section 2 discusses related work, section 3 describes the components of the olfactory and video media display system used during the subjective testing and section 4 outlines the assessment methodology employed. Section 5 presents the test results and analysis, and section 6 concludes the paper and sets directions for future research.

2. RELATED WORK

Lately, research on synchronization between different media components, a fundamental requirement of multi-sensory media applications, has been highly active [3][8].

One focus has been on studying user perception of the inter-stream synchronization of olfactory data with other media - such as video [12], audio and video [3][11] and haptic [8]. The methodology used in these works was originally documented in [13] and [14]. In [13], inter-stream skews were artificially introduced between audio and video (lip synchronization) to determine the acceptable user perceived temporal synchronization boundary.

Research reported in [10] used a novel head mounted olfactory-enhanced display and found that the addition of smell attracted the assessor's attention. A guideline relevant to the presentation of smell with audiovisual information was proposed. Among others, it recommended that the

period between different smells should be longer than 5s.

In [5], the methodology of [13] was employed to define a user perceived temporal boundary within which audiovisual data is synchronized with olfactory data. The paper also analyzed the impact of asynchrony in terms of annoyance, distraction, enjoyment, sense of reality and sense of relevance. The authors found that assessors were more tolerant of olfactory stimuli ahead of audiovisual content as opposed to olfaction behind audiovisual content. In [12], artificial skews were introduced between olfactory and visual media only to define a general user perceived temporal boundary.

During initial testing, using the same video as in [5][11] some assessors noted that the context of the audio content made large skews between the video and olfactory media acceptable. As such, this paper analyses perceived synchronization between olfactory and video media components, while considering participants' age and gender profiles.

3. EXPERIMENTAL SET-UP

This section presents the olfactory-enhanced video presentation equipment, laboratory test bed, assessors and video and scents used in this work.

3.1. Olfactory-Enhanced Video Presentation Equipment

Fig. 1 shows the olfactory-enhanced video presentation system used. It consists of a laptop with Intel Core™ 2 Duo CPU @ 1.66GHz with 2GB RAM, a 21 inch display screen with a resolution of 1024*768 and an Exhalia SBi4 – radio v2 scent emitter (Fig. 1, item Y) [15]. SBi4 presents scents by blowing air (using in-built fans) through scent cartridges placed in the device. SBi4 can store up to four scent cartridges at any one time. The video content was played using the VLC media player 1.0.1 Goldeneye. SBi4 is controlled using the Exhalia java-based SDK. A special

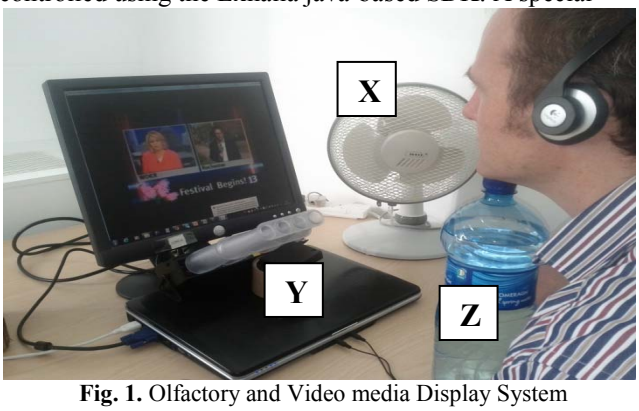

Fig. 1. Olfactory and Video media Display System

program was developed that controlled the synchronized presentation of olfactory data and video, including the introduction of artificial skews between the two media components presented in step sizes as per section 4.2.

During testing, assessors were seated at the testing booth shown in Fig. 1, in a dedicated testing room. A bottle of water was placed under participants' chin during testing (Fig. 1, item Z) to ensure consistency across all assessors in terms of the location of their olfactory fields regardless of posture or physical size [12]. A fan (Fig. 1, item X) was employed to remove scents faster in between test sequences.

3.2. Laboratory and Assessors

The design of the test laboratory was in accordance with ISO/IEC standard 8589 [16]. A sign restricting access to the test room was posted outside the door. No ventilation system exists per se, but the test lab is large and has multiple windows to remove scent from the test area. In addition, between clips, the fan (Fig. 1, item X) was turned on. A total of 43 assessors (20 female: 9 (20-30 yrs), 8 (30-40 yrs), 3 (40+ yrs.) and 23 male: 11 (20-30 yrs), 10 (30-40 yrs), 2 (40+ yrs)) took part in the study. This group included people aged between 19 to 56 years, of multiple nationalities and from a wide variety of backgrounds: post graduate researchers, academic staff, professionals and members of the public.

3.3. Videos Sequences and Scents

Six video clips, kindly provided by the authors of [5], are each 90s duration. These clips were in the form of documentaries, cookery programs and news shows. Each of the video clips can be divided into three 30 second blocks. The middle 30s block in each clip contains content related specifically to the scent being presented [5]. The following six scents, each presented continuously for 30s, were selected based on the video content. Flowery, foul, fruity, burned, resinous and spicy reflect a fair distribution between pleasant and unpleasant smell categories. These scents match those used in [5] and are widely used in olfactory research [17][18].

4. ASSESSMENT METHODOLOGY

Assessors were presented with each olfaction enhanced clip twice. The reference sample was delivered first (synchronized presentation of scent and video), then the sample under test, which contained either synchronized or artificially skewed scent and video media. The assessors were asked to answer a questionnaire on their experience related to the sample under test. This is consistent with the methodology of ITU-T P.910 [19].

4.1. Questionnaire and Rating Scale

The questionnaire required assessors to respond to five questions/statements. The questionnaire was reviewed by a Psychologist and underwent a reliability assessment with a subset of assessors.

The assessors were asked to select one of the five possible answers per question as per table 1 relative to their experience of the stimulus under test. The first statement aimed to determine assessor ability to detect the existence of a synchronization error. Question 2 aimed to identify how tolerant assessors were to different levels of skew. The final

three statements were included to analyze the impact of inter-media skew on the user experience (relevance, enjoyment and sense of reality). The questions were ordered from general to being more specific.

Table 1. Rating scales for each of the five statement/questions

Score	Statement 1	Question 2	Statements 3, 4, 5
5	Too Late	Imperceptible	Strongly Agree
4	Late	Perceptible but not annoying	Agree
3	Neither Early or Late	Slightly Annoying	Neither Agree or Disagree
2	Early	Annoying	Disagree
1	Too Early	Very Annoying	Strongly Disagree

4.2. Introduction of Artificial Skews

In order to determine the perceptible and tolerant levels of inter-media skew between olfactory and video media components, assessors were presented with varying levels of skew (including no skew) and queried about their perception of the experience. For olfactory media to be in sync with the video (0s skew), it should be presented during the 30s to 60s block of the video clip. Olfactory data ahead of video content is represented by skew times of -30s, -25s, -20s, -15s, -10s and -5s and olfactory data after video content is represented by skews of +5s, +10s, +15s, +20s, +25s and +30s.

5. SUBJECTIVE TEST RESULTS

The results of the subjective testing, which involved responses to the five questions/statements presented above and expressed in terms of mean opinion score (MOS), are analyzed next.

5.1. Preliminary Experiment: Odor Detection Latency

Because of the slow moving nature of olfactory data compared with audio or video media, it was critical for the synchronization study to determine how long it took assessors to detect the presence of odors once emitted. 15 participants (9 male, 6 female) were presented with the 6 scents twice in random order. Assessors clicked on the mouse once they detected a scent. As we considered it took 1 second for assessors' reaction and click on the mouse we determined, on average, but per scent, how long in advance the olfactory display device's fans should be started in order to ensure timely presentation to the users. Based on the SBi4 being 0.5 meters from the assessor, it was found that it took assessors between 2.7s - 3.7s to detect the scents depending on the scent.

5.2. Study: Perception of Synchronization Error Considering Gender and Age

Detection of Skew

Fig. 2 and Fig. 3 present responses to statement 1, indicating male and female assessors' ability to detect inter-media skew. Fig. 2 shows how all the participants, regardless of gender, have identified the existence of inter-stream skew. Worth noting is that both sexes better accept olfactory stimuli after, rather than before, the corresponding video event. This conclusion is based on the fact that the MOS scores are closer to 3 (neither early nor late) in the former case (MOS scores for +5s is 3.62) than in the latter situation (MOS score for -5s is 1.95).

Direct comparison of MOS scores at various skews indicates that the female group is marginally more sensitive to skew than their male counterparts with the exception of -20s and +15s. Interestingly based on MOS comparison, male assessors viewed skews of +15s and -5s similarly in terms of being late or early. A similar result was obtained for female assessors when scoring skews of +20 and -5s.

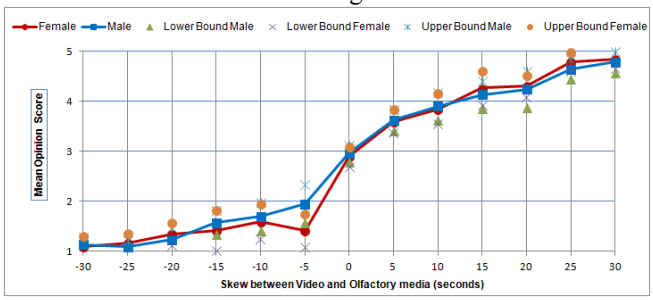

Fig. 2. Gender analysis detection of skew with confidence intervals based on a 95% confidence level.

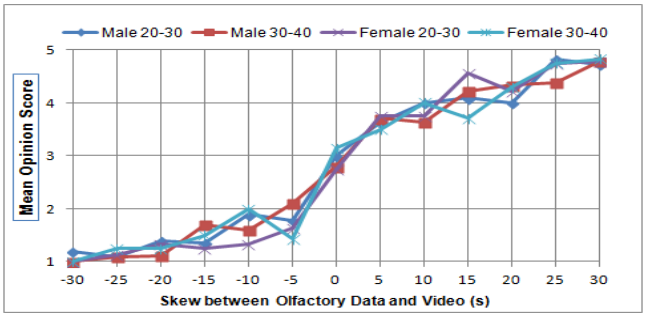

Fig. 3. Gender/age analysis of detection per skew

Fig. 3 shows that there is a general consistency across all age groups in terms of their ability to identify inter-media skew. However, the females are more (or equally) sensitive to inter-media skew than men, and younger users than older assessors for all skew levels.

Perception of Skew

Question 2 assessed the effect of the inter-media skew has in terms of annoyance level. This is key to determining temporal boundaries, as works involving other media have shown that users can tolerate certain levels of skew [11] [13]. Fig. 4 and Fig. 5 show the level of annoyance different inter-media skews have on users of various gender and age groups. Significant changes in MOS between olfaction after and before the video for both sexes occur between skews of -15s and +15s. Assessors perceived skew of -15s as being between annoying and slightly annoying. Comparing this with the equivalent skew when olfaction is after the video, the MOS score is in the range between slightly annoying and perceptible but not annoying. This is confirmed further

when analysis of the -10 s and +10 s skews is performed. Another interesting result occurs at skews of 5s. Scent presented -5s ahead of video has the same scoring rating as when scent is presented +10s after the video, still in the perceptible but not annoying to slightly annoying range. However, scent presented at +5s is in the "perceptible but not annoying" to "imperceptible" range. Hence, it is valid to conclude that assessors are less tolerant to olfaction ahead of video than they are of olfaction after the corresponding video event. Particularly interesting is the difference in perception between males and females at skews of -10s and -5s, respectively. The female group members are much more sensitive and less tolerant at these skew levels than their male equivalents. Fig. 5 presents the analysis considering gender and age within the 20-30 years and 30-40 years male and female categories. Skew levels of -5s and -10s are between perceptible not annoying to slightly annoying for three of the categories (20-30 male and 30-40 male/female), but is clearly in the annoying to slightly annoying range for the 20-30 females.

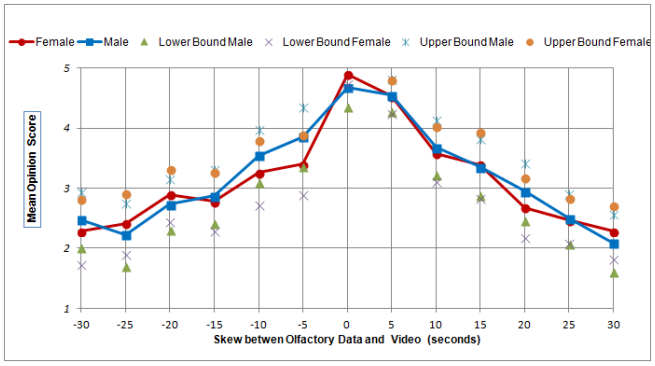

Fig. 4. Gender analysis of annoyance level per skew with confidence intervals based on a 95% confidence level.

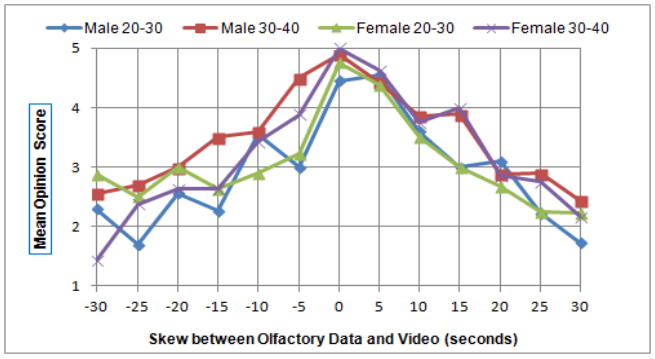

Fig. 5. Gender/Age Analysis of Annoyance Level per Skew

From Fig. 5, it is clear that the 30-40 males are more acceptable to scent before video than other groups, and even found -15 s skews somewhat acceptable, while all other groups scored between slightly annoyed and annoyed. Skew levels of +10s were acceptable to all age groups and genders, while +15s skews were acceptable to male and female 30-40 age groups, but not to the corresponding younger groups. The results indicate how the two younger participant categories are less tolerant to and more easily negatively affected by skew than the two older groups.

5.2.1. Age and Gender-based Temporal Boundary Definition

Findings from Fig. 2 and Fig. 3 support the literature that recounts human sensitivity to scent and the ability to detect it. The MOS values for corresponding skews indicate greater sensitivity to olfaction ahead of video, than olfaction after video. Considering Fig. 4 and Fig 5 it is plausible to conclude that olfaction before video is more annoying than olfaction after video.

In terms of defining the temporal boundaries for synchronizing olfactory and video media, these results reveal "in-sync" and "out-of-sync" regions as well as revealing certain characteristics based on gender and age. The in-synch region is based on the range where assessors perceive errors not to be annoying. Based on the assessor tolerance to olfaction after video as opposed to olfaction ahead of video, the span is larger for olfaction after video.

These boundaries are based on (1) the above results that indicate assessors are more tolerant to skews when olfaction is after video, (2) Differences in perception based on gender and age and (3) An impairment rating of above 3.5 (i.e. between "perceptible but not annoying" and "slightly annoying") is minimum for synchronized presentation of olfactory and video media.

- The in-synch region spans between a maximum skew of 0s to -5s/-10s when olfaction is ahead of video, and a maximum skew of 0s to +10s/+15s when olfaction is after video depending on the age and gender of the user.
- The out-of-synch region for olfaction ahead of video spans beyond the skew of -10 s and skew of greater than +15s when olfaction is after video depending on the age and gender of the user.

5.3. Study: Synchronization Skew Influence on User Experience

This section analyzes the influence of inter-media skew on the quality of experience of olfaction enhanced video.

5.3.1. Influence on the Sense of Relevance

Fig. 6 shows the assessors' level of agreement with statement 3 in the presence of varying degrees of inter-stream skew. Subjected to synchronized presentation (0s skew), both female and male assessors found the smell relevant to what they were watching. The female group consider skews with olfaction ahead of video having a more negative impact on relevance as they have awarded consistently lower MOS scores for skews of -15s, -10s and -5s. For olfaction after video, females are more tolerant of skew than their male counterparts, at all skew levels between +5s and +30s. For both sexes, olfaction after the video is more relevant than olfaction ahead of video. Olfaction 15s after the video is still more relevant than scent 5 seconds ahead. The slow gradual decrease in relevance for olfaction after video is another interesting finding of this work. Considering the age categories (Fig. 7), both female age categories demonstrate comparatively lower relevance

values when olfaction is before video and comparatively higher relevance values when olfaction is after video than males. The male 30-40 years category clearly presents lower values of relevance when in synch (0s) and across a wide range of skews for olfaction behind video (+5s to +20s). The male 20-30 years category is broadly similar to the same age category for females with the exception of +10s.

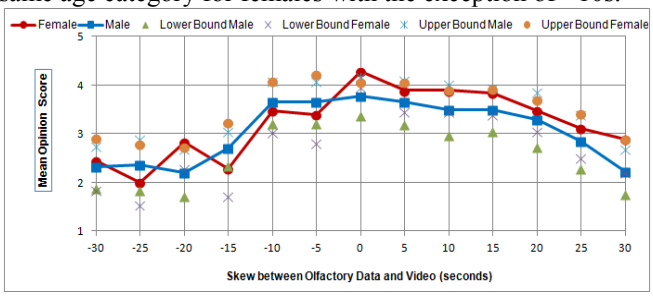

Fig. 6 Gender analysis of relevance level per skew with confidence intervals based on a 95% confidence level.

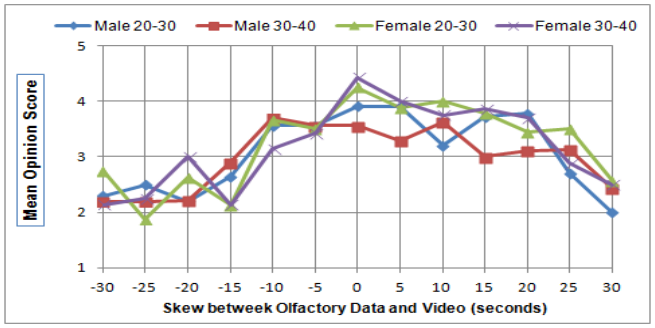

Fig. 7 Gender/age analysis aware of relevance level per skew

5.3.2. Influence on the Sense of Enjoyment

Fig. 8 and Fig. 9 illustrate assessors' level of agreement with statement 5. When synchronized presentation takes place, the female assessors enjoyed watching the video clip more than the males. As was the case for relevance, olfaction presented before the video has a greater impact on sense of enjoyment for the females compared to the males for skews from -30s to -5s. With in synch presentation and also where olfaction was after the video up to +10s, the female group indicated higher sense of enjoyment compared to their male counterparts. For skew values greater than +10s there are minimal differences between the groups.

Fig. 9 presents an illustrative analysis considering both gender and age. A salient pattern across the four age/gender categories is that skew size affects the sense of enjoyment of olfaction enhanced media. In comparing the views of the various age groups, the impact of skew is greatest for the female 30-40 years category, especially with olfaction ahead of the video, as their MOS scores are the lowest. The 20-30 years groups are broadly consistent in their opinion regardless of gender when scents are presented ahead of the video. Additionally both female categories indicate higher MOS scores than their male counterparts in the presence of zero skew. For olfaction stimuli after the video, the female 20-30 years category are least impacted by small skews.

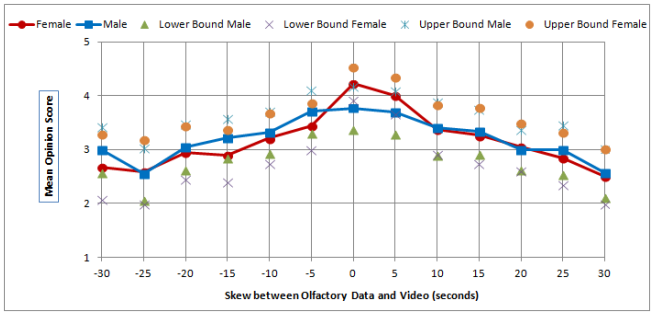

Fig. 8. Gender analysis of enjoyment level per skew with confidence intervals based on a 95% confidence level.

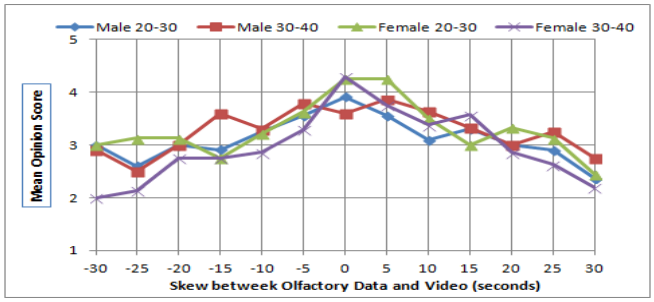

Fig. 9. Gender/age analysis of enjoyment level per skew

5.3.3. Influence on the Sense of Reality

Fig. 10 shows MOS scores reflecting assessor's level of agreement with statement 4 in the presence of varying degrees of inter-media skew. When synchronized presentation takes place, assessors agreed that the smell heightened the sense of reality of what they were watching. This response was particularly strong for the females. In the presence of large skews (e.g. -30 s or +30 s), most assessors from both male and female groups stated that "Neither agree or Disagree" or "Disagree" with statement 4. For olfaction presented before the video, with small skews of -10s and -5s, the male groups perceived the olfaction as contributing to an enhanced sense of reality. The equivalent female MOS scores indicate less heightened sense of reality at these skew levels. The most interesting finding was the slow reduction in heightened sense of reality for the scent presented after the video with increasing skews for both male and female groups. In particular, when comparing the ratings of the female group when scent was distributed ahead or after the video, the impact on reality was higher at skews of +20s then it was for skews of -5s. A similar trend, although not as significant, exists for males rating the sense of reality for -10s and +15s. Analyzing assessor opinions considering both their age and gender from Fig. 11 again indicates important differences across age and gender categories. The female 20-30 years category perceived a heightened sense of reality up to skew levels of +15s in comparison with the other groups. The male 30-40 years group is more sensitive to the impact of skews sizes of greater than +10s. Interestingly the same group appears less sensitive to skews of -15s and -10s than all other groups, which seem to be relatively consistent in their evaluation.

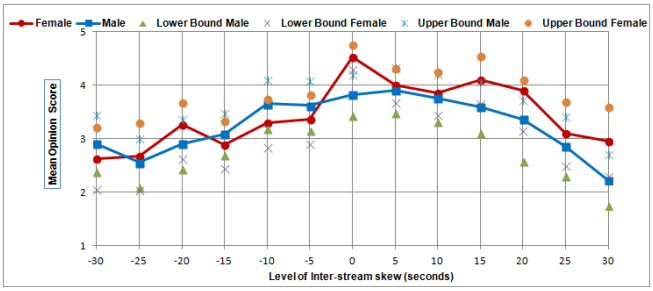

Fig. 10. Gender analysis of the level of reality per skew with confidence intervals based on a 95% confidence level.

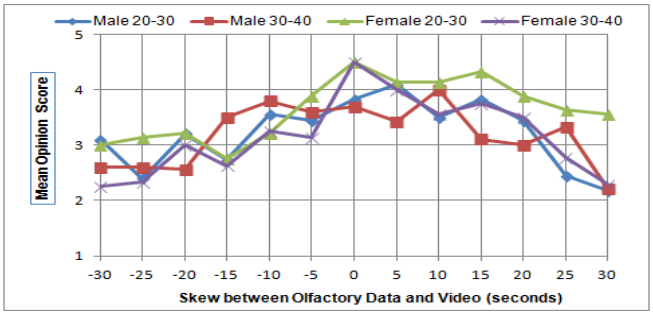

Fig. 11. Gender/age analysis of the level of reality for various skews

6. CONCLUSION

This work investigated the perception of inter-stream synchronization error between olfaction and video media and its effect on user quality of experience for different genders. It defined user temporal boundaries for user perceived synchronization. Synchronization is tolerated differently by various gender and age groups with up to a maximum skew of between 5-10 seconds when olfaction is presented before the video, and a maximum skew of 10-15 seconds when olfactory stimuli are released after the video. Skews outside this range make the media components out-of-sync. This is in contrast with the work reported in [5] (which included contextual audio), where a temporal boundary up to a maximum skew of between 30 seconds when olfaction is presented before the video, and a maximum skew of 20 seconds when olfaction stimuli are released after the video. The authors conclude the significant difference between the results is based on the removal of contextual audio in this work. Contextual audio in [*5*] made larger skews of olfaction before and after the video more acceptable. Arguably the most interesting difference between the results was that here, olfaction after the video was more tolerant for assessor's than olfaction before video. Again, the authors conclude that the removal of contextual audio accounts for this. In fact it was noted from assessor feedback that when contextual audio was present, the scent presented early provided more information on what they were going to see.

In addition, younger female assessors are the most sensitive to inter media skew in the tested groups and older males the least sensitive. Considering the results between the different groups reported above, these results also highlight the potential requirement for profile based presentation of olfaction enhanced multimedia. It also highlights the requirement for research into the relationship between olfaction and audio, as from this work, it is clear that the semantics of the audio has a significant impact on user perceived temporal boundaries for olfaction enhanced multimedia.

Finally, this work also analyzed the impact of inter-stream skew on assessors QoE. It was found that skews beyond a particular range between olfaction and video impact negatively the sense of relevance, reality and enjoyment when compared with the case when synchronized presentation took place.

7. REFERENCES

[1] Cha, J., Eid, M., Barghout, A., Mahfujur Rahman, ASM., and El Saddik, A. 2009 "HugMe: Synchronous Haptic Teleconferencing," In *Proceedings of 17th ACM International Conference on Multimedia*.

[2] Narumi, T., Nishizaka, S., Kajinami, T., Tanikawa, T., and Hirose, M. 2011 "Augmented Reality Flavors: Gustatory Display Based on Edible Marker and Cross-Modal Interaction," *Proc. of the ACM CHI Conference on Human Factors in Computing Systems*. (CHI '11).

[3] Ghinea, G. and Ademoye, O. A. "The Sweet Smell of Success: Enhancing Multimedia Applications with Olfaction." *ACM TOMCAPP,* 8, 1 (Jan. 2012).

[4] Timmerer, C., Waltl, M., Rainer, B., Hellwagner, H. "Assessing the quality of sensory experience for multimedia presentations". In *Signal Processing: Image Communication* (Feb. 2012).

[5] Ademoye, O. A., Ghinea, G., "Synchronization of Olfaction-Enhanced Multimedia". In *IEEE Transactions on Multimedia 11(3): 561-565,* 2009.

[6] Calvert, G., Spence, C., Stein, B.E. "The Handbook of Multisensory Processes". In MIT Press. 2004

[7] Ghinea, G., Ademoye, O. A., "Olfaction-enhanced multimedia: perspectives and challenges". *Multimedia Tools Appl* 55(3): 601-626 2011.

[8] Hoshino, S., Ishibashi, Y., Fukushima, N., and Sugawara, S., "QoE Assessment in Olfactory and Haptic Media Transmission: Influence of Inter-Stream Synchronization Error", *IEEE International Communications Quality and Reliability Workshop (CQR)*. 2011

[9] Ramic, B., Hasic, J., Rizvic, S., Chalmers, A. "Selective Rendering in a multimodal environment: Scent and graphics" In *Spring Conference on Computer Graphics, ACM SIGGraph*. 2006

[10] Nakamoto, T., Yoshikawa, K., "Movie with Scents Generated by Olfactory Display Using Solenoid Valves". *IEICE Trans. Fund. Electr. Comm. Comput.* Sci. E89-A, 11, 3327–3332

[11] Ghinea, G. and Ademoye, O. A.,. "Perceived Synchronization of Olfactory Multimedia". In *IEEE Transactions on Systems, MAN, AND CYBERNETICS – PART A: SYSTEMS AND HUMANS* 40, 4 (July. 2010), 657-663.

[12] Murray, N., Qiao, Y., Lee, B., Karunakar, AK, Muntean, G.-M., "Subjective Evaluation of Olfactory and Visual Media Synchronization" In Proceedings of *ACM Multimedia Systems conference*. Feb 26 - March 1, Oslo, Norway. 2013.

[13] Steinmetz, R. "Human Perception of Jitter and Media Synchronization". *IEEE Journal on Selected Areas in Comms* 1996.

[14] Steinmetz, R., Nahrstedt, K.,. "Multimedia: computing, communications and applications". In *Prentice Hall.* 1995

[15] www.exhalia.com

[16] ISO/IEC 8589 "Sensory analysis –Guidance for design of test rooms".

[17] Chastrette, Maurice, 2002. "Classification of Odors and Structure – Odor Relationships". In *Olfaction, Taste, and Cognition* pp. 100-116, Cambridge University Press.

[18] Kaye, N. "Symbolic olfactory display". M.S. thesis, Massachusetts Inst. Techonology, Cambridge, MA, 2001.

[19] ITU-T P.910. "Subjective video quality assessment methods for multimedia applications", 2008.