User Profile Based Perceived Olfactory and Visual Media Synchronization

NIALL MURRAY, YUANSONG QIAO, BRIAN LEE, Athlone Institute of Technology GABRIEL-MIRO MUNTEAN, Dublin City University

As a step towards enhancing users' perceived multimedia quality levels, the authors present the results of a study which looked at user's perception of inter-stream synchronization between scent and video. The ability to detect, the perception of and impact of skew on user's quality of experience is analyzed considering user's age, sex and culture (user profile). The results indicate that skews beyond a certain level between olfaction and video have a negative impact on user perceived experience. Olfaction before video is more noticeable to users than olfaction after video and assessors are more tolerable of olfactory data presented after video.

Categories and Subject Descriptors: H.1.2 [**Information Systems**]: User/Machine Systems - *Human Factors*; H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems - *Artificial, augmented, and virtual realities*; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces - *Evaluation/methodology*

General Terms: Human Factors, Design, Experimentation

Additional Key Words and Phrases: Olfaction, Multimedia Synchronization, Subjective Quality Assessment, Quality of Experience

ACM Reference Format:

Murray, N., Qiao, Y., Lee, B., and Muntean, Gabriel-Miro, C. 2014. User Profile Based Perceived Olfactory and Visual Media Synchronization. ACM Trans. Multimedia Comput. Commun. Appl. x, y, Article 1 (May 2014), 20 pages

DOI=10.1145/0000000.0000000 http://doi.acm.org/10.1145/0000000.0000000

1. INTRODUCTION

Multimedia systems have been characterized by the integration, combination, presentation, storage and communication of independent discrete and continuous media such as: text, animation, graphics, images, audio and video. Today the research community is extending this list with so-called new media like e-touch (Cha et al., 2009), e-taste (Narumi et al., 2011) and e-smell (Ghinea and Ademoye, 2012). The result is the emergence of multisensory media communication and experiences. The rationale of enhancing multimedia applications to stimulate more than audiovisual senses is to increase the user's Quality of Experience (QoE) (Timmerer et al., 2012). With a significant demand already placed on the audiovisual senses, another avenue to increase QoE is through the stimulation of the other senses. Olfaction is the sense of smell. Recently, scents have been used in multimedia, in particular with movies as it is assumed that presenting the scent according to the scenes would deepen the viewer's understanding and sense of reality (Tomono et al., 2004). In addition, we can now find the use of olfaction across other industries in the literature; gaming (Nakamoto et al., 2008), health (Spencer 2006)(Gerardi 2008)(Pair 2006), education (Shams and Seitz, 2008), training (Washburn, 2003) and tourism (Dann and Jacobsen, 2003).

Research on modeling and analyzing the human perception of multimedia experiences is an active topic (Ghinea and Ademoye, 2012)(Timmerer et al., 2012)(Haung et al., 2012)(Gulliver and Ghinea, 2007)(Lee et al., 2011). It is widely accepted that objective measures alone do not reflect the end user perception of a multimedia experience. Humans perceive smell differently based on a number of factors including age, culture (Ayabe-Kanamura et al), life experiences, mood and gender (Ghinea and Ademoye, 2011). In the multimedia domain, little research has been carried out analyzing the perception of olfactory data with other media (Ghinea and Ademoye, 2010, 2012) (Hoshino et al., 2011)(Ramic et al., 2006)(Nakamoto and

Author's address: N Murray, Athlone Institute of Technology, Dublin Road, Athlone, Ireland; email: nmurray@research.ait.ie

DOI10.1145/0000000.0000000 http://doi.acm.org/10.1145/0000000.0000000

Yoshikawa, 2006). With olfaction, how the user perceives the experience is particularly important, considering the number of characteristics that affect its perception.
This paper reports 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). The use of the term video reflects the use of visual media only, the term audiovisual refers to the combination of audio and video. When compared with (Ghinea and Ademoye, 2010), where the relationship between olfactory and audiovisual media was studied, the results here show that the removal of contextual audio has a significant impact on user detection of skew, the scale of acceptable skew, as well as impacting reality, relevance and enjoyment. 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 (Calvert et al., 2004). The contribution of this work is to define the temporal relations between olfaction and video using subjective studies whilst considering age, sex and culture. In our previous work (Murray et al., 2013) the authors reported a study of olfaction enhanced multimedia based on smaller sample size (43 assessors) and analysis of sex and age range with limited ranges (assessors were mainly in range of 20-30 yrs (years) and 30-40 yrs). Here, this work reports a subjective study involving 84 assessors with a stronger balance of age, gender and culture as discussed in section 4. The additional assessors were specifically recruited based on their culture (required assessors outside of Europe) and their age (required male and female assessors greater than 40 yrs). With the extra assessors of specific culture and age profile, the contribution here is defining a user profile for olfaction enhanced multimedia synchronization which considers age, sex and culture. With little research carried on perceived synchronization of olfactory data integrated with other media, no works document such results considering these three variables.

The remainder of this paper is organized as follows: Section 2 discusses related work, Section 3 presents challenges and phenomena associated with olfactory data as a media and Section 4 describes the components of the olfactory and video media display system used during the subjective testing. Section 5 outlines the assessment methodology employed, Section 6 presents the results and analysis of the completed subjective testing and Section 7 discusses our conclusions and directions for future research.

2. RELATED WORKS

A fundamental requirement of any multimedia application, including those enhanced with olfaction, is the synchronized display of multiple media streams. Research on synchronization of multi-sensory media applications is an active research area (Haung et al., 2011)(Arefin et al., 2011)(Eid et al., 2011)(Ghinea and Ademoye, 2010, 2012). In the context of standardization, MPEG-V defines metadata representations for olfactory data among other sensory effects as part of its Sensory Effects Description Language (SEDL) within Sensory Information (part 3) (ISO/IEC, 2010). Little work has been documented on user perceived inter-stream synchronization of olfactory data with other media, with (Ghinea and Ademoye, 2010, 2012) for audiovisual and olfactory, (Hoshino et al., 2011) haptic and olfactory being the exceptions. The methodology used in these works was originally documented in (Steinmetz, 1996)(Steinmetz and Nahrstedt, 1995). In those works, inter-stream skews were artificially introduced between audio and video (lip synchronization) to determine the acceptable user perceived temporal synchronization boundary.

Works attempting to address the issue of scent lingering, approach this from the scent emitter perspective (Nakamoto and Minh, 2007)(Ariyakul and Nakamoto, 2011)(Sugimoto et al., 2010)(Noguchi et al., 2011). These works focus on the hardware that enables controlled emission of minute amounts of scent. The aim is to minimize scent lingering and "enable the instantaneous switching of scents" (Sugimoto et al., 2010) through the precisely controlled presentation of olfactory data. It is arguable that these works are dealing with olfactory data from an intra-stream perspective. In (Ghinea and Ademoye, 2010), the methodology of (Steinmetz, 1996) was employed to define a user perceived temporal boundary within which audiovisual data is synchronized with olfactory data. Artificial inter-media skews were introduced between olfactory and audiovisual media and assessors qualified their experience. In addition to defining audiovisual and olfactory temporal synchronization boundaries, they analyzed the impact of asynchrony in terms of annoyance, distraction, enjoyment, sense of reality and sense of relevance. They found that olfaction before audiovisual content is more tolerable than olfaction after audiovisual content. Their work (Ghinea and Ademoye, 2009, 2010) is the closest to our work found in the literature. In the next section, we

introduce and discuss some important characteristics that require considering when working with olfactory data.

3. CHARACTERISTICS OF OLFACTORY DATA AS A MEDIA

Adding olfactory data as a media brings a number of challenges not common with text, graphics, audio or video media. Smell has a tendency to linger, it is slow moving media, unlike the transitory nature of audio and video. In addition, it is important to recognize the existence of a number of phenomena associated with the olfactory sense. Unlike video or audio, smell is a chemical media. Humans detect odors based on the interaction of odor molecules with smell receptors. Olfactory adaptation occurs when assessors are subjected to continuous olfactory stimulation. The sensory nerve activity decreases to a level where assessors find it difficult to perceive stimuli or don't perceive at all. With the removal of scents, perception is generally restored within a few minutes. Anosmia is another olfactory related phenomenon whereby there exists a lack of sensitivity to olfactory stimuli. It can be total, partial, permanent or temporary. It may result in an inability to perceive one or many different odors. Olfactory thresholds are values that express the amount of scent stimulus required to give an olfactory sensation. A number of sensory related thresholds are described in [ISO 5492:2008]. The detection threshold for any media is the minimum value of a sensory stimulus needed to give rise to a sensation without the sensation needing to be defined. The olfactory detection threshold "has strong appeal because it measures a feature of perception and performance in physical units of concentration" [Lawless, 1998]. The recognition threshold is the minimum physical intensity of a stimulus for which an assessor will assign the same descriptor each time it is presented. The terminal threshold is the minimum value of an intense sensory stimulus above which, no difference in intensity can be perceived. In this work, the term detection instant is defined as the time at which assessors recognize the existence of an odor. This work analyses this instant in terms of the assessor perception of the synchronization between olfactory and video media.

4. EXPERIMENTAL SET-UP

This section outlines the olfactory and video display, laboratory design, assessors as well as video and scents used in this work.

4.1 Olfactory-Enhanced Video Presentation Equipment

As per Fig. 1, the olfactory and video display system consists of the SBi4 – radio v2 scent emitter (item Y sitting on laptop) from Exhalia [Exhalia]. It presents scents by blowing air (using 4 in-built fans) through scent cartridges. In version 2 of the SBi4, it is possible to control the intensity of the scent emitted by altering the fan speed. All scents were presented at full intensity during the tests reported here. SBi4 can store up to four scent cartridges at any one time. Fig. 2 shows the SBi4, scent cartridges and the bespoke extension that was designed and added to the SBi4 as shown in Fig 1 (Item Y). The purpose of this extension was to facilitate an accurate presentation of the scent to the users' olfactory field as opposed to a more general presentation. 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. Further discussion of how this was determined is documented in section 6.1. The cartridges of SBi4, exposed during operation, are made from scented polymer balls. Initially with the SBi4 cartridges, it is possible to detect odors in advance of any fans running (due to natural vaporization); however after 2-3 days, detectable odors are minimal and for most scents, not possible to detect at all when fans are not running. The SBi4 system is controlled using the Exhalia java-based SDK. It is connected to the laptop via a USB port. The video content was played using the VLC media player 1.0.1 Goldeneye. A special control 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 table 4 section 5.1. The laptop is windows 7 professional, Intel Core™ 2 Duo CPU @ 1.66GHz with 2GB RAM. The display screen was 21 inches with a resolution of 1024*768. During the testing, assessors were seated at the testing booth shown in Fig. 1, in the experimentation room as shown in Figure 3. In addition, Fig. 1 also includes a bottle of water that the assessors placed under their chin during testing (Fig. 1, item Z). The purpose of this was to have consistency across all assessors in terms of the location of their olfactory fields regardless of posture or physical size. The fan in Fig 1. (Item X) is turned on between test sequences to remove any lingering scent.

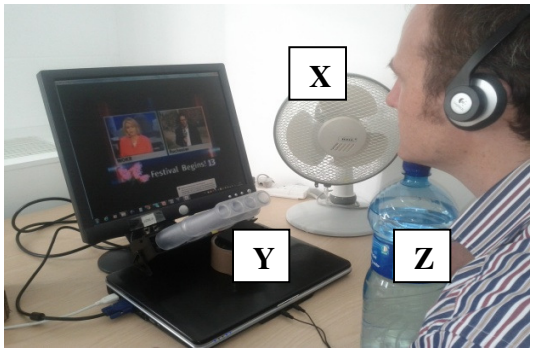

Fig. 1. Olfactory and Video media display system.

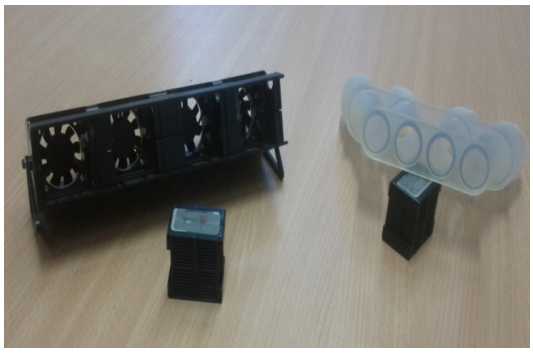

Fig. 2 SBi4 V2, scent cartridges and bespoke extension.

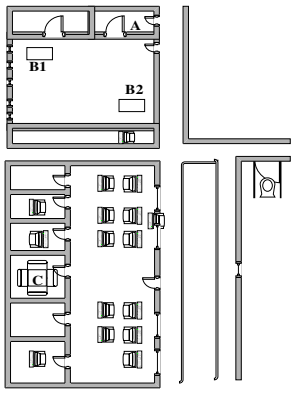

Fig. 3 Plan view of experimentation room (B), preparation room (A) and meeting room (C).

4.2 Laboratory Design

The design of the test laboratory is in accordance with ISO standard (ISO/IEC 8589), “Sensory analysis – General guidance for the design of test rooms”. The aim of this standard is to design test rooms such that it is possible (1) to conduct sensory evaluations under known and controlled conditions with minimum distraction and (2) to reduce the effects that psychological factors and physical conditions can have on human judgment. The minimum requirement for the creation of test rooms are (1) a testing area in which work may be carried out individually in testing booths and (2) a preparation area. Fig. 3 shows a plan view of the preparation room and storage room for samples (room A), experimentation room (room B), and waiting room for assessors (room C). Walls in the test room are Matt off-white. One temporary testing booth (Fig. 3, B1) is situated in the corner of the test room to minimize distraction. Assessors complete questionnaires (see section 5.1) in the furthest point form the testing booth (B2). This allows time for scent to diffuse, minimizes adaptation, gives assessors a break between each judgement and avoids assessor being influenced by lingering scent in the air. A sign restricting access to the test room is posted outside the door. The preparation area is located adjacent to the test room as shown. Assessors do not have access to this room. Whilst no ventilation system exists per se, the test lab is large, has 3 doors and has multiple windows to remove scent from the test area. Between viewing clips, the fan was turned on to remove lingering scent.

4.3 Assessors

A total of 84 assessors took part in the study. This group included assessors between the ages of 19 to 60 yrs from a wide variety of backgrounds: students, post graduate researchers, academic staff, health care professionals, members of defence and police forces, accountants, farmers, teachers, IT industry professionals, persons from medical and construction industry and also persons unemployed. In order to be eligible, assessors could not be involved in any sensory analysis testing in the twenty minutes preceding the tests. In an attempt to provide contamination free results, assessors must be free from colds or flu’s; must avoid wearing perfume, aftershave or scented deodorants on the day of the testing. In

addition they were requested to avoid chewing gum, eating food, drinking tea or coffee in the 30 minutes prior to the test. Assessors were also screened for anosmia as per [ISO 5496:2006], with assessors who could not detect particular scents not included for the test sequences. Based on this prescreening, two assessors were deemed ineligible and did not take part in the tests reported here. A detailed tutorial on the execution of testing involving olfaction enhanced multimedia is available in [Murray et al., 2013].

Table 1 Video Categories and Scents Used (Ghinea and Ademoye, 2012) © ACM 2012

Smell Category	Burnt	Flowery	Fruity	Foul	Resinous	Spicy
Video Clip #	**Clip1**	**Clip2**	**Clip3**	**Clip4**	**Clip5**	**Clip6**
Video Description	Documentary on bush fires in Oklahoma	News broadcast featuring perfume launch	Documentary about rotting fruits	Cookery show on how to make a fruit cocktail	Documentary on Spring allergies & cedar wood	Cookery show on how to make chicken curry
Smell Used	Burning Wood	Wallflower	Strawberry	Rancid Acrid	Cedar Wood	Curry
						

Table 2 Breakdown of assessors based on age, gender and culture.

Gender	Age	African	European	Asian	South American	Australian	Totals
Female	20-30	1	8	3	1	0	13
Female	30-40	2	5	2	0	0	9
Female	40-60	0	9	2	0	1	12
Total Female		3	22	7	1	1	34
Male	20-30	4	9	9	1	0	23
Male	30-40	2	8	4	0	0	14
Male	40-60	2	11	0	0	0	13
Total Male		8	28	13	1	0	50
Totals		11	50	20	2	1	84

4.3.1 Limitations

Considering table 2, the authors identify some limitations in terms of the number of assessors from South America and Australia and some minor limitations in the African and Asian groups. For this reason, statistical analysis reported hereafter considers just the African, European and Asian groups. Hence, the authors report this work as an exploratory study to investigate the influence of culture on olfaction-enhanced multimedia. A more balanced study will be carried out as outlined later in future work.

4.4 Video Sequences and Scents

Six videos used (kindly provided by the authors of [Ghinea and Ademoye, 2010]) were of 90s duration (audio was removed). Each of the video clips can be divided into three 30 second blocks whereby the middle 30s block contains content related specifically to the scent being presented. The clips are in the form of documentaries, cookery programs and news shows, and were chosen and altered such that the middle 30s segment corresponded to the content relating to the olfactory media. Each of the six scents chosen also matched those used in the work of [Ghinea and Ademoye, 2010]. The scents of flowery, foul, fruity, burnt, resinous and spicy reflect a “fair distribution ration between what can be termed as pleasant and unpleasant smell categories”. These scents are widely used in olfactory research (Chastrette, 2002)(Kaye, 2001). A detailed comparison between the work of (Ghinea and Ademoye, 2010) and this work is discussed in detail in section 6.4.

5. ASSESSMENT METHODOLOGY

On arrival to the meeting room (room C in Fig. 3), assessors were provided with an information sheet on the tests. Any questions were addressed and assessors were required to sign a consent form. From here they were brought to the experimentation room (room B in Fig. 3) where testing was carried out. Assessors were asked to engage for the duration of each test sequence. On completion of tests, windows in the room were opened and the fan (item X in Fig. 1) was turned to remove any lingering scents. There was always a minimum of fifteen minutes between consecutive executions of tests between assessors. This gave ample time for removal of any lingering scents, collection of questionnaire sheets and preparation for subsequent assessor testing. It also included time for the new assessor to read the questionnaire sheets and sign consent forms, ask questions etc. The entire testing time took for a single subject was approximately 1 hour. This comprised of 250 seconds per test sequence (i.e. reference sample, break, sample under test and voting) as shown in Fig. 4 below. In addition, at the mid-point of the test, each assessor was given a ten-fifteen minute break to address any concerns over olfactory adaptation or assessor fatigue. Assessors were permitted to drink water at any time during the testing period.

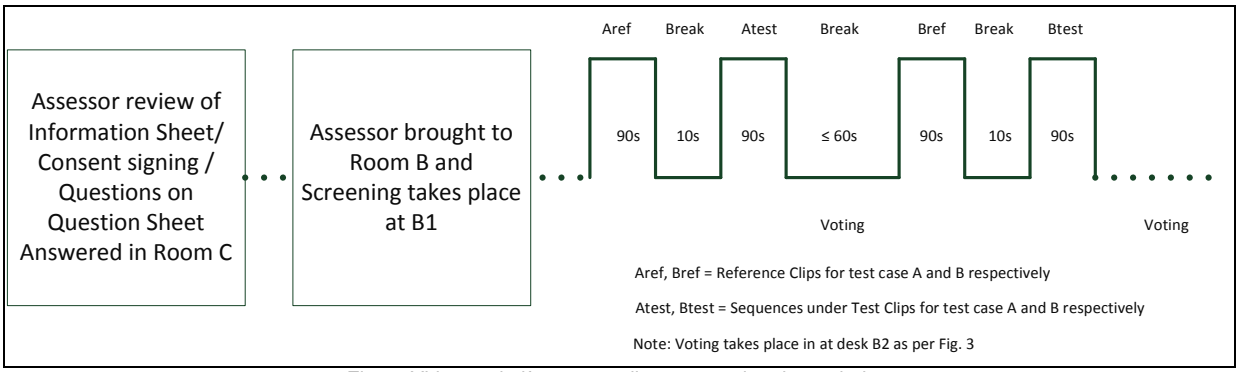

Fig. 4. Video and olfactory media presentation times during

5.1 Questionnaire and rating scale

A number of approaches exist in the literature for offline subjective evaluations of multimedia applications. The absolute category rating (ACR) method proposed in BT.500 (ITU-T BT.500, 2002) requests participants to provide a rating score from 1 to 5 (5 being best) after observing a single sample. With this approach, there is no reference sample and scores are given based on user expertise. This leads to non-uniform distributions of rating scores, which can invalidate subjective results (Huang and Nahrstedt, 2012). Specifically in relation to olfactory media, considering the variable perception of olfactory media, this issue is exaggerated. In addition, feedback from assessors during preliminary testing indicated that the "novelty" of olfactory media made even large errors temporarily acceptable. (ITU-T P. 910, 2008) proposes an alternative assessment method to address the reliance on assessor expertise by exposing participants to two media samples of different qualities and giving a comparative rating score. The first stimulus presented in each pair is always the source reference, while the second stimulus is the stimulus under test. This method is known as Degradation Category Rating (DCR) or Double Stimulus Impairment Scale method. To address the two issues highlighted above, this method was selected for the subjective testing. The reference sample was always a synchronized presentation of olfactory and video media. As per [ITU-T P.9.10], assessors were told that they would be presented with each olfaction enhanced clip twice and that the first time they saw each clip it was "the reference sample". They were told that the second time they saw each clip it was the "sample under test". They were requested to "answer the questionnaire on their experience of the sample under test", and to base their judgments on their overall experience of the olfaction enhanced clips compared to the reference clip using the wordings available on each of the scales in table 3.

The samples being tested included inter-media skew of varying degrees (shown in table 4) as well as the synchronized presentation of olfactory and video media. For all questions, assessors chose one answer from the Likert scale, shown in table 3. The questions used have evolved from those asked in (Steinmetz, 1996) and (Ghinea and Ademoye, 2010). As part of the preliminary testing, a reliability assessment was performed on the questionnaire to ascertain if the purpose and phraseology were clear and comprehendible to assessors. Discussion with each assessor was undertaken, feedback was recorded and

necessary amendments were made to the draft questions. The questions explained in the remainder of this section are the final versions updated after feedback comments were considered and following a review by a Psychologist. Assessors were asked to select one of the five possible answers per question as per table 3 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, "*Relative to the content of the video clip, the smell was released:*". Assessors answered by selecting one of the five possible answers as shown under statement 1 in table 3. Question 2 aimed to determine how **tolerant** assessors were to different levels of skew. Hence they were asked to qualify their annoyance of the inter-media skew by answering; "*In the event that you may have perceived the video clip and smell being out of sync, please indicate the extent to which it impacted upon you. Please select the appropriate option below that reflects how you would qualify it?*" As per answers for question 2 in table 3, assessors had the option of selecting one of five values that reflected how they perceived the synchronization error (if it existed) in terms of its annoyance. The mean opinion score (MOS) of respondents was used to determine the tolerable level of skew as well as deriving a level of annoyance graph considering age, sex and culture as shown in Fig. 11-14.

Table 3 Rating scales for each of the statements/questions (Likert Scale)

Score	Statement 1	Question 2	Statement 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

The final three statements were included to analyze the impact of inter-media skew on the user experience. Assessors were asked to select one of five possible answers in terms of their agreement with the statements. The statements were ordered from general to being more specific. To determine the impact of inter-stream skew, assessors' agreement with "*You enjoyed watching the video clip*" evaluates assessor level of **enjoyment** of olfactory data as a media when in sync and explores any deterioration in this perception with the introduction of inter-media skew. "*The smell when presented, was relevant to what I was watching"* queried the **relevance** olfactory media had to the video when skews existed as opposed to synchronized presentation. By examining the assessors' agreement with "*The smell contributed to a heightened sense of reality whilst watching the video clip*", the aim was to determine the impact the level of skew has on assessors' sense of **reality** of an olfaction enhanced multimedia clip.

5.2 Introduction of skews between olfactory and visual media

In order to determine the perceptible and tolerable levels of inter-media skew between olfactory and video media, assessors were presented with varying levels of skew (including no skew) and queried about their perception of the experience. The audio from these video clips was removed using Windows Live Moviemaker and replaced with the sound of a blowing fan. The sound was added to negate the influence of the noise of blowing from the SBi4 v2 (which differed depending on which fan was running). Table 4, shows the skews introduced for each of the video clips and how it was divided across participants. Once the presentation of the olfactory media was complete, a SBi4 fan with no odor cartridge was turned on to address scent lingering. Fig. 5 shows how olfactory media is presented at different times relative to the video time axis. For olfactory media to be in sync (0s skew) with the video, it should be presented for the middle 30s block (i.e. from time 30s to time 60s on the video presentation time axis). Olfactory data before 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.

Table 4. Case 1 applies to participants 1, 14, 27, etc., case 2 applies to participants 2, 15, 28 and so on

Case	Clip 1 Skew	Clip 2 Skew	Clip 3 Skew	Clip 4 Skew	Clip 5 Skew	Clip 6 Skew	Clip 1 Skew	Clip 2 Skew	Clip 3 Skew	Clip 4 Skew	Clip 5 Skew	Clip 6 Skew
1	0s	-30s	-25s	-20s	-15s	-10s	-5s	+5s	+10s	+15s	+20s	+25s
2	+30s	0s	-30s	-25s	-20s	-15s	-10s	-5s	+5s	+10s	+15s	+20s
3	+25s	+30s	0s	-30s	-25s	-20s	-15s	-10s	-5s	+5s	+10s	+15s
4	+20s	+25s	+30s	0s	-30s	-25s	-20s	-15s	-10s	-5s	+5s	+10s
5	+15s	+20s	+25s	+30s	0s	-30s	-25s	-20s	-15s	-10s	-5s	+5s
6	+10s	+15s	+20s	+25s	+30s	0s	-30s	-25s	-20s	-15s	-10s	-5s
7	+5s	+10s	+15s	+20s	+25s	+30s	0s	-30s	-25s	-20s	-15s	-10s
8	-5s	+5s	+10s	+15s	+20s	+25s	+30s	0s	-30s	-25s	-20s	-15s
9	-10s	-5s	+5s	+10s	+15s	+20s	+25s	+30s	0s	-30s	-25s	-20s
10	-15s	-10s	-5s	+5s	+10s	+15s	+20s	+25s	+30s	0s	-30s	-25s
11	-20s	-15s	-10s	-5s	+5s	+10s	+15s	+20s	+25s	+30s	0s	-30s
12	-25s	-20s	-15s	-10s	-5s	+5s	+10s	+15s	+20s	+25s	+30s	0s
13	-30	-25s	-20s	-15s	-10s	-5s	+5s	+10s	+15s	+20s	+25s	+30s

6. RESULTS AND DISCUSSION

The results of the subjective testing are presented and the definition of temporal boundaries for olfactory and video synchronization is explained.

6.1 Preliminary Experiment: Measurement of the Detection Instant

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. Each

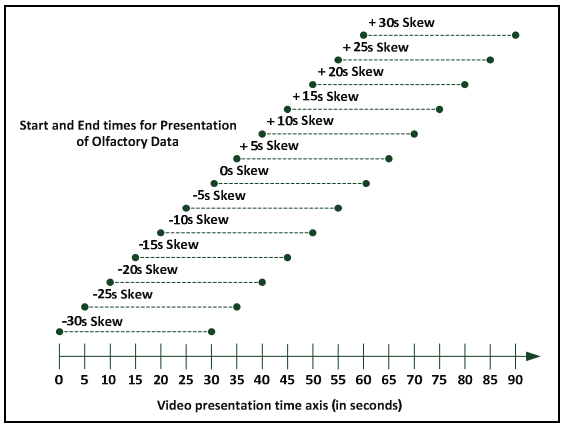

Fig. 5. Video and olfactory media presentation times during subjective testing.

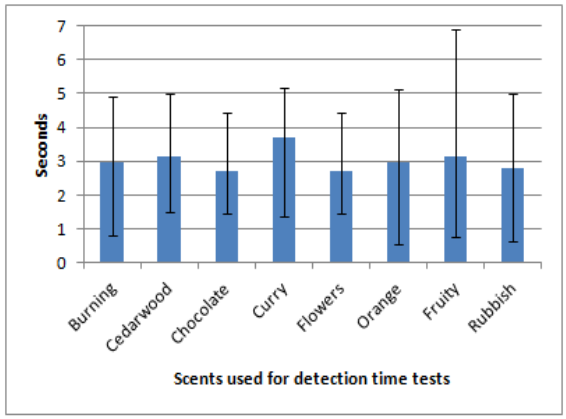

Fig. 6. Detection instant per scent average and maximum/minimum detection instants per scent.

of these 15 participants later took part in the full tests. 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 olfaction device's fans should be started in order to ensure timely presentation to the users. With 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 as per Fig. 6. For each scent, the average time it takes an assessor to detect it is taken into account in terms of presenting the scent according to the

above mentioned skews i.e. if it takes 3 seconds for a burning scent to reach the assessor, the fan to emit the scent is turned on at time 27s such that the scent reaches the assessor at time, t = 30s and as such is said to be synchronized with the video.

6.2 Detection and Perception of Error considering Gender, Age and Culture

Fig. 7 gives a general overview of the results of statement 1, to determine users' ability to detect levels of inter-media skew. The vertical axis shows ratings related to the five possible answers to question one i.e. when the scent arrived relative to the video. The horizontal axis indicates the level of skew artificially introduced between the olfactory and video media with the negative values representing olfactory media before video media. Fig. 7 shows assessors were able to identify the existence of inter-stream skew very well. It also indicates that assessors were more sensitive to scent that was early rather than late based on the comparison of skews before and after time 0s. Direct comparison of MOS scores at skews of +5s and -5s show that the MOS for +5s of 3.35 was much closer to being at the "correct time" (represented by a value of 3) as opposed to the value of 2.05 for -5s. Interestingly based on MOS comparison, assessors viewed skews of +10s and -5s similarly in terms of being Late or Early respectively. In order to analyze if significant differences existed in participants' perception between synchronized and unsynchronized scent and video, the data collected was analyzed using paired sample t-test. With 99% confidence level, the t-tests showed for all levels of skew between the olfactory data and video that the significant two tailed p values were less than 0.01 ($p<0.01$) (+5s had a two tailed p value of 0.00032, -5s had a two tailed p value of 0.0000002) and hence it can be concluded that there is a statistically significant difference between the mean of the synchronized and mean of participant responses for the "skewed" release times.

Fig. 8 and Fig. 9 and Fig. 10 present responses to statement 1 considering culture, gender and gender\age respectively. Fig. 8 present the results of comparison based on assessor culture. All three main groups (Europe, Africa and Asia) detect the presence of inter-media skew between olfaction with the European group showing the greatest sensitivity to skew for both olfaction before and after video. Comparing sensitivity between the African and Asian group shows the African group the least sensitive for 8 out of the 13 skews tested with minimal difference for 3 out of the remaining 5 skew. To determine if statistical significant results exist between each of the three culture groups, a one way ANOVA between-groups with post-hoc tests were performed. When olfaction was presented before video, at skew level of -5s, analysis revealed that a statistically significant difference exists between the European and African group ($p=0.049$, $p<0.05$). With skew of -10s, the European and Asian groups reported statistically significant results ($p=0.007$, $p<0.05$). At skew level of -15s and -25, there were statistically significant differences between the African and both the European ($p=0.005$, $p<0.05$) ($p=0.030$, $p<0.05$) and Asian ($p=0.024$, $p<0.05$) ($p=0.025$, $p<0.05$) groups. For olfaction presented after video, again a number of statistically significant results are reported at various skews between the three cultural groups. At skew level of +5s, there were statistically significant differences between the African and both the European ($p=0.012$, $p<0.05$) and Asian ($p=0.003$, $p<0.05$) groups, whilst at +10s significant differences exist between the European and both the African ($p=0.047$, $p<0.05$) and Asian ($p=0.045$, $p<0.05$) groups. At +20s, a statistically significant difference exists between the European and the African groups with two tailed p value of $p=0.034$, $p<0.05$. Finally at +30s, there were statistically significant differences between the African and both the European ($p=0.002$, $p<0.05$) and Asian ($p=0.004$, $p<0.05$) groups. For the remaining combinations, the results were found to not be statistically significant between the culture groups. It is valid to conclude considering Fig. 8, 9 and 10 that the females are more (or equally) sensitive to inter-media skew than men, and younger users than older assessors for all skew levels, with the European group the most sensitive of skew, followed by the Asian and African groups.

Fig. 9 shows both gender groups identify the existence of inter-stream skew accurately. Worth noting is that both sexes better accept olfactory stimuli after, rather than before, the corresponding video event. Considering the MOS scores are closer to 3 (at the correct time) when olfaction is presented after video (MOS scores for +5s is 3.28(male), 3.42(female)) than olfaction before video (MOS score for -5s is 2.21 (male), 1.71 (female)), we conclude that both groups prefer olfaction after video than before. It is also clear from Fig. 9 that the female group is more sensitive to skew with olfaction before and after video. 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 -10s and +10s where the values were very similar. Interestingly based on MOS comparison, male assessors viewed skews of between +15s to +10s and -5s similarly in terms of being late or early. A similar result was obtained for female assessors when scoring

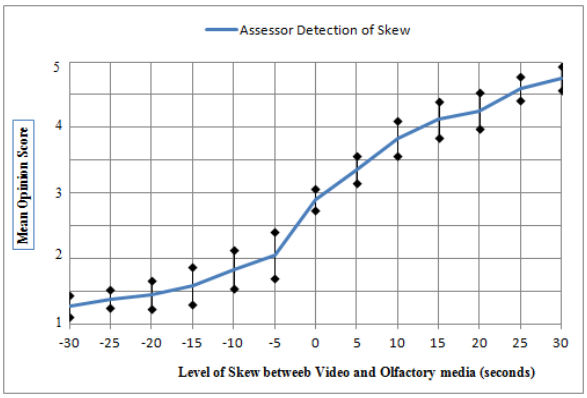

Fig. 7. Analysis of skew detection with confidence interval based on 99% confidence level.

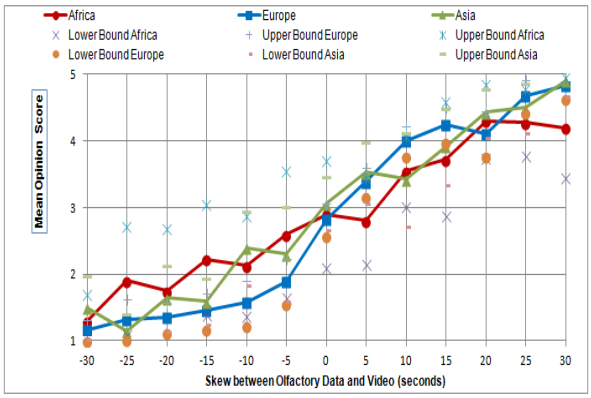

Fig. 8. Culture based analysis of detection per skew with confidence intervals based on a 99% confidence level.

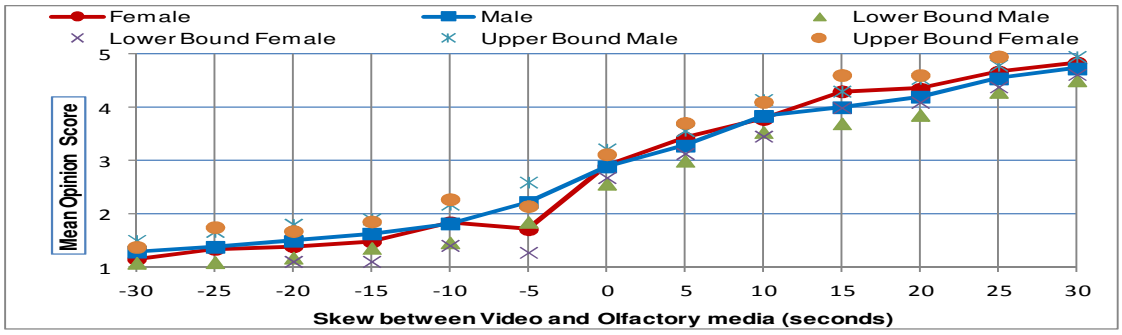

Fig. 9. Gender analysis detection of skew with confidence intervals based on a 99% confidence level.

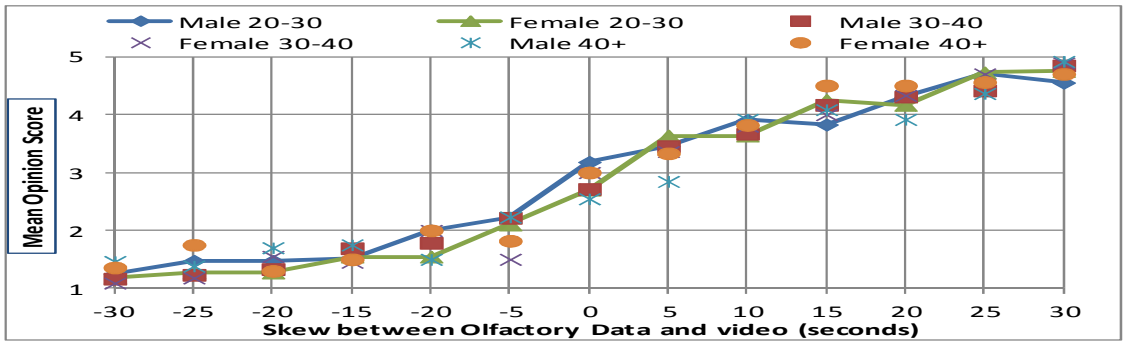

Fig. 10. Gender/age analysis of detection per skew

skews of +15s and -5s. To determine if statistically significant differences existed between genders an independent-samples t-test was conducted. Based on 95% confidence interval, statistically significant differences existed for skews of -5s ($p=0.027$, $p<0.5$). For all other test sequences, results were not statistically significant between gender groups.

Fig. 10 shows that there is a general consistency across all age groups in terms of their ability to identify inter-media skew. Comparison of the male and female groups the 20-30 yrs, the female group are generally more sensitive to skew, whether olfaction is before or after video (with the exception of +10s and +20s). The results show that both these groups generally are noticeably more sensitive than their corresponding 40+ yrs male group at all levels of skew (Male 40+ yrs group are the least sensitive). In terms of comparison within the 30-40 yrs male female group, the female groups is equally or more sensitive. Again these two groups report a greater sensitivity to skew than the corresponding 40+ yrs male/female group, with the male 30-40 yrs and male 40+ alternating in terms of this sensitivity across various skew levels. To determine if statistical significant results exist between each of the six gender-age groups, a one way ANOVA between-groups with post-hoc test was performed. Analysis revealed that a

statistically significant difference exists between the 20-30 yrs and 40+ yrs male groups when synchronized presentation took place ($p=0.020$, $p<0.05$) and when olfaction was presented +5s after video ($p=0.013$, $p<0.05$). The 40+ male group also reported statistically significant results at for the same skew level with the 20-30 female ($p=0.006$, $p<0.05$). At skews of +15s, the male 20-30 yrs group reported statistically significant results from the 40+ yrs female group ($p=0.041$, $p<0.05$).

The task of question 2 was to determine the effect that the skews had on the perception of the olfactory-video clip. The effect an error has is key to determine temporal boundaries, as works involving other media have shown that users can tolerate certain levels of skew (Steinmetz, 1996) (Ghinea and Ademoye, 2010). Hence, assessors were asked to qualify the level of impairment the inter-media skew had on the experience when comparing it to the synchronized reference sample. Fig. 11 shows the MOS for level of annoyance for inter-media skew. Scores of 1 or 2 or 3 represent the perception of the skew being annoying with scores of 4 and 5 representing a tolerable skew level and imperceptible error respectively. Based on the comparison of MOS scores with olfaction before and after video, that assessors are **less tolerant** to olfaction **before** of video than they are of olfaction **after** video between skews of -20s to +20s. Skews of -5s and +10s are viewed similarly by assessors with more gradual reduction in MOS for olfaction after video than olfaction before video up to skews of -20s to +20s. As was the case with data collected for statement one, a paired sample t-test was run with a confidence interval of 99% assessing significant differences in assessors' perceptions during the perfectly synchronized case and the cases where skews existed. For skew size of +5s, the results between an in-sync and skewed presentation were found not to be statistically significant, with the significant two-tailed p value of 0.813 which is greater than 0.01 ($p>0.01$). For all other levels of skew between the olfactory data and video that the significant two tailed values were less than 0.01 ($p<0.01$) (-5s had a two tailed p value of .000008, +10s had a two tailed p value of .000005), hence it can be concluded that there is a statistically significant difference between the mean of the case of synchronized presentation in comparison with the "skewed" presentations except skew of +5s.

Fig. 12 and Fig. 13 show the level of annoyance different inter-media skews have on users of various gender and gender-age groups. Significant changes in MOS between olfaction after and before the video for both sexes occur between skews of -15s and +15s. Female assessors perceived skew of -15s as being between annoying and slightly annoying. Comparing the male rating for 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 -10s and -5s skews is performed. The female group members are much more sensitive and less tolerant at these skew levels than their male equivalents. Another interesting result occurs at skews of 5s. For the female group, scent presented -5s ahead of video has the same scoring rating as when scent is presented +15s 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. Although not as exaggerated, the male group are also more sensitive to olfaction before video than olfaction after video. 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. To determine if statistically significant differences existed between gender groups an independent-samples t-test was conducted. Based on 95% confidence interval, the analysis reported no statistically significant differences existed for skews between the groups, for -5s ($p=0.071$, $p>0.05$). Fig. 13 presents the analysis considering gender and age within the 20-30 yrs, 30-40 yrs and 40+ yrs 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 closer to slightly annoying range for the 20-30 females. From Fig. 12, it is clear that the 40+ males and females are more acceptable to scent before video than other groups, and found -10 s skews perceptible but not annoying, 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 negatively affected by skew than the two older groups.

To determine if statistical significant results exist between each of the six gender-age groups, we performed a one way ANOVA between groups with post-hoc tests. Analysis revealed that statistically significant difference exists between the male 30-40 yrs and female 20-30 yrs group at a skew level of -5s ($p=0.033$, $p<0.05$). At skews of olfaction -10s before video, the male and female 40+ yrs groups differed statistically ($p=0.038$, $p<0.05$) and for olfaction -30s before video, the two tailed p-value was $p=0.025$,

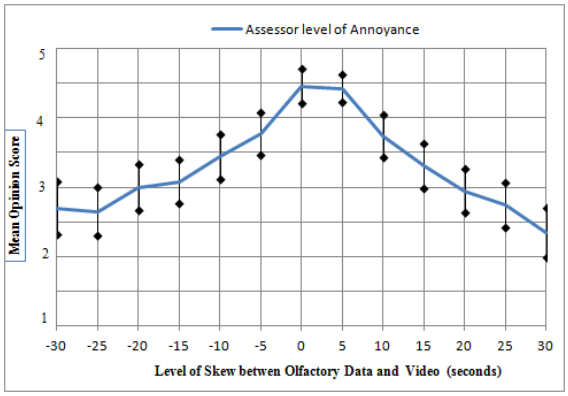

Fig. 11. Analysis of annoyance level per skew with confidence based on 99% confidence level.

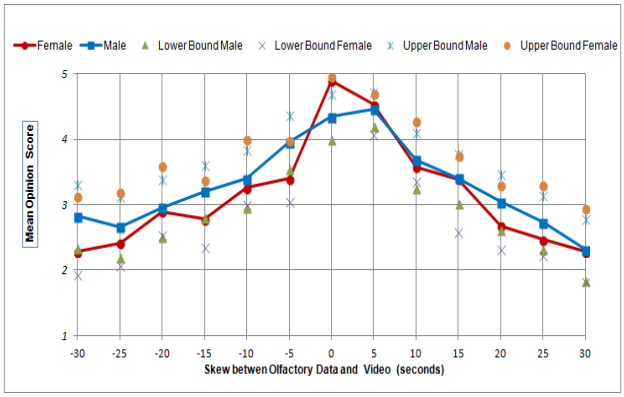

Fig. 12. Gender analysis of annoyance level per skew with interval Confidence intervals based on a 99% confidence level.

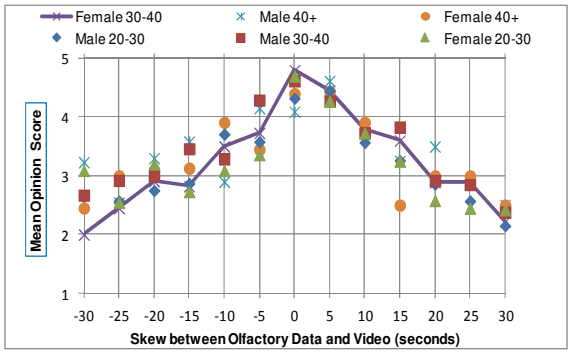

Fig. 13. Gender/Age Analysis of Annoyance Level per Skew

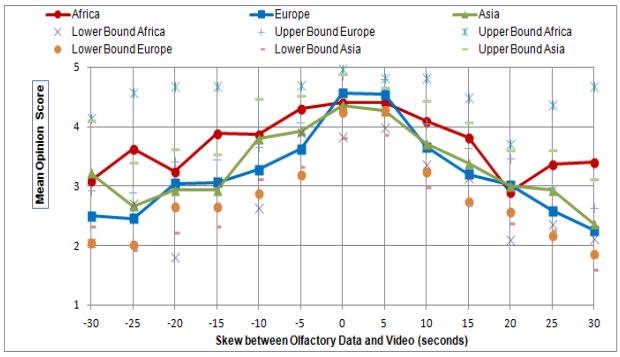

Fig. 14. Culture based analysis of Annoyance Level per Skew with confidence intervals based on a 99% confidence level

between the 30-40 female group and the male 40+ yrs group. For olfaction after video, at skew level of +15s, the male 30-40 yrs group had a statistically significant difference from the female 40+ age group ($p=0.009$, $p<0.05$). At the same skew size the female 30-40 yrs group were also different from the female 40+ group ($p=0.035$, $p<0.05$). At skew sizes of +20s, the male 40+ group mean was statistically different from the 20-30 female group ($p=0.040$, $p<0.05$). For all other skew sizes the differences between the gender-age groups in terms of their perception of the olfaction enhanced multimedia were not statistically significant.

From Fig. 14, there are clear differences again between the groups considering culture. The African group deemed skews in the range of -15s to +15s as perceptible but not annoying whereas the Europe group rated skews outside of the -5s to +10s range to be annoying. The Asian group scored skews in the 10s to +10s to be between perceptible but not annoying to annoying. These results for annoyance are broadly consistent with the detection of skew ratings discussed above. To determine if statistical significant results exist between perceptions of each of the three culture groups, a one way ANOVA between-groups with post-hoc test was performed. With skew of -15s and -25s, the African groups reported statistically significant results with both European ($p=0.031$, $p<0.05$) ($p=0.004$, $p<0.05$) and Asian ($p=0.014$, $p<0.05$) ($p=0.024$, $p<0.05$). For olfaction presented after video, again a number of statistically significant results are reported at various skews between the three cultural groups. At skew level of +25s, there were statistically significant differences between the African and the European ($p=0.037$, $p<0.05$) group. At +30s, a statistically significant difference exists between the African and both the European and Asian groups ($p=0.005$, $p<0.05$) and ($p=0.011$, $p<0.05$) respectively. For the remaining test sequences, the results were found to not be statistically significant between the culture groups.

6.2.1 Detection and perception of synchronization error discussion

Findings from Fig. 7 support the plethora of literature that recounts human sensitivity to, and ability to, detect scent. If we consider and compare the MOS values for corresponding skews i.e. -30s with +30s, -25s with +25s, -20s with +20s, -15s with +15s, -10s with +10s and -5s with +5s generally and in terms of

age, sex and culture, interesting conclusions can be drawn. Assessors are accurately able to detect the presence of inter media skew between olfaction and video. The female group is more sensitive to olfaction before and after video than the male group. The younger male groups had broadly similar scores with levels of greater sensitivity with both being noticeably more sensitive than their corresponding 40+ male group at all levels of skew. A similar trend exists between the younger female groups and the older group. Considering culture, the European group shows the greatest sensitivity to skew for both olfaction before and after video with the African group the least sensitive. All groups indicate that olfaction **before** video is more noticeable than olfaction **after** video.

In terms of rating the impairment caused by the existence of a synchronization error, Fig. 11-14 details the assessor annoyance at varying levels of skews. Assessors rate olfaction before video more annoying than olfaction after video. The female group is more sensitive to skew with olfaction before video than the male group, with both groups reporting similar annoyance to skew with olfaction after video. The younger female group is the most sensitive to skew, with the male (20-30 yrs and 30-40 yrs) and female (30-40 yrs) group similar in terms of the their rating of skews. The two older groups are the most tolerant to skew. In terms of defining the temporal boundaries for synchronizing olfactory and video media, these results reveal "in-sync" and "out-of-sync" regions. 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, age and culture (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-sync region spans between a maximum skew of 0s to -5s/-15s 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 and culture 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 and culture of the user.

The in-synch region is based on the range where assessors perceive errors to be not annoying (rating above 3.5) as per Fig. 11-14. Considering the assessor tolerance to olfaction **after** video as opposed to olfaction before video, the span is larger for olfaction after video. This is plausible, in our everyday lives; we see first, smell later. Based on these findings, user perceived synchronization of olfaction enhanced multimedia OeM[s] can be expressed as per equation 1, Where $S_{s,i}$ represents skew size (s) and also whether olfaction is before or after video (i), P represents perception of the assessor based on age of the assessor (A), gender of the assessor (G) and culture of the assessor (C). The age category A, is defined as having 3 sets: 20-30 yrs, 30-40 yrs and 40+ yrs. The gender category G, has two groups: male and female. The culture group is defined based on continent. OeM[s] in its current form represents results to date that the level of user perceived experience is variable based on inter-media skew as well as age, sex and culture of assessors.

$$OeM\ [s] = f\left(S_{s,i}, P(A, G, C)\right) \qquad (1)$$

6.3 Impact of Synchronization Error on User Experience considering age, sex and culture

This section analyzes the impact inter-media skews have on the QoE of an olfaction-enhanced video clip. Assessors were asked to choose one of five levels of agreement with statements 3, 4 and 5 already outlined in section 5.1. The MOS results presented in the following sections compare assessors' experience of an olfaction-enhanced video clip with the existence of skew as against the case of synchronized presentation in terms of assessor enjoyment, sense of relevance and sense of reality.

6.3.1 Impact on Sense of Enjoyment considering age, sex and culture

Figure 15 shows for MOS reflecting assessor's level of agreement with statement 3 in the presence of varying degrees of inter-media skew. When synchronized presentation takes place, assessors agreed that they enjoyed watching the video clip. In the presence of large skews (e.g. -30 s or +30 s), the MOS values show that generally participants were between "Neither agree or Disagree" and "Disagree" with statement 3 in the presence of large skew. Assessors indicated higher levels of enjoyment for olfaction presented after video compared with olfaction before video. Interestingly, olfaction presented at skew size of +5s

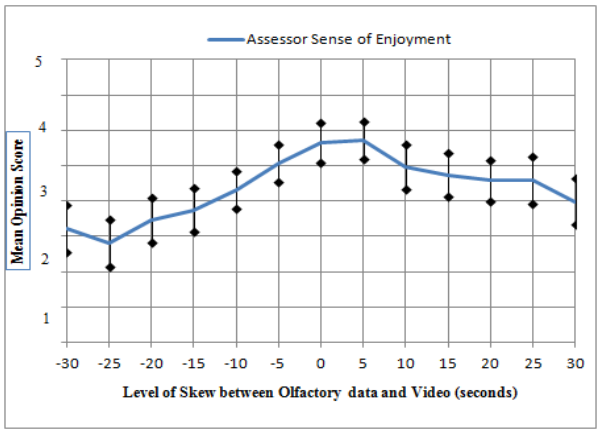

Fig. 15. Analysis of sense of enjoyment per skew with confidence interval based on 99% confidence level.

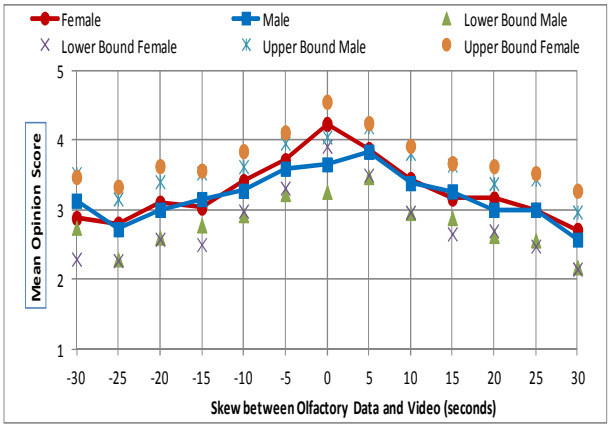

Fig. 16. Gender analysis of enjoyment level per skew with confidence intervals based on a 99% confidence level

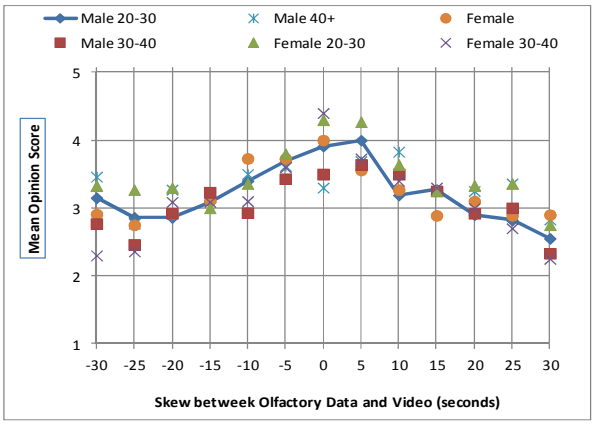

Fig. 17. Gender/age analysis of enjoyment level per skew

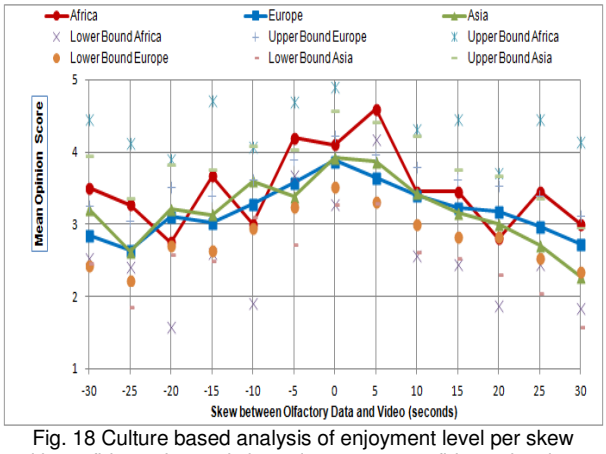

Fig. 18 Culture based analysis of enjoyment level per skew with confidence intervals based on a 99% confidence level.

achieved a higher enjoyment score that olfaction presented “in-sync”. Assessors’ opinions were compared to determine if statistically significant differences exist in responses between means when skews were present as against the case when olfaction and video were synchronized (via a paired sample t test with a confidence interval of 99%). For skew size of +5s and -5s, the results between an in-sync and skewed presentation were found not to be statistically significant, with the significant two tailed value of 0.713 and 0.038 respectively, both greater than 0.01 ($p>0.01$). For all other skew sizes the results were statistically significant from the synchronized case with the significant two tailed values of 0.001 and 0.001 for skews of +10s and -10s respectively which are less than 0.01 ($p>0.01$). From Fig. 16, when synchronized presentation takes place, the female assessors enjoyed watching the video clip more than the males. Interestingly, the female group reports a higher sense of enjoyment for olfaction before video compared to the male group. They scored higher in terms enjoyment for every skew from -25s to +5s with the exception of -15s. Both the male and female group report similar experiences of enjoyment for skews of +5s and +10s with the male group the least sensitive to olfaction after videos for skews from +15s to +30s. To determine if statistically significant differences existed between genders an independent-samples t-test was conducted. Based on 95% confidence interval, the analysis reported statistically significant differences existed for synchronized presentation between the groups with a two tailed $p=0.008$, $p<0.05$. For skewed presentations at all other levels, the differences were not statistically significant. Fig. 17 presents an illustrative analysis considering both gender and age. Skew size affects the sense of enjoyment of olfaction enhanced media for all groups with the male 40+ yrs group least affected. In comparing the views of the various age groups, the impact of skew is greatest for the female 30-40 yrs category, especially with olfaction ahead of the video, as their MOS scores are the lowest. The 20-30 yrs female group reported consistently higher scores for enjoyment with the male 30-40 yrs groups the lowest. 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 yrs category are least impacted by small skews. To determine if statistically significant results exist between each of the six gender-age groups in terms of their enjoyment, we performed a one way ANOVA between-groups with post hoc tests. An analysis revealed that a statistically significant difference exists between the 30-40 yrs male group and the 20-30 yrs female group ($p=0.037$, $p<0.05$) and the 30-40 yrs female group ($p=0.020$, $p<0.05$) and between the 40+ yrs male groups and the 20-30 yrs female group ($p=0.016$, $p<0.05$) and the 30-40 yrs female group ($p=0.009$, $p<0.05$) when synchronized presentation takes place. At skew levels of olfaction -10s before video, statistically significant differences exist between the male 30-40 yrs and the female 40+ yrs group ($p=0.033$, $p<0.05$). At skew levels of olfaction -30s before video, statistically significant differences exist between the male 40+ yrs and the female 30-40 yrs group ($p=0.017$, $p<0.05$) and between the female 20-30 and 30-40 yrs groups ($p=0.037$, $p<0.05$). For olfaction after video, no statistically significant differences existed. For all other skew sizes the differences between the groups were not statistically significant.

Fig. 18 compares sense of enjoyment considering culture with the African group reporting the highest sense of enjoyment across numerous skews with the European group consistently scoring the lowest sense of enjoyment rating. To determine if statistical significant results exist between enjoyment of each of the three culture groups, a one way ANOVA between-groups with post-hoc test was performed. With skew of +5s, the African group reported statistically significant results with the European group ($p=0.002$, $p<0.05$). For the remaining test sequences, the results were found to not be statistically significant between the culture groups.

6.3.2 Impact on Sense of Relevance considering age, sex and culture

To determine the impact of skew on the perceived sense of relevance, assessors were required to express their level of agreement with statement 4 as outlined in section 5.1. Fig. 19 shows the MOS reflecting assessors' 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 was relevant to what they were watching. Also, it is clear that in the presence of large skews, e.g. -30 s or +30 s the sense of relevance is affected as the MOS values show that generally participants neither agreed nor disagreed with statement 4. Interestingly, the level of agreement deteriorated much more quickly as the levels of skew increased with olfaction before video as opposed olfaction after video. Fig. 19 shows a sharp increase in sense of relevance between skews of -15s to -10s. Olfaction after the video is more relevant than olfaction before video. Olfaction with a skew of +15s (15s after video) is viewed as being very similar to olfaction presented with a skew of -5s. The slow gradual decrease in relevance for olfaction after video is particularly interesting when it is compared with the sharp decrease the larger the skew size when scent is before video. A paired sample t-test with a 99% confidence interval was used to determine if statistically significant differences existed in responses between synchronized and unsynchronized presentations of scent and video. For skew size of +5s, +10s, the results were not significantly different with the significant two tailed values of 0.556, 0.092, respectively ($p>0.01$). For skew size of -5s and -10s the results were not significantly different with the significant two tailed values of 0.103 and 0.014 respectively ($p>0.01$). For all other skews the significant two tailed values were less than 0.01 ($p<0.01$), (-15s had a two tailed p value of 0.0000001, and +15s had a value of 0.008).

Fig. 20 shows the male and female assessors' level of agreement with statement 4. Subjected to synchronized presentation (0s skew), both female and male assessors found the smell relevant to what they were watching. The female group considers skews with olfaction ahead of video having a more negative impact on relevance as they have awarded consistently lower MOS scores up to skews of -10s. Thereafter with olfaction after video, the female group found the scent more relevant for all skew values between -10s to +30s. For both sexes, olfaction after the video is more relevant than olfaction ahead of video. For both male and female groups, they view olfaction +20s after the video and olfaction -10 seconds ahead of video with the same ratings. The slow gradual decrease in relevance for olfaction after video is another interesting finding of this work. To determine if statistically significant differences existed between genders an independent-samples t-test was conducted. Based on 95% confidence interval, the analysis reported statistically significant differences between the male and female groups existed for -20s skew with two tailed value for $p=0.016$, $p<0.05$. For skewed presentations at all other levels, the findings were not statistically significant ($p=0.053$ for +30s skew, $p=0.070$ for skew of -15s being the only other sequences where differences were close to being statistically significant). Considering the age categories (Fig. 21), all female age categories demonstrate comparatively higher relevance values for all skew levels. The male

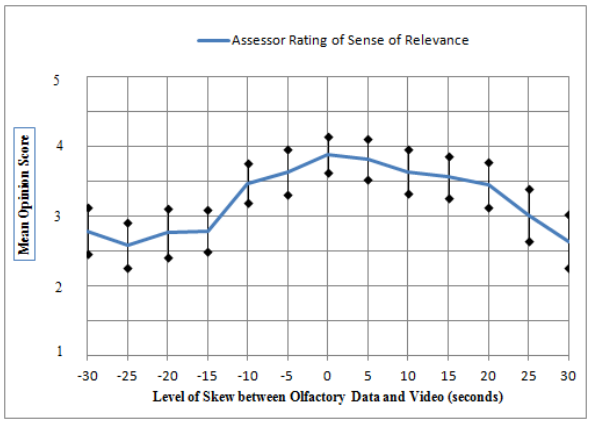

Fig. 19. Analysis of relevance per skew with confidence interval based on 99% confidence level.

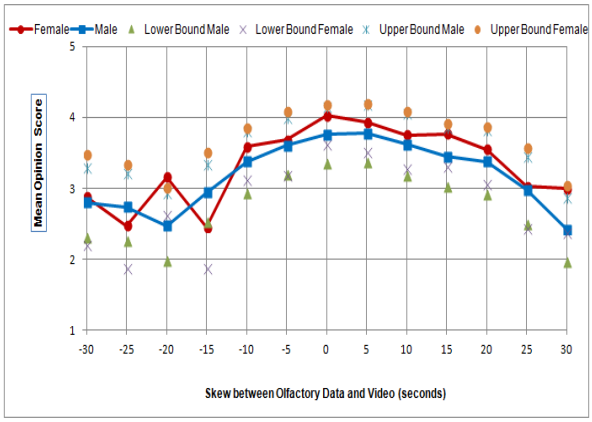

Fig. 20 Gender analysis of relevance level per skew with confidence intervals based on a 99% confidence level.

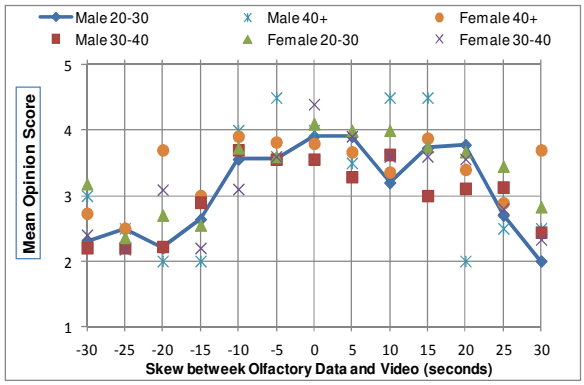

Fig. 21 Gender/age analysis aware of relevance level per skew

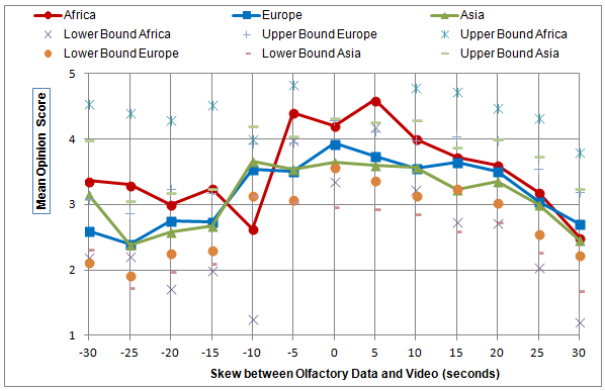

Fig. 22 Culture based analysis of relevance level per skew with confidence intervals based on a 99% confidence level.

30-40 yrs category presents lower values of relevance across a wide range of skews than the male 20-30 yrs group and furthermore the male 40+ group reports very high relevance scores at skew levels -5s, +10s and +15s. To determine if statistically significant results exist between each of the six gender-age groups in terms of their reality, we performed a one way ANOVA between-groups analysis with post hoc test. Analysis revealed statistically significant differences between the 20-30 yrs and 30-40 male groups ($p=0.032$, $p<0.05$) and the 30-40 yrs male and female groups ($p=0.009$, $p<0.05$) when synchronized presentation takes place. At skew levels of -20s, statistically significant differences exist between both the male 20-30 and 30-40 yrs groups and the female 40+ yrs group ($p=0.016$, $p<0.05$), ($p=0.002$, $p<0.05$). In addition, the male and female 30-40 yrs groups also report statistically significant results at this skew level ($p=0.045$, $p<0.05$). At skew levels of olfaction -30s before video, statistically significant differences exist between the male 40+ yrs and the male 30-40 yrs group ($p=0.021$, $p<0.05$). For olfaction after video, statistically significant differences exist at skews of +5s ($p=0.04$, $p<0.05$). At skews of +30s, the 20-30 and 30-40 male groups were statistically significant from the female 40+ group ($p=0.03$, $p<0.05$), ($p=0.008$, $p<0.05$). Also at this skew level, the female 30-40 yrs and 40+ yrs groups reported statistically significant results ($p=0.017$, $p<0.05$). For all other skew sizes the differences between the groups were not statistically significant.

Considering cultures, the African group reports the greatest sense of relevance for the olfaction-enhanced clips across all skew levels bar two. The Asian and European groups indicate similar levels of relevance with the European group scoring slightly higher in terms of relevance considering 7 out of the 13 skews tested. To determine if statistical significant results exist between relevance of each of the three culture groups, a one way ANOVA between-groups with post-hoc test was performed. For olfaction presented before video, with skew of -5s, the African group reported statistically significant results with the European ($p=0.015$, $p<0.05$) and Asian ($p=0.051$, $p<0.05$). With skew size of -10s and -25s, the African group reported statistically significant results with the European ($p=0.031$, $p<0.05$) ($p=0.036$, $p<0.05$). For

olfaction presented after video, with skew sizes of +5s the African group mean was statistically significant from the European ($p=0.013$, $p<0.05$) and Asian ($p=0.006$, $p<0.05$) groups. For the remaining test sequences, the results were found to not be statistically significant between the culture groups.

6.3.3 Impact on Sense of Reality considering age, sex and culture

Fig. 23 shows MOS scores reflecting assessor's level of agreement with statement 5 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. However, again in the presence of large skews (e.g. -30 s or +30 s), the MOS values show that generally participants were between "Neither agree or Disagree" and "Disagree" with statement 5 in the presence of skew. For olfaction before video, with skews of -10s and -5s, assessors perceived the olfaction somewhat contributing to an enhanced sense of reality with a MOS between "Neither agree or Disagree" and "Agree". With no skew, 0s, and +5s skew assessors "Agree" that the scent contributed to a heightened sense of reality. The most interesting finding from analysis of the MOS from this question was the **slow reduction** in heightened sense of reality for olfaction **after** scent with skews of +5s, +10s and +15s. Just 0.15 separates the rating of +5s from +15s. When comparing the ratings of scent before or after the video, the opinions in terms of impact on reality were that not much difference existed between scent -5s before video and +15s. Assessors' opinions were compared to determine if statistically significant differences existed in responses when a synchronization error was present with the case when olfaction and video were perfectly synched via a paired sample t test. For skew size of +5s, +10s and +15s, -5s, -10s the results were not significantly different with the significant two tailed values were 0.485, 0.045, 0.050, 0.012 and 0.011 respectively which are greater than 0.01 ($p>0.01$). For all other skews the significant two tailed values were less than 0.01 ($p<0.01$) (+20s had a significant two tailed value of 0.001, -15s had a significant two tailed value of 0.00001), hence the results for skewed presentation were statistically significant when compared with ratings for in-sync presentation. Fig. 24 shows MOS scores reflecting male and female assessor's level of agreement with statement 5 in the presence of varying degrees of inter-media skew. When synchronized presentation takes place, the female group in particular strongly agreed that the smell heightened the sense of reality of what they were watching. 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 5. 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. For skew levels between -5s and +30s the female group indicated a higher sense of reality. 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. The trend is not as exaggerated for the male group, but interestingly the male group reported a higher sense of reality with a skew of +5s as opposed to 0s. To determine if statistically significant differences existed between genders an independent-samples t-test was conducted. Based on 95% confidence interval, the analysis reported statistically significant differences between the male and female groups existed for synchronized presentations with two tailed value for $p=0.001$, $p<0.05$. For skewed presentations at all other levels, the findings were not statistically significant ($p=0.064$ for -20s skew, $p=0.086$ for skew of +15s being the only other sequences where differences were close to being statistically significant). Analyzing assessor opinions considering their age and gender from Fig. 25, all female age groups indicated a greater sense of reality compared to their male equivalents. The female and male 30-40 yrs group are most sensitive to the impact of large skews sizes. To determine if statistically significant results exist between each of the six gender-age groups in terms of reality, we performed a one way ANOVA between-groups analysis with post hoc tests. The results reveal statistically significant differences between the male 30-40 yrs and female 20-30, 30-40 and 40+ groups ($p=0.021$, $p<0.05$), ($p=0.013$, $p<0.05$), ($p=0.040$, $p<0.05$) when synchronized presentation of olfaction enhanced multimedia was presented. At skew levels of -10s, the statistically significant differences existed between the 30-40 and 40+ female groups ($p=0.032$, $p<0.05$). At skews of -25s, the 20-30 female and 30-40 female groups ($p=0.026$, $p<0.05$) and at -30s, the analysis reported significant differences 30-40 female and 40+ male groups ($p=0.033$, $p<0.05$). For olfaction presented after video, at skews of -15s, statistically significant results were reported between the male 30-40 and female 20-30 yrs groups ($p=0.007$, $p<0.05$). At +25s statistically significant results were reported determined between the male 20-30 and 40+ yrs groups ($p=0.017$, $p<0.05$).

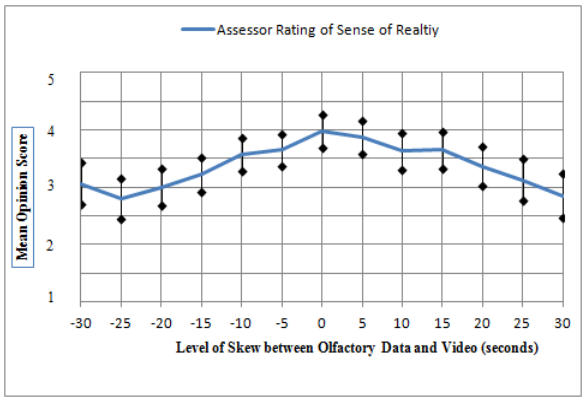

Fig. 23. Analysis of Sense of reality per Skew with confidence interval based on 99% confidence level.

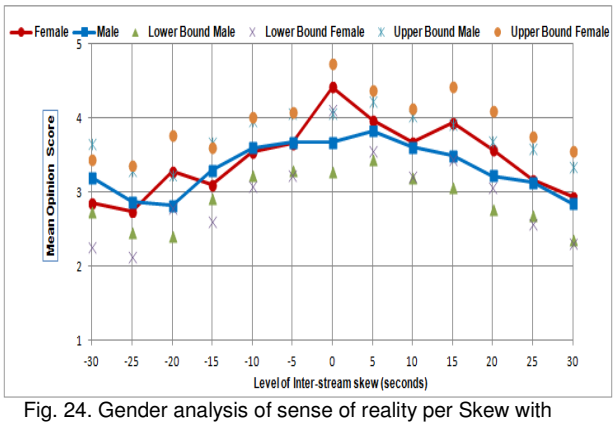

Fig. 24. Gender analysis of sense of reality per Skew with confidence interval based on 99% confidence level.

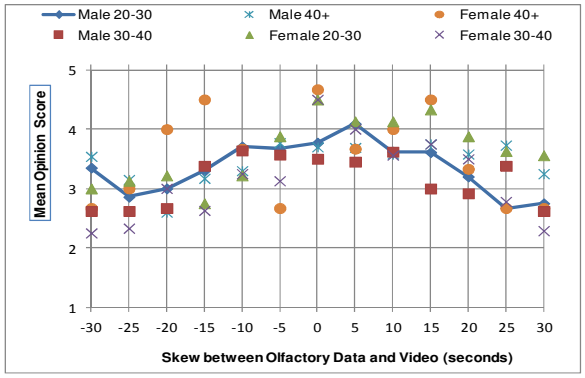

Fig. 25. Gender/age analysis aware of reality level per skew.

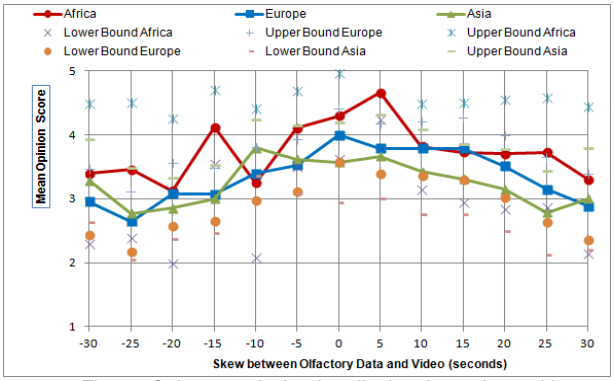

Fig. 26 Culture analysis of reality level per skew with confidence intervals based on a 99% confidence level.

Finally considering culture, the African group reported the greatest sense of reality across the majority of the skew sizes. The European group reports the next highest level of reality with the final Asian group scoring the lowest sense of reality. What is consistent across all groups is the slow gradual decrease in reality for olfaction after video (-15s for the African group being the exception). Consistently across all skews, olfaction after video contributes to a greater sense of reality compared to equivalent skews of olfaction before video. To determine if statistical significant results exist between sense of reality of each of the three culture groups, a one way ANOVA between-groups with post-hoc test was performed. For olfaction presented before video, with skew of -15s, the African group reported statistically significant results with the European ($p=0.008$, $p<0.05$) and Asian ($p=0.020$, $p<0.05$). For olfaction presented after video, with skew sizes of +5s the African group mean was statistically significant from the European ($p=0.018$, $p<0.05$) and Asian ($p=0.013$, $p<0.05$) groups. For the remaining test sequences, the results were found to not be statistically significant between the culture groups.

6.4 Comparison of audiovisual & olfactory with visual & olfactory user perceived synchronization

This section compares this work with current state of art research on olfactory and audiovisual media synchronization, highlighting the similarities and differences between them. As already mentioned, the same video clips were used in both works. In (Ghinea and Ademoye, 2010), (Murray et al., 2013) and here the middle block of 30s clip presented content related to the scent being presented. In (Murray et al., 2013) and here the videos were edited such that the audio was replaced with the sound of a blowing fan and not contextual audio data associated with the meaning of the clip. This, the authors conclude is the reason for the different results reported in each of the works. In fact, when executing preliminary work, assessors noted that the presence of scent before video when contextual data was presented assisted assessors in forecasting what was going to be visually provided and this was a likeable characteristic. This explains the results in (Ghinea and Ademoye, 2010) where assessors were more tolerant to olfaction before video than olfaction after video. Without contextual audio, the presence of olfaction before video was found to be annoying and confusing. The same scents were used although provided by different vendors. (Ghinea and Ademoye, 2010) states that it took “about 2 seconds” for emitted scent to reach users using the Vortex

Active emitter from Dale Air. As discussed above, it was experimentally evaluated that it took between 2.7s – 3.7s using the SBi4 emitter depending on the scent. Artificial skew values ranging from -30s to +30s were tested in (Ghinea and Ademoye, 2010) in step sizes of 10s (hence each assessor tested 6 samples), whereas here the step sizes were of magnitude of 5s (each assessor tested 12 samples). In both works the playback order was randomized to minimize ordering effects. 84 participants were tested in this work (42 tested in (Ghinea and Ademoye, 2010)).

In terms of conclusions, (Ghinea and Ademoye, 2010) define that the temporal relationship between olfactory and audiovisual media is from -30s to +20s. With the removal of contextual audio, the results show an "in-sync" region of -5s/-15s to +10/+15s depending on age, sex and culture. Even considering the largest boundary reported here, the difference is surprisingly large, the presence of contextual audio has a significant effect. Assessors noted that the scent, when complemented by relevant and contextual audio, provided extra information even with large skews compared to the video content. Without the presence of contextual audio, assessors, as shown by this work, assessors found it confusing, annoying and it had negative impacts on relevance, sense of reality and enjoyment. Finally, (Ghinea and Ademoye, 2010) found that olfaction before audiovisual information was more tolerable than olfaction after audiovisual media. As shown, the **reverse** is true for olfaction and video only i.e. olfaction after video is more tolerable.

7. CONCLUSION AND FUTURE WORK

This work investigated the perception of inter-stream synchronization error between olfactory and video media and its affect on the quality of experience for the user considering the users age, gender and culture. It defined general user temporal boundaries for user perceived synchronization as well as boundaries based on each of these factors. Depending on age, sex and culture, an in sync region is supported up to a maximum skew of 5s/15s seconds when olfaction is before video, and a maximum skew of 5s/15s seconds when olfaction is after video. Skews outside this range are in the out-of-sync region. Interestingly, this region is significantly smaller than the equivalent boundary for synchronization of olfaction, video and audio as reported in the literature. In addition, the results indicate that assessors were more aware of olfaction before video as opposed to olfaction after video. It was also identified that assessors are less tolerant to olfaction before video than they are of olfaction after video. 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 negatively impact the sense of relevance, reality and enjoyment when compared with feelings achieved when synchronized presentation takes place. Again, the impact of age, sex and culture are key considerations when analyzing the impact of skew on QoE. This contribution has demonstrated the requirement for profile based classification of assessors in terms of their perception of olfaction enhanced multimedia synchronization to ensure highest possible quality level. Future work will encompass subjective testing involving more than one olfactory stream as part of an olfaction enhanced multimedia clip. In addition, the overall aim for this work will be the definition of a utility model that reflects users perception of olfaction enhanced multimedia experiences.

8. ACKNOWLEDGEMENTS

The authors would like to gratefully acknowledge the generosity of Dr. Gheorghita Ghinea of Brunel University, UK and Dr. Oluwakemi A. Ademoye, Swansea Metropolitan University, UK (previously, of Brunel University, UK) for providing the videos for this work and their availability to answer questions. I would like to acknowledge the contribution of Mr. Oliver Hegarty, Psychologist and Head of Dept. of Humanities in Athlone Institute of Technology for his assistance with the questionnaire content, Prof. Klara Nahrstedt and Dr. Zixia Huang of University of Illinois at Urbana-Champaign and Prof. Ralf Steinmetz of Technische Univ. Darmstadt, Germany for answering queries in relation to their works and Mr. Yvan Régeard of Exhalia. The authors also recognize the assistance of Tunku Salha Tunku Ahmad of School of Science, as well as Dr. Austin Hanley and Dr. Marcus Rahilly of the School of Engineering, AIT. Finally, we recognize the contribution of Enterprise Ireland through its Applied Research Enhancement (SUNAT) and Technology Gateway (COMAND) funding programs.

REFERENCES

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

Ariyakul, Y., Nakamoto, T., "Olfactory display using a miniaturized pump and a SAW atomizer for presenting low-volatile scents". In *IEEE Virtual Reality* pp 193-194, (March 2011).

Ayabe-Kanamura, S., Schiker, I., Laska, M., Hudson, R, Distel, H, Kobayakawa, T., Saito, S. “Differences in Perception of Everyday Odors: A Japanese-German Cross-Cultural Study” In *Chem. Senses* 23:31-38, (1998).

Calvert, G., Spence, C., Stein, B.E. “The Handbook of Multisensory Processes”. In MIT Press. 2004

Cha, J., Eid, M., Barghout, A., Mahfujur Rahman, ASM., and El Saddik, A. “HugMe: Synchronous Haptic Teleconferencing”. In *Proceedings of 17th ACM International Conference on Multimedia* (ACM MM 2009).

Chastrette, M., “Classification of Odors and Structure – Odor Relationships”. In *Olfaction, Taste, and Cognition* pp. 100-116, Cambridge University Press (2002).

Dann, G., Jacobsen J., K., S., “Tourism smellscapes” In *Tourism Geographies*, 5, 1, pp 3-25 (Feb. 2003)

Eid, M., Cha, J., and Saddik, A. El., “Admux: An Adaptive Multiplexer for Haptic-Audio-Visual-Data Communication” In *IEEE Trans. Instrum. Meas*, 60 1, 21-31 (Jan. 2011).

Exhalia www.exhalia.com (accessed 03.01.2013)

Ghinea, G. and Ademoye, O. A. 2012. “The Sweet Smell of Success: Enhancing Multimedia Applications with Olfaction”. In *ACM Transactions on Multimedia Computing, Communications and Applications* 8, 1 (Jan. 2012).

Ghinea, G. and Ademoye, O. A., 2010. “Perceived Synchronization of Olfactory Multimedia”. In *IEEE Trans. on SYSTEMS, MAN, AND CYBERNETICS – PART A: SYSTEMS AND HUMANS* 40, 4, 657-663 (July. 2010).

Gulliver, S. R., Ghinea, G., “Defining the users perception of distributed multimedia quality”. In *ACM Trans. on Multimedia Computing, Communications and Applications (TOMCCAP)* 2, 4, pp 241-257 (Nov. 2006).

Ghinea, G., Ademoye, O.A., 2011. “Olfaction-enhanced multimedia: perspectives and challenges”. In *Multimedia Tools Appl* 55, 3, pp. 601-626 (Dec. 2011)

Huang, Z., Nahrstedt, K., 2012. “Perception-based playout scheduling for high-quality real-time interactive multimedia”. In *Proceedings of IEEE International Conf. on Computer Communications (INFOCOM)* pp. 2786-2790 (Mar. 2012)

Huang**, Z.,** Arefin, A., Agarwal, P., Nahrstedt, K and Wu, W., “Towards the understanding of human perceptual quality in tele-immersive shared activity”. In *Proc. of ACM Multimedia Systems Conf. (MMSYS),* pp 29-34 (Feb. 2012).

ISO/IEC FDIS 23005-3 Information Technology – Media context and control- Part 3: Sensory Information, ISO Publication (2010)

ISO 5492:2008 Sensory analysis – Vocabulary, International Standards Organization (ISO)

ISO/IEC 8589 Sensory analysis – General guidance for the design of test rooms.

ISO 5496:2006 – Sensory Analysis – Methodology – Initiation and training of assessors in the detection and recognition of odours

ITU-T BT.500. Methodology for the subjective assessment of the quality of television pictures, 2002.

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

Kaye, N., “Symbolic olfactory display” M.S. thesis, Massachusetts Inst. Techonology, Cambridge, MA, 2001.

Lee, J.-S., Goldmann, L., and Ebrahimi, T., “A new analysis method for paired comparison and its application to 3D quality assessment”. In *Proc. of ACM International Conference on Multimedia* pp1281-1284 (MM’11).

Narumi, T., Nishizaka, S., Kajinami, T., Tanikawa, T., and Hirose, M., “Augmented Reality Flavors: Gustatory Display Based on Edible Marker and Cross-Modal Interaction”. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems.*

Hoshino, S., Ishibashi, Y., Fukushima, N., and Sugawara, S. “QoE Assessment in Olfactory and Haptic Media Transmission: Influence of Inter-Stream Synchronization Error”. In *IEEE Int. Workshop Technical Committee on Communications Quality and Reliability 2011.*

Lawless, H. T. 1997. “Olfactory Psychophysics”. In *Tasting and Smelling. Handbook of Perception and Cognition, 2nd Edition* pp125-174. Beauchamp & Bartoshuk (Eds), Academic Press, San Diego, CA (1997).

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.

Murray, N., Qiao, Y., Lee, B., Karunakar, AK, Muntean, G.-M., “Age and Gender Influence on Perceived Olfactory and Visual Media Synchronization” In Proceedings of *IEEE International Conference on Multimedia and Expo* (ICME 2013). San Jose, California, USA.

Murray, N., Qiao, Y., Lee, B., Karunakar, AK, Muntean, G.-M., “Olfaction enhanced multimedia: A survey of application domains, displays and research challenges”. IEEE Communications Surveys and Tutorials, 2014

Nakamoto, T., Minh, H.P.D., 2007. “Improvement of Olfactory Display Using Solenoid Valves”. In *IEEE Virtual Reality* pp 179-186.

Nakamoto, T., Otaguro, S., Kinoshita, M., Nagahama, M., Ohinishi K., Ishida T., 2008. “Cooking up an Interactive Olfactory Game Display”. In *IEEE Computer Graphics Applications* 28, 1 pp 75-78. (Jan-Feb 2008).

Noguchi, D, Sugimoto, S., Bannai, Y., Okada, K., 2011. “Time characteristics of olfaction in a single breath”. In *Proceedings of the ACM SIG CHI Conference on Human Factors in Computing Systems* pp 83-92 (CHI ’11).

Pair, J., Allen, B., Dautricourt, M., Treskunov, A., Liewer, M., Graap, K., Reger, G., and Rizzo, A. “A virtual reality exposure therapy application for Iraq war post-traumatic stress disorder”. In IEEE Virtual Reality 2006.

Ramic, B., Hasic, J., Rizvic, S., Chalmers, A. 2006. “Selective Rendering in a multimodal environment: Scent and graphics,” In *Spring Conference on Computer Graphics, ACM SIGGraph,* (2006).

Spencer, B. S., “Incorporating the Sense of Smell Into Patient and Haptic Surgical Simulators” In *IEEE TRANSACTIONS on Information Technology in Biomedicine* 10, 1, 168-173 (Jan. 2006).

Shams, L., Seitz, AR, (2008). “Benefits of Multisensory learning”. In *Trends in Cognitive Sciences* 12,11 pp 411-417 (Nov. 2008)

Steinmetz, R. 1996. “Human Perception of Jitter and Media Synchronization”. In *IEEE Journal on Selected Areas in Communications* 14, 1 61-72 (Jan. 1996).

Steinmetz, R., Nahrstedt, K., 1995. “Multimedia: computing, communications and applications”. In *Prentice Hall.* (Sept. 1995).

Sugimoto, S., Noguchi, D., Bannai, Y., Okada, K., 2010. “Ink jet olfactory display enabling instantaneous switches of scents”. In *Proceedings of ACM International Conference on Multimedia* pp 301-310 (October 2010).

Timmerer, C., Waltl, M., Rainer, B., Hellwagner, H. 2012. “Assessing the quality of sensory experience for multimedia presentations”. In *Signal Processing: Image Communication* 27, 8, pp 909-916 (Sept. 2012).

Tomono, A., Yamamoto, S., Utsunomiya, M., Ikei, D., Yanagida, Y., Hosaka, K.. 2004. “Effect that the Image Media with scent Gives to Contents Understanding”. In *Human Interface Symposium* pp. 249-254 (March 2004).

Washburn, D.A., “Olfactory Use in Virtual Environment Training”. In *Modelling and Simulation* 2, 3, pp 19-25 (2003).