

Audio Masking Effect on Inter-Component Skews in Olfaction-Enhanced Multimedia Presentations

OLUWAKEMI A. ADEMOYE, University of Wales Trinity Saint David, UK
NIALL MURRAY, Athlone Institute of Technology, Ireland
GABRIEL-MIRO MUNTEAN, Dublin City University, Ireland
GHEORGHITA GHINEA, Brunel University, UK

Media-rich content plays a vital role in consumer applications today, as these applications try to find new and interesting ways to engage their users. Video, audio, and the more traditional forms of media content continue to dominate with respect to the use of media content to enhance the user experience. Tactile interactivity has also now become widely popular in modern computing applications, while our olfactory and gustatory senses continue to have a limited role. However, in recent times, there have been significant advancements regarding the use of olfactory media content (i.e., smell), and there are a variety of devices now available to enable its computer-controlled emission. This paper explores the impact of the audio stream on user perception of olfactory-enhanced video content in the presence of skews between the olfactory and video media. This research uses the results from two experimental studies of user-perceived quality of olfactory-enhanced multimedia, where audio was present and absent, respectively. Specifically, the paper shows that the user Quality of Experience (QoE) is generally higher in the absence of audio for nearly perfect synchronized olfactory-enhanced multimedia presentations (i.e., an olfactory media skew of between $\{-10, +10\text{s}\}$); however, for greater olfactory media skews (ranging between $\{-30\text{s}; -10\text{s}\}$ and $\{+10\text{s}, +30\text{s}\}$) user QoE is higher when the audio stream is present. It can be concluded that the presence of the audio has the ability to mask larger synchronization skews between the other media components in olfaction-enhanced multimedia presentations.

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

General Terms: Design, Experimentation, Human Factors

Additional Key Words and Phrases: Audio, olfaction, masking effect, quality of experience, synchronization

ACM Reference Format:

Oluwakemi A. Ademoye, Niall Murray, Gabriel-Miro Muntean, and Gheorghita Ghinea. 2016. Audio masking effect on inter-component skews in olfaction-enhanced multimedia presentations. *ACM Trans. Multimedia Comput. Commun. Appl.* 12, 4, Article 51 (August 2016), 14 pages.
DOI: <http://dx.doi.org/10.1145/2957753>

Authors' addresses: O. A. Ademoye, School of Applied Computing, University of Wales Trinity Saint David, Mount Pleasant, Swansea, SA1 6ED; email: kemi.ademoye@uwtsd.ac.uk; N. Murray, Dept. of Electronics, Computer Software Engineering, Athlone Institute of Technology, Dublin Road, Athlone, Ireland; email: nmurray@research.ait.ie; G.-M. Muntean, School of Electronic Engineering, Dublin City University, Glasnevin, Dublin 9, Ireland; email: gabriel.muntean@dcu.ie; G. Ghinea, Department of Computer Science, Brunel University, Kingston Lane, Uxbridge, UB8 3PH, UK; email: george.ghinea@brunel.ac.uk.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701 USA, fax +1 (212) 869-0481, or permissions@acm.org.

© 2016 ACM 1551-6857/2016/08-ART51 \$15.00

DOI: <http://dx.doi.org/10.1145/2957753>

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. The past century has witnessed generations of multimedia applications, including transitions from analogue modulation to digital media, single-audio, single-video playback to multimodal multichannel presentation, and two-party bidirectional communication to large-scale multiparty sharing using the Internet [Huang et al. 2007]. With new user interface devices like haptic devices, olfactory displays, and other associated sensors now commercially available, research and industry are working on extending the list of media and applications 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 communication and experiences—mulsemedia (multiple sensorial media [Ghinea et al. 2011, 2014]). The embedding of technology into our everyday lives on our bodies, in our clothes, in our appliances, in our cars, homes, and workplaces means that future multimedia systems will involve multisensory media integrated into everyday objects. Each of these aspects will place significant challenges on existing multimedia systems and synchronization approaches.

Synchronization is an important issue both in mulsemedia and multimedia research. Any timing artefacts introduced by the transport network will result in time lags between the media streams. These time lags have to be corrected at the destination in order to provide an “in-sync” presentation and achieve good user Quality of Experience (QoE) levels [Timmerer et al. 2014; Yuan et al. 2015]. Various approaches have been put forward to address synchronization issues [Huadong and Shin 2004]; for instance, Montagud et al. [2013] propose a solution for Inter-Destination Media Synchronization (IDMS) using an extension of the RTP/RTCP mechanisms, yet no QoE studies are undertaken to evaluate it. A solution based on a global clock for synchronized multimedia delivery is presented and tested by Concolato et al. [2012] over two real-life scenarios (a DVB and FM network, respectively, mixed over a broadband IP network); however, no QoE results are reported. QoE results are reported by Mekuria et al. [2012], which detail the QoE impact of delays on watching football (soccer) on digital TV. Other multimedia synchronization studies [Steinmetz 1996] claim that audio might be played up to 120ms ahead of video and, conversely, video played up to 240ms ahead of audio. Both temporal skews can be accepted by users without significant impact on experience.

User-perceived synchronization between olfactory media and multimedia sequences is studied in Ademoye and Ghinea [2009] and Ghinea and Ademoye [2014], which show that olfaction ahead of video content is more tolerable than olfaction behind content. Therefore, synchronization is needed to ensure a temporal ordering of events in a mulsemedia system. Consequently, this article reports on and compares the results of two separate research studies, which explored the user QoE of olfaction-enhanced multimedia, in the presence, respectively absence, of its audio component.

2. OLFACTION-ENHANCED MULTIMEDIA

Olfaction-enhanced multimedia concerns itself with associating computer-generated smell with other media to enrich the users’ experience and perception of a multimedia presentation. In this way, olfactory media is already being used towards a variety of goals in multimedia applications. However, the concept of computer-generated smell, that is, olfactory media data, is relatively new in the field, and there are still a number of challenges to overcome with regards to its use, and specifically relating to achieving quality olfactory-enhanced multimedia applications. Achieving multimedia quality can be considered from both a technical perspective and a user perspective. However,

as multimedia applications are inherently designed for users, certainly, determining quality factors from a user perspective, which directly relates to the user-perceived experience of the multimedia, is more important than from a technical perspective.

Olfaction-enhanced media research has previously been predominantly application-led. Two extensive surveys, Ghinea and Ademoye [2011] and Murray et al. [2016a], have highlighted encouraging results in the domains of entertainment, virtual reality, film industry, multimodal applications, tourism, health, and education. Briefly highlighting work in some of these domains. In entertainment, olfaction-enhanced applications have a long history, starting off with the use of olfaction to enhance the cinematic viewer experience: Smell-O-Vision (cf. Smith and Kiger [2006]) and AromaRama (cf. Kaye [2001]) were two such early attempts. Currently, olfaction-enhanced media are being employed in 4D and 5D experiences in theme parks worldwide.

In tourism, for instance, with the motivation for visitors to move beyond the “browse mode” and learn, museums are searching for avenues that include multisensory experiences, to further engage their consumers [Schauble et al. 2002]. Hall and Bannon [2006] outline a learning environment based in a museum and reflect that the addition of olfaction adds “authenticity to the experience and helped create the scene of an old and musty room.” Fernstrom and Bannon [1997] mentioned the concept of “hybrids” between museums and arcades. They discuss the use of interactive multimedia within museums. The aim is to ensure that technology adds value, as part of the visitor experience. Pan and Ryan [2009] discuss the importance of olfaction as part of tourism experience to facilitate a link with the past as part of the user experience. They also specify the requirement to address tourism as a corporeal experience.

Last but not least, under the broad umbrella of education, Mikropoulos and Natsis [2011] reviewed the use of education in virtual environments but only made reference to two works with respect to olfaction: Tijou et al. [2006] and Richard et al. [2006a]. Richard et al. [2006b] described a VR application that investigated “the effect of olfaction on learning” and information retention. Richard et al. [2006b] described a VR platform that provides haptic, olfactory, and auditory experiences. The “Nice-smelling Interactive Multimedia Alphabet” project (www.exhalia.com) used the modalities of vision, sound, and olfaction to assist with learning the letters of the alphabet. Speech and language therapists use scent as part of the Picture Exchange Communication System (PECS) (www.pecs.com) as a reinforcer. Dinh et al. [1999] evaluated the importance of olfaction among other modalities “on memory and sense of presence in virtual environments.” Hughes et al. [2005] included olfactory modality within their mixed reality training environments and applications, while Ludvigson and Rottman [1989] described and showed how the use of lavender and clove can improve cognitive processing. The interested reader can acquire a review of commercially available olfactory displays among other immersive multisensory media equipment in Kovacs et al. [2015].

3. OLFACTION SYNCHRONIZATION

Studying olfactory synchronization requirements in mulsemmedia systems constitutes a challenge from two very important (and different) perspectives. First, olfactory media has the tendency to linger, as well as its emission usually happening without hurried movement, as it relies on atmospheric breeze to move it through the air. Smell is, therefore, different from the media objects multimedia applications have become accustomed to, that is audio, visual, and traditional media (text, images, animations), whose component media streams have a transitory nature and do not experience this natural delay. In addition, the occurrence of data loss with olfactory data is something to be expected, as sensitivity to the presence of a smell tends to decrease with prolonged exposure to it (i.e., olfactory adaptation), and, moreover, the presence of atmospheric breeze may cause the olfactory data to waft off in the wrong direction and miss the intended target.

Based on these prior work outputs, we know that olfactory synchronization boundaries are less clearly defined than those associated with traditional media and the challenge is to identify them.

Second, in order to explore the acceptable perceptual boundaries for olfaction, existing models, such as electronic noses [Nagle et al. 1998; Rowe and Jain 2005], capable just of detection and identification of odors, are sorely lacking in this respect. The only alternative available to answer this question is to conduct an empirical, participant-based study. This is indeed the approach taken by us in our work and, in so doing, closely follows the example set by Steinmetz [1996], who has, in empirical work, explored the human perceptual boundaries for lip and pointer synchronization, as well as undertaking a survey of synchronization constraints that exist between traditional media objects [Blakowski and Steinmetz 1996].

3.1. Olfaction Skews

Perceived synchronization of olfaction-enhanced multimedia was investigated in detail in Ademoye and Ghinea [2009], which specifies a range of olfaction skews. Accordingly, artificial skews were introduced at intervals of 10s ranging from -30 s to $+30$ s between olfactory stimuli and video content. Steps of 10s were chosen based on human reactions to the smell. For instance, the average time for humans to sniff a scent typically ranges from 0.5s to 2.5s [Laing 1983], and research in Washburn et al. [2003] suggests that, based on spectrum range, exposure to short varying intervals should be limited between 20s and 30s.

3.2. Impact of Inter-Media Skews on Perceived Olfaction

According to our work [Ademoye and Ghinea 2009], artificially created skews, ranging in value from -30 s to $+30$ s between olfaction media and video media were used to achieve inter-media skew variation. The impact of the introduced skew on user satisfaction and enjoyment of the olfaction-enhanced multimedia experience is studied. It was concluded that significant differences of participants' perception of the synchronization effect between olfactory media and video content occurred at inter-media skew values of -30 s for olfaction ahead of video content and $+20$ s for olfaction behind video content. In this context, beyond a skew of $+20$ s, the olfaction media no longer had the impact to improve the sense of reality of the multimedia/mulsemmedia experience; it is yet to be established, however, if skews beyond -30 s are still acceptable to the viewing public. The two studies described in this article use the same set of videos as in Ademoye and Ghinea [2009] and consequently the same range of media skews applies.

In follow-up work, synchronization issues were discussed considering a single scent and its relationship with the associated video content in Murray et al. [2014a], the influence on this relationship considering participant age and gender [Murray et al. 2013a], of the influence of olfaction on information recall [Ademoye and Ghinea 2013], and of multiple scents and their complex interaction in the presence of synchronized video content in Murray et al. [2014b]. Perceptual testing of olfaction-enhanced multimedia content using the same video clips as in Ademoye and Ghinea [2009] (albeit with the soundtrack removed) and a methodology based on the same work was also reported in detail in Murray et al. [2013b]. Finally, the effect of scent type on the user ability to detect skew as well as its impact on user QoE was presented in Murray et al. [2016b].

Nonetheless, one issue that has not hitherto been explored is the influence of the audio stream on the user QoE of olfaction-enhanced multimedia. This is the precise focus of this article, which reports on and compares the results of two experimental studies, respectively, incorporating/withholding audio content in olfaction-enhanced multimedia presentations, and whose experimental methodology we proceed to describe next.



Fig. 1. Olfactory devices: Exhalia (l) and Vortex Active (r).

4. EXPERIMENTAL METHODOLOGY

This section discusses the design and setup of the two experimental studies being compared. It provides an outline of the scents and devices used, the olfactory and visual media displays, laboratory setup and participants of the studies. Although the setup of both studies was similar, there were a few differences and these will be appropriately highlighted in the subsequent sections of this article. In order to distinguish between the two studies, they are identified as follows: AV (audiovisual) – study 1 [Ademoye and Ghinea 2009], and VO (video only) – study 2 [Murray et al. 2013b].

4.1. Participants

A total of 42 participants (14 female and 28 male) took part in AV, while VO had a total of 43 participants (20 female and 23 male). The experimental groups included participants from multiple nationalities and a variety of backgrounds, including undergraduate and postgraduate students from different universities, academic staff, professionals, and members of the public. In both studies, the participants were aged from 18 to 56 years.

4.2. Olfactory Stimuli and Scent Emission Devices

The olfactory dataset used by both studies consisted of six scents that could individually be described as having a burnt, flowery, foul, fruity, resinous, or spicy smell. Olfactory-enhanced visual media displays were created by synchronizing each smell with a 90-second video containing content specifically related to the smell, which have been described in the following section. The videos used in the two studies were exactly the same, but in the case of VO, the audio stream was removed from the videos [Murray et al. 2013b]. The videos were selected such that the emission of in-sync olfactory media content will coincide with the middle 30s interval of the 90s video clips used.

Scent emission in olfactory-enhanced multimedia can be broadly categorized according to whether it is ambient or directed in nature. In the former, the olfactory data completely surrounds the user's environment, intending, for example, to relax the user, whilst the latter entails emitting a scent directed towards the target user. In the case of ambient scents, users are most times not consciously aware of the scents [Ademoye and Ghinea 2013]. Accordingly, in the two studies described in this article, we have used directed scents, as the aim of using multiple media objects in multimedia applications is to convey informational content via the multiple media objects used.

In both studies, scent emission was controlled via a USB device connected to the computer and programmed to synchronize the presentation of smell with video data. AV used the Vortex Activ USB Device by DaleAir (www.daleair.com), while VO used the Exhalia SBi4 radio v2 scent emitter (www.exhalia.com). The devices (Figure 1) are quite similar, both having four scent cartridges and using in-built miniature fans (of similar emitting power) to send out air to diffuse the scent being emitted. The SBi4 had the additional ability to control the intensity of the emitted scent by varying the

Table I. Video Categories and Scents Used

SMELL	Fruity	Resinous
Image Clip Frame		
SMELL	Burnt	Flowery
Image Clip Frame		
SMELL	Spicy	Foul
Image Clip Frame		

speed of the fan, although this feature was not used in the experiment discussed in this article.

4.3. Video Clips

The multimedia video clips were chosen so that their content related to the smells they were combined with, and hence the naming convention of the video clips below and in Table I:

- **Burnt Video:** This video clip is an extract from a documentary about using bush fires to control the spread of the invasive Eastern Red Cedar tree on prairie lands in Oklahoma.
- **Flowery Video:** The Flowery video clip is an extract from a news broadcast reporting on the launch of a perfume product made from the scent of lilac flowers. In this video clip, people attending the perfume launch festival sample the new perfume being launched by participating in a sniff test of the perfume product.
- **Foul Video:** This video clip is an extract from a documentary report about the ethylene gas produced by fruit and vegetables as they age and ripen. The documentary is combined with the advertisement of a product that can be used to eliminate this gas as it is produced by fruits and vegetables kept in refrigerators.
- **Fruity Video:** This video clip is an extract from a cookery show program teaching viewers how to prepare a particular fruit cocktail recipe. The fruit cocktail is made from several fruits including strawberries, peaches, pears, grapes, pineapple.
- **Resinous Video:** The Resinous video clip is an extract from a documentary report on how the pollens from the invasive Eastern Red Cedar tree makes spring-time allergy suffering worse in America.
- **Spicy Video:** This video clip is an extract from a cookery show program teaching viewers how to prepare a chicken curry dish. The clip shows the cook combining the different ingredients, including spices and chicken, as he prepares the dish.

Table II. Question 1: Detection of the Synchronization Error

Question 1	AV	<i>The smell was released</i>				
	VO	<i>Relative to the video content in the clip, the smell was released</i>				
Response Options	AV	<i>Too Early</i>	<i>Early</i>	<i>At an appropriate time</i>	<i>Late</i>	<i>Too Late</i>
	VO	<i>Too Late</i>	<i>Late</i>	<i>At the correct time</i>	<i>Early</i>	<i>Too Early</i>
	VO					

4.4. Olfactory Skews

The experimental studies were each conducted by introducing artificially created inter-media skews between the olfactory data and video data content. The artificially introduced skews for olfactory media synchronized with video were created at a continuous 10s interval range starting from $-30s$ (olfaction ahead) to $+30s$ (olfaction behind). Pilot experiments conducted in both studies determined that it took an average of 2s for the emitted scent to reach the user for AV, and between 2.7s to 3.7s for VO. To compensate for this, each study offset the start time of the synchronized olfactory data accordingly. However, while VO explicitly includes the offset as part of the skewed start time in the interpretation of their results, AV uses the skewed start time without the applied offset. Thus, for VO, the reported artificially introduced skews at 10s intervals range from $-32.91s$ to $+27.09s$. Furthermore, in both studies, assessors were assigned randomly into groups based on the test input combinations they would experience. In Ghinea and Ademoye [2014], there were seven groups in total. Assessors in each of these groups experienced six of the seven possible skew levels (one skew for each video clip). Although a larger number of groupings were defined in Murray et al. [2014a], these also included the same test input combinations as [Ghinea and Ademoye 2014]. For the interested reader, a detailed breakdown of the different test inputs and associated assessor groupings is provided in Ghinea and Ademoye [2014] and Murray et al. [2014a]. Finally, for both groups, the playback order was randomized to minimize ordering effects.

4.5. Experimental Set-Up

The experiments were carried out in controlled laboratory environments in Athlone Institute of Technology, Ireland, and Brunel University, United Kingdom. Both environments were generally similar and in compliance with ITU-T R. P.913 [2014]. Each had one testing booth placed in the corner of the test room to minimize distraction. Also, the facility to open windows to remove ambient scent was possible and taken advantage of to minimize any impact on user perception.

The olfactory displays were placed approximately 0.5m from the participants and were tilted at an angle in line with the participant's nose. In addition, in both labs, a desktop fan was switched on to further eliminate the possibility of smell-lingering effects between successive viewings of the six olfaction-enhanced multimedia video excerpts. The interested reader can obtain additional detailed information in Ademoye and Ghinea [2009] and Murray et al. [2013b].

4.6. Questionnaire

To measure user perception of the inter-stream synchronization between olfactory data and video content in the studies, participants were asked to complete a questionnaire after watching each olfactory-enhanced video clip. The four common questions to the two studies are presented in Tables II and III. For analysis purposes in this article, the responses of participants of VO were mirrored around the mid-point.

Table III. Questions 2–4 of the User Experience Questionnaire

Question 2	AV	<i>The smell was relevant to what I was watching.</i>				
	VO	<i>The smell was relevant to what I was watching.</i>				
Question 3	AV	<i>The smell heightened the sense of reality whilst watching the video clip.</i>				
	VO	<i>The smell contributed to a heightened sense of reality whilst watching the video clip.</i>				
Question 4	AV	<i>I enjoyed watching the video clip.</i>				
	VO	<i>I enjoyed watching the video clip.</i>				
Response Options	AV	Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree
	VO	Strongly Disagree	Disagree	Neither Agree or Disagree	Agree	Strongly Agree

Table IV. User-Perceived Skews: Difference between Mean Opinion Scores in the Presence/Absence of Audio Based on a 95% Confidence Level

Skew s1	-30s	-20s	-10s	0s	10s	20s	30s
Skew s2	-32.91s	-22.91s	-12.92s	-2.91s	7.09s	17.09s	27.09s
Mean s1	2.0385	2.7500	2.8276	2.9394	3.1613	3.3214	3.8333
Mean s2	1.2564	1.4400	1.8312	2.8961	3.8312	4.2500	4.7500
T	4.2848	6.7943	-6.1367	-0.3381	-3.7440	5.5531	-.2850
sig.	0.0002	0.0000	0.0000	0.7362	0.0003	0.0000	0.0000

5. RESULTS

In this section, the results from the two experiments conducted are presented and compared with each other in order to establish the influence of the audio component on user QoE in olfaction-enhanced multimedia in the presence or, respectively, absence of artificially introduced skews between olfactory and video content.

5.1. Detection of Synchronization Error

To determine the detection of the synchronization error, participants in the study were asked the questions shown in Table III. The results (Table IV and Figure 1) show that there was a significant difference in users' ability to detect any synchronization error when artificial inter-media skews were introduced between olfactory data content combined with audio-visual content compared to when it was only combined with visual content. The mean opinions were significantly higher when smell was combined with audio-visual content. However, in the case of perfect synchronization (no-skew), the difference between the means was not found to be statistically significant. We can conclude, however, that in the absence of sound, users are more likely to notice synchronization errors between the olfactory and video media content. Thus, the introduction of the audio stream acts like a masking effect that disguises olfactory synchronization skews.

5.2. User Sense of Enjoyment

As can be seen in Figure 3 and Table V, users displayed a positive disposition toward the olfaction-enhanced videos, irrespective of whether a soundtrack was involved. Even though there were no statistically significant differences in enjoyment levels for skews between -20s and +20s (Figure 2), one can detect that the results display an almost symmetric aspect where olfactory-enhanced videos with soundtrack have higher enjoyment levels than their nonaudio counterparts with the exception of the {-10s, 10s} range. It can be noticed, however, that at the +30s instance the difference in user enjoyment levels between the two scenarios is statistically significant, with user enjoyment levels polling higher for olfaction enhanced with sound video clips. The fact that differences in enjoyment levels only become significant at this point is in line with previous research of ours [Ademoye and Ghinea 2009], which has highlighted that olfaction ahead of audiovisual content is more tolerable than olfaction behind the same content.

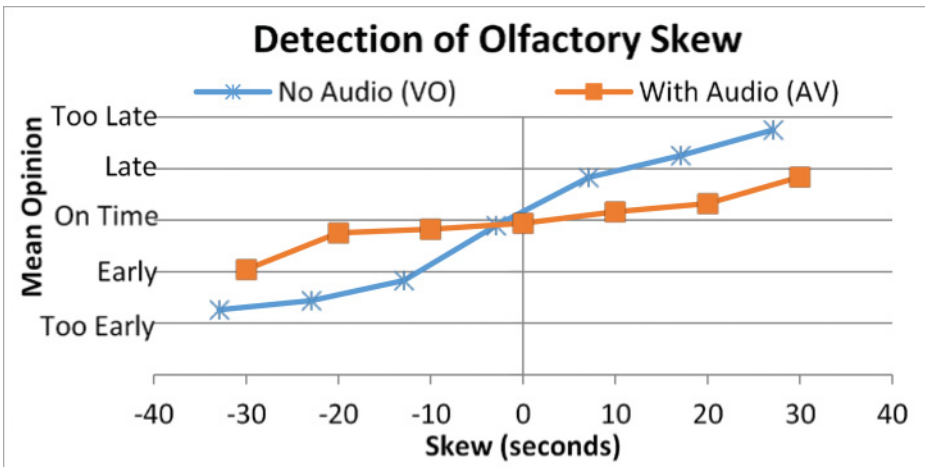


Fig. 2. User-perceived detection of olfactory skews.

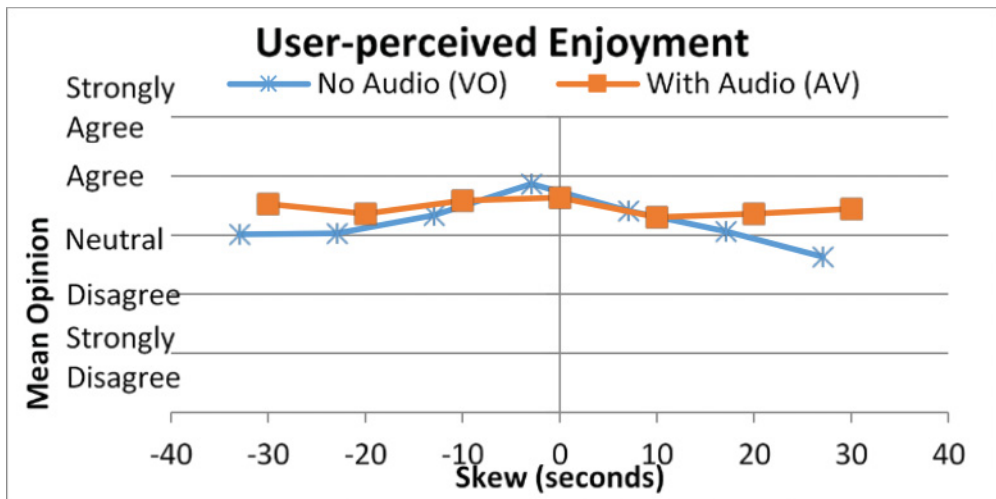


Fig. 3. User-perceived enjoyment of olfactory-enhanced multimedia.

Table V. User-Perceived Enjoyment of Olfactory-enhanced Multimedia: Difference between Mean Opinion Scores in the Presence/Absence of Audio Based on a 95% Confidence Level

Skew s1	-30s	-20s	-10s	0s	10s	20s	30s
Skew s2	-32.91s	-22.91s	-12.92s	-2.91s	7.09s	17.09s	27.09s
Mean s1	3.5278	3.3611	3.5833	3.6389	3.3056	3.3611	3.4444
Mean s2	3.0128	3.0274	3.3333	3.8667	3.4079	3.0658	2.6351
t	-2.3369	-1.6291	-1.3613	1.2434	0.4726	-1.5546	-3.8097
sig.	0.0212	0.1062	0.1762	0.2164	0.6374	0.1229	0.0002

5.3. User Sense of Relevance

The same qualitative aspect of the results as previously observed in the enjoyment scenario can also be noticed in terms of results asking users about the degree to which olfactory content was relevant to what they were watching (Figure 4 and Table VI). Thus, olfactory-enhanced video content with sound again polled higher scores than

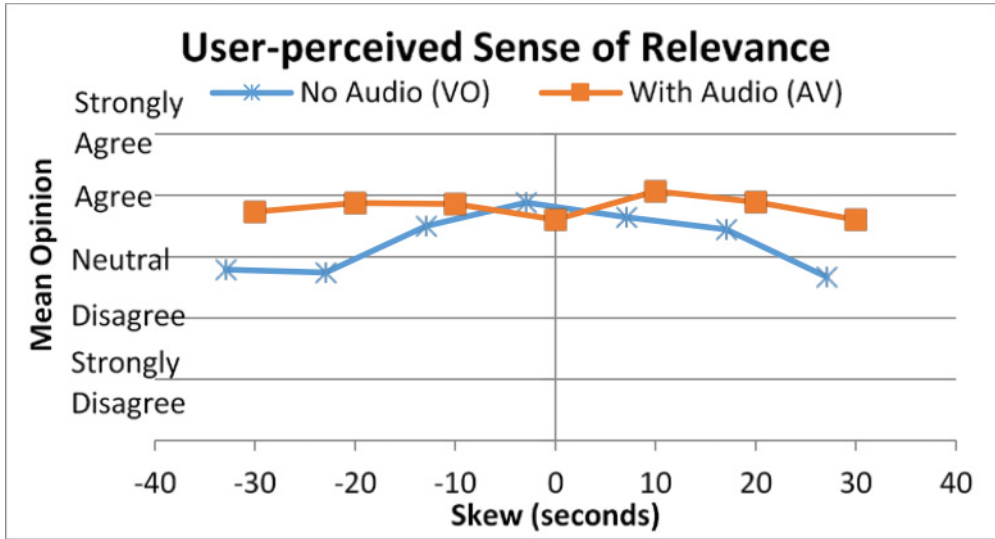


Fig. 4. User-perceived relevance of olfactory-enhanced multimedia.

Table VI. User-Perceived Sense of Relevance of Olfactory-Enhanced Multimedia: Difference between Mean Opinion Scores in the Presence/Absence of Audio Based on a 95% Confidence Level

Skew s1	-30s	-20s	-10s	0s	10s	20s	30s
Skew s2	-32.91s	-22.91s	-12.92s	-2.91s	7.09s	17.09s	27.09s
Mean s1	3.7308	3.8750	3.8621	3.6061	4.0645	3.8929	3.6000
Mean s2	2.7922	2.7397	3.5000	3.8800	3.6447	3.4474	2.6667
t	-3.8445	-5.3740	-1.6473	-1.2547	-1.8799	-1.7750	-3.6059
sig.	0.0003	0.0000	0.1025	0.2124	0.0629	0.0789	0.0005

the case when there was no sound involved, with the exception of $\{-5s, 5s\}$ —roughly what would be considered as near-perfect synchronization. However, the differences between the two scenarios (with and without sound) do become statistically significant more rapidly (compared to the enjoyment scores) at $\{-30s, -20s\}$ and at $\{30s\}$ (the same point as in the enjoyment scenario above), with the “with sound” olfaction-enhanced video scoring higher than its “no sound” counterpart. It is also to be remarked that in the “with sound” scenario, the case of perfect sync is judged to be less relevant than cases when skew was involved—this could be due to the tendency of smell to linger while propagating through the environment.

5.4. User Sense of Reality

Results show that users generally agreed that olfaction contributed to an enhanced realism of the watched videos, whether or not accompanied by sound (Figure 5 and Table VII). Again, scores were higher in the case when sound complemented video, with the exception of the $\{-5s, +5s\}$ region—similarly to users’ perception of relevance, as discussed earlier. Unlike the aforementioned case, however, differences in user scores between the two scenarios (with and without sound) for the sense of reality conveyed do not get to be statistically significant at the extremities of the studied skews, with one noticeable exception ($-20s$).

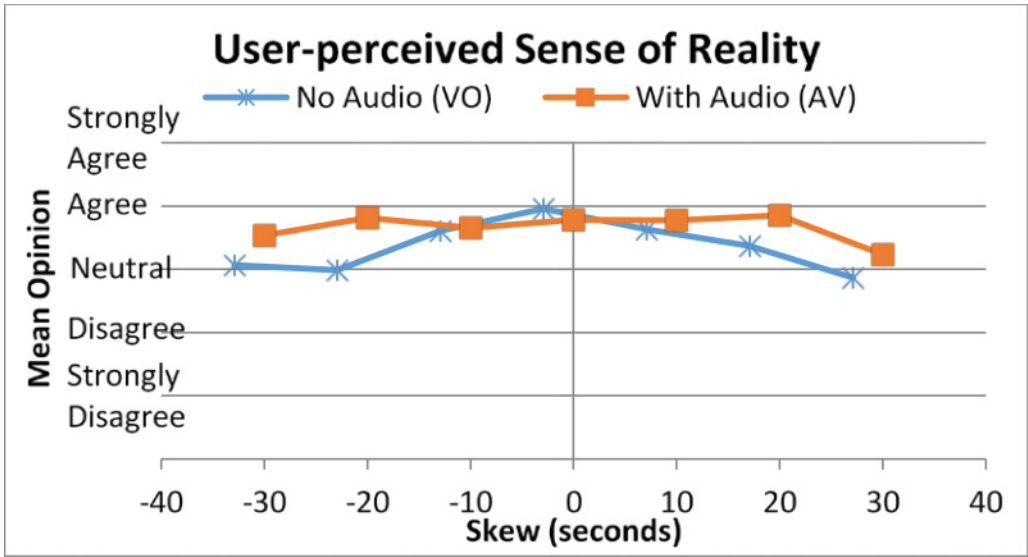


Fig. 5. User-perceived sense of reality of olfactory-enhanced multimedia.

Table VII. User-Perceived Sense of Reality of Olfactory-Enhanced Multimedia: Difference between Mean Opinion Scores in the Presence/Absence of Audio Based on a 95% Confidence Level

Skew s1	-30s	-20s	-10s	0s	10s	20s	30s
Skew s2	-32.91s	-22.91s	-12.92s	-2.91s	7.09s	17.09s	27.09s
Mean s1	3.5385	3.8125	3.6552	3.7879	3.7742	3.8571	3.2333
Mean s2	3.0649	2.9863	3.6053	3.9600	3.6364	3.3684	2.8667
t	-1.7180	-3.7064	-0.2410	0.8731	0.6010	-1.9352	-1.3617
sig.	0.0889	0.0003	0.8100	0.3846	0.5492	0.0557	0.1763

5.5. Discussion

When analyzing the results of the AV and VO tests performed, a more acute sense of synchronization error detection in the absence of the audio media than when the audio components is present is noted. Additionally, user enjoyment levels are lower when the audio is absent and there are olfaction-video synchronization issues, in both situations which involve olfaction presentation after the associated video and before the actual video sequence, respectively. Similar consistent results are obtained also when analyzing the sense of reality and sense of relevance, respectively. Lower levels of these metrics are experienced when there is no audio media component in olfaction-enhanced multimedia presentation than when audio is also present which experiences synchronization issues. However, the results favor the no audio component presentation when the inter-media presentation is in synch.

These results show how the presence of the audio media component of a multisensorial stream acts as a mask for potential synchronization issues between the olfaction media component and video, hiding some of their negative effects from the user. These results support and complement the findings in Brkic at al. [2009], although here is the first work that has demonstrated the masking effect of audio. In Brkic at al. [2009], a reduction in video quality was masked by the presence of an olfactory component. In this article, the effects were expressed in terms of skew detection, user enjoyment, sense of reality and sense of relevance. This is a very important result and may help in

the design of the future multiple sensorial sequences that will be delivered to audiences with heterogeneous operational environments (network, device, etc.).

6. CONCLUSION

Multiple sensorial media—mulsemedia for short—has the potential to radically change the way that users experience digital content. However, little is known about the impact that the interplay between the various media has on the user experience. In particular, user-perceived synchronization between olfactory media and other traditional media types needs to be studied, especially given the lingering nature of smell. Accordingly, in this article, we have reported results of a study that examined the olfactory impact on the user experience, when smell accompanied video content, in both the presence and absence of sound.

Our results highlight that olfaction-enhanced multimedia enhanced the user experience. However, olfaction generally has a higher potential to do this when it complements video incorporating an audio soundtrack; however, our results also consistently show that around the case of near-perfect synchronization ($\{-10s, 10s\}$ or less), the absence of sound results in a better user experience of olfaction-enhanced video.

We have reported a novel masking effect in mulsemedia in respect of audio and olfaction. In the future, it would be worthwhile to explore what other masking effects new media such as e-taste and e-smell produce when used in conjunction with the traditional ones of video and audio. Also worthwhile is the exploration of whether video clip content impacts olfactory (and gustatory) perceived synchronization boundaries.

REFERENCES

- O. A. Ademoye and G. Ghinea. 2009. Synchronization of olfaction-enhanced multimedia. *IEEE Transactions on Multimedia* 11, 3, 561–565.
- O. A. Ademoye and G. Ghinea. 2013. Information recall task impact in olfaction-enhanced multimedia. *ACM Transactions on Multimedia Computing, Communications and Applications* 9, 3, Article 17.
- G. Blakowski and R. Steinmetz. 1996. A media synchronization survey: Reference model, specification, and case studies. *IEEE Journal on Selected Areas in Communications* 14, 5–35.
- B. R. Brkic, A. Chalmers, K. Boulanger, S. Pattanaik, and J. Covington. 2009. Cross-modal effects of smell on the real-time rendering of grass. In *Proceedings of the Spring Conference on Computer Graphics (SCCG)*. 161–166.
- J. Cha, M. Eid, A. Barghout, A. M. Rahman, and A. El Saddik. 2009. HugMe: Synchronous haptic teleconferencing. In *Proceedings of the 17th ACM International Conference on Multimedia (MM'09)*. 1135.
- C. Concolato, S. Thomas, R. Bouqueau, and J. Le Feuvre. 2012. Synchronized delivery of multimedia content over uncoordinated broadcast broadband networks. In *Proceedings of the 3rd Multimedia Systems Conference (MMSys'12)*. 227–232.
- H. Q. Dinh, N. Walker, L. F. Hodges, C. S. C. Song, and A. Kobayashi. 1999. Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. In *Proceedings of the IEEE Virtual Reality Symposium*. 222–228.
- M. Fernstrom and L. Bannon. 1997. Enabling technology for museum visitors: Issues and experiences. In *Proceedings of the International Conference on Museums and the Web*. 191–199.
- G. Ghinea, S. R. Gulliver, and F. Andres. 2011. *Multiple Sensorial Media Advances and Applications: New Developments in Mulsemedia*. IGI Global, Hershey, PA.
- G. Ghinea and O. A. Ademoye. 2011. Olfaction-enhanced multimedia: Perspectives and challenges. *Multimedia Tools and Applications* 55, 3, 601–626.
- G. Ghinea and O. A. Ademoye. 2012. The sweet smell of success. *ACM Transactions on Multimedia Computing Communications and Applications* 8, 1, 2012, 1–17.
- G. Ghinea and O. A. Ademoye. 2014. Perceived synchronization of olfactory multimedia. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 40, 4, 2014, 657–663.
- G. Ghinea, C. Timmerer, W. Lin, and S. R. Gulliver. 2014. Mulsemedia: State of the art, perspectives, and challenges. *ACM Transactions on Multimedia Computing Communications and Applications* 11, 1s, 1–23.

- T. Hall and L. Bannon. 2006. Designing ubiquitous computing to enhance children's learning in museums. *Journal of Computer Assisted Learning* 22, 231–243.
- M. Huadong, K. G. Shin. 2004. Checking consistency in multimedia synchronization constraints. *IEEE Transactions on Multimedia* 6, 4, 565–574
- Z. Huang, K. Nahrstedt, and R. Steinmetz. 2013. Evolution of temporal multimedia synchronization principles: a historical viewpoint. *ACM Transactions on Multimedia Computing Communications and Applications* 9, 1–23.
- C. E. Hughes, C. B. Stapleton, D. E. Hughes, and E. M. Smith. 2005. Mixed reality in education, entertainment, and training. *IEEE Computer Graphics and Applications* 25, 24–30.
- ITU-T R. P.913 2014. Methods for the subjective assessment of video quality, audio quality and audiovisual quality of Internet video and distribution quality television in any environment.
- P. T. Kovacs, N. Murray, G. Rozinaj, Y. Sulema, and R. Rybarova. 2015. Application of immersive technologies for education: State of the art. In *Proceedings of the International Conference on Interactive Mobile Communication, Technologies and Learning (IMCL15)*. 283–288.
- D. G. Laing. 1983. Natural sniffing gives optimum odour perception for humans. *Perception* 12, 2, 99–117.
- H. W. Ludvigson and T. R. Rottman. 1989. Effects of ambient odors of lavender and cloves on cognition, memory, affect and mood. *Chemical Senses* 14, 525–536.
- R. Mekuria, P. Cesar, and D. Bulterman. 2012. Digital TV: The effect of delay when watching football. In *Proceedings of the 10th European conference on Interactive TV and Video (EuroITV12)*. 71–74.
- T. A. Mikropoulos and A. Natsis. 2011. Educational virtual environments: A ten-year review of empirical research (1999-2009). *Computers and Education* 56, 769–780.
- M. Montagud, F. Boronat, and H. Stokking. 2013. Early event-driven (EED) RTCP feedback for rapid IDMS. In *Proceedings of the 21st ACM International Conference on Multimedia (MM13)*. 323–332.
- N. Murray, Y. Qiao, A. K. Karunakar, B. Lee, and G. M. Muntean. 2013a. Age and gender influence on perceived olfactory & visual media synchronization. In *Proceedings of IEEE International Conference on Multimedia & Expo*. 1–6. DOI: 10.1109/ICME.2013.6607467
- N. Murray, Y. Qiao, B. Lee, A. K. Karunakar, and G. M. Muntean. 2013b. Subjective evaluation of olfactory and visual media synchronization. In *Proceedings of the 4th ACM Multimedia Systems Conference*. 162–171. DOI: 10.1145/2483977.2483999
- N. Murray, Y. Qiao, B. Lee, and G. M. Muntean. 2014a. User-profile-based perceived olfactory and visual media synchronization. *ACM Trans. Multimed. Comput. Commun. Appl.* 10, 1s, article no. 11.
- N. Murray, B. Lee, Y. Qiao, and G. M. Muntean. 2014b. Multiple-scent enhanced multimedia synchronization. *ACM Trans. Multimed. Comput. Commun. Appl.* 11, 1s, article no. 12.
- N. Murray, B. Lee, Y. Qiao, and G. M. Muntean. 2016a. Olfaction-enhanced multimedia: A survey of application domains, displays, and research challenges. *ACM Comput. Surv.* 48, 4, Article 56 (May 2016), 34 pages. DOI: <http://dx.doi.org/10.1145/2816454>
- N. Murray, B. Lee, Y. Qiao, and G. M. Muntean. 2016b. The impact of scent type on olfaction-enhanced multimedia quality of experience. In *IEEE Transactions on Systems, Man, and Cybernetics: Systems PP*, 99 (2016), 1–13. DOI: 10.1109/TSMC.2016.2531654
- H. T. Nagle, S. S. Schiffman, and R. Gutierrez-Osuna. 1998. How and why of electronic noses. *IEEE Spectrum* 35, 22–34.
- T. Narumi, S. Nishizaka, T. Kajinami, T. Tanikawa, and M. Hirose. 2011. Augmented reality flavors: Gustatory display based on edible marker and cross-modal interaction. In *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems (CHI11)*. 93–102.
- S. Pan and C. Ryan. 2009. Tourism sense-making: The role of the senses and travel journalism. *Journal of Travel and Tourism Marketing* 26, 7, 625–639.
- Picture Exchange Communication System. Retrieved from <http://www.pecs.com/>.
- L. Schauble, M. Gleason, R. Lehrer, K. Bartlett, A. Petrosino, A. Allen, K. Clinton, E. Ho, M. Jones, Y. S. Yee, J. E. Phillips, J. Siegler, and J. Street. 2002. Supporting science learning in 266 museums. In *Learning Conversations in Museums*. 425–452.
- M. J. Smith and P. J. Kiger. 2006. The lingering reek of smell-o-vision. *Los Angeles Times*, Feb-2006.
- R. Steinmetz. 1996. Human perception of jitter and media Synchronization. *IEEE Journal on Selected Areas in Communication* 14, 1, 61–72.
- J. N. Kaye. 2001. *Symbolic Olfactory Display*. Massachusetts Institute of Technology.
- E. Richard, A. Tijou, P. Richard, and J. L. Ferrier. 2006a. Multi-modal virtual environments for education with haptic and olfactory feedback. *Virtual Reality* 10, 3–4, 207–225.

- E. Richard, A. Tijou, and P. Richard. 2006b. Multi-modal virtual environments for education: From illusion to immersion. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* LNCS, 3942, 1274–1279.
- L. A. Rowe and R. Jain. 2005. ACM SIGMM retreat report on future directions in multimedia research. *ACM Transactions on Multimedia Computing, Communications, and Applications* 1, 3–13.
- SBi4 radio v2, Exhalia. Retrieved from <http://exhalia.com/fr/>.
- A. Tijou, E. Richard, and P. Richard. 2006. Using olfactive virtual environments for learning organic molecules. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, LNCS, 3942, 1223–1233.
- C. Timmerer, M. Walzl, B. Rainer, and N. Murray. 2014. Sensory experience: quality of experience beyond audio-visual. In *Quality of Experience, Ed: Springer*, 351–365.
- Vortex Activ USB. 2014. *Dale Air*. Retrieved from <http://www.daleair.com/dispensers/machines/vortex-activ>.
- D. A. Washburn, L. M. Jones, R. V. Satya, C. A. Bowers, and A. Cortes. 2003. Olfactory use in virtual environment training. *Modeling & Simulation* 2, 3, 19–25.
- Z. Yuan, G. Ghinea, and G M. Muntean. 2015 Beyond multimedia adaptation: quality of experience-aware multi-sensorial media. *IEEE Transactions on Multimedia* 17, 1, 104–117.

Received January 2016; revised May 2016; accepted May 2016