Kent Academic Repository

Downloaded from

The version of record is available from

This document version

Author's Accepted Manuscript

DOI for this version

Licence for this version

UNSPECIFIED

Additional information

Versions of research works

Is Multimedia Multisensorial? - A Review of Mulsemedia Systems

ALEXANDRA COVACI, Brunel University London
LONGHAO ZOU, Shenzhen University
IRINA TAL, Dublin City University
GABRIEL-MIRO MUNTEAN, Dublin City University
GHEORGHITA GHINEA, Brunel University London

Mulsemedia - multiple sensorial media - makes possible the inclusion of layered sensory stimulation and interaction through multiple sensory channels. The recent upsurge in technology and wearables provides mulsemedia researchers a vehicle for potentially boundless choice. However, in order to build systems that integrate various senses, there are still some issues that need to be addressed. This review deals with mulsemedia topics remained insufficiently explored by previous work, with a focus on *multi-multi* (multiple media - multiple senses) perspective, where multiple types of media engage multiple senses. Moreover, it addresses the evolution of previously identified challenges in this area and formulates new exploration directions.

1 INTRODUCTION

You are perched atop of a motorcycle, speeding up through Brooklyn with the wind whipping through your hair. The rattle of the engine rumbles your backside, the handlebars shake to the beat of the road, the sound of the engine and of the surroundings are delivered in full stereo, and an indefinable city whiff hits your nose. Then your tokens run out. You have just experienced the Sensorama Simulator, a machine invented in 1962 by Heilig [100]. In the spirit of Aldous Huxley's Brave New World, where sensory elements were used to heighten the "feely effects" of a movie, Heilig created an immersive environment by playing a 3D film along with stereo sound, scents and

wind. Sensorama stands out as one of the early examples of using technology to create alternate realities. There is an obvious correlation between Sensorama and today's multimedia and virtual reality (VR) projects.

In order for systems like Sensorama to successfully simulate reality, they need to integrate several sensory channels into a unique multisensory experience by following the brain's perceptual mechanisms [111]. VR is becoming a promising framework to immerse people into realistic environments and triggers the development and evaluation of various prototypes, displays, interfaces and interactions with virtual objects [41, 74, 134]. However, many multimedia systems lack the inclusion of layered sensory stimulation and of interaction through multiple sensorial channels. Alas, over the past years, physically cumbersome technologies have generally restrained the delivery of truly immersive experiences.

In multimedia applications, research and technology developments are generally focused on improving image and/or sound in audiovisual (AV) systems. This limitation of stimuli [61] maintains a level of disconnect between the user and the represented scene. The potential impact on enriching the interaction through combinations of AV with one or several different types of feedback (e.g., haptic, olfactory, gustatory) remains underused in digital media. Moreover, recent advancements in ultra-high definition video technologies are not sufficient for higher immersion because the simulated senses remain visual and audio. This, notwithstanding the fact that complementary pieces of multisensory information could push the sensations of immersion way beyond anything experienced so far [35]. Indeed, the importance of sensorial feedback in user immersion has increasingly been studied over the last years, with contributions focused on olfactory [85] and haptic aspects [163].

Mulsemedia was coined as a type of multimedia involving more than three senses in [61]. Back then, just as Sensorama envisaged in the 60s, the leap towards this transportive world was still difficult because of the effort behind incorporating other senses and the lack of multisensoriality on the devices market. Nowadays, the exponential upsurge in technology and wearables is providing mulsemedia researchers a vehicle for potentially boundless choice. The underpinning technology is becoming incredibly widespread determining us to head towards a future economic "golden age" of technological convergence in 2020s. Roto VR chair[1] offers a Sensorama-like experience where the user can feel wind, heat, scent and force feedback. Wearables are permanently evolving from their single function of counting the number of daily steps to measuring brain activity and correlate it with specific sounds (Muse headband[2]). Kickstarter projects that promise to engage in various ways all the senses have emerged, trying to find pilot clients and development partners.

The dynamics of these changes motivate the current review that puts multisensory media challenges in an actualized context and illustrates technologies and research problems involved in building systems that integrate various senses. Existent surveys [61, 194] offer a useful introduction to the field of mulsemedia with a focus on applications that have a multisensory dimension. Other works were focused on specific subsets of senses: olfaction [58, 130], or haptics [35]. These surveys presented detailed aspects of the production-distribution-rendering workflow, with a scope narrowed down on the addition of one sense. However, this paper has a wider remit (illustrated in Figure 1) than the aforementioned surveys and brings a new perspective on mulsemedia, situating the technical problem of communicating via multiple senses within the underlying affordances of human sensory and cognitive systems.

In the present review, we extend the work carried out in previous multisensory surveys and we address aspects that should be taken into consideration when starting to develop mulsemedia

---

[1]Roto VR: http://www.rotovr.com
[2]Muse Headband: http://www.choosemuse.com/
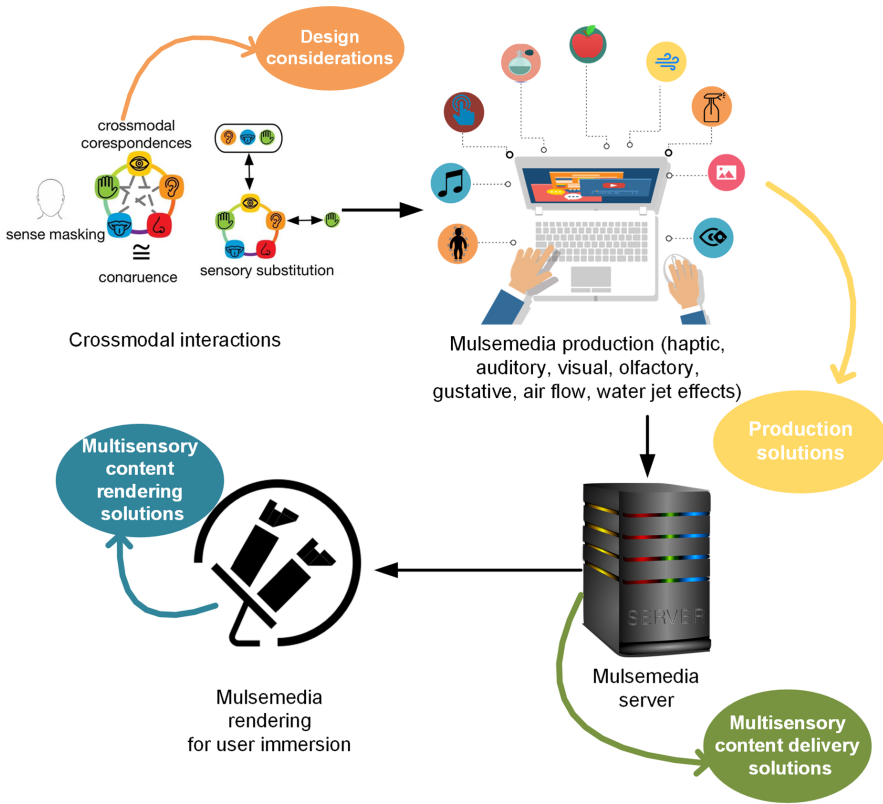
Fig. 1. The main directions of the survey

systems. Moreover, we show how our perception is influenced by senses, how we can substitute a sense with a combination of other senses or what happens when a sensory stimulus is stronger than the others. We believe that crossmodal correspondences should be considered alongside semantic, temporal and spatial congruency in the design of products from food experiences to mulsemedia systems because they solve the crossmodal binding problem, generate our unique perceptual experience and could aid in the rehabilitation of people with sensory impairments through the use of sensory substitution devices. Importantly, whilst models of multisensory perception and processing have been proposed in the literature [9, 17], there is relatively little work focusing on how one maps such models into the digital realm of mulsemedia, and this is the main remit of the work reported in this paper. This survey follows four main directions, as depicted in Figure 1, such as: (1) offering a set of information about different types of congruences that people who design mulsemedia systems could reflect on. In the remaining of the paper, we focus on design issues related to the (2) production - (3) distribution - (4) rendering chain as we synthesize available software and hardware solutions for building mulsemedia systems. Thus, we supplement the aforementioned surveys by focusing on specific challenges and solutions characteristic to all phases of the mulsemedia work flow and we identify some potential areas for future research in the area.

This paper is thus a comprehensive, up-to-date survey of research on mulsemedia systems that aims to provide the research community with a detailed insight on this relatively new research field. Additionally, our aim is to provide fundamental guidelines and knowledge for allowing readers to

start designing or performing experiments with such systems. This work comprises the following main contributions:

- A thorough analysis of the current state-of-the-art in mulsemedia systems with the emphasis on their challenges;
- An overview of different opportunities that provide designers and developers a starting point in creating meaningful mulsemedia systems where they make use of a wide spectrum of sensory experiences;
- A discussion of the challenges of enriching digital content with multisensory features in different phases of the production-distribution-rendering work flow;
- Identification and discussion of future challenges/research directions.

2 MULTISENSORY PERCEPTION AND INTERACTION

We rarely experience senses in isolation. Real world events produce different sensory signals that comprise our perceptual experience as a coherent percept through the integration of information from various independent channels. Multisensory interactions occur at the intersection of two or more sensory modalities and are essential in constructing a meaningful representation of the environment. Humans are equipped with multiple senses in order to sense the available information; thus it is likely that our brain has evolved to learn, adapt and operate optimally in multisensory settings. The process of **multisensory integration** describes the synergy among the senses and the fusion of their information content. Multisensory integration is more likely to occur under time and space coincidence constraints and it is enhanced by semantic and synaesthetic congruency on multisensory information processing [186]. The key idea is that we can use the findings and research methods from the field of cognitive science to benefit the process of design and evaluation of crossmodal mappings for multimodal user interaction and information display. Exploring the multisensory phenomena in human computer interaction (HCI) might lead to the development of more intelligent and adaptive user interfaces. In this section: (i) we look at multisensory processing with a focus on the benefits of multisensory stimulation, and (ii) we explore the relevance of crossmodal correspondences in the design process with a focus on how multisensory technologies could transform the way we interact and experience.

2.1 The Selection of Human Senses

When it comes to the number of senses that could be mixed in building a multisensory system, the answer could be anything between five (the Aristotelian senses) and twenty-one [46]. Exteroceptive senses, corresponding to a sense organ, are that of sight, hearing, taste, smell, and touch (tactile/haptic). Sensation is carried out by these sense organs that convert energy from the environment to electrical pulse. This electrical signal is transmitted to the brain, where it is processed and interpreted creating a perception. However, humans have also interoceptive capabilities that make them aware of the internal state of the body. These can be broken down into the following categories [64]:

- **Equilibrioception** contributes in helping us maintain our balance. Although vision plays a main role in this sense, the vestibular system of the internal ear provides the leading contribution to the sense of balance and spatial orientation.
- **Nociception** is the sense of pain. Seen initially as an experience related to touch, recent research showed that this phenomenon corresponds to a specific area of the brain [45].
- **Proprioception** is the awareness of the position of our body. This kinaesthetic sense is responsible for conveying information of where our body parts are, even if we cannot see them.

- **Thermoception** is the sense of heat and cold that relies on the temperature sensors in our skin.
- **Temporal perception** is related to the perception of time. Although this can be subjective, research shows that our basal ganglia and other parts of the brain are responsible for it [53].
- **Interoception** conveys information about our visceral sensory receptors found in our internal organs.

Consistency between proprioceptive information and sensory feedback is essential when building effective environments, increasing the sense of immersion. VR systems like "Amphibian" SCUBA simulator, developed by the MIT Media Lab, engage sensory modalities additional to AV like thermoception, equilibrioception, and proprioception [88]. Users wear an Oculus Rift head-mounted display and headphones that allow them to see and hear the underwater environment. Buoyancy, drag, and temperature are simulated through various sensors. The participants have Peltier modules attached to their wrists to simulate temperature changes as they dive deeper into the water and use an inflatable airbag placed under their torso that allows them to control ascent or descent through breathing while the movements of their virtual representation are synchronized. However, despite the efforts of the community, interoceptive senses have made little progress.

Through the synthesis and fusion of different sensory signals, multisensory integration creates a coherent and unified perceptual evaluation of our surroundings [190]. To successfully combine signals from different sensory channels, the brain needs to solve a correspondence problem, integrate the related information and dynamically adapt to spatial or temporal conflicts across senses [189]. Each sense has its own sphere and it is important in a different context. Choosing the appropriate sensing channels should be determined by environmental constraints and by the tasks or information the user has to handle or her/his abilities.

2.2 Multisensory Processing

HCI studies, most specifically in cognitive load theory, Gestalt theory or Baddeley's model of working memory illustrate many advantages of designing multisensory interfaces [8, 149, 150]. Findings in cognitive science related to multimodal interaction show that providing an individual with multiple sensory cues expands processing capabilities and leads to faster acquisition and retention of information [179]. Research demonstrated that the presentation of non-visual cues (e.g., tactile, auditory, audiotactile) has a beneficial impact on the performance in visual tasks [205]. Moreover, it was shown that in certain circumstances, non-visual warning signals provided benefits that visual cues cannot offer [205].

Little research has investigated how multisensory information is represented in working memory. In [200], the authors used crossmodal stimuli in their study, and found an improvement of free recall of crossmodal audiovisual stimuli compared to modality-specific, audio or visual stimuli. These findings suggest that the the combination of information from different modalities (not the redundancy) leads to an improved memory performance. The advantage of crossmodal object recall can be explained through the dual-coding theory.

Having as starting point the generative and the dual-coding theory, Mayer proposed a model of multimedia learning, where visual animations that were accompanied with audio information led to a split-attention effect in which students learned better [115]. This model consists of three analogous processes that take place in the dual planes: **Selection** - here the user selects relevant information from the sense receptors into visual and verbal short-term memory; **Organization** - involves a dual process of knowledge transformation within the respective short-term memories; **Integration** - where one builds one-to-one references between the verbal and textual facets of

the information. This process takes place in *working memory*, which is part of the short-term memory. However, it must be borne in mind that, due to the limited capacity of short-term memory, integration of the visual and verbal models is constrained by the memory load. At this stage, information has the potential to be absorbed into long-term storage, but it usually requires effort on the part of the learner.

Hypermedia learning depends strongly on individuals' spatial ability, i.e. their ability to mentally map out the information presented to them, complete with the locations and routes of the hypertext structure. Very often, individuals with high spatial ability create interior maps, which are superior to the ones provided externally. This prevents disorientation and reduces the cognitive overload that might potentially be experienced [154]. Learners attend differently to various stimuli and invest intellectual effort accordingly. Thus, learners invest more mental effort in media that they perceive to be difficult, than in those perceived as being easy. This would suggest the use of contrived methods such as highlighting in order to grab and direct the attention of the viewer [154]. Video and audio can be successfully used in a multimedia environment to improve the quality of learning as long as they do not compete for the same processing resources. Indeed, concomitant video events can help users filter out unnecessary sound effects and focus on relevant audio information [70]. These findings were applied to educational interface design based on user-centered principles and were shown to minimize the cognitive load and to lead to a better performance [149].

Providing an individual with multiple sensory cues should facilitate a representation; thus, numerous educational programs have advocated the benefits of mutually supportive multisensory information. Indeed, psychologists and neurologists have advanced the field of multisensory perception over recent decades. Research has shown that the facilitative effect in information processing is significantly greater for multisensory stimulus combinations than within-modal combinations [62]. Multisensory processes facilitate memory [126], as well as perceptual and implicit learning and training outcomes [6]. In [68], the authors have shown that multisensory audio/tactile cueing improves the performance (speed and accuracy) of visual search and reduces the amount of mental workload. Multisensory intervention has been successfully used also as a therapy for correcting and improving literacy skills [155] or for increasing sustained and selective attention of people with autism [49]. Moreover, in this case, the multisensory experience was even more effective by increasing students' engagement with people and objects. Technology advances now provide the opportunity to combine and connect neuroscience knowledge with game design thinking, creating toolkits for learning, training, and therapy.

2.3 Cognitive Foundations

Multisensory integration combines information from different sensory channels with positive effects on accuracy, precision and reaction times. Some of the elements that govern multisensory processing are: **the temporal and spatial aspects** (two signals should happen at the same time and come from the same location), **semantic congruency** - sensory information should share the same identity (this happens for both contexts and objects), **crossmodal or synaesthetic correspondences** - information should be cross-sensory compatible. These factors act in conjunction during perception and considering them might guide the process of multisensory and, by implication, mulsemedia design.

2.3.1 Semantic congruency

Semantic congruency is one of the factors that underlies the integration effect and refers to situations where the identity and/or meaning of pairs of stimuli match or mismatch. In the presence

Table 1. Crossmodal correspondences between different sensory dimensions

		Sound			Sight		Touch	
		Pitch	Timbre	Spectral Balance	Shape	Color	Temperature	Softness
Smell		high-fruity [31, 182]	piano-fruity [31]	n/a	angular-lemon [69]	paired [110]	warm-cinnamon [69]	soft-lemon [37]
Taste		low-bitter, high-sour, staccato - salty [66, 120]	trombone-bitter, piano-sweet [10, 28–30]	low-sweet high-sour [14]	symmetry-sweet, pleasant [202] angular-bitterness [40]	pink-sweet, white-salty green-sour, black-bitter [188]	n/a	rough-sour [184]
Touch		variation-direction [38], high-cold water [212]	n/a	n/a	n/a	light color-soft smooth, thin, light [213]	n/a	n/a

of redundant sensory information, their integration can increase signal strength and reliability, thus semantic congruency becomes a critical factor in multisensory behavioral performance. Semantic congruency was shown to increase visual attention through both spatial and object-based interactions in a visual search paradigm presented in [84], where participants found a cat in a masked image faster when hearing a meow sound. Positive affective effects when using semantical congruent stimuli were illustrated in olfactory-audio or olfactory-visual combinations in [55, 178, 185]. In [178], participants were exposed to a scent (e.g., odor of an orange) while they were asked to explore a picture containing a corresponding odor-related visual cue (e.g., image of an orange) embedded among other objects. The results showed that the odor-related visual cue was explored faster and for a shorter time in the presence of the congruent odor, highlighting that olfaction can affect visual processing by capturing people's attention. Similar results were found in the case of an information recall task or of a word search game when cue-related odors were emitted [3, 56]. Another study, where associations between the odorants and motion directions were arbitrarily established by means of classical Pavlovian conditioning, consolidated the functional linking between the olfactory information processing and the visual information processing are functionally [97]. Castiello et al. [19] showed the effects of olfactory information on the process of selection for the control of a goal-directed action illustrated by the task of grasping. The authors administered an odor of an object requiring a hand aperture similar in size to (congruent) or different in size from (incongruent) the one of a certain visual target and demonstrated that the size of the object evoked by the smell influenced the subsequent grasping kinematics. These findings show that in a sensorium dominated by vision and audition, olfaction influences perception and cognition in adults and opens up new interaction possibilities.

2.3.2 *Synaesthetic congruency*

Multisensory binding is enhanced also when crossmodally congruent stimuli are presented [186]. Synaesthetic congruency refers to correspondences between basic stimuli features (e.g., pitch, timbre, shape, color) in different modalities. Crossmodal correspondences were shown to influence people's performance under different experimental paradigms: direct crossmodal matching, faster classification tasks, faster simple detection tasks, Implicit Association Tests, spatial localization tasks, and perceptual discrimination tasks [186]. In Table 1, we summarize a set of crossmodal correspondences that have been shown to influence participants' performance. Crossmodal correspondences were demonstrated between various pairs of stimuli: pitch was shown to change the perception on different gustatory dimensions (high pitch was matched with sweet/sour taste, while low pitch was matched with bitter taste [10, 29, 30]). High pitch was associated also with fruity smells [31, 182] and in the haptic dimension with cold water [212].

Table 1 presents the pairs of stimuli that have a positive effect on performance and that should be considered in any mulsemedia design process.

2.3.3 Congruences and HCI

Semantic and crossmodal mappings have a great potential in the design of meaningful computer interfaces and interaction methods. This is because they support comprehension and retention of information through the accommodation of users' sensorimotor skills. Moreover, computer interfaces that explore the usage of crossmodal effects and correspondences have the potential to be more effective and to increase immersion. Promising applications were illustrated in [43], where the authors explore the efficiency of olfaction in introducing a new semantic layer into interaction design and HCI. The study analyzed different mappings between driving-relevant messages and scents, and proposed ways of using olfactory simulation to transfer specific information to the user, based on the congruence between visual and olfactory information. Strong associations were found between the "Slow down" message and the scent of lemon, the "Fill gas" message and the scent of peppermint and between "Passing by a point of interest" message and the scent of rose. In [198], the authors explored the crossmodal correspondences between haptic and audio output through "atmoSphere", a sphere that provides haptic feedback designed to augmented focus during mindfulness training by guiding the users into a particular rhythm of breathing.

Because all our senses interact to influence our experiences, different types of sensory cues can be used to guide or modify sensory expectations, search and augmentation. Since some of our senses have currently a limited use in designing experiences (e.g., taste is currently explored via digitally controlled electric or thermal sensations), there is also potential in designing applications, built on top of crossmodal mappings, that target the user's flavor expectation. For instance, extrinsic cues of flavor can be augmented based on the crossmodal correspondence between roundness and sweetness [207] or between musical articulation and the four basic taste groups [66]. Thus, data can be encoded with relevant parameters that can be perceived as synonymous in different sensory modalities. Inspired by findings from research on crossmodal perception, technology was used to augment the way we experience the flavor in a variety of applications. Visual-taste crossmodal stimuli were used in a mixed reality installation where a colored animated shape was projected on yogurt, changing its taste [15] or in a study that proposes an augmented reality system that changes the flavor of chips through increasing the chroma of a plate [143]. Crossmodal correspondences can be exploited by interactive systems to reinforce or improve different sensory perceptions. Thus, we suggest that crossmodal correspondences could play an important role in designing HCI from research to more commercial settings.

Sensory Substitution Devices (SSDs) are most of the time complicated to use, since they require intensive training. Crossmodal correspondences could be useful in designing such devices that semantically translate cues of a specific type (e.g., visual information: brightness, shape) into another dimension (e.g., audio information: pitch, loudness). Thus, a more efficient alternative to existing SSDs could be to engage the sensory substitution via multiple sensory channels using different crossmodal mappings. As a result, the process becomes more intuitive, training time is shortened and the participant performance is improved [192]. Encouraging results were found for color to sound [67], tactile to sight [16], and sounds to smell substitutions [26].

3 PRODUCTION, DISTRIBUTION AND RENDERING IN MULSEMEDIA SYSTEMS

Mulsemedia systems are usually obtained by adding feedback to visual and/or audio stream which triggers other sensorial channels (e.g., haptic, olfactory, gustatory). This process consists of three stages: (i) Production, (ii) Distribution and (iii) Rendering. In [34], the authors propose a workflow for adding haptic feedback to audiovisual (AV) content. In this section, we extend this workflow
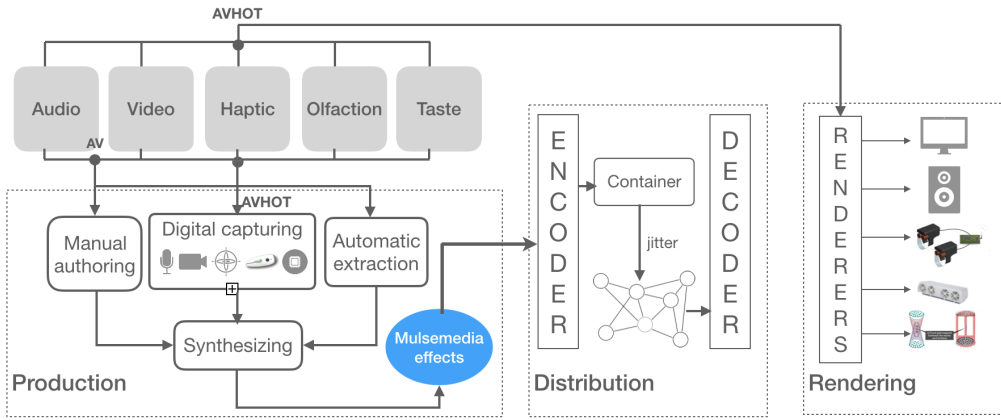
Fig. 2. Production-distribution-rendering workflow for creating multisensory systems. AV = Audio Video, AVHOT = Audio Video Haptic Olfaction Taste

with details regarding production and rendering of systems that stimulate also the chemical senses (taste and smell), as illustrated in Figure 2. Accordingly, **production** deals with how different sensory effects can be created or generated in synchonization with the AV content. Multisensory effects can be produced through different techniques: **digital capturing** and processing of data acquired from sensors; **automatic extraction** from video or AV content; or through **manual authoring** of mulsemedia effects. Once the mulsemedia effects are captured, they are **encoded** for transport, processed and then emitted for **distribution** to providers, distributed to the end users and then **decoded**, so that the mulsemedia effects can be **rendered** by different devices and experienced by the end users. In order to be effective, mulsemedia effects have to be synchronized temporally.

Since [35] provides an extensive overview of the haptic audio visual (HAV) aspects, in this section we will focus on new research in the haptic area and on the addition of chemical sensory modalities, which differ in the production/rendering mechanism from the physical ones.

While we understand sight, hearing and touch good enough to build machines to mimic them, we know less about the chemical basis of taste and smell. Cameras are electronic eyes, mechanical resonators transmute sounds into signals our brain recognizes, touch is a pressure sensor, but smell and taste are subtle senses, complex mixtures of different molecules. Both sensing and actuation of smells are hard to realize because of their organic nature. Thus, despite their influence on the human behavior and the interesting possibilities that they raise (for instance we have stereo smell and can distinguish the origin vector of a scent to approximately 10 degrees of acuity [209]), there is a surprisingly paucity of use of scent and taste in VR, HCI techniques and mulsemedia. In the following subsections, we will offer informative insights that extend the *ad hoc* knowledge about mulsemedia design into a formal framework. We expect that academia or industry practitioners who are not experts in mulsemedia design will have a starting point that will allow them to accelerate the generation of multisensory tools, techniques and test scenarios.
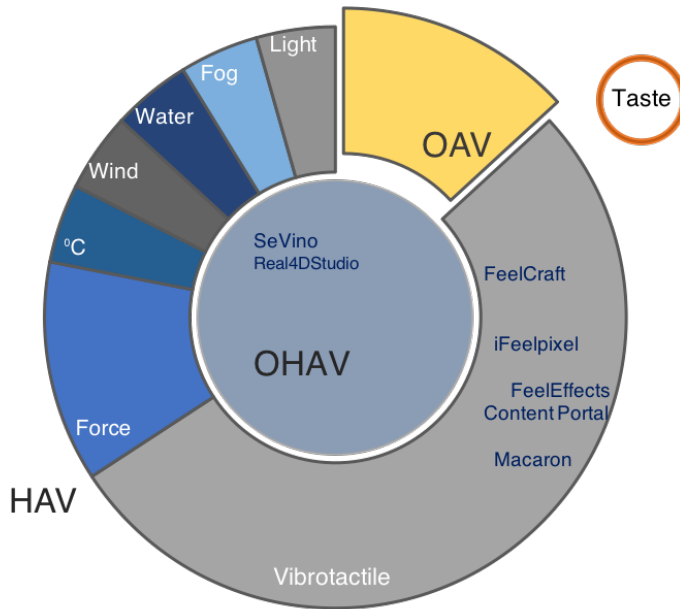
Fig. 3. A Histogram of Different Existing "Sense" Editors. Haptic Audio Video (HAV) editors are the most popular with several open-source tools. Olfaction Haptic Audio Video (OHAV) are less represented. Olfaction Audio Video (OAV) and Audio Video tools enhanced with Temperature, Light, Fog, Wind, Water, Light, Force are mostly experimental, while Taste editors are not yet implemented.

3.1 Production

Although most human-technology interactions are still based on traditional desktop/mobile interfaces that involve primarily the visual and audio senses, in recent years we have witnessed a progress in multisensory experiences involving haptics. Though much development is needed, the perspective is that the senses we use in our interactions with technology will not be restricted and we will be able to create richer experiences.

Notwithstanding the fact that audio and video content can be easily produced with a variety of tools and methods, development of new interaction methodologies and the use of new interaction devices is more problematic since it involves the rapid prototyping of user interfaces [152]. Designers still manually create effects using in-house authorization tools in an event trigger scheme, resulting in a labor-intensive work. Moreover, often, guiding principles for creating such effects that are associated to the content watched, heard or read are missing [87]. Commercial companies like D-Box[3] or CJ 4Dplex[4] create haptic effects using proprietary tools, but they do not make available the lessons from their releases. Even if this technique does not require a change in the current media production process, a better placement of effects should be provided when a designer has the possibility to tune the result with authoring tools.

In this subsection, we will focus on how to create content that addresses various sensory channels (visual, auditory, haptic, olfactory, gustatory) in order to produce new body experiences through novel types of interactions. We will present libraries that could be used along with MPEG-V

---

[3]D-Box: http://www.d-box.com/
[4]CJ 4Dplex: http://www.cj4dx.com/

based protocols and editing tools for different types of sensory effects to be rendered with the AV stream. In order to offer an overview of the existing software-based editing tools that could be used to accelerate the prototyping of mulsemedia systems and the exploration of mulsemedia design, we looked at different solutions provided by the literature. In Figure 3 we illustrate the repartition of these editors based on the targeted sense. We found that existing collections of sensory sensations and programming toolkits are focused mostly on haptic effects (especially vibrotactile (VT) stimuli) [20, 86, 87, 92, 93, 104, 112, 152, 166, 173–175, 177, 197], while the olfaction field remains insufficiently explored [24, 92, 196, 211] and there is no tool for editing gustatory effects. These tools are presented in more detail in Table 2, Table 3 and Table 4, respectively.

3.1.1   *Production of AV Stream*

Visual content has evolved from traditional 2D videos to stereoscopic 3D panoramic images. However, 3D content is often expensive to produce. 2D to 3D conversion techniques have the potential to make the process faster and several approaches have been proposed for this: manual, human-assisted and automatic conversion. Moreover, 3D visual information can be also produced using capturing devices like 3D cameras and 3D scanners [151]. Mobile phones and tablets can be used for stereo-based depth measurement opening up new possibilities for 3D reconstruction [94]. 3D printers started to be used to provide a tactile dimension to traditional visual representations: for cosmic representations [25] or for visualizing anatomy [1].

Sound had until recently a marginal role in the design of virtual environments or mulsemedia systems. With the development of VR, immersive audio became something crucial because it needs to match up 3D visuals that convey an enhanced sense of realism. Recent research started to evaluate the influence of using 3D sound in improving performance or spatial localization in VR systems, showing that it enhances the immersion level of players [23][144]. Binaural recording systems is a specific approach of the stereo recording where two microphones are located in place of the two ears on an artificial head [95]. Binauralisation engines that utilize Head Related Transfer Functions (HRTFs) are available from plugins such as 3DCeption from Two Big Ears[5] which captures interaural level and time differences as well as spectral cues. 3DCeption ensures that the audio is synchronized with the video, regardless the playback mechanism and ensures all kinds of audio: mono, stereo, multi-channel and ambisonics across platforms and devices.

3.1.2   *Production of Haptic Effects*

In this paper, we use the HCI definition for *haptic* that involves one or more perceived sensations of touch (tactile and proprioceptive feedback, active human touch, and passive experience of actuated technology).

Haptic effects occur within a scene with the aim to augment the user's presence, to enhance ambiance or emotion or for training purpose. A comprehensive classification of the types of haptic effects that can be used in multisensorial systems is presented in [210]. They include: *temperature, water sprayer, air-flow, whole body vibration, passive kinaesthetic motion and force, active kinaesthetic, tactile and rigid body*. Haptic feedback can be produced in free space in several ways: through conventional direct-contact wearable devices (e.g., gloves [109], haptic vest [106], manipulation of the location of the actuators [171] or contactless haptic feedback through air-jet or ultrasonic radiation pressure [7].

Haptic technology can produce stimuli to feel in diverse ways, the most common approach being the vibrotactile (VT) feedback. VT arrays (gaming chairs or tactile vests [158]) are a haptic technology that uses parametric models of sensory illusions to render moving haptic patterns. In

---

[5]Two Big Ears and 3Dception: http://twobigears.com/spatworks/index.html#features

fact, VT signals exist in many everyday devices and a study presented in [176] confirms users' interest in customizing these effects. However, despite this potential and the rise of new haptic interfaces, there is still a gap between haptics and mainstream communication caused by the lack of haptic authoring tools, the production infrastructure, the standardized playback protocol or the design of engaging, understandable effects. The design of intuitive and effective VT cues requires multiple development iterations and user evaluations. Haptic technology, already used to create immersive user experiences in VR or video games, seems to display a great potential to also enhance user experiences in mulsemedia [111]. However, mainstream media does not yet use the richness of the haptic modality because of the lack of haptic authoring tools, standardized playback protocols and trained workers.

Haptic feedback is mostly manually customized through dedicated editors with effects created during the post-production phase. Haptic information can also be produced through data acquired and processed directly from the sensors, or through automatic extraction from AV content or from metadata [168] (e.g., effects from a football game displayed with a vibrating device [203]). For instance, activity in a video feed [90] can be mapped to haptic transducers that are arranged along a chair; movements from the visual scene are thus mapped with events from 4D movies or rides. Similarly, sound can also be used to extract haptic cues for music or video gaming.

In embedding haptic information in media, designers need to have access to haptic effects tools and libraries that enable rapid prototyping and easy authoring. As seen in Figure 3, haptic editors are favored over authoring tools for other senses. Several interfaces that support different types of devices are currently available. Some of them are based on direct representation of the vibration signal as editable waveforms [152]. The Hapticon Editor [48] is used for creation and editing of haptic icons for a 1-DOF[6] force feedback device. VT icons can be prototyped with posVibEditor[166] that supports also the creation of a library of patterns using XML-based pattern files. Other editors also support media augmentation through the predefinition of a library of haptic patterns. The Touch Effects Studio (Immersion Corporation) and Vivitouch Studio [197] are focused on devices with either a temporal or spatial component through tools for enhancing video and audio with effects from a library of haptic icons. Schneider, Israr and MacLean in [174] introduced the tactile animation metaphor and proposed Mango to allow users to efficiently create dynamic 2D haptic patterns, associate them to multimedia and apply them to a variety of devices. Other editors are built for authoring haptic content using accessible touchscreen interactions. Simplistic patterns, that allow users to control only the vibration duration can be created using Apple's vibration editor (since *iOS 5*). Moreover, in [79], Apple's approach was extended on an Android platform to support complex vibration waveform design in an intuitive way (sequence of taps and drags).

3.1.3 *Production of Chemical Effects (Olfactory and Gustatory)*

Unlike the other senses, smell and taste are complicated processes that result from the biomechanical reaction between human receptor(s) and a binding site on a molecule. In fact, an aroma is not identified due to one single molecule, but due to the interaction of many molecules and biochemical reactions. Although there are many challenges related to studying taste and smell, no other sensory modality is as emotionally potent as olfaction [148]. It has been shown that smell reinforced the way a viewer enjoys an artwork [101] or that smell and taste have the potential to facilitate the acquisition of new knowledge based on previous experiences in education [73]. With the advance of technology, and the understanding achieved in psychology and neuroscience, researchers started to shape the development of technologies based on smell and taste [141]. 'Virtual Cocoon' is a helmet that addresses all the five senses: hearing, sight and touch senses are stimulated digitally,

---

[6]DOF = Degrees of Freedom

while taste and smell are stimulated analogically, by chemicals emitted by the helmet [21]. This analog solution, applied for chemical senses, has several drawbacks related to its manageability and scalability.

Taste. There are five basic tastes: sweet, salty, bitter, sour, umami [36]. When we taste any food or beverage, we perceive each type of taste as a qualitative data analyzed by our brain indicating effects on the sensory organs called taste buds. Within HCI, taste stimulation is mainly achieved through the use of **chemical stimulation** or through **electrical and/or thermal stimulation** of the user's tongue. This invasive nature of taste and the fact that taste-based experiences are volatile represent important hurdles for HCI researchers [125].

Traditionally, users experience taste through chemical components (liquid or solid form): glucose (sweet), citric acid (sour), caffeine/quinine (bitter), sodium chloride (salty), and monosodium glutamate (unami) [146]. This approach has been used in [116], where the author proposed the new paradigm of Edible User Interfaces and evaluated its effectiveness in a HCI context. The "painted bits" of a computer monitor are replaced by tangible "edible bits", with potential applications for network monitoring and profiling. This paradigm is used also for 'Taste Screen', which consists of small transparent cartridges that are placed on the screen and release a flavoring agent when users lick it [116]. In [125], the researchers proposed an example of gustatory interface called 'LLOLio', an interactive lollipop that changes its taste. The authors focused on the sweet taste of the lollipop in combination with sour liquid to deliver a taste and haptic input modality. Results showed that gustatory games could benefit from the interaction between sweet and sour. While the chemical approach is less invasive than the electrical and thermal simulation of the tongue, it requires the continuous preparation of the taste stimuli that are difficult to store and manipulate. Moreover, the chemical stimulation is analogous, thus impractical for digital interactions.

Taste qualities can be perceived also by stimulating the tongue papillae through thermal or electrical or stimulation. Sweet sensation was shown to be produced by warming the anterior (front) edge of the tongue, while the sourness and/or saltiness were created through the cooling of the tongue in [32]. Sweet, sour, and bitter tastes were stimulated electrically by placing a silver wire on the tip of the tongue and a reference electrode on the left wrist of the subject in [156]. In [160], a team from the University of Singapore presented a device with a 2×2 grid of peltier elements that deliver heating and cooling stimuli to the tongue simulating the sensation of sweetness. Using this approach, the team proposed a "Vocktail" (Virtual Cocktail) that utilizes three sensory modalities (taste, smell, and visual - color) to create virtual flavors [162] and "Digital lollipop" that digitally simulated the sour sensation on the human tongue at three intensity levels. In [170], the authors presented a taste interface for future meaningful Internet multisensory communication through the digital actuation of sweet sensations. They created the sweet taste by manipulating the temperature on the tongue. Texture, another important factor in the eating experience, was simulated through the Electric Food Texture System presented in [142]. This device, placed on the masseter muscle can not only mimic the texture, but also modify the texture of real food and its creators are contemplating to combine it with chewing sounds. An advantage of the electrical/thermal approach is that they allow for more control and replication of taste stimuli over a long period of time and over distance (e.g., Taste/IP [159]). Currently, non-chemical taste stimulation methods are the least explored in interactive systems.

Another approach consists of pseudo-gustatory displays based on the the virtual color of a real drink [140]. "MetaCookie" is a system that creates customized tastes of the same cookie by dispensing different scents to users' noses with the motivation to investigate crossmodal effects of visual feedback and the interpretation of the flavor. "Affective Tumbler" uses the same crossmodal principles and applies thermal sensations on the skin to stimulate skin temperature changes [195].

In [208], the authors highlight through "TasteBud" how the complete spectrum of five basic tastes can be employed and propose a novel and innovative delivery mechanism that uses acoustic levitation - "TastyFloats". "Virtual Lemonade" introduces a new method and a digital platform for sharing lemonade using crossmodal principles that consists of: sensing (RGB color and pH value are captured), distribution (the XML protocol to encode the data) and stimulation through a tumbler that overlays color over the beverage and stimulates the tongue using electric impulses [161].

Commercially, *Planet Licker*[7] is a first step toward incorporating taste into game play by licking flavored ice pops on a USB-connected controller.

Smell. Natural odor molecules with the exception of a few molecules like hydrogen sulphide are small carbon-based entities that float into the air and form an invisible plume of odor. Although it is not clear how we detect molecules, there are some theories about this process. One idea refers to the "vibrational modes" of an odorant that are its signature [54]. This study suggests that an odor molecule can be seen as a collection of atoms on springs. The quantum effect causes the spring to vibrate and thus the molecule to be identified.

Smell in an environment can be captured manually by sucking the air across an automated thermal desorption tube. The odor molecules stick to the fine granular material in the tube. These molecules can be determined after they are separated by a gas-liquid chromatography instrument. A mass-spectrometer then produces a histogram of the present molecules [63]. Olfactory information can be produced through a system consisting of an electronic nose or an odor recorder. The classical electronic nose is based on an array of sensors that provides a pattern of a diagnostic given odor. Classical electronic noses utilize an array of chemical sensors of different specificities that respond to volatile compounds present in gases. This array of chemical sensors is coupled with a pattern recognition mechanism that allows the identification of complex smells [220]. The electronic nose can be used for olfactory recording if combined with a smell reproducing technique [133]. In [132], the authors present an odor recorder for dynamic changes in an environment and in [214], the authors proposed a method to record odors that consists of many components. In addition to the classical sensor-array-based approach, electronic noses based on other technologies have also been exploited. There are examples where mass and ion mobility spectrometers or flash gas chromatographs are used to detect the components of a gas mixture. Electronic noses have been successfully used in cancer diagnosis [107] or in food industry [51]. However, in spite of having different approaches and thus providing different input, the limits of electronic nose technology resulting from the fundamental sensing components are obvious [167]. Commercially, *NeOse*[8] combines different technologies to directly analyze and quantify volatile organic compounds and tests them against a database of known smells. *NeOse* allows companies to record their own scents in the database and has as main targets the food and cosmetic industries, but it can also test air quality.

Despite these attempts in developing gustatory and olfactory input devices, the main focus of the digital olfaction community in mulsemedia has been nonetheless on olfactory output modalities through studies on timing and multimedia processing [57, 129].

3.1.4 *Creating Multisensory Content; Ready-made Libraries and Editors*

As shown, multisensory feedback has the potential to convey meaningful information to users. However, despite the availability of a broad range of technologies there are no guiding principles for associating multisensory feedback with AV content. Therefore, different types of tools should be provided to a designer who wants to create engaging experiences, to meaningfully harness touch,

---

[7]Planet Licker: http://a-o.in/games/pl/
[8]NeOse: http://aryballe-technologies.com

Table 2. Multisensory Editors Based on Direct Signal Graphical Representation

Data Format	Effects	Tools	Type	About	Users
XML	Audio, Video, Vibrotactile	posVibEditor [166]	Standalone - edit and play	Drag-and-drop design paradigms. Waveform editing, multi-channel support for multiple actuators placed on body sites. Suitable for mobile devices and virtual reality applications.	Gen.
	3D shapes, Vibrotactile	VITAKI [112]	Standalone (C++ API) - device independent	Waveform editing, prototyping platform for decoding vibrotactile feedback (i.e. Morse signals) on the haptic gloves driven by Eccentric Rotating Mass (ERM) motors (e.g. object fall detection in VR gaming).	Dev.
	Audio, Vibrotactile	Beadbox [138]; Emoti-Chair [89]	Standalone	Timeline editing, interface design with "Bead" based on sensory substitution models (color-frequency).	Art.
MPEG-V Part 3 (XML)	Audio, Video, Vibrotactile, Wind, Water, Temperature, Light, Fog, Scent	SeVino[a] [211]	Standalone, Web service	Tool chain for video annotation and synchronization of different effects - device independent (writes MPEG-V based SEM and performs 3D simulation).	Dev.
		Real4DStudio[b] - authoring tool	Standalone		Gen.
		SMURF (extension of SeVino) [92]	Web service		Dev.
		RoSE studio [24]	Framework		Gen.
	Audio Video Vibrotactile Wind, Force	H-Studio authoring tool [34], PlaySEM [169]	Standalone	Edition of motion effects and synchronization with AV content.	Dev.
		Web haptic player [44]	Web service	Import 3D models and add haptic properties - haptic device independent.	Gen.
MPEG-2MPEG-4 BIFS	Audio, Video, Vibrotactile, Force	Authoring framework for haptic broadcasting [20][93]	Framework	Passive haptic interactions.	Dev.
	Video, Scent	Smell video system [196]	Standalone	Image processing and text analysis to extract smells and authoring tool.	Dev.

Note: Gen. - General area; Dev. - Development; Art. - Artists

[a]SeVino: http://selab.itec.aau.at/software-and-services/
[b]Real4Dstudio: http://cast.real4dhub.or.kr/about/authoring/download

smell, taste in mulsemedia interaction, to determine the meaningful design space for multisensory interactive experiences, to study and understand different sensory modalities along with their cross-sensory associations, etc. Using a framework or a library that includes fine-grained descriptions with their experiential correlates can speed up the design process [148]. While the libraries of effects might still be useful for end user customization, they do not support the capability for

Table 3. Metaphor-based Multisensory Editors

Data Format	Effects	Tools	Type	About	Users
MPEG-4	Video, Audio, Vibrotactile	Tactile movie authoring [93]	Standalone	Visual media metaphor (first, third person and background effect) for tactile authoring.	Dev.
		TactiPEd [152]	Standalone	Actuator sequencing metaphor for visualisation and adjustment of amplitude, frequency and sequence duration (separate control of each actuator).	Dev.
XML	Audio, Vibrotactile	VibScoreEditor [104]		Score-based editing using the musical score and VT clef metaphors.	Dev.
	Video, Vibrotactile	Mango [174]		Animation metaphor that allows the manipulation of vibrotactile sensations continuously in space and time.	PA.

Note: Gen. - General area; Dev. - Development; Art. - Artists; PA. - Professional Animators.

deconstruction and re-composition, they are opaque in construction and immutable. For experts, who want access to the source code, editable model and external examples would be necessary in the design process.

Table 2, Table 3 and Table 4 present different types of multisensory design tools starting from VT libraries and continuing with editors based on variations of MPEG-V format [201] that support different types of haptic effects (vibration, temperature, wind, water-spray, motion) as well as scent, shadow or fog. Some of these editors are track-based with a graphical representation to edit waveforms (Table 2), some are based on different types of metaphors (Table 3), some address additional features like mobile interfaces while others are example-based tools. In Table 4 we present libraries of effects that can be used for keyframe based behavior editing and refinement. Based on the level of required expertise and on the multisensory channels that need to be stimulated, these tools could be used to make the prototyping process faster.

3.2 Distribution

In the mulsemedia content distribution, the formalized sensorial effects are stored, transmitted and synchronized through authoring tools/platforms (see Table 2, 3 and 4) over the heterogeneous network environment [35]. In networked systems there are many concurrent processes making use of the same resources. When simultaneous signals are evoked from more than two sensory modalities, the user's attention is quickly directed towards the source and the Quality of Experience (QoE) is enhanced [165, 216]. These sensory inputs become coherent through crossmodal integration. As a result, we make better temporal or spatial judgments [12], or we have a richer experience [47] of the event that generated those signals. In mulsemedia systems, sensory inputs are integrated according to the temporal interval between the stimuli[190]. When the real-time data streams are transmitted over a network, the delay jitter disturbs the temporal relationship between the media streams. Moreover, a temporal ordering of sensorial events has to be guaranteed in the mulsemedia synchronization systems [2].

Scant research in the mulsemedia domain has focused on identifying the thresholds that allow different senses to be separated while remaining integrated. The procedure to identify the point where the bond between sensorial channels breaks, consists of separating the events in time and space and varying the increments until they are no longer perceived as an entity [42]. These

Table 4. Multisensory Editors based on Libraries and Databases

Data Format	Effects	Tools	Type	About	Users
Libraries	Vibrotactile	VibViz [177]	Standalone	120-items VT library (annotated by 5 taxonomies: physical, sensory, emotional, usage examples, metaphor).	Dev.
		FeelEffects [87] and FeelCraft plugin[a] [173]		50-items VT library (6 types of metaphors - rain, travel, etc.).	Gen.
		Macaron[b] [175]	Web service	Open-source interactive design editor that facilitates remixes of existing vibrations.	Gen.
		Feel Messenger [86]	Standalone (Android)	Communication API, haptic vocabulary (haptic patterns stored as feelgits with preset parameters - feelbits), and interface for receiving and authoring haptic messages.	Gen.
	Vibrotactile, Force	TouchSense SDK[c] (Android), ViviTouch Studio [197]	Standalone/ plugin	Haptic prototyping alongside video and audio (pause, seek, mute, synchronize effects).	Gen.
	Vibrotactile, Force, Audio	iFeelpixel[d]	Standalone	Enables users to interact with the computer GUI using vision, touch and hearing and allows them to mediate graphic structures.	Gen.
Databases	Video, Audio, Vibrotactile	Content Portal [e]	Standalone (Android, iOS)	Showcase of video, media, and games enhanced with tactile effects.	Gen.
		Demonstration-based editor [80], Apple's vibration editor (only binary on/off information)		Generates a VT pattern by imitating the touch input pattern of the user.	Gen.

Note: Gen. - General area; Dev. - Development; Art. - Artists; PA. - Professional Animators.

[a]FeelCraft: https://github.com/Litarvan/Feelcraft
[b]Macaron: https://www.cs.ubc.ca/labs/spin/macaron
[c]TouchSense SDK: https://www.immersion.com/products-services/touchsense-sdk-mobile-apps/
[d]iFeelpixel: http://www.ifeelpixel.com/download/
[e]Content Portal: https://play.google.com/store/apps/details?id=com.immersion.tsengageportal&hl=en

temporal in sync/out of sync boundaries were explored for audio, video, olfactory [57, 131], haptic and airflow [215] combinations as shown in Table 5, where we present different asynchrony thresholds between media channels.

Media synchronization has been evaluated as the relationship between two types of media streams. The investigation of intermodal asynchrony was initially focused on audiovisual events. In [42], the authors observed that speech and non-speech stimuli had different thresholds (non-speech: {-75ms, 175ms} speech: {-130ms, 250ms}). Reinforcing this, [119] proposed a lip-reading task and showed that the performance of users dropped significantly when the soundtrack was delayed by 160ms. In [191], the users' perception evaluations of lip synchronization for audiovisual events

Table 5. Quality of Service for Synchronization Purposes

Media 1	Media 2	Synchronization Type	Application	Tolerable Range
Video	Scent	Audio media dependent [57]	In the absence of **audio media** [129]	[-5s, +10s]
			In the presence of **audio media** [4]	[-30s, +20s]
		Scent type dependent [128]	Foul scent	[0s, +15s]
			Spicy, fruity, flowery scents	[-10s, 10s]
			Burning scents	[-5s, +10s]
	Haptic	Tightly coupled	Gaming Vest [215]	[0s, 1s]
	Airflow	Tightly coupled	Fan [215]	[-5s, +3s]
	Taste	Tightly coupled	Difference between taste-visual latencies [135]	[13ms - 837ms]
Haptic	Scent	Video media dependent [83]	Good	[-1380ms, -250ms]
			Allowable	[-1380ms, 400ms]

showed that the skews could be more easily sensed by users when the video is ahead of the audio, and the "in sync" region of the lip synchronization is {-80ms, 80ms}.

Recently, additional media streams (e.g., addressing touch) were added as a new dimension to the user experience [169]. Haptic technology has become accessible and started to be used in modern devices interfaces like teleoperators (e.g., haptic mouse and grip for video gaming, medical simulation or art design), touchable screens of smartphones and some PCs, and some contactless devices in Augmented/VR systems. Haptic applications require high demand of bandwidth to enable the data transmissions over a high reliable and responsive network [183]. In applications like teleoperation, the accuracy of timing between haptic and other media streams is vital for the QoE. Thus, it is significant to consider the network restrictions (i.e. delays, jitter and losses) in the mulsemedia distributions and to identify the thresholds of inter-stream synchronization. In [181], a high degree of sensitivity (100ms) in perceiving asynchrony between media channels (HA and HV) has been observed in a multimedia presentation with a haptic reference stream resembling a ping pong game. In [199], the authors investigated to what extent QoE is affected by network delay across the transmission of haptic media, audio and video through a haptic-transferred system.

In this study, a simulated network delay for each media stream has been introduced, while monitoring changes in the QoE. The results indicated that Mean Opinion Score (MOS) values are bound up with the differences of network delay jitters among the media streams (i.e. haptics, audio and videos)[199]. However, the authors did not identify the threshold of perception of asynchrony. In [219], this study was extended with a networked musical ensemble composed out of a tambourine and a keyboard harmonica with results showing that the inter-stream synchronization error rate in this situation was decreased against the works in [199]. Interestingly, this asynchrony can be masked in tri-stimuli systems by one of the present sensory modalities (e.g., audio has been shown to mask the synchronization skews in olfaction-enhanced presentations [4]).

Another important aspect of the distribution is supporting high quality services for mulsemedia in the context where there is an exponential growth in user population and in the resource requirements of the networked applications.

In [217], the authors propose an adaptive mulsemedia framework (*ADAMS*) that has been shown to improve user perception and QoE in variable network delivery conditions. Using a MPEG-7[9] description scheme compatible with MPEG codecs, *ADAMS* designs two adaptation schemes on the server side: a flow-based steam-switch adaptation scheme and a mulsemedia-based packet priority scheduling scheme which differentiate the mulsemedia (i.e. haptic, air-flow and olfaction) and multimedia (i.e. audiovisual) content distribution depending on the available network bandwidth and human preferences.

Network Function Virtualization (NFV) and Software Defined Networking (SDN) are technologies that provide tools to design networks with greater degree of abstraction, increasing the network flexibility [5]. SDN architecture decouples network control and data forwarding functions and enables direct programmability by upper layer applications. SDN paradigms can be used to dynamically allocate network resources to different DASH clients, while a SDN manager would be used to dynamically modify video flows and/or network resources in order to achieve a QoE fairness between users [50]. Thus SDN techniques will play a crucial role in the design of the Tactile Internet, which is a combination of multi-systems and multi-technologies which need to share the frequency spectrum as well as the physical infrastructure. In addition, existing multimedia distribution technologies have successful experiences on supporting the adaptive multimedia streaming, which offers smooth real-time audiovisual services and solves the issues of network restrictions, energy consumption and user perceived quality [123, 193, 221]. Considerable effort has been made in this research direction and a large plethora of adaptive multimedia delivery solutions have been proposed in the literature. We argue that these solutions could teach us important lessons that can be further used in mulsemedia delivery: mulsemedia services can also be tailored to users based on the different network conditions and multisensorial devices.

To this end, one promising endeavor is that afforded by the Tactile Internet in 5G, which envisages an ultra-responsive and ultra-reliable network connectivity that enables the delivery of physical haptic experiences remotely and allows the building of real-time interactive systems [52]. Using predictive analysis together with artificial intelligence and changing the air interface and architecture design at the wireless edge, the Tactile Internet thus has the challenge to achieve 1ms round trip latency and to provide an enhanced haptic perception. If the Tactile Internet delivers on its potential, then we surmise that mulsemedia applications of the future can harness its abilities to deliver seamless synchronization between the different media components.

3.3 Rendering - From Prototypes to Off-the-shelf Devices

After mulsemedia content has been transmitted over a network, different types of sensorial devices have to retrieve and translate the content to represent

the appropriate effects [35]. An extensive review of olfactory displays following different classifications (location, how the scent is delivered) is presented in [130]. The authors analyze different setups and discuss limitations of different technologies (both experimental and commercial) used in building olfactory displays. Haptic devices and their possible effects and applications are presented in [34]. A more recent effort in this direction is presented in [124], where the authors present the concept of altered touch through an integrated fingertip haptic display that uses integrated force, tactile and thermal feedback and the appropriate effects [35]. An extensive review of olfactory displays following different classifications (location, how the scent is delivered) is presented in [130]. The authors analyze different setups and discuss limitations of different technologies (both experimental and commercial) used in building olfactory displays. Haptic devices and their possible effects and applications are presented in [34]. A more recent effort in this

---

[9]MPEG-7: http://mpeg.chiariglione.org/standards/mpeg-7

Table 6. Rendering Devices

	Prototype	Theme	Interesting Unexplored Directions	Off-the-shelf Devices
Haptic Gloves	Exoskeleton glove with pin arrays [111]	Force and tactile feedback that allows people to touch, stretch and pull virtual fabrics	Distinction between different fabrics	*Dexmo* - exoskeleton glove that captures 11 DOF of users' hand motion and provides force feedback of variable stiffness [65]
	Mood glove [117]	Enrich moods in films/music through tactile sensations (eight vibration patterns)	Emotional experience through cross-modality	*Gloveone/Avatar VR* - tracks finger and arm motions and produces haptic sensations through vibrations
Haptic Chairs	Feel effects - Haptic chair (24 tactors on the back of the chair). Effect libraries: stroke, pulse, motor, rain, strike, travel [87].	Storytelling, story listening	Training for disabled (wheel chair users)	*Roto VR* - Haptic chair with rumble shakers fixed to the underside and back of the chair (effect libraries: crashes, flaps, turbulence - plane, race car, rollercoaster), VR amplifier (filters audio signals and creates tactile feedback), base and foot pedals (walk in VR).
	Haptic chair [139], Emoti-chair [13]	Enrich the music experience for the deaf by the touchable vibrations		
Other Haptic Devices	HALUX [204]	Interactive skin for digital sports based on projections	Investigate the level of immersion	*Rez Infinite* - whole body vibrotactile interaction
	Jorro Beat [82] - shower head that controls the water flow and provides a tactile stimulation synchronized with music	Enrich the music experience in the bathroom	Playing live music with an input interface	n/a
Olfactory Displays	Wearable	Product evaluation in visual-olfactory environment [18]; Low cost wearable display [75]; Miniaturised olfactory display [71]	Generate multiple odors simultaneously; Create scents in the vapor state	*Aroma Join* - messenger app, notification, odor card, DaleAir , Exhalia AromaStick - personal mood changer for office workers
	Desktop	Odor emanated by a specific region of the screen in an immersive environment [114]; Cooking game with odors [134]		
Wind Display	WindCube [122] MSF [113] Treadpod Active Wind Tunnel TPAWT [99] [98]	Create stable flow patterns for directional wind display; Wind direction perception [136]	Crossmodal effects with other sensory modalities; Precision of wind perception	DC fans; Feelreal mask - Two powerful microcoolers (cool and hot air)

direction is presented in [124], where the authors present the concept of "altered touch" through an integrated fingertip haptic display that uses integrated force, tactile and thermal feedback and can be used with augmented reality applications.

One of the aims of this paper is to offer guidance in the aspects that should be considered by researchers who explore the potential or the particularities of multisensory interaction. Thus, in Table 6 we focus on a list of prototype rendering devices presented in research papers by describing the context in which they were used, the feedback they provide, the research directions we identified as interesting to explore and potential off the shelf commercial devices that anyone can use in replicating or extending the initial studies.

In this section we have focused on different particularities related to production, distribution and rendering of multisensory systems. Most of the existent studies in mulsemedia delivery or in

creating new ways of interaction with technology are considering bi-sensory settings. However, as presented in previous subsections, the creating multisensorial systems started to be populated with tri or multiple-sensory options from creation to rendering. All these tools and devices allow researchers to further address the set of challenges presented in the next section.

4   FUTURE DIRECTIONS

Media production and consumption play an important role in our everyday life. Research on interactive experiences with mulsemedia is a growing field of study subdivided into topics related to: crossmodal processes, content production, system design and architecture, application and evaluation of user's experience and development of new interaction paradigms. While in the 20th century the interest in the media industry was on the visual, the present demands an investigation of touch, taste and smell as sensory interaction modalities. Starting from this, a set of questions arises: which mulsemedia experience to design for, how can this be achieved, or how to evaluate the QoE. Recent advances in experimental psychology (described partially in Section 2) and HCI [147] are two important sources of inspiration [145].

Previous work showed that researchers focused on rendering and production of haptic, sound, olfaction or taste, from the point of view that they represent a single-medium engaging a single-sense [61][130][35]. In this article, and indeed, in our work, we are interested in the **multi-multi** perspective where **multiple** types of media engage **multiple** senses.

4.1   Research Topics

Rendering continues to present challenges in creating a holistic multisensory percept in a digital context. An important characteristic of the rendering devices should be *adaptability* that can happen automatically or through the help of the designer or end-user. In order to ensure correct rendering, devices that are operable by non-experts need easy methods for troubleshooting. In Table 6 we present interesting research prototypes and their corresponding off-the-shelf devices that could be considered isolated or combined in building mulsemedia setups.

Design guidelines need to be further developed. Like in the case of the standardization of HTML/CSS, and Blu-Ray versus HD DVD, diverse file formats and infrastructures will emerge. Given the diversity of multisensory technologies and experiences, we expect these to be centered around paradigms that can propose ways to work with mulsemedia content and which can be applied to multiple devices in a class. Best practices in working with multisensory content have yet to be formulated. An interesting study about the haptic design process was presented in [172], where the authors identify current obstacles in the design of haptic user experiences and present the results of intensive interviews with six professional haptic designers. The authors make valuable recommendations for accelerating the development of HaXD that can be applied also in a mulsemedia context. Another paper that provides a tutorial and recommendations on the key steps to conduct olfactory-based evaluation of the QoE is presented in [127]. This represents an important step in proposing a standardized methodology to conduct subjective mulsemedia quality assessment. However, there is still a need for more guidelines since they will enable the support of mulsemedia researchers and HCI designers alike. It is important to develop a set of best practices to guide people in the design of meaningful mulsemedia experiences.

Non-Aristotelian senses are not yet intensively targeted by digital mulsemedia systems, where the focus is overwhelmingly, if not exclusively on the traditional five Aristotelian senses. Whilst interoceptive senses like proprioception have been studied in the non-digital world, their stimulation and integration through traditional and/or non-traditional media types is yet to be explored in the digital. However, whilst steps in this direction have been made in different virtual environments

for the treatment of phobias or for the investigation of the contribution of proprioception , this has been done by calling upon intermediate modalities, such as force feedback, which stimulate the interoceptive senses. However, such incipient work did not make any formal conclusions since follow-up research is necessary to validate their findings.

Taste was identified in several research papers as one of the least investigated senses in terms of sensory interaction modalities [61][148]. So far, things have not changed substantially. Although gustation has been extensively studied to understand the functionality of the taste system, there is still a lack of understanding related to how taste can be exploited in developing new ways of interaction with technology. An interesting work in this context is presented in [146]. After a thorough analysis that is centered on user experience, the authors of this work identified three main types of characteristics that define taste experiences: temporality, affective reactions and embodiment. These characteristics are presented as a framework for designing gustatory experiences. As such, one of the main contribution of the work in discussion is that it provides a rich understanding and a vocabulary of the gustatory experience and its influence on user experience as well as new tools to design such experiences. Both electrical and thermal stimulation of taste require more focused experimental studies considering aspects such as: material characteristics (i.e. gold vs. silver), position on the tongue, the waveform of the electrical signal. Moreover, it is very important to experiment different physiological and psychological aspects of the sensation of taste at the intersection with other sensory cues. As highlighted in [206], there is a growing interest on enhancing HCI and food interaction design by understanding the multisensory influences on flavor perception. Scenarios where visual, auditory, and haptic cues were used for modulating flavor are presented, while the authors also summarize important challenges related to the need for long-term follow-up investigations that could add value to flavor augmentation.

Crossmodal correspondences and multisensory substitution. Crossmodal correspondences still present a number of unanswered questions related to sensory cue integration and to how correspondences affect metaphorical understanding, feelings of 'knowing', behavior, learning and perceptual experiences [39]. Practical guidelines for the design of experiments that might shed new light in crossmodal correspondences are presented in [153]. Continuously investigating this direction could also benefit the development of more effective sensory substitution devices, since it has been shown that crossmodal correspondences enhance the performance of color to sound systems [67].

Attention modeling. There are often situations when an interface designer wants to capture the attention of the user. Attention is a cognitive concept that determines how particular sensory input, perceptual objects, thoughts, or courses of action are selected for further processing. Typically, researchers tended to focus on an individual sensory modality, ignoring the other senses [187], and noticed that each sense presents different capabilities: vision is precise for spatial information, sound is effective in the perception of temporal information, haptics can be used in processing both temporal and spatial [180]. However, it has been shown that multisensory audio/tactile cueing improves the performance (speed and accuracy) of visual search and reduces the amount of mental workload [68]. Thus, crossmodal cuing represents a feasible approach for enhancing task performance without causing an increase in the attention attributed to a singular sensory modality [81]. The magnitude of cueing effects was measured in a laboratory environment for HCI or for driving. A multisensory driver described in [78] provides interesting interface design guidelines supported by multisensory human perception research and has the potential to help the development of ergonomic interfaces. In designing a car interface, implications for VT warning signals of the crossmodal links between touch and vision in motion processing have also been presented in [76]. However, the same authors show that laboratory-based studies on spatial attention do not capture

all the key factors of real world situations and present a set of recommendations that could help researchers without compromising the value of their experiments [77].

Building databases, libraries of effects, examples and rapid prototyping tools The use of multisensory media is nontrivial, thus it needs to be supported through a variety of tools that help the design process. Most of the existing libraries (media interactions [173][33], *TouchSense*; for accessible wayfinding [218]) do not offer examples, or indeed the possibility to decompose these effects. Latest features allow parameters to be adjusted [87] and provide faceted library search and browsing [177], but most times designers need to choose predefined effects or to build them from scratch. An interesting approach for VT effects is the *Macaron* tool [175] that includes a web-based editor and a gallery for interactive design [102]. The authors' observations present implications for other editing tools.

4.2 Application Areas

Mulsemedia and performing arts. Interactive art started to be increasingly present not only in art galleries, but also in events like ISEA and Prix Ars Electronica or in demo tracks at academic conferences like CHI or SIGGRAPH. Multisensory interactive art installations perform an open-end exploration, questioning the role of multisensory interfaces for communication. Art offers an alternative approach and allows researchers to gain insights into the interactive process, facilitating the collaboration between artists and technologists. In [137], the authors identify a theoretical framework of interactive installations and highlight their relevance to HCI through relevant examples. Art installations could lead to a better understanding of under-exploited senses like taste and smell. In [103] authors propose an installation exhibited at *SIGGRAPH ASIA Art Gallery* that allows the user draw the lollipop on a virtual screen and print it with fragrance inks. Tate *Sensorium*[10] was a winning project of *IK Prize 2015 Award*. It featured an immersive display with four paintings from the Tate Collection that allowed viewers to experience sounds, smells, tastes, and physical forms inspired by these artworks while having their psychological responses recorded.

An ongoing project funded by the European Commission's Horizon 2020 Framework, *Accessible Resources for Cultural Heritage EcoSystems (ARCHES)*[11] aims to exploit multisensorial devices and media among other technologies in order to create more inclusive environments at museums and cultural heritage sites for people having difficulties associated with perception, memory, cognition and communication. We argue that mulsemedia systems can have a high impact on helping people with different disabilities not only in art, but also in other aspects of human life, such as education. Therefore, we consider this to be a challenging but worthwhile exploring research direction in the context of mulsemedia systems.

We are at the point where technology offers affordable tools that allow us to create new experiences and gather knowledge. It is important to understand how these tools should be used and in this process, keeping this experimental mindset often met in creating art, would be key to give insights and clarify aspects that can be used also in other fields.

Wearable mulsemedia. The evolution of multisensory technology embedded into wearables represents a new modality for people to experience life enhancing all the human senses. Haptics cannot be anymore dismissed as boring vibrations that alert you to messages. Now, prototype touchscreens can use ultrasonics to create the 'feel' of different objects while haptic motors are becoming smaller and more flexible. People can send hugs over distance with the "Hug Shirt" [12],

---

[10]Sensorium: http://www.tate.org.uk/whats-on/tate-britain/display/ik-prize-2015-tate-sensorium
[11]ARCHES: http://arches-project.eu/
[12]The Hug Shirt: http://cutecircuit.com/the-hug-shirt/

feel vibrational cues that help them keeping up the rhythm with "Soundbrenner"[13] or experience what a player feel during games through Alert Shirt[14]. With the upcoming release of "watchOS 3", Apple is pushing things further. Breathe[15], its new application designed to coach people through stress-relieving breathing patterns, uses the Taptic engine to deliver subtle physical prompts on when to inhale and exhale. Moreover, these advancements in affordable wearable technology have the potential to also help people with special needs. Lechal[16] is a haptic footwear that help the visually impaired navigate. Unlike smartphones and tables which rely on people looking at them, the lack of display space and digital interfaces make multisensory feedback an even more powerful tool. Olfactory displays are still in a prototype status [71, 118], but it would be interesting to see how the market evolves in this direction. Interesting horizons can be envisaged through devices like Aromastic, a mobile scent dispenser that aims to act as a personal mood changer for the busy professionals. These type of initiatives seem promising also in respect of multisensory therapy (e.g. Snoezelen), which has been successfully applied to alleviate the symptoms associated with a wide range of pathologies [22].

Mulsemedia and e-Learning. With the rapid growth and development of the information and communication technologies, e-learning has seen a fast evolution over the past decade. Mobile and multimedia technologies have also changed radically the online learning landscape [59, 60]. The advances in wearables and mobile technologies such as improved network speeds, improved processing power, improved graphics and higher-resolution displays enable enhanced, more complex mobile learning experiences. Olfaction and haptic interfaces have started to be evaluated in the context of education [108, 164], predominantly in science and mathematics but also in social sciences [121]. However, since the majority of educational virtual environments are focused on visual representation [157], there is still a need to investigate the capabilities of a multisensory system and its effects on learning outcomes and knowledge retention. In [179], the authors discussed the benefits of multisensory learning over a unisensory paradigm. However, the authors emphasize the fact that the magnitude of this benefit is dependent on the congruency of the stimuli e.g., the learning environment must closely map what users have experienced in nature, supporting the findings in [72]. Crossmodal correspondences and crossmodal substitutions can also play an important role in the learning process (also for people with disabilities) that has to be analyzed. Over the last few years, the European Union Framework programs and Horizon 2020 have supported a number of projects in the multimodal space. NEWTON[17] (Networked labs for training sciences and technologies) is of relevance for the present study since it focuses on multisensorial aspects in learning STEM subjects [27].

Mulsemedia in e-Commerce and advertising. With the current upsurge in technology and wearables on one hand and the increasing preference of customers towards on-line shopping on the other hand, mulsemedia has the potential to influence and change consumer behavior. Several studies have shown an increase of satisfaction and a better emotional state when using environmental cues. Ambient cues (colors and scents) have an impact on user's behavior with cool visual and olfactory cues (blue or citrus-mint) rated higher than warm cues. Also, the influence of touch on on-line shopping attitudes was demonstrated in [105] for a Korean population. Using devices like the smelling screen [114] or the FeelReal Helmet, digital marketers can improve the

---

[13]Sounbrenner: http://www.soundbrenner.com/
[14]Alert Shirt: http://wearableexperiments.com/alert-shirt/
[15]Breathe: https://support.apple.com/en-gb/HT206999
[16]Lechal: http://www.lechal.com/
[17]EU H2020 NEWTON: http://www.newtonproject.eu

store environment with both design and ambient factors [91] by proposing a new interface that can be applied to a variety of visual contents such as ads and website.

Sensory marketing engages customers' senses and can influence advertisement design and effectiveness because of its influence on perception, judgment and behavior. An extensive review discusses the role of sensory marketing in driving advertisement effectiveness, showing that product evaluation is enhanced by touch and smell, attention and persuasiveness are affected by music, whilst memory is enhanced by smell and taste perception is improved under multisensory stimulation [96]. Well-implemented sensory experiences have positive effects on brand equity, studies showing that participants agreed that sensory stimulators are congruent with the brand identity [11].

5 CONCLUSION

Whilst multisensory interaction and displays have been studied in quite some depth in the non-digital world (e.g.in psychology, education, therapy, performing arts, etc.), they are relative newcomers to the digital world. Moreover, whilst multimodal systems have a long track record of being studied in the digital world, digital mulsemedia systems are again relative newcomers on the research scene, and, in the current paper, we have reviewed research efforts targeting this extra dimension of the multimedia experience.

It is already obvious that mulsemedia systems have the potential to make more efficient use of the human perceptual and cognitive capabilities because the human brain has evolved in a multisensory environment. In the current review we tried to offer a full picture of the spectrum of multisensory systems starting with the neurological processes involved (from crossmodal correspondences to multisensory integration), continuing with the workflow for production-distribution-rendering and ending with challenges that still need to be addressed. We made sure to emphasize important aspects related to multi-stream synchronization, adaptation, fast prototyping, off the shelf devices and we identified areas that could have a major benefit from this multisensorial approach. All are worthy of future endeavors.

Mulsemedia systems are multi-multi systems, where multiple (more than three) senses are engaged by multiple media. In contrast to multimodal systems, mulsemedia systems are all about the media and not the modality; the focus here is on the diverse new, *non-traditional, media types*, their production, distribution, and rendering, and the issues raised when integrating non-traditional and traditional digital media in mulsemedia systems. As our paper has shown, progress is being made in our understanding of mulsemedia systems, yet there are challenges and opportunities galore to be explored and overcome in the quest to transcend the overwhelmingly bisensorial nature of digital multimedia into the multisensorial one of mulsemedia.

REFERENCES

[1] Justin W Adams, Lisa Paxton, Kathryn Dawes, Kateryna Burlak, Michelle Quayle, and Paul G McMenamin. 2015. 3D printed reproductions of orbital dissections: a novel mode of visualising anatomy for trainees in ophthalmology or optometry. *British Journal of Ophthalmology* 99, 9 (2015), 1162–1167. DOI:http://dx.doi.org/10.1136/bjophthalmol-2014-306189 arXiv:http://bjo.bmj.com/content/99/9/1162.full.pdf

[2] O. A. Ademoye and G. Ghinea. 2009. Synchronization of Olfaction-Enhanced Multimedia. *IEEE Transactions on Multimedia* 11, 3 (April 2009), 561–565. DOI:http://dx.doi.org/10.1109/TMM.2009.2012927

[3] Oluwakemi A Ademoye and Gheorghita Ghinea. 2013. Information recall task impact in olfaction-enhanced multimedia. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* 9, 3 (2013), 17.

[4] 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 (Aug. 2016), 14 pages. DOI:http://dx.doi.org/10.1145/2957753

[5]   A. Aijaz, M. Dohler, A. H. Aghvami, V. Friderikos, and M. Frodigh. 2017. Realizing the Tactile Internet: Haptic
      Communications over Next Generation 5G Cellular Networks. *IEEE Wireless Communications* PP, 99 (October 2017),
      2–9. DOI:http://dx.doi.org/10.1109/MWC.2016.1500157RP

[6]   David Alais and John Cass. 2010. Multisensory perceptual learning of temporal order: audiovisual learning transfers
      to vision but not audition. *PloS one* 5, 6 (2010), e11283.

[7]   F. Arafsha, L. Zhang, H. Dong, and A. E. Saddik. 2015. Contactless haptic feedback: state of the art. In *2015 IEEE
      International Symposium on Haptic, Audio and Visual Environments and Games (HAVE)*. 1–6. DOI:http://dx.doi.org/10.
      1109/HAVE.2015.7359447

[8]   Alan Baddeley. 1992. Working memory. *Science* 255, 5044 (1992), 556–559.

[9]   Philip J. Barnard and John D. Teasdale. 1991. Interacting cognitive subsystems: A systemic approach to cognitive-
      affective interaction and change.   *Cognition and Emotion* 5, 1 (1991), 1–39.   DOI:http://dx.doi.org/10.1080/
      02699939108411021 arXiv:http://dx.doi.org/10.1080/02699939108411021

[10] Kira Belkin, Robyn Martin, Sarah E Kemp, and Avery N Gilbert. 1997. Auditory pitch as a perceptual analogue to odor
      quality. *Psychological Science* 8, 4 (1997), 340–342.

[11] Bethan Alexander and Karinna Nobbs. 2016. Multi-Sensory Fashion Retail Experiences: The Impact of Sound,
      Smell, Sight and Touch on Consumer Based Brand Equity. In *Handbook of Research on Global Fashion Management
      and Merchandising*, Alessandra Vecchi and Chitra Buckley (Eds.). IGI Global, Hershey, PA, USA, 420–443. http:
      //services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-5225-0110-7.ch017 DOI: 10.4018/978-1-5225-0110-
      7.ch017.

[12] Nadia Bolognini, Francesca Frassinetti, Andrea Serino, and Elisabetta Làdavas. 2005. "Acoustical visio" of below
      threshold stimuli: interaction among spatially converging audiovisual inputs. *Experimental brain research* 160, 3 (2005),
      273–282.

[13] C. Branje, M. Karam, D. Fels, and F. Russo. 2009. Enhancing entertainment through a multimodal chair interface.
      In *2009 IEEE Toronto International Conference Science and Technology for Humanity (TIC-STH)*. 636–641. DOI:http:
      //dx.doi.org/10.1109/TIC-STH.2009.5444421

[14] K Bronner, K Frieler, H Bruhn, R Hirt, and D Piper. 2012. What is the sound of citrus? Research on the correspondences
      between the perception of sound and taste/flavour. *Manuscript submitted for publication* (2012).

[15] Merijn Bruijnes, Gijs Huisman, and Dirk Heylen. 2016. Tasty Tech: human-food interaction and multimodal interfaces.
      In *Proceedings of the 1st Workshop on Multi-sensorial Approaches to Human-Food Interaction*. ACM, 4.

[16] Leandro Cancar, Alex Díaz, Antonio Barrientos, David Travieso, and David M Jacobs. 2013. Tactile-sight: A sensory
      substitution device based on distance-related vibrotactile flow. *International Journal of Advanced Robotic Systems* 10, 6
      (2013), 272.

[17] S. K. Card. 1981. The Model Human Processor: A Model for Making Engineering Calculations of Human Perfor-
      mance. *Proceedings of the Human Factors Society Annual Meeting* 25, 1 (1981), 301–305. DOI:http://dx.doi.org/10.1177/
      107118138102500180 arXiv:http://dx.doi.org/10.1177/107118138102500180

[18] M. Carulli, M. Bordegoni, and U. Cugini. 2015. Visual-olfactory immersive environment for product evaluation. In
      2015 IEEE Virtual Reality (VR). 161–162. DOI:http://dx.doi.org/10.1109/VR.2015.7223345

[19] Umberto Castiello, Gesualdo M Zucco, Valentina Parma, Caterina Ansuini, and Roberto Tirindelli. 2006. Cross-modal
      interactions between olfaction and vision when grasping. *Chemical senses* 31, 7 (2006), 665–671.

[20] J. Cha, Y. Seo, Y. Kim, and J. Ryu. 2007. An Authoring/Editing Framework for Haptic Broadcasting: Passive Haptic
      Interactions using MPEG-4 BIFS. In *Second Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual
      Environment and Teleoperator Systems (WHC'07)*. 274–279. DOI:http://dx.doi.org/10.1109/WHC.2007.20

[21] Alan Chalmers, David Howard, and Christopher Moir. 2009. Real Virtuality: A Step Change from Virtual Reality. In
      Proceedings of the 25th Spring Conference on Computer Graphics (SCCG '09). ACM, New York, NY, USA, 9–16. DOI:
      http://dx.doi.org/10.1145/1980462.1980466

[22] Sally Wai-chi Chan, David R Thompson, Janita PC Chau, Wilson WS Tam, Ivy WS Chiu, and Susanne HS Lo. 2010. The
      effects of multisensory therapy on behaviour of adult clients with developmental disabilitiesfi!?A systematic review.
      International journal of nursing studies 47, 1 (2010), 108–122.

[23] T. Chatzidimitris, D. Gavalas, and D. Michael. 2016. SoundPacman: Audio augmented reality in location-based games.
      In *2016 18th Mediterranean Electrotechnical Conference (MELECON)*. 1–6. DOI:http://dx.doi.org/10.1109/MELCON.2016.
      7495414

[24] B. Choi, E. S. Lee, and K. Yoon. 2011. Streaming Media with Sensory Effect. In *2011 International Conference on
      Information Science and Applications*. 1–6. DOI:http://dx.doi.org/10.1109/ICISA.2011.5772390

[25] DL Clements, S Sato, and A Portela Fonseca. 2016. Cosmic sculpture: a new way to visualise the cosmic microwave
      background. *European Journal of Physics* 38, 1 (2016), 015601.

[26] Ed Cooke and Erik Myin. 2011. Is trilled smell possible? How the structure of olfaction determines the phenomenology
      of smell. *Journal of Consciousness Studies* 18, 11-12 (2011), 59–95.

[27] Alexandra Covaci, Gheorghita Ghinea, Chang-Hsin Lin, Shu-Hsien Huang, and Ju-Ling Shih. 2018. Multisensory games-based learning-lessons learnt from olfactory enhancement of a digital board game. *Multimedia Tools and Applications* (2018), 1–19.

[28] Anne-Sylvie Crisinel and others. 2010. As bitter as a trombone: Synesthetic correspondences in nonsynesthetes between tastes/flavors and musical notes. *Attention, Perception, & Psychophysics* 72, 7 (2010), 1994–2002.

[29] Anne-Sylvie Crisinel and Charles Spence. 2009. Implicit association between basic tastes and pitch. *Neuroscience letters* 464, 1 (2009), 39–42.

[30] Anne-Sylvie Crisinel and Charles Spence. 2010. A sweet sound? Food names reveal implicit associations between taste and pitch. *Perception* 39, 3 (2010), 417–425.

[31] Anne-Sylvie Crisinel and Charles Spence. 2012. A Fruity Note: Crossmodal associations between odors and musical notes. *Chemical Senses* 37, 2 (2012), 151–158. DOI:http://dx.doi.org/10.1093/chemse/bjr085

[32] Alberto Cruz and Barry G Green. 2000. Thermal stimulation of taste. *Nature* 403, 6772 (2000), 889.

[33] H. Culbertson, J. Unwin, and K. J. Kuchenbecker. 2014. Modeling and Rendering Realistic Textures from Unconstrained Tool-Surface Interactions. *IEEE Transactions on Haptics* 7, 3 (July 2014), 381–393. DOI:http://dx.doi.org/10.1109/TOH.2014.2316797

[34] Fabien Danieau, Jérémie Bernon, Julien Fleureau, Philippe Guillotel, Nicolas Mollet, Marc Christie, and Anatole Lécuyer. 2013. H-Studio: An Authoring Tool for Adding Haptic and Motion Effects to Audiovisual Content. In *Proceedings of the Adjunct Publication of the 26th Annual ACM Symposium on User Interface Software and Technology (UIST '13 Adjunct)*. ACM, New York, NY, USA, 83–84. DOI:http://dx.doi.org/10.1145/2508468.2514721

[35] F. Danieau, A. Lecuyer, P. Guillotel, J. Fleureau, N. Mollet, and M. Christie. 2013. Enhancing Audiovisual Experience with Haptic Feedback: A Survey on HAV. *IEEE Transactions on Haptics* 6, 2 (April 2013), 193–205. DOI:http://dx.doi.org/10.1109/TOH.2012.70

[36] J. Delwiche. 1996. Are there fibasicfi tastes? *Trends in Food Science & Technology* 7, 12 (1996), 411 – 415. DOI:http://dx.doi.org/10.1016/S0924-2244(96)20010-X

[37] M Luisa Dematte, Daniel Sanabria, Rachel Sugarman, and Charles Spence. 2006. Cross-modal interactions between olfaction and touch. *Chemical Senses* 31, 4 (2006), 291–300.

[38] Ophelia Deroy, Irène Fasiello, Vincent Hayward, and Malika Auvray. 2016. Differentiated audio-tactile correspondences in sighted and blind individuals. *Journal of experimental psychology: human perception and performance* 42, 8 (2016), 1204.

[39] Ophelia Deroy and Charles Spence. 2016. Crossmodal correspondences: four challenges. *Multisensory Research* 29, 1-3 (2016), 29–48.

[40] Ophelia Deroy and Dominique Valentin. 2011. Tasting liquid shapes: investigating the sensory basis of cross-modal correspondences. *Chemosensory Perception* 4, 3 (2011), 80.

[41] H. Q. Dinh, N. Walker, L. F. Hodges, Chang Song, and A. Kobayashi. 1999. Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. In *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)*. 222–228. DOI:http://dx.doi.org/10.1109/VR.1999.756955

[42] Norman F Dixon and Lydia Spitz. 1980. The Detection of Auditory Visual Desynchrony. *Perception* 9, 6 (1980), 719–721. DOI:http://dx.doi.org/10.1068/p090719 arXiv:http://dx.doi.org/10.1068/p090719 PMID: 7220244.

[43] Dmitrijs Dmitrenko, Emanuela Maggioni, Chi Thanh Vi, and Marianna Obrist. 2017. What did I sniff? Mapping scents onto driving-related messages. In *AutomotiveUI'17 Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM, 154–163.

[44] H. Dong, Y. Gao, H. A. Osman, and A. E. Saddik. 2015. Development of a Web-Based Haptic Authoring Tool for Multimedia Applications. In *2015 IEEE International Symposium on Multimedia (ISM)*. 13–20. DOI:http://dx.doi.org/10.1109/ISM.2015.71

[45] Adrienne E Dubin and Ardem Patapoutian. 2010. Nociceptors: the sensors of the pain pathway. *The Journal of clinical investigation* 120, 11 (2010), 3760–3772.

[46] Bruce Durie. 2017. Senses Special: Doors of Perception. https://www.newscientist.com/article/mg18524841-600-senses-special-doors-of-perception/. (2017). Accessed: 2017-03-19.

[47] Marcus Eldridge, Elliot Saltzman, and Amir Lahav. 2010. Seeing what you hear: Visual feedback improves pitch recognition. *European Journal of Cognitive Psychology* 22, 7 (2010), 1078–1091. DOI:http://dx.doi.org/10.1080/09541440903316136 arXiv:http://dx.doi.org/10.1080/09541440903316136

[48] M. J. Enriquez and K. E. MacLean. 2003. The hapticon editor: a tool in support of haptic communication research. In *11th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2003. HAPTICS 2003. Proceedings*. 356–362. DOI:http://dx.doi.org/10.1109/HAPTIC.2003.1191310

[49] Lizbeth Escobedo, Monica Tentori, Eduardo Quintana, Jesus Favela, and Daniel Garcia-Rosas. 2014. Using augmented reality to help children with autism stay focused. *IEEE Pervasive Computing* 13, 1 (2014), 38–46.

[50] C. f. Lai, R. h. Hwang, H. c. Chao, M. M. Hassan, and A. Alamri. 2015. A buffer-aware HTTP live streaming approach for SDN-enabled 5G wireless networks. *IEEE Network* 29, 1 (Jan 2015), 49–55. DOI:http://dx.doi.org/10.1109/MNET.2015.7018203

[51] M Falasconi, I Concina, E Gobbi, V Sberveglieri, A Pulvirenti, and G Sberveglieri. 2012. Electronic nose for microbiological quality control of food products. *International Journal of Electrochemistry* 2012 (2012).

[52] G. P. Fettweis. 2014. The Tactile Internet: Applications and Challenges. *IEEE Vehicular Technology Magazine* 9, 1 (March 2014), 64–70. DOI:http://dx.doi.org/10.1109/MVT.2013.2295069

[53] Chris Forsythe, Huafei Liao, Michael Christopher Stefan Trumbo, and Rogelio E Cardona-Rivera. 2014. *Cognitive Neuroscience of Human Systems: Work and Everyday Life.* CRC Press.

[54] Simon Gane, Dimitris Georganakis, Klio Maniati, Manolis Vamvakias, Nikitas Ragoussis, Efthimios M. C. Skoulakis, and Luca Turin. 2013. Molecular Vibration-Sensing Component in Human Olfaction. *PLOS ONE* 8, 1 (01 2013), 1–7. DOI:http://dx.doi.org/10.1371/journal.pone.0055780

[55] Gheorghita Ghinea and Oluwakemi Ademoye. 2012. User Perception of Media Content Association in Olfaction-enhanced Multimedia. *ACM Trans. Multimedia Comput. Commun. Appl.* 8, 4, Article 52 (Nov. 2012), 19 pages. DOI:http://dx.doi.org/10.1145/2379790.2379794

[56] Gheorghita Ghinea and Oluwakemi Ademoye. 2015. Olfactory media impact on task performance: The case of a word search game. In *Interactive Mobile Communication Technologies and Learning (IMCL), 2015 International Conference on.* IEEE, 296–300.

[57] G. Ghinea and O. A. Ademoye. 2010. Perceived Synchronization of Olfactory Multimedia. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 40, 4 (July 2010), 657–663. DOI:http://dx.doi.org/10.1109/TSMCA.2010.2041224

[58] Gheorghita Ghinea and Oluwakemi A Ademoye. 2011. Olfaction-enhanced multimedia: perspectives and challenges. *Multimedia Tools and Applications* 55, 3 (2011), 601–626.

[59] Gheorghita Ghinea and Sherry Y Chen. 2006. Perceived quality of multimedia educational content: A cognitive style approach. *Multimedia systems* 11, 3 (2006), 271–279.

[60] Gheorghita Ghinea and Sherry Y Chen. 2008. Measuring quality of perception in distributed multimedia: Verbalizers vs. imagers. *Computers in Human Behavior* 24, 4 (2008), 1317–1329.

[61] Gheorghita Ghinea, Christian Timmerer, Weisi Lin, and Stephen R. Gulliver. 2014. Mulsemedia: State of the Art, Perspectives, and Challenges. *ACM Trans. Multimedia Comput. Commun. Appl.* 11, 1s, Article 17 (Oct. 2014), 23 pages. DOI:http://dx.doi.org/10.1145/2617994

[62] Guy Gingras, Benjamin A Rowland, and Barry E Stein. 2009. The differing impact of multisensory and unisensory integration on behavior. *Journal of Neuroscience* 29, 15 (2009), 4897–4902.

[63] Mark Grimshaw. 2014. *The Oxford handbook of virtuality.* Oxford University Press.

[64] K. Gsöllpointner, R. Schnell, and R.K. Schuler. 2016. *Digital Synesthesia: A Model for the Aesthetics of Digital Art.* Gruyter, Walter de GmbH.

[65] Xiaochi Gu, Yifei Zhang, Weize Sun, Yuanzhe Bian, Dao Zhou, and Per Ola Kristensson. 2016. Dexmo: An Inexpensive and Lightweight Mechanical Exoskeleton for Motion Capture and Force Feedback in VR. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16).* ACM, New York, NY, USA, 1991–1995. DOI:http://dx.doi.org/10.1145/2858036.2858487

[66] Rachel Guetta and Psyche Loui. 2017. When music is salty: The crossmodal associations between sound and taste. *PLOS ONE* 12, 3 (03 2017), 1–14. DOI:http://dx.doi.org/10.1371/journal.pone.0173366

[67] Giles Hamilton-Fletcher, Thomas D Wright, and Jamie Ward. 2016. Cross-modal correspondences enhance performance on a colour-to-sound sensory substitution device. *Multisensory Research* 29, 4-5 (2016), 337–363.

[68] Peter A Hancock, Joseph E Mercado, James Merlo, and Jan BF Van Erp. 2013. Improving target detection in visual search through the augmenting multi-sensory cues. *Ergonomics* 56, 5 (2013), 729–738.

[69] Grant Hanson-Vaux, Anne-Sylvie Crisinel, and Charles Spence. 2013. Smelling shapes: Crossmodal correspondences between odors and shapes. *Chemical Senses* 38, 2 (2013), 161–166.

[70] Kevin Hapeshi and Dylan Jones. 1992. Interactive multimedia for instruction: A cognitive analysis of the role of audition and vision. *International Journal of Human-Computer Interaction* 4, 1 (1992), 79–99.

[71] Kazuki Hashimoto and Takamichi Nakamoto. 2016. Tiny olfactory display using surface acoustic wave device and micropumps for wearable applications. *IEEE Sensors Journal* 16, 12 (2016), 4974–4980.

[72] Michael Haverkamp. 2013. *Synesthetic design: Handbook for a multi-sensory approach.* Walter de Gruyter.

[73] C.H. Hawkes. 2004. Olfaction Taste And Cognition. *Brain* 127, 1 (2004), 231–231.

[74] Vincent Hayward, Oliver R. Astley, Manuel CruzfifftHernandez, Danny Grant, and Gabriel RoblesfifftDefifftLafifft-Torre. 2004. Haptic interfaces and devices. *Sensor Review* 24, 1 (2004), 16–29. DOI:http://dx.doi.org/10.1108/02602280410515770

[75] Nicolas S. Herrera and Ryan P. McMahan. 2014. Development of a Simple and Low-Cost Olfactory Display for Immersive Media Experiences. In *Proceedings of the 2Nd ACM International Workshop on Immersive Media Experiences (ImmersiveMe '14)*. ACM, New York, NY, USA, 1–6. DOI : http://dx.doi.org/10.1145/2660579.2660584

[76] C. Ho, R. Gray, and C. Spence. 2014. Reorienting Driver Attention with Dynamic Tactile Cues. *IEEE Transactions on Haptics* 7, 1 (Jan 2014), 86–94. DOI : http://dx.doi.org/10.1109/TOH.2013.62

[77] C. Ho, R. Gray, and C. Spence. 2014. To What Extent do the Findings of Laboratory-Based Spatial Attention Research Apply to the Real-World Setting of Driving? *IEEE Transactions on Human-Machine Systems* 44, 4 (Aug 2014), 524–530. DOI : http://dx.doi.org/10.1109/THMS.2014.2316502

[78] Cristy Ho and Charles Spence. 2012. *The multisensory driver: Implications for ergonomic car interface design*. Ashgate Publishing, Ltd.

[79] Kyungpyo Hong, Jaebong Lee, and Seungmoon Choi. 2013. Demonstration-based Vibrotactile Pattern Authoring. In *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction (TEI '13)*. ACM, New York, NY, USA, 219–222. DOI : http://dx.doi.org/10.1145/2460625.2460660

[80] Kyungpyo Hong, Jaebong Lee, and Seungmoon Choi. 2013. Demonstration-based Vibrotactile Pattern Authoring. In *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction (TEI '13)*. ACM, New York, NY, USA, 219–222. DOI : http://dx.doi.org/10.1145/2460625.2460660

[81] Kevin Hopkins, Steven J Kass, Lisa Durrance Blalock, and J Christopher Brill. 2016. Effectiveness of auditory and tactile crossmodal cues in a dual-task visual and auditory scenario. *Ergonomics* (2016), 1–9.

[82] Keisuke Hoshino, Masahiro Koge, Taku Hachisu, Ryo Kodama, and Hiroyuki Kajimoto. 2015. Jorro Beat: Shower Tactile Stimulation Device in the Bathroom. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '15)*. ACM, New York, NY, USA, 1675–1680. DOI : http://dx.doi.org/10.1145/2702613.2732841

[83] Sosuke Hoshino, Yutaka Ishibashi, Norishige Fukushima, and Shinji Sugawara. 2011. QoE assessment in olfactory and haptic media transmission: Influence of inter-stream synchronization error. In *Communications Quality and Reliability (CQR), 2011 IEEE International Workshop Technical Committee on*. IEEE, 1–6.

[84] Lucica Iordanescu, Emmanuel Guzman-Martinez, Marcia Grabowecky, and Satoru Suzuki. 2008. Characteristic sounds facilitate visual search. *Psychonomic Bulletin & Review* 15, 3 (2008), 548–554.

[85] Matthieu Ischer, Naëm Baron, Christophe Mermoud, Isabelle Cayeux, Christelle Porcherot, David Sander, and Sylvain Delplanque. 2014. How incorporation of scents could enhance immersive virtual experiences. *Applied Olfactory Cognition* 119, 3 (2014), 4.

[86] Ali Israr, Siyan Zhao, and Oliver Schneider. 2015. Exploring Embedded Haptics for Social Networking and Interactions. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '15)*. ACM, New York, NY, USA, 1899–1904. DOI : http://dx.doi.org/10.1145/2702613.2732814

[87] Ali Israr, Siyan Zhao, Kaitlyn Schwalje, Roberta Klatzky, and Jill Lehman. 2014. Feel Effects: Enriching Storytelling with Haptic Feedback. *ACM Trans. Appl. Percept.* 11, 3, Article 11 (Sept. 2014), 17 pages. DOI : http://dx.doi.org/10.1145/2641570

[88] Dhruv Jain, Misha Sra, Jingru Guo, Rodrigo Marques, Raymond Wu, Justin Chiu, and Chris Schmandt. 2016. Immersive Scuba Diving Simulator Using Virtual Reality. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. ACM, 729–739.

[89] Maria Karam, Carmen Branje, Gabe Nespoli, Norma Thompson, Frank A. Russo, and Deborah I. Fels. 2010. The Emoti-chair: An Interactive Tactile Music Exhibit. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems (CHI EA '10)*. ACM, New York, NY, USA, 3069–3074. DOI : http://dx.doi.org/10.1145/1753846.1753919

[90] Myongchan Kim, Sungkil Lee, and Seungmoon Choi. 2012. *Saliency-Driven Tactile Effect Authoring for Real-Time Visuotactile Feedback*. Springer Berlin Heidelberg, Berlin, Heidelberg, 258–269. DOI : http://dx.doi.org/10.1007/978-3-642-31401-8_24

[91] So-Jeong Kim and Dong-Hee Shin. 2016. The Multisensory Effects of Atmospheric Cues on Online Shopping Satisfaction. In *International Conference on HCI in Business, Government and Organizations*. Springer, 406–416.

[92] Sang-Kyun Kim. 2013. Authoring multisensorial content. *Signal Processing: Image Communication* 28, 2 (2013), 162–167.

[93] Y. Kim, J. Cha, J. Ryu, and I. Oakley. 2010. A tactile glove design and authoring system for immersive multimedia. *IEEE MultiMedia* 17, 3 (July 2010), 34–45. DOI : http://dx.doi.org/10.1109/MMUL.2010.5692181

[94] Kalin Kolev, Petri Tanskanen, Pablo Speciale, and Marc Pollefeys. 2014. Turning Mobile Phones into 3D Scanners. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

[95] P. T. Kovcs, N. Murray, G. Rozinaj, Y. Sulema, and R. Rybrov. 2015. Application of immersive technologies for education: State of the art. In *2015 International Conference on Interactive Mobile Communication Technologies and Learning (IMCL)*. 283–288. DOI : http://dx.doi.org/10.1109/IMCTL.2015.7359604

[96]   Aradhna Krishna, Luca Cian, and Tatiana Sokolova. 2016. The power of sensory marketing in advertising. *Current Opinion in Psychology* 10 (2016), 142–147.

[97]   Shenbing Kuang and Tao Zhang. 2014. Smelling directions: olfaction modulates ambiguous visual motion perception. *Scientific reports* 4 (2014), 5796.

[98]   S. D. Kulkarni, C. J. Fisher, P. Lefler, A. Desai, S. Chakravarthy, E. R. Pardyjak, M. A. Minor, and J. M. Hollerbach. 2015. A Full Body Steerable Wind Display for a Locomotion Interface. *IEEE Transactions on Visualization and Computer Graphics* 21, 10 (Oct 2015), 1146–1159. DOI:http://dx.doi.org/10.1109/TVCG.2015.2424862

[99]   S. D. Kulkarni, M. A. Minor, M. W. Deaver, and E. R. Pardyjak. 2007. Output Feedback Control of Wind Display in a Virtual Environment. In *Proceedings 2007 IEEE International Conference on Robotics and Automation*. 832–839. DOI:http://dx.doi.org/10.1109/ROBOT.2007.363089

[100]   Heilig Morton L. 1962. Sensorama Simulator. (Aug. 28 1962). US Patent 3,050,870.

[101]   Mei-Kei Lai. 2015. Universal Scent Blackbox: Engaging Visitors Communication Through Creating Olfactory Experience at Art Museum. In *Proceedings of the 33rd Annual International Conference on the Design of Communication (SIGDOC '15)*. ACM, New York, NY, USA, Article 27, 6 pages. DOI:http://dx.doi.org/10.1145/2775441.2775483

[102]   Brian Lee, Savil Srivastava, Ranjitha Kumar, Ronen Brafman, and Scott R. Klemmer. 2010. Designing with Interactive Example Galleries. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 2257–2266. DOI:http://dx.doi.org/10.1145/1753326.1753667

[103]   Jaejoong Lee and Jin Wan Park. 2013. A Media Art Study Using Multi-Sensory Elements. In *International Conference on Human-Computer Interaction*. Springer, 277–281.

[104]   Jaebong Lee, J. Ryu, and S. Choi. 2009. Vibrotactile score: A score metaphor for designing vibrotactile patterns. In *World Haptics 2009 - Third Joint EuroHaptics conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*. 302–307. DOI:http://dx.doi.org/10.1109/WHC.2009.4810816

[105]   Seung-Hee Lee, Jane E Workman, and Kwangho Jung. 2016. The influence of need for touch and gender on Internet shopping attitudes among Korean consumers. *International Journal of Fashion Design, Technology and Education* (2016), 1–10.

[106]   P. Lemmens, F. Crompvoets, D. Brokken, J. van den Eerenbeemd, and G. J. de Vries. 2009. A body-conforming tactile jacket to enrich movie viewing. In *World Haptics 2009 - Third Joint EuroHaptics conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*. 7–12. DOI:http://dx.doi.org/10.1109/WHC.2009.4810832

[107]   Nicoline Leunis, Marie-Louise Boumans, Bernd Kremer, Sinh Din, Ellen Stobberingh, Alfons G. H. Kessels, and Kenneth W. Kross. 2014. Application of an electronic nose in the diagnosis of head and neck cancer. *The Laryngoscope* 124, 6 (2014), 1377–1381. DOI:http://dx.doi.org/10.1002/lary.24463

[108]   Chang-Hsin Lin, Shu-Hsien Huang, Ju-Ling Shih, Alexandra Covaci, and Gheorghita Ghinea. 2017. Game-Based Learning Effectiveness and Motivation Study between Competitive and Cooperative Modes. In *Advanced Learning Technologies (ICALT), 2017 IEEE 17th International Conference on*. IEEE, 123–127.

[109]   CyberGlove Systems LLC. 2016. Home. http://www.cyberglovesystems.com/. (2016). Accessed: 19-Sep-2016.

[110]   M Luisa Dematté, Daniel Sanabria, and Charles Spence. 2006. Cross-modal associations between odors and colors. *Chemical senses* 31, 6 (2006), 531–538.

[111]   N. Magnenat-Thalmann and U. Bonanni. 2006. Haptics in virtual reality and multimedia. *IEEE MultiMedia* 13, 3 (July 2006), 6–11. DOI:http://dx.doi.org/10.1109/MMUL.2006.56

[112]   Jonatan Martnez, Arturo S. Garca, Miguel Oliver, Jos P. Molina, and Pascual Gonzlez. 2014. VITAKI: A Vibrotactile Prototyping Toolkit for Virtual Reality and Video Games. *International Journal of HumanfiComputer Interaction* 30, 11 (2014), 855–871. DOI:http://dx.doi.org/10.1080/10447318.2014.941272 arXiv:http://dx.doi.org/10.1080/10447318.2014.941272

[113]   H. Matsukura, T. Nihei, and H. Ishida. 2011. Multi-sensorial field display: Presenting spatial distribution of airflow and odor. In *2011 IEEE Virtual Reality Conference*. 119–122. DOI:http://dx.doi.org/10.1109/VR.2011.5759448

[114]   H. Matsukura, T. Yoneda, and H. Ishida. 2013. Smelling Screen: Development and Evaluation of an Olfactory Display System for Presenting a Virtual Odor Source. *IEEE Transactions on Visualization and Computer Graphics* 19, 4 (April 2013), 606–615. DOI:http://dx.doi.org/10.1109/TVCG.2013.40

[115]   Richard E Mayer and Roxana Moreno. 1998. A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of educational psychology* 90, 2 (1998), 312.

[116]   Dan Maynes-Aminzade. 2005. Edible bits: Seamless interfaces between people, data and food. In *Conference on Human Factors in Computing Systems (CHI'05)-Extended Abstracts*. Citeseer, 2207–2210.

[117]   Antonella Mazzoni and Nick Bryan-Kinns. 2016. Mood Glove: A haptic wearable prototype system to enhance mood music in film. *Entertainment Computing* 17 (2016), 9 – 17. DOI:http://dx.doi.org/10.1016/j.entcom.2016.06.002

[118]   David McGookin and Dariela Escobar. 2016. Hajukone: Developing an Open Source Olfactory Device. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. ACM, New York, NY, USA, 1721–1728. DOI:http://dx.doi.org/10.1145/2851581.2892339

[119] Matthew McGrath and Quentin Summerfield. 1985. Intermodal timing relations and audio-visual speech recognition by normal-hearing adults. *The Journal of the Acoustical Society of America* 77, 2 (1985), 678–685.

[120] Bruno Mesz, Mariano Sigman, and Marcos Trevisan. 2012. A composition algorithm based on crossmodal taste-music correspondences. *Frontiers in Human Neuroscience* 6 (2012), 71.

[121] Tassos A Mikropoulos and Antonis Natsis. 2011. Educational virtual environments: A ten-year review of empirical research (1999–2009). *Computers & Education* 56, 3 (2011), 769–780.

[122] Taeyong Moon and Gerard J. Kim. 2004. Design and Evaluation of a Wind Display for Virtual Reality. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST '04)*. ACM, New York, NY, USA, 122–128. DOI : http://dx.doi.org/10.1145/1077534.1077558

[123] G. M. Muntean, P. Perry, and L. Murphy. 2004. A new adaptive multimedia streaming system for all-IP multi-service networks. *IEEE Transactions on Broadcasting* 50, 1 (March 2004), 1–10. DOI : http://dx.doi.org/10.1109/TBC.2004.824745

[124] Takaki Murakami, Tanner Person, Charith Lasantha Fernando, and Kouta Minamizawa. 2017. Altered touch: miniature haptic display with force, thermal and tactile feedback for augmented haptics. In *ACM SIGGRAPH 2017 Posters*. ACM, 53.

[125] Martin Murer, Ilhan Aslan, and Manfred Tscheligi. 2013. LOLLio: Exploring Taste As Playful Modality. In *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction (TEI '13)*. ACM, New York, NY, USA, 299–302. DOI : http://dx.doi.org/10.1145/2460625.2460675

[126] Micah M Murray, Christoph M Michel, Rolando Grave de Peralta, Stephanie Ortigue, Denis Brunet, Sara Gonzalez Andino, and Armin Schnider. 2004. Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging. *Neuroimage* 21, 1 (2004), 125–135.

[127] Niall Murray, Oluwakemi A Ademoye, Gheorghita Ghinea, and Gabriel-Miro Muntean. 2017. A Tutorial for Olfaction-based Multisensorial Media Application Design and Evaluation. *ACM Computing Surveys (CSUR)* 50, 5 (2017), 67.

[128] Niall Murray, Brian Lee, Yuansong Qiao, and Gabriel Miro-Muntean. 2017. The impact of scent type on olfaction-enhanced multimedia quality of experience. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 47, 9 (2017), 2503–2515.

[129] Niall Murray, Brian Lee, Yuansong Qiao, and Gabriel-Miro Muntean. 2014. Multiple-scent enhanced multimedia synchronization. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* 11, 1s (2014), 12.

[130] Niall Murray, Brian Lee, Yuansong Qiao, and Gabriel-Miro Muntean. 2016. Olfaction-enhanced multimedia: A survey of application domains, displays, and research challenges. *ACM Computing Surveys (CSUR)* 48, 4 (2016), 56.

[131] Niall Murray, Yuansong Qiao, Brian Lee, and Gabriel-Miro Muntean. 2014. User-profile-based Perceived Olfactory and Visual Media Synchronization. *ACM Trans. Multimedia Comput. Commun. Appl.* 10, 1s, Article 11 (Jan. 2014), 24 pages. DOI : http://dx.doi.org/10.1145/2540994

[132] Takamichi Nakamoto. 2005. Study of Odor Recorder for Dynamical Change of Odor. *Chemical Senses* 30 (2005), i254. DOI : http://dx.doi.org/10.1093/chemse/bjh211

[133] T Nakamoto, Y Nakahira, H Hiramatsu, and T Moriizumi. 2001. Odor recorder using active odor sensing system. *Sensors and Actuators B: Chemical* 76, 1fi?!3 (2001), 465 – 469. DOI : http://dx.doi.org/10.1016/S0925-4005(01)00587-1 Proceeding of the Eighth International Meeting on Chemical Sensors IMCS-8 - Part 1.

[134] T. Nakamoto, S. Otaguro, M. Kinoshita, M. Nagahama, K. Ohinishi, and T. Ishida. 2008. Cooking Up an Interactive Olfactory Game Display. *IEEE Computer Graphics and Applications* 28, 1 (Jan 2008), 75–78. DOI : http://dx.doi.org/10.1109/MCG.2008.3

[135] Hiromi Nakamura and Homei Miyashita. 2012. Development and evaluation of interactive system for synchronizing electric taste and visual content. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 517–520.

[136] Takuya Nakano, Yuya Yoshioka, and Yasuyuki Yanagida. 2014. Effects of wind source configuration of wind displays on property of wind direction perception. In *Proc. Int. Conf. Adv. Comput.–Human Interactions*. Citeseer, 365–370.

[137] Hye Yeon Nam and Michael Nitsche. 2013. Interactive Installations As Performance: Inspiration for HCI. In *Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction (TEI '14)*. ACM, New York, NY, USA, 189–196. DOI : http://dx.doi.org/10.1145/2540930.2540976

[138] Somang Nam and Deborah Fels. 2016. *Design and Evaluation of an Authoring Tool and Notation System for Vibrotactile Composition*. Springer International Publishing, Cham, 43–53. DOI : http://dx.doi.org/10.1007/978-3-319-40244-4_5

[139] Suranga Nanayakkara, Elizabeth Taylor, Lonce Wyse, and S H. Ong. 2009. An Enhanced Musical Experience for the Deaf: Design and Evaluation of a Music Display and a Haptic Chair. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. ACM, New York, NY, USA, 337–346. DOI : http://dx.doi.org/10.1145/1518701.1518756

[140] Takuji Narumi. 2016. Multi-sensorial virtual reality and augmented human food interaction. In *Proceedings of the 1st Workshop on Multi-sensorial Approaches to Human-Food Interaction*. ACM, 1.

[141] Takuji Narumi, Shinya Nishizaka, Takashi Kajinami, Tomohiro Tanikawa, and Michitaka Hirose. 2011. Augmented Reality Flavors: Gustatory Display Based on Edible Marker and Cross-modal Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 93–102. DOI: http://dx.doi.org/10.1145/1978942.1978957

[142] Arinobu Niijima and Takefumi Ogawa. 2016. Study on Control Method of Virtual Food Texture by Electrical Muscle Stimulation. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16 Adjunct)*. ACM, New York, NY, USA, 199–200. DOI: http://dx.doi.org/10.1145/2984751.2984768

[143] Masahiro Nishizawa, Wanting Jiang, and Katsunori Okajima. 2016. Projective-AR system for customizing the appearance and taste of food. In *Proceedings of the 2016 workshop on Multimodal Virtual and Augmented Reality*. ACM, 6.

[144] Ryuichi Nisimura, Kazuki Hashimoto, Hideki Kawahara, and Toshio Irino. 2014. *Proposal for an Interactive 3D Sound Playback Interface Controlled by User behavior*. Springer International Publishing, Cham, 446–450. DOI: http://dx.doi.org/10.1007/978-3-319-07857-1_78

[145] Marianna Obrist, Pablo Cesar, David Geerts, Tom Bartindale, and Elizabeth F. Churchill. 2015. Online Video and Interactive TV Experiences. *interactions* 22, 5 (Aug. 2015), 32–37. DOI: http://dx.doi.org/10.1145/2799629

[146] Marianna Obrist, Rob Comber, Sriram Subramanian, Betina Piqueras-Fiszman, Carlos Velasco, and Charles Spence. 2014. Temporal, Affective, and Embodied Characteristics of Taste Experiences: A Framework for Design. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 2853–2862. DOI: http://dx.doi.org/10.1145/2556288.2557007

[147] Marianna Obrist, Alexandre N. Tuch, and Kasper Hornbaek. 2014. Opportunities for Odor: Experiences with Smell and Implications for Technology. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 2843–2852. DOI: http://dx.doi.org/10.1145/2556288.2557008

[148] Marianna Obrist, Carlos Velasco, Chi Vi, Nimesha Ranasinghe, Ali Israr, Adrian Cheok, Charles Spence, and Ponnampalam Gopalakrishnakone. 2016. Sensing the Future of HCI: Touch, Taste, and Smell User Interfaces. *interactions* 23, 5 (Aug. 2016), 40–49. DOI: http://dx.doi.org/10.1145/2973568

[149] Sharon Oviatt. 2006. Human-centered Design Meets Cognitive Load Theory: Designing Interfaces That Help People Think. In *Proceedings of the 14th ACM International Conference on Multimedia (MM '06)*. ACM, New York, NY, USA, 871–880. DOI: http://dx.doi.org/10.1145/1180639.1180831

[150] Sharon Oviatt, Rachel Coulston, Stefanie Tomko, Benfang Xiao, Rebecca Lunsford, Matt Wesson, and Lesley Carmichael. 2003. Toward a Theory of Organized Multimodal Integration Patterns During Human-computer Interaction. In *Proceedings of the 5th International Conference on Multimodal Interfaces (ICMI '03)*. ACM, New York, NY, USA, 44–51. DOI: http://dx.doi.org/10.1145/958432.958443

[151] Panocam3D.com. 2017. Stereoscopic and Panoramic Video Cameras - 3D 360 Videos Production. http://www.panocam3d.com. (2017). Accessed in 2017-03-19.

[152] S. A. Panels, J. C. Roberts, and P. J. Rodgers. 2010. HITPROTO: a tool for the rapid prototyping of haptic interactions for haptic data visualization. In *2010 IEEE Haptics Symposium*. 261–268. DOI: http://dx.doi.org/10.1109/HAPTIC.2010.5444647

[153] Cesare V. Parise. 2016. Crossmodal Correspondences: Standing Issues and Experimental Guidelines. *Multisensory Research* 29, 1-3 (2016), 7–28. DOI: http://dx.doi.org/10.1163/22134808-00002502

[154] Innwoo Park and Michael J Hannafin. 1993. Empirically-based guidelines for the design of interactive multimedia. *Educational Technology Research and Development* 41, 3 (1993), 63–85.

[155] Robert A Pieretti, Sandra D Kaul, Razi M Zarchy, and Laureen M OfiHanlon. 2015. Using a multimodal approach to facilitate articulation, phonemic awareness, and literacy in young children. *Communication Disorders Quarterly* 36, 3 (2015), 131–141.

[156] K-H Plattig and J Innitzer. 1976. Taste qualities elicited by electric stimulation of single human tongue papillae. *Pflügers Archiv* 361, 2 (1976), 115–120.

[157] Veljko Potkonjak, Michael Gardner, Victor Callaghan, Pasi Mattila, Christian Guetl, Vladimir M Petrović, and Kosta Jovanović. 2016. Virtual laboratories for education in science, technology, and engineering: A review. *Computers & Education* 95 (2016), 309–327.

[158] Manoj Prasad, Paul Taele, Daniel Goldberg, and Tracy A. Hammond. 2014. HaptiMoto: Turn-by-turn Haptic Route Guidance Interface for Motorcyclists. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 3597–3606. DOI: http://dx.doi.org/10.1145/2556288.2557404

[159] Nimesha Ranasinghe, Adrian David Cheok, and Ryohei Nakatsu. 2012. Taste/IP: the sensation of taste for digital communication. In *Proceedings of the 14th ACM international conference on Multimodal interaction*. ACM, 409–416.

[160] Nimesha Ranasinghe and Ellen Yi-Luen Do. 2016. Virtual Sweet: Simulating Sweet Sensation Using Thermal Stimulation on the Tip of the Tongue. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16 Adjunct)*. ACM, New York, NY, USA, 127–128. DOI: http://dx.doi.org/10.1145/2984751.2985729

[161] Nimesha Ranasinghe, Pravar Jain, Shienny Karwita, and Ellen Yi-Luen Do. 2017. Virtual Lemonade: Let's Teleport Your Lemonade!. In *Proceedings of the Eleventh International Conference on Tangible, Embedded, and Embodied Interaction (TEI '17)*. ACM, New York, NY, USA, 183–190. DOI:http://dx.doi.org/10.1145/3024969.3024977

[162] Nimesha Ranasinghe, Thi Ngoc Tram Nguyen, Yan Liangkun, Lien-Ya Lin, David Tolley, and Ellen Yi-Luen Do. 2017. Vocktail: A Virtual Cocktail for Pairing Digital Taste, Smell, and Color Sensations. In *Proceedings of the 2017 ACM on Multimedia Conference (MM '17)*. ACM, New York, NY, USA, 1139–1147. DOI:http://dx.doi.org/10.1145/3123266.3123440

[163] Shahzad Rasool and Alexei Sourin. 2016. Real-time haptic interaction with RGBD video streams. *The Visual Computer* 32, 10 (2016), 1311–1321.

[164] Emmanuelle Richard, Angèle Tijou, Paul Richard, and J-L Ferrier. 2006. Multi-modal virtual environments for education with haptic and olfactory feedback. *Virtual Reality* 10, 3-4 (2006), 207–225.

[165] Lars A Ross, Dave Saint-Amour, Victoria M Leavitt, Daniel C Javitt, and John J Foxe. 2007. Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. *Cerebral Cortex* 17, 5 (2007), 1147–1153.

[166] J. Ryu and S. Choi. 2008. posVibEditor: Graphical authoring tool of vibrotactile patterns. In *2008 IEEE International Workshop on Haptic Audio visual Environments and Games*. 120–125. DOI:http://dx.doi.org/10.1109/HAVE.2008.4685310

[167] Frank Rck, Nicolae Barsan, and Udo Weimar. 2008. Electronic Nose:fift Current Status and Future Trends. *Chemical Reviews* 108, 2 (2008), 705–725. DOI:http://dx.doi.org/10.1021/cr068121q arXiv:http://dx.doi.org/10.1021/cr068121q PMID: 18205411.

[168] Jamal Saboune and Juan Manuel Cruz-Hernandez. 2015. Method and apparatus to generate haptic feedback from video content analysis. (June 23 2015). US Patent 9,064,385.

[169] Estêvão Bissoli Saleme, Jordano R. Celestrini, and Celso A. Saibel Santos. 2017. Time Evaluation for the Integration of a Gestural Interactive Application with a Distributed Mulsemedia Platform. In *Proceedings of the 8th ACM on Multimedia Systems Conference (MMSys'17)*. ACM, New York, NY, USA, 308–314. DOI:http://dx.doi.org/10.1145/3083187.3084013

[170] Nur Amira Samshir, Nurafiqah Johari, Kasun Karunanayaka, and Adrian David Cheok. 2016. Thermal Sweet Taste Machine for Multisensory Internet. In *Proceedings of the Fourth International Conference on Human Agent Interaction*. ACM, 325–328.

[171] K. Sato, H. Kajimoto, N. Kawakami, and S. Tachi. 2007. Electrotactile Display for Integration with Kinesthetic Display. In *RO-MAN 2007 - The 16th IEEE International Symposium on Robot and Human Interactive Communication*. 3–8. DOI:http://dx.doi.org/10.1109/ROMAN.2007.4415044

[172] Oliver Schneider, Karon MacLean, Colin Swindells, and Kellogg Booth. 2017. Haptic experience design: What hapticians do and where they need help. *International Journal of Human-Computer Studies* 107 (2017), 5–21.

[173] Oliver Schneider, Siyan Zhao, and Ali Israr. 2015. *FeelCraft: User-Crafted Tactile Content*. Springer Japan, Tokyo, 253–259. DOI:http://dx.doi.org/10.1007/978-4-431-55690-9_47

[174] Oliver S. Schneider, Ali Israr, and Karon E. MacLean. 2015. Tactile Animation by Direct Manipulation of Grid Displays. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology (UIST '15)*. ACM, New York, NY, USA, 21–30. DOI:http://dx.doi.org/10.1145/2807442.2807470

[175] O. S. Schneider and K. E. MacLean. 2016. Studying design process and example use with Macaron, a web-based vibrotactile effect editor. In *2016 IEEE Haptics Symposium (HAPTICS)*. 52–58. DOI:http://dx.doi.org/10.1109/HAPTICS.2016.7463155

[176] H. Seifi, C. Anthonypillai, and K. E. MacLean. 2014. End-user customization of affective tactile messages: A qualitative examination of tool parameters. In *2014 IEEE Haptics Symposium (HAPTICS)*. 251–256. DOI:http://dx.doi.org/10.1109/HAPTICS.2014.6775463

[177] H. Seifi, K. Zhang, and K. E. MacLean. 2015. VibViz: Organizing, visualizing and navigating vibration libraries. In *2015 IEEE World Haptics Conference (WHC)*. 254–259. DOI:http://dx.doi.org/10.1109/WHC.2015.7177722

[178] Alix Seigneuric, Karine Durand, Tao Jiang, Jean-Yves Baudouin, and Benoist Schaal. 2010. The nose tells it to the eyes: crossmodal associations between olfaction and vision. *Perception* 39, 11 (2010), 1541–1554.

[179] Ladan Shams and Aaron R Seitz. 2008. Benefits of multisensory learning. *Trends in cognitive sciences* 12, 11 (2008), 411–417.

[180] Roland Sigrist, Georg Rauter, Robert Riener, and Peter Wolf. 2013. Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review. *Psychonomic bulletin & review* 20, 1 (2013), 21–53.

[181] Juan M. Silva, Mauricio Orozco, Jongeun Cha, Abdulmotaleb El Saddik, and Emil M. Petriu. 2013. Human Perception of Haptic-to-video and Haptic-to-audio Skew in Multimedia Applications. *ACM Trans. Multimedia Comput. Commun. Appl.* 9, 2, Article 9 (May 2013), 16 pages. DOI:http://dx.doi.org/10.1145/2457450.2457451

[182] Julia Simner, Christine Cuskley, and Simon Kirby. 2010. What sound does that taste? Cross-modal mappings across gustation and audition. *Perception* 39, 4 (2010), 553–569.

[183] M. Simsek, A. Aijaz, M. Dohler, J. Sachs, and G. Fettweis. 2016. 5G-Enabled Tactile Internet. *IEEE Journal on Selected Areas in Communications* 34, 3 (March 2016), 460–473. DOI:http://dx.doi.org/10.1109/JSAC.2016.2525398

[184] BG Slocombe, DA Carmichael, and J Simner. 2016. Cross-modal tactile–taste interactions in food evaluations. *Neuropsychologia* 88 (2016), 58–64.

[185] Eric R. Spangenberg, Bianca Grohmann, and David E. Sprott. 2005. It's beginning to smell (and sound) a lot like Christmas: the interactive effects of ambient scent and music in a retail setting. *Journal of Business Research* 58, 11 (2005), 1583 – 1589. DOI:http://dx.doi.org/10.1016/j.jbusres.2004.09.005 Retailing Reseach:things change, things stay the sameSixth {SMA} Retailing Research Symposium.

[186] Charles Spence. 2011. Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics* 73, 4 (2011), 971–995.

[187] Charles Spence and Cristy Ho. 2015. Multisensory information processing. In *APA handbook of human systems integration*, D. A. Boehm-Davis, F. T. Durso, and J. D. Lee (Eds.). American Psychological Association, Washington, DC, US, 435–448. DOI: 10.1037/14528-027 DOI:10.1037/14528-027.

[188] Charles Spence, Xiaoang Wan, Andy Woods, Carlos Velasco, Jialin Deng, Jozef Youssef, and Ophelia Deroy. 2015. On tasty colours and colourful tastes? Assessing, explaining, and utilizing crossmodal correspondences between colours and basic tastes. *Flavour* 4, 1 (08 Jul 2015), 23. DOI:http://dx.doi.org/10.1186/s13411-015-0033-1

[189] Barry E Stein. 2016. *The new handbook of multisensory processing.* Mit Press Cambridge, MA.

[190] Barry E Stein, Terrence R Stanford, and Benjamin A Rowland. 2014. Development of multisensory integration from the perspective of the individual neuron. *Nature Reviews Neuroscience* 15, 8 (2014), 520–535.

[191] R. Steinmetz. 1996. Human perception of jitter and media synchronization. *IEEE Journal on Selected Areas in Communications* 14, 1 (Jan 1996), 61–72. DOI:http://dx.doi.org/10.1109/49.481694

[192] NR Stiles and Shinsuke Shimojo. 2014. Auditory Sensory Substitution is Intuitive and Automatic with Texture Stimuli. *Scientific reports* 5 (2014), 15628–15628.

[193] Thomas Stockhammer. 2011. Dynamic Adaptive Streaming over HTTP –: Standards and Design Principles. In *Proceedings of the Second Annual ACM Conference on Multimedia Systems (MMSys '11)*. ACM, New York, NY, USA, 133–144. DOI:http://dx.doi.org/10.1145/1943552.1943572

[194] Yevgeniya Sulema. 2016. Mulsemedia vs. Multimedia: State of the art and future trends. In *Systems, Signals and Image Processing (IWSSIP), 2016 International Conference on*. IEEE, 1–5.

[195] Chie Suzuki, Takuji Narumi, Tomohiro Tanikawa, and Michitaka Hirose. 2014. Affecting Tumbler: Affecting Our Flavor Perception with Thermal Feedback. In *Proceedings of the 11th Conference on Advances in Computer Entertainment Technology (ACE '14)*. ACM, New York, NY, USA, Article 19, 10 pages. DOI:http://dx.doi.org/10.1145/2663806.2663825

[196] Risa Suzuki, Shutaro Homma, Eri Matsuura, and Ken-ichi Okada. 2014. System for Presenting and Creating Smell Effects to Video. In *Proceedings of the 16th International Conference on Multimodal Interaction (ICMI '14)*. ACM, New York, NY, USA, 208–215. DOI:http://dx.doi.org/10.1145/2663204.2663269

[197] C. Swindells, S. Pietarinen, and A. Viitanen. 2014. Medium fidelity rapid prototyping of vibrotactile haptic, audio and video effects. In *2014 IEEE Haptics Symposium (HAPTICS)*. 515–521. DOI:http://dx.doi.org/10.1109/HAPTICS.2014.6775509

[198] Benjamin Tag, Takuya Goto, Kouta Minamizawa, Ryan Mannschreck, Haruna Fushimi, and Kai Kunze. 2017. atmo-Sphere: mindfulness over haptic-audio cross modal correspondence. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*. ACM, 289–292.

[199] A. Tatematsu, Y. Ishibashi, N. Fukushima, and S. Sugawara. 2010. QoE assessment in haptic media, sound and video transmission: Influences of network latency. In *2010 IEEE International Workshop Technical Committee on Communications Quality and Reliability (CQR 2010)*. 1–6. DOI:http://dx.doi.org/10.1109/CQR.2010.5619913

[200] Valerie A Thompson and Allan Paivio. 1994. Memory for pictures and sounds: Independence of auditory and visual codes. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale* 48, 3 (1994), 380.

[201] Christian Timmerer, Jean Gelissen, Markus Waltl, and Hermann Hellwagner. 2009. Interfacing with virtual worlds. *Network and Electronic Media Summit* (2009).

[202] Nora Turoman, Carlos Velasco, Yi-Chuan Chen, Pi-Chun Huang, and Charles Spence. 2017. Symmetry and its role in the crossmodal correspondence between shape and taste. *Attention, Perception, & Psychophysics* (2017), 1–14.

[203] S. u. Rehman, J. Sun, L. Liu, and H. Li. 2008. Turn Your Mobile Into the Ball: Rendering Live Football Game Using Vibration. *IEEE Transactions on Multimedia* 10, 6 (Oct 2008), 1022–1033. DOI:http://dx.doi.org/10.1109/TMM.2008.2001352

[204] Haruya Uematsu, Daichi Ogawa, Ryuta Okazaki, Taku Hachisu, and Hiroyuki Kajimoto. 2016. HALUX: Projection-based Interactive Skin for Digital Sports. In *ACM SIGGRAPH 2016 Emerging Technologies (SIGGRAPH '16)*. ACM, New York, NY, USA, Article 10, 2 pages. DOI:http://dx.doi.org/10.1145/2929464.2929479

[205] Erik Van der Burg, Christian NL Olivers, Adelbert W Bronkhorst, and Jan Theeuwes. 2008. Pip and pop: nonspatial auditory signals improve spatial visual search. *Journal of Experimental Psychology: Human Perception and Performance* 34, 5 (2008), 1053.

[206] Carlos Velasco, Marianna Obrist, Olivia Petit, and Charles Spence. 2018. Multisensory technology for flavor augmentation: a mini review. *Frontiers in psychology* 9 (2018), 26.

[207] Carlos Velasco, Andy T. Woods, Olivia Petit, Adrian D. Cheok, and Charles Spence. 2016. Crossmodal correspondences between taste and shape, and their implications for product packaging: A review. *Food Quality and Preference* 52 (2016), 17 – 26. DOI:http://dx.doi.org/10.1016/j.foodqual.2016.03.005

[208] Chi Thanh Vi, Damien Ablart, Daniel Arthur, and Marianna Obrist. 2017. Gustatory interface: the challenges of fihowfito stimulate the sense of taste. In *Proceedings of 2nd ACM SIGCHI International Workshop on Multisensory Approaches to Human-Food Interaction (MHFIfi17)*. ACM, 29–33.

[209] Georg von Békésy. 1964. Olfactory analogue to directional hearing. *Journal of Applied Physiology* 19, 3 (1964), 369–373. arXiv:http://jap.physiology.org/content/19/3/369.full.pdf http://jap.physiology.org/content/19/3/369

[210] Markus Waltl. 2010. *Enriching Multimedia with Sensory Effects: Annotation and Simulation Tools for the Representation of Sensory Effects*. VDM Verlag, Saarbrücken, Germany, Germany.

[211] Markus Waltl, Benjamin Rainer, Christian Timmerer, and Hermann Hellwagner. 2013. An end-to-end tool chain for Sensory Experience based on MPEG-V. *Signal Processing: Image Communication* 28, 2 (2013), 136 – 150. DOI:http://dx.doi.org/10.1016/j.image.2012.10.009 MPEG-V.

[212] Qian Janice Wang and Charles Spence. 2017. The Role of Pitch and Tempo in Sound-Temperature Crossmodal Correspondences. *Multisensory Research* 30, 3-5 (2017), 307–320.

[213] Oliver Wright, Yasmina Jraissati, and Dila ..zelik. 2017. Cross-Modal Associations Between Color and Touch: Mapping Haptic and Tactile Terms to the Surface oftheMunsell Color Solid. *Multisensory Research* 30, 7-8 (2017), 691–715. DOI:http://dx.doi.org/10.1163/22134808-00002589

[214] T. Yamanaka, R. Matsumoto, and T. Nakamoto. 2003. Fundamental study of odor recorder for multicomponent odor using recipe exploration method based on singular value decomposition. *IEEE Sensors Journal* 3, 4 (Aug 2003), 468–474. DOI:http://dx.doi.org/10.1109/JSEN.2003.815778

[215] Z. Yuan, T. Bi, G. M. Muntean, and G. Ghinea. 2015. Perceived Synchronization of Mulsemedia Services. *IEEE Transactions on Multimedia* 17, 7 (July 2015), 957–966. DOI:http://dx.doi.org/10.1109/TMM.2015.2431915

[216] Zhenhui Yuan, Gheorghita Ghinea, and Gabriel-Miro Muntean. 2015. Beyond multimedia adaptation: Quality of experience-aware multi-sensorial media delivery. *IEEE Transactions on Multimedia* 17, 1 (2015), 104–117.

[217] Z. Yuan, G. Ghinea, and G. M. Muntean. 2015. Beyond Multimedia Adaptation: Quality of Experience-Aware Multi-Sensorial Media Delivery. *IEEE Transactions on Multimedia* 17, 1 (Jan 2015), 104–117. DOI:http://dx.doi.org/10.1109/TMM.2014.2371240

[218] John S Zelek, Sam Bromley, Daniel Asmar, and David Thompson. 2003. A haptic glove as a tactile-vision sensory substitution for wayfinding. *Journal of Visual Impairment and Blindness* 97, 10 (2003), 621–632.

[219] Q. Zeng, Y. Ishibashi, N. Fukushima, S. Sugawara, and K. E. Psannis. 2013. Influences of inter-stream synchronization errors among haptic media, sound, and video on quality of experience in networked ensemble. In *2013 IEEE 2nd Global Conference on Consumer Electronics (GCCE)*. 466–470. DOI:http://dx.doi.org/10.1109/GCCE.2013.6664891

[220] Syeda Erfana Zohora, A. M. Khan, and Nisar Hundewale. 2013. *Chemical Sensors Employed in Electronic Noses: A Review*. Springer Berlin Heidelberg, Berlin, Heidelberg, 177–184. DOI:http://dx.doi.org/10.1007/978-3-642-31600-5_18

[221] L. Zou, R. Trestian, and G. M. Muntean. 2015. E2DOAS: User experience meets energy saving for multi-device adaptive video delivery. In *2015 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. 444–449. DOI:http://dx.doi.org/10.1109/INFCOMW.2015.7179425