Chapter 2
Novel Approaches to Immersive Media: From Enlarged Field-of-View to Multi-sensorial Experiences

Iris Galloso, Claudio Feijóo, and Asunción Santamaría

Abstract This chapter presents a review of current evidence on the influence of immersion (defined in terms of the technical features of the system) on the user experience in multimedia applications. Section 2.1 introduces the concepts of media enjoyment, presence, and Quality of Experience (QoE) that frame our analysis from the user perspective. Section 2.2 discusses the bounding effects of multimodal perception on the previously defined metrics. Section 2.3 analyses the influence of relevant technical factors on presence, enjoyment, and QoE, with emphasis on those characterizing the level of immersion delivered by system across four dimensions: inclusiveness, extensiveness, surrounding, and vividness. Section 2.4 presents recent works integrating some of these factors into multi-sensorial media experiences and highlights open issues and research challenges to be tackled in order to deliver cost-effective multi-sensorial media solutions to the mass market.

2.1 Conceptualizing User Experience with Entertaining Technologies

2.1.1 Media Enjoyment

Consistent results across more than seven decades of mass media effects research (in particular, under the uses and gratifications approach) identify enjoyment as the primary gratification sought from media [1]. Considered a direct predictor of audience, **media enjoyment** has been in the focus of media effects research for almost 40 years [1, 2].

I. Galloso (✉) • C. Feijóo • A. Santamaría
CeDInt-UPM. Campus de Montegancedo, Pozuelo de Alarcón 28223, Madrid, Spain
e-mail: iris@cedint.upm.es; cfeijoo@cedint.upm.es; asun@cedint.upm.es

A. Kondoz, T. Dagiuklas (eds.), *Novel 3D Media Technologies*,
DOI 10.1007/978-1-4939-2026-6_2

The encyclopedia of positive psychology defines enjoyment as "*engagement in a challenging experience that either includes or results in a positive affective state*" [3]. Csikszentmihalyi, the father of the theory of the optimal experience, conceptualizes the term beyond pleasure, arguing that enjoyment is characterized by "*forward movement that accomplishes something novel or challenging, resulting in a growth experience*" [4]. Indeed, an enjoyable media experience presents several features inherent to the state of flow, such as: intense and focused concentration, merging of action and awareness, loss of reflective self-consciousness, distortion of temporal experience, and experience of the activity as intrinsically rewarding [5].

The components and dynamics underlying media enjoyment have been studied across a great variety of genres as a dependent variable of personality traits, individual differences, mood, content characteristics, social context, or a combination of these. As a result, it has been characterized as a multidimensional construct conditioned by affective components in a first place but also by cognitive and behavioral factors [6–11]. In particular, emotional enjoyment has been found closely linked to entertainment as a media effect, which at the same time correlates with some of the more frequently reported motivations for media use: arousal, to pass time, relaxation, and to escape. In this sense, media provides a mean to "*escape to a fantasy world where emotions can be experienced*" [1].

2.1.2 Presence

The desire of escaping from reality (or to some extent, of being "transported" to a different place) leads to the concept of **presence**, which is defined as "*the subjective experience of being in one place*", even when the person is physically located in another [12].

The sense of presence has been widely analyzed as a mean to describe the psychological mechanisms underlying user experience with entertaining technologies, with particular emphasis on interactive computer-generated applications. Presence has been found strongly related to the capability of mediated environments—including 3DTV, videogames, and artistic and cultural heritage Virtual Environments (VEs)—to elicit emotions [13, 14] and in particular, enjoyment [15, 16]. In consequence, an enhanced sense of presence is considered to have a direct impact on the adoption potential of these entertaining applications.

The factors influencing the subjective sense of presence can be classified as those related to the media form, the media content and the media users [17]. Media form is related to the *extent and fidelity of sensory information* and to the consistency between the sensory inputs/outputs presented through the different modalities [12, 17, 18]. In other words, it encompasses those features characterizing in an objective manner the capability of a specific software and hardware solution to deliver a rich and consistent multi-sensorial media content in a transparent manner (i.e., as an invisible medium). The influence of relevant media form and content factors in the sense of presence is analyzed in Sect. 2.3.

Media content is a very broad category concerning issues as the story, messages, objects, activities, or characters presented as part of the dynamic experience. Among the content characteristics identified as determinants of presence are: social realism, representation of virtual body, autonomous behavior and appearance of characters and objects, the ability to modify the physical environment, and to anticipate the effect of an action and possible interactions between the type, nature, and complexity of tasks or activities [17–20].

As regards to the characteristics of the media user, the sense of presence has been found significantly influenced by emotional, cognitive, and motivational-behavioral factors, such as: immersive tendency (measured in terms of absorption and emotional involvement, which at the same time correlate with openness to experience and with neuroticism and extraversion in the last case), attention, relevance, skill, perceived control, challenge, cognitive capabilities (e.g., spatial intelligence), and personality traits [21–26]. In particular, works as [24] provide evidence on the impact of user features as competence, challenge, and ability to explore the VE as well as of media form variables as interaction speed, mapping, and range on the spatial presence. The study also supports previous findings on the relation between the emotional response (in terms of arousal) and the levels of attention, spatial presence, and situational involvement. Such results point to a significant influence of the individual's cognitive–affective assessment of the immersive media form and content on the emotional response and the perceived level of physical presence.

2.1.3 Quality of Experience

The factors and mechanisms that influence the subjective quality assessment of a multimedia asset (i.e., the content quality as perceived by an individual) are analyzed by researchers in the field of user experience in multimedia applications. The study of these phenomena has been encompassed into the concept of "**Quality of Experience (QoE)**," which is defined as *"the degree of delight or annoyance of the user of an application or service (…) which results from the fulfillment of [his/her] expectations with respect to the utility and/or enjoyment of the application or service in the light of the his/her personality and current state"* [27]. In this context, the user experience has been found influenced by a combination of interrelated factors of contextual, technical, and human nature.

Contextual factors have been defined as those *"that embrace any situational property to describe the user's environment"* [28]. These not only concern the physical context, but also other dynamic or static features of economic, social, or technical nature [27]. Research on the influence of contextual factors is out of the scope of this chapter.

Technical factors (also known as system factors) refer to those conditioning the resulting technical quality of an application or service [28]. Different categories of system factors have been proposed in literature, both from a technical perspective, in which they are divided according to the related components of the service/architecture chain and from a user perspective, considering their final influence/manifestation during the experience [29]. Relevant findings on the influence of system factors on the QoE are analyzed in Sect. 2.3.

Human factors comprise those features that characterize the user and have an influence on his/her perception of quality. Quality perception is framed by the human perception mechanism, which flows at two main levels: the early sensory processing level, aimed at extracting relevant features from the incoming multimodal sensory information, and the high-level cognitive processing level, focused on conscious interpretation and judgment [28, 30]. This dynamic is supported from a psychological perspective by Lazarus' theory of appraisal [31]; in which primary appraisal involves an appraisal of the outside world and secondary appraisal involves an appraisal of the individual themselves.

Although this classification has been useful for analysis purposes, the boundaries between the two processing levels are not clearly established. In contrast, there is strong evidence pointing to a modulating effect of high level factors as knowledge, emotions, expectations, attitudes, and goals on the relative importance of sensory modalities and their attributes, as well as on the orientation of attentional resources accordingly [32–34]. These changes in early sensory processing can be subject to a specific domain of expertise (e.g., image-based diagnosis) [35, 36] or can be eventually consolidated as a general ability [37, 38]. Furthermore, in case of discrepancies between the individual knowledge schema (built from past experiences and from abstract expectations and representations of the external reality) and the sensory input, the structure of the schema can be modified to integrate the contradictory stimuli (i.e., absorption of new knowledge) [39].

2.2 The Bounding Effect of Multimodal Perception

An extended belief in the presence research community is that the more extensive an immersive system is (i.e., in terms of its capability to stimulate a greater number of human senses), the greater its capability to evoke presence (see [17] and citations thereof). This hypothesis is supported by works as [40], where the addition of tactile, olfactory, and auditory cues showed to have a positive impact on the sense of presence. Likewise, in [41] an inverse correlation between the mental processing times (i.e., simple detection times) and the number of sensory modalities presented (unimodal, bimodal, or trimodal) was found. However, these results can't be generalized in a straightforward manner to the quality perception context.

Multimodal perception is a complex phenomenon dealing with the integration of two or more sensory inputs into a single, coherent, and meaningful stimulus. Although the factors influencing the perceptual experience have not been entirely characterized yet, there is strong evidence on the integration and sharing of perceptual information since the very early sensory processing stages [42, 43] and on the bounding effects of cognitive, emotional, and personality factors [10].

The presence of a given modality can distort or modulate (either intensifying or attenuating) the perception in other modality. Cross-modal interaction processes—as for example, synesthesia—underline the relative importance of the different sensory modalities presented (see [43] and citations thereof). This phenomenon has been widely analyzed, from an empirical perspective, in terms of the relative influence of vision and sound on task-related performance [44–48]. Findings reveal the potential of vision to alter the perception of speech and spatial location of audio sources and the influence of audio on vision in terms of temporal resolution, intensity, quality, structure, and interpretation of visual motion events. Concerning other modalities, a form of synesthesia—defined as "crossmodal transfer"—has been reported between vision and touch. The phenomenon is characterized by the appearance of a perceptual illusion in one modality induced by a correlated stimulus on other sensory modality (e.g., illusion of physical resistance induced by the manipulation of a virtual object in a mediated environment) [49]. Interestingly (although not surprisingly), this cross-modal illusion was found correlated with the sensation of spatial and sensory presence in the displayed environment. In [41], participants reacted faster (i.e., lower simple detection times were measured) to auditory and haptic stimulus than to visual stimulus when only one of them was presented (unimodal condition). In coherence, the bimodal auditory–haptic combination resulted in even faster reactions in comparison to those reported for each unimodal component and for the other two bimodal combinations (visual-haptic and visual-auditory). These results suggest a highly relevant influence of auditory and haptic stimuli on processing times at the initial perceptual stage, which according to the authors allows users more time in the consequent cognitive stages, enabling them better integration and filling in of missing information. Similarly, the authors in [50] found that haptic feedback can led to an improved task performance and feeling of "sense of togetherness" in shared VEs.

The majority of these empirical findings support the "modality appropriateness" hypothesis, which argues that *the modality that is most appropriate or reliable with respect to a given task dominates the perception in the context of that task* [51]. However, this and other approaches still require further elaboration to better explain complex effects as the wide variety of responses to inter-sensory divergent events reported in literature.

2.3 The Influence of Immersion

The effectiveness of a mediated environment to evoke cognitive and emotional reactions in a similar way to non-mediated experiences is heavily conditioned by the consistency between the displayed environment and an equivalent real environment as regards to the user experience [52, 53]. Two main components contributing to this realistic response are identified in [54]. These are: *place illusion*, defined as *the subjective sensation of being in a real place* (i.e., presence); and, *plausibility illusion,* referred as the *illusion that the scenario being depicted is actually occurring,* even when the person is cognitively aware of the fact that it isn't. In this sense, the plausibility judgment is highly related to the capability of the system as a whole to produce events that are meaningful and credible in comparison to the individual's expectations [54].

The capability of a technical system "*to deliver an inclusive, extensive, surrounding and vivid illusion of reality to the senses of a human participant*" has been defined as *immersion* [12, 55]. At this point, it should be emphasized a conceptual difference observed along this chapter between *immersion*, describing the capabilities of the system in an objective manner, and *presence*, considered a state of consciousness derived from the subjective perception of the experience [55].

An immersive system can be characterized in terms of four major dimensions as: *inclusive*, the extent to which it is able to isolate the physical reality; *extensive*, the range of sensory modalities addressed; *surrounding*, the extent to which the user is physically surrounded by the displayed environment; and, *vivid*, the resolution, fidelity, and variety of the sensorial stimuli delivered through each sensory modality. Each of these dimensions can be present at different levels and scales according to the correlating psycho–physiological responses and to the extent of their realization, respectively [12, 56].

2.3.1 Breakdown of System Factors

The independent and combined influence of system factors (including media form and content variables) on the emotional response, on the subjective assessment of presence and on quality judgment (in terms of QoE) has been analyzed extensively in scientific literature. In this section, we present and discuss relevant findings illustrating the complexity and wide variety of approaches to these fields on a non-exhaustive basis. Table 2.1 summarizes in a schematic way the facts analyzed as follows.

The influence of factors such as image motion, stereoscopy, and screen size has been studied in [57]. Image motion and stereoscopy showed to have, in that order, a great influence on presence. A large effect of screen size on presence was also observed, but only for the video stimulus that contained motion. High motion content has also shown an impact on the relative quality of video and audio perceived by the user, being the video quality weighted significantly higher than the audio quality when high motion content is presented [58].

A relationship between motion-based interaction and the perceived field-of-view (FOV) is reported in [59]. The perceived FOV for a small-hand held device was found around 50 % greater than the actual value when motion-based interaction was used. Coherently, the sense of presence under this condition was higher than or comparable to those in VR platforms with larger displays. The effects of head tracking, visual cues (including stereoscopic and motion parallax cues), and

Table 2.1 Influence of system factors on the emotional response, the sense of presence, and the QoE

System factors	Influence on the perceived... (correlation sign in parenthesis)	Sensory modality(ies) addressed	Main dimension (s) involved
Image motion [57, 58]	Presence (+), relative quality of video and audio (trade-off)	Sight	Vividness
Interaction between image motion and screen size [57]	Presence (+)	Sight	Vividness
Stereoscopy and stereoscopic and motion parallax cues [57, 60, 62]	Presence (+)	Sight	Vividness
Visual cues (spatial and object cues) [22]	Presence (+)	Sight	Vividness
Screen size, geometric field-of-view, omnidirectional video [57, 59–61, 63]	Presence (+), simulator sickness (+), enjoyment (−)	Sight	Inclusiveness, surrounding
Interaction between motion-based interaction and perceived field-of-view [59]	Presence (+)	Sight, proprioception	Extensiveness, inclusiveness
Interaction between natural (hand-based) interaction and narrative [63]	Presence (trade-off)	Sight, proprioception	Extensiveness
Pictorial realism [64, 65]	Presence (+)	Sight	Vividness
Delay of visual feedback [12, 65]	Presence (−)	Sight	Vividness
Presence or absence of interactivity [65]	Presence (+)	Sight, touch	Extensiveness
Frame rate [66, 67]	Presence (+)	Sight	Vividness
Passive haptic feedback [67]	Presence (+)	Sight, touch	Extensiveness
Presence or absence of spatialized sound, addition of spatialized versus non-spatialized sound to a stereoscopic display [14, 60, 68]	Presence (+), QoE (+), emotional response (+), emotion recognition (+), and emotional realism (+)	Hearing	Vividness
Image quality [69]	Audio quality (+), audiovisual quality (+)	Sight	Vividness
Audio quality [48, 69]	Presence (+), enjoyment (+), visual quality (+)	Hearing	Vividness
Natural physical interactions: head tracking [60], walk in place [70]	Presence (+)	Sight, touch, proprioception, equilibrioception	Extensiveness
Sensory effects (wind, vibration, light effects) and genre [71, 72]	QoE (+), enjoyment (+)	Sight, touch, thermoception	Extensiveness, surrounding, inclusiveness

(continued)

Table 2.1 (continued)

System factors	Influence on the perceived... (correlation sign in parenthesis)	Sensory modality(ies) addressed	Main dimension (s) involved
Olfactory effects [73–75]	QoE (+), relevance (+), reality (+), and enjoyment (+)	Olfacception	Extensiveness, surrounding, inclusiveness
Synchronization errors (outside the tolerance range) between video + audio or video without audio and olfaction [73, 74]	QoE (−), relevance (−), reality (−), and enjoyment (−)	Sight, hearing, olfacception	Vividness
Audio–video asynchrony (in particular, audio-led asynchrony) [28, 77, 78]	Clarity of the message (−), distraction (+)	Sight, hearing	Vividness
Stereoscopic disparities: large disparity at short convergence distances [79, 81]	Presence (−), enjoyment (−), QoE (−)	Sight	Vividness
(In stereoscopy) spatial distortions: shifts, magnification, rotation, keystone [80, 81]	Presence (−), enjoyment (−), QoE (−)	Sight	Vividness
(In stereoscopy) photometric asymmetries: luminance, color, contrast, crosstalk [80, 81]	Presence (−), enjoyment (−), QoE (−)	Sight	Vividness
Immersive technology (PC, big screen, HMD) [82]	Simulation sickness (+)	Sight	Inclusiveness, surrounding

geometric field-of-view are also explored in [22, 60–62]. The reported level of presence was positively influenced by the use of tracking, stereoscopic, and spatial and object cues [22, 60]. Presence was also correlated with the geometric field-of-view, showing an asymptotic behavior for field-of-view values beyond 140° [60, 61].

The experience of a theatrical performance and television using interactive omnidirectional video is qualitatively explored in [63]. Participants referred to the experience—in cognitive and physical terms—as being *discovering and exploring the (mediated) environment*. They also described transitions between the real and the displayed environment as disturbing and therefore, requiring a recalibration of the senses. Under this engaging experience, narrative was pushed to a second place and the hand-based interaction put in place was qualified as highly intuitive. The authors conclude that interactivity may influence the perception of narrative and therefore, these factors need to be carefully balanced to maximize presence.

Pictorial realism, observer interactivity, and delay of visual feedback are analyzed in [64, 65]. Realism and interactivity were shown to have a positive impact on presence while delay of visual feedback had an opposite effect. Participants reported a relative low influence of pictorial realism on presence in comparison to the other two components considered. The influence on presence of other screen variables as the frame rate has been shown in works as [66, 67].

As regards to the influence of auditory features, the audio quality, a realistic aural rendering of events, and the presence of auditory cues are considered to have a significant impact on the sense of presence [14, 48, 60]. The influence of realistic aural rendering, measured in terms of the number of audio channels (mono, stereo, and six-channel reproduction), on presence, emotional response and emotion recognition is analyzed in [14]. Stereo and six-channel reproduction had a significantly stronger impact in emotional response than the mono condition. Similarly, six-channel reproduction resulted in the highest ratings of presence and emotional realism. In coherence, an enhanced sense of presence and QoE are reported in [60] and [68] respectively, in response to the addition of spatialized audio. In [48], the relative influence of image quality (high definition vs. standard definition) and sound quality (Dolby 5.1 surround sound vs. Dolby stereo) on presence and enjoyment is studied. No significant effects of image quality were found. In contrast, the impact of sound quality on presence and enjoyment was shown to be significant. Furthermore, a significant cross-modal influence of audio on visual quality and vice versa has been reported in [69], although video quality seemed to dominate the overall perceived audiovisual quality in the context of the study.

The introduction of interaction has been also found to be significant [65]. In particular, interactions entailing natural physical movements—e.g., head movement [60] or walking in place [70]—and leading to a coherent system response (as regards to the individual's expectations) have shown a great impact on presence. Likewise, a significant influence of passive haptic feedback on presence has been reported in [67].

Less traditional stimuli as wind, vibration, and light effects have also shown a significant impact on the user experience (both in terms of enjoyment [71] and QoE), in particular with genres as action movies, sports, news, documentary, and commercials [72]. Likewise, olfactory effects have shown a positive influence on the perceived quality, relevance, and reality and on the reported enjoyment of a multimedia experience [73–75]. A potential exception to these positive effects may be given by synchronization errors producing a mismatch between video + audio and olfaction that is outside the temporal range of −30s (olfaction ahead of video + audio) to +20s (video + audio ahead of olfaction) [73]. However, in the case of video without audio, the tolerance to synchronization errors with olfaction decreases [74].

Technological breakdowns significantly reduce the potential of mediated environments to elicit presence and emotions [76]. For instance, an asynchronous reproduction of audio and video in the context of an audiovisual experience has shown a negative impact on the clarity of the message, distracting the viewer from the intended content [28]. In particular, users are more sensitive and report higher annoying effects in the case of audio-led asynchrony [77, 78]. Concerning stereoscopy, the variables influencing visual comfort in a negative fashion can be classified as: those introducing spatial distortions as shifts, magnification, rotation, and keystone; those leading to photometric asymmetries as luminance, color, contrast, and crosstalk; and, those leading to stereoscopic disparities as the disparity level, which shows a larger effect at short convergence distances [79–81]. Other meaningful studies show, for instance, that the level of simulator sickness is positively correlated with the geometric field-of-view [61]. Interestingly, presence was positively correlated with simulator sickness while enjoyment showed the opposite behavior. Similarly, a relationship between the immersive technology used and the severity of the negative effects reported was found in [82]. From the three immersive technologies analyzed (PC, Head Mounted Display (HMD), and big screen), HMD was the one producing more negative effects.

2.4 Implementing Multi-sensorial Media: Current Issues and Future Challenges

In an attempt to deliver a more immersive experience (i.e., more extensive, inclusive, surrounding, and vivid and in consequence, more enjoyable), several works propose the integration of sensory effects (beyond the conventional audiovisual content) into a multimedia asset. In particular, the concept seems to have the potential to bring actual immersive experiences to the home in a non-disruptive manner. That is, presenting sensory effects as a complement to current display technology that can be progressively adopted in transparent way.

An early initiative introducing meaningful lighting effects as a mean to complement the main audiovisual content is illustrated in [71]. Using their HomeLab research facility, the authors installed the Philips Living Light system. The system comprised four LightSpeakers (left–right front–back), a CenterLight, and a SubLight (situated underneath the couch). Ad-hoc light scripts were developed, with the support of light designers, theatre lighting experts, filmmakers, and musicians, for selected pieces of film and music. In the qualitative interview conducted participants expressed that *lighting effects made watching movies or listening to music a very enjoyable and more immersive experience*. The concept was also found appealing for creating personalized ambiances at home in the context of other social or personal activities.

The authors in [83] present sensory effects as a new dimension contributing to the QoE. The sensory effects are defined by the Sensory Effect Metadata (SEM) which should accompany or be retrieved together with the media content. The media processing engine is responsible for playing the audiovisual content and the corresponding sensory effects in a synchronized manner, considering the capabilities of the rendering devices. In one of their experiments, the authors analyze the influence of wind, vibration, and light effects in the user experience across different genres [72]. They found that the QoE was positively influenced by the introduction of sensory effects in the case of documentary, sports, and action genres. A less noticeable but still positive influence was found for commercials. As future research, Timmerer et al. [83] outline the need to establish a quality/utility model for sensory experiences and to develop (semi-)automatic annotation techniques for the generation and integration of sensory effects into media assets.

In [68] an end-to-end solution integrating sensory effects and interactive components into a hybrid (internet-broadcast) 3DTV system is presented. In the experimental setup deployed the main audiovisual content (showing an extended report of a football match) is complemented with binaural audio, cut grass scent, ambient lighting effects, and main lighting and shutter controllers (immersion dimension), and with interactive 3D objects and meaningful content delivered through a second screen (interaction dimension). A combination of broadcast–broadband transmission mechanisms is implemented to transmit this complementary content. At the user's premises, the content is delivered using the private IP network that connects the receiver gateway with the visualization terminals and sensory devices. The resulting system is compatible with current transmission (DVB-T), coding (AVC), multiplexing (MPEG-2), signaling (DVB), and automation (KNX) standards.

The development and official release of the MPEG-V standard by the Moving Picture Expert Group (MPEG) (and in particular, of its Part 3—Sensory Information [84]) represents an important step in the consolidation of the sensory experience concept. The standard establishes the architecture and associated information representations for the interaction and interoperability between virtual worlds (i.e., multimedia content) and real worlds through various sensors and actuators. The Part 3 defines a set of sensory effects (e.g., light, temperature, wind, vibration, touch) and associated semantics to deliver multi-sensorial content in association with multimedia.

A recent Special Issue on MPEG-V, released on February 2003, gathers several contributions proposing end-to-end frameworks that implement the standard for the creation and delivery of sensory effects synchronized with audiovisual content. Three relevant examples are those provided in [85–87]. In [85] an authoring tool called SEVino is used for the generation of the SEM descriptions corresponding to the different sensory effects introduced. The annotated content can be delivered over various distribution channels and visualized in any MPEG-V-compliant device. The SEM descriptions enable sensory effects to be rendered on off-the-shelf hardware synchronized with the main audiovisual content, either in a stand-alone application or in a web browser. Concerning the user experience, the authors confirmed the hypotheses that sensory effects have a positive impact on the QoE and on the intensity of emotions like happiness or fun.

The framework presented in [86] delivers sensory effects for home theaters based on MPEG-V standard via the broadcast network. The paper discusses thoroughly the technical choices provided by the MPEG-V standard (and those adopted in the targeted implementation) for the description, encoding, synchronization, transport, adaptation, and rendering of sensory effects. The work in [87] also exploits the broadcasting network capabilities to deliver a haptic-enabled system based on the MPEG-V standard. The paper illustrates the data flow within the system, which comprises four main stages: the creation of haptic contents using the MPEG-V standard, their encoding/decoding using BIFS encoders/decoders, their transmission through the MPEG-4 framework, and the rendering of the haptic effects using the MPEG-V standard in the rendering stage.

Important challenges remain to deliver a cost-effective implementation of multi-sensorial media solutions. A major issue identified by several authors is the need to establish a quality/utility model for sensory experiences. At the content creation stage, the development of effective (semi)automatic video annotation tools is a common challenge to the majority of multi-sensorial media implementations reviewed. Semantic video analysis seems a suitable strategy to identify those relevant events that should trigger sensory effects and/or interactive actions. A significant challenge is posed also by the use of computer vision algorithms to recognize specific scene features, objects, elements, or characters as a way to boost the visualization of additional content (i.e., sensory effects) associated to the recognized element/character. However, the cost-intensity of these algorithms needs to be decreased to enhance their deployment feasibility.

Other issues that should be subject to further analysis and/or improvement are: the identification of more efficient encoding/decoding methods (in particular for large SEM), the configuration of suitable transport mechanisms and the effective management of the various types of delays introduced along the transmission chain.

At the receiver side, automatic techniques are required for enabling the discovery, feature detection, and remote configuration of sensory devices. Likewise, effective automatic mechanisms shall be developed to adapt sensory effects to the capabilities of the specific rendering devices available at the user's premises (e.g., specific protocol, resolution, or time constraints).

From a market perspective, sensory-enhanced video has the potential to support the development and deployment of immersive media services targeting the wide domestic segment. The high consumption of action movies, sports, and documentaries in this context might favor the adoption of these solutions.

References

1. Sherry JL (2004) Flow and media enjoyment. Commun Theory 14:328–347. doi:10.1111/j.1468-2885.2004.tb00318.x
2. Ruggiero TE (2000) Uses and gratifications theory in the 21st century. Mass Commun Soc 3:3–37. doi:10.1207/S15327825MCS0301_02
3. Kapsner JC (2009) The encyclopedia of positive psychology. doi: 10.1111/b.9781405161251.2009.x
4. Csikszentmihalyi M (1990) Flow: the psychology of optimal experience. Harper & Row, New York
5. Nakamura J, Csikzentmihalyi M (2003) The construction of meaning through vital engagement. In: Keyes CLM, Haidt J (eds) Flourishing Posit. Psychol. Life well-lived. American Psychological Association, Washington, pp 83–104
6. Nabi RL, Krcmar M (2004) Conceptualizing media enjoyment as attitude: implications for mass media effects research. Commun Theory 14:288–310. doi:10.1111/j.1468-2885.2004.tb00316.x

7. Slater MD (2003) Alienation, aggression, and sensation seeking as predictors of adolescent use of violent film, computer, and website content. J Commun 53:105–121. doi:10.1111/j.1460-2466.2003.tb03008.x
8. Zillmann D, Vorderer P (2000) Media entertainment: the psychology of its appeal. Lawrence Erlbaum Associates, Mahwah, p 282
9. Raney AA, Bryant J (2002) Moral judgment and crime drama: an integrated theory of enjoyment. J Commun 52:402–415. doi:10.1111/j.1460-2466.2002.tb02552.x
10. Wechsung I, Schulz M, Engelbrecht K-P et al (2011) All users are (not) equal – the influence of user characteristics on perceived quality, modality choice and performance. In: Kobayashi T, Delgado RL-C (eds) Proceedings of the paralinguistic information and its integration in spoken dialogue systems workshop. Springer, New York, pp 175–186
11. Tamborini R, Bowman ND, Eden A et al (2010) Defining media enjoyment as the satisfaction of intrinsic needs. J Commun 60:758–777. doi:10.1111/j.1460-2466.2010.01513.x
12. Slater M, Wilbur S (1997) A framework for immersive virtual environments (FIVE): speculations on the role of presence in virtual environments. Presence-Teleop Virtual Environ 6:603–616
13. Riva G, Mantovani F, Capideville CS et al (2007) Affective interactions using virtual reality: The link between presence and emotions. Cyberpsychology Behav 10:45–56. doi:10.1089/cpb.2006.9993
14. Västfjäll D (2003) The subjective sense of presence, emotion recognition, and experienced emotions in auditory virtual environments. Cyberpsychology Behav 6:181–188. doi:10.1089/109493103321640374
15. Sylaiou S, Mania K, Karoulis A, White M (2010) Exploring the relationship between presence and enjoyment in a virtual museum. Int J Hum Comput Stud 68:243–253
16. Skalski P, Tamborini R, Shelton A et al (2010) Mapping the road to fun: Natural video game controllers, presence, and game enjoyment. New Media Soc 13:224–242
17. Lombard M, Ditton T (1997) At the heart of it all: the concept of presence. J Comput Commun 3:0. doi: 10.1111/j.1083-6101.1997.tb00072.x
18. IJsselsteijn W (2000) Presence: concept, determinants, and measurement. In: Proc. SPIE. Spie. pp 520–529
19. Slater M, Steed A, McCarthy J, Maringelli F (1998) The influence of body movement on subjective presence in virtual environments. Hum Factors 40:469–477
20. Schuemie MJ, van der Straaten P, Krijn M, van der Mast CAPG (2001) Research on presence in virtual reality: a survey. Cyberpsychol Behav 4:183–201. doi:10.1089/109493101300117884
21. Darken RP, Bernatovich D, Lawson JP, Peterson B (1999) Quantitative measures of presence in virtual environments: the roles of attention and spatial comprehension. Cyberpsychol Behav 2:337–347
22. Lee S, Kim GJ (2008) Effects of visual cues and sustained attention on spatial presence in virtual environments based on spatial and object distinction. Interact Comput 20:491–502
23. Novak T, Hoffman D, Yung Y (2000) Measuring the customer experience in online environments: a structural modeling approach. Mark Sci 19:22–44
24. Takatalo J, Nyman G, Laaksonen L (2008) Components of human experience in virtual environments. Comput Human Behav 24:1–15. doi:10.1016/j.chb.2006.11.003
25. Weibel D, Wissmath B, Mast FW (2010) Immersion in mediated environments: the role of personality traits. Cyberpsychol Behav Soc Netw 13:251–256
26. Alsina-Jurnet I, Gutierrez-Maldonado J (2010) Influence of personality and individual abilities on the sense of presence experienced in anxiety triggering virtual environments. Int J Hum Comput Stud 68:788–801
27. Le Callet P, Möller S, Perkis A (2012) Qualinet white paper on definitions of quality of experience. In: European network on quality of experience in multimedia systems and services (COST Action IC 1003), Lausanne, Switzerland, Version 1.1, June 3, 2012

28. Jumisko-Pyykkö S (2011) User-centered quality of experience and its evaluation methods for mobile television. Tampere University of Technology
29. Bracken C, Pettey G, Wu M (2011) Telepresence and attention: secondary task reaction time and media form. In: Proc. Int. Soc. Presence
30. Goldstein EB (2010) Sensation and perception. p 496
31. Lazarus RS (1993) From psychological stress to the emotions: a history of changing outlooks. Annu Rev Psychol 44:1–21. doi:10.1146/annurev.ps.44.020193.000245
32. Bey C, McAdams S (2002) Schema-based processing in auditory scene analysis. Percept Psychophys 64:844–854
33. Jennings JR, Van der Molen MW, Van der Veen FM, Debski KB (2002) Influence of preparatory schema on the speed of responses to spatially compatible and incompatible stimuli. Psychophysiology 39:496–504
34. Cui LC (2003) Do experts and naive observers judge printing quality differently? In: Miyake Y, Rasmussen DR (eds) Electron. Imaging 2004. International Society for Optics and Photonics, pp 132–145
35. Werner S, Thies B (2000) Is "Change Blindness" attenuated by domain-specific expertise? An expert-novices comparison of change detection in football images. Vis Cogn 7:163–173. doi:10.1080/135062800394748
36. Sowden PT, Davies IR, Roling P (2000) Perceptual learning of the detection of features in X-ray images: a functional role for improvements in adults' visual sensitivity? J Exp Psychol Hum Percept Perform 26:379–390
37. Curran T, Gibson L, Horne JH et al (2009) Expert image analysts show enhanced visual processing in change detection. Psychon Bull Rev 16:390–397. doi:10.3758/PBR.16.2.390
38. Sowden PT, Rose D, Davies IRL (2002) Perceptual learning of luminance contrast detection: specific for spatial frequency and retinal location but not orientation. Vision Res 42:1249–1258. doi:10.1016/S0042-6989(02)00019-6
39. Neisser U (1976) Cognition and reality: principles and implications of cognitive psychology. p 230.
40. Dinh HQ, Walker N, Hodges LF, Kobayashi A (1999) Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. In: Proc. IEEE virtual real (Cat. No. 99CB36316). IEEE Comput. Soc., pp 222–228
41. Hecht D, Reiner M, Halevy G (2006) Multimodal virtual environments: response times, attention, and presence. Presence-Teleop Virtual Environ 15:515–523
42. Coen M (2001) Multimodal integration-a biological view. In: Proc. Fifteenth Int. Jt. Conf. Artif. Intell. Seattle, WA, pp 1417–1424
43. Shimojo S, Shams L (2001) Sensory modalities are not separate modalities: plasticity and interactions. Curr Opin Neurobiol 11:505–509
44. Mcgurk H, Macdonald J (1976) Hearing lips and seeing voices. Nature 264:746–748
45. Rock I, Victor J (1964) Vision and touch: an experimentally created conflict between the two senses. Science 143:594–596
46. Scheier C, Nijhawan R, Shimojo S (1999) Sound alters visual temporal resolution. Invest Ophthalmol Vis Sci 40:4169
47. Chandrasekaran C, Ghazanfar AA (2011) When what you see is not what you hear. Nat Neurosci 14:675–676. doi:10.1038/nn.2843
48. Skalski P, Whitbred R (2010) Image versus sound: a comparison of formal feature effects on presence and video game enjoyment. Psychology J 8:67–84, doi: Article
49. Biocca F, Kim J, Choi Y (2001) Visual touch in virtual environments: an exploratory study of presence, multimodal interfaces, and cross-modal sensory illusions. Presence-Teleop Virtual Environ 10:247–265
50. Basdogan C, Ho C, Srinivasan MA, Slater MEL (2001) An experimental study on the role of touch in shared virtual environments. ACM Trans Comput Interact 7:443–460
51. Welch RB, Warren DH (1980) Immediate perceptual response to intersensory discrepancy. Psychol Bull 88:638–667

52. De Kort YAW, Ijsselsteijn WA, Kooijman J, Schuurmans Y (2003) Virtual laboratories: comparability of real and virtual environments for environmental psychology. Presence-Teleop Virtual Environ 12:360–373
53. Bowman DA, McMahan RP (2007) Virtual reality: how much immersion is enough? Computer (Long Beach Calif) 40:36–43. doi:10.1109/MC.2007.257
54. Slater M (2009) Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. Philos Trans R Soc Lond B Biol Sci 364:3549–3557. doi:10.1098/rstb.2009.0138
55. Slater M (2003) A note on presence terminology. Presence-Connect, pp 1–5
56. Steuer J (1992) Defining virtual reality: dimensions determining telepresence. J Commun 42:73–93. doi:10.1111/j.1460-2466.1992.tb00812.x
57. Ijsselsteijn W, de Ridder H, Freeman J et al (2001) Effects of stereoscopic presentation, image motion, and screen size on subjective and objective corroborative measures of presence. Presence-Teleop Virtual Environ 10:298–311
58. Hands DS (2004) A basic multimedia quality model. IEEE Trans Multimedia 6:806–816. doi:10.1109/TMM.2004.837233
59. Hwang J, Kim GJ (2010) Provision and maintenance of presence and immersion in hand-held virtual reality through motion based interaction. Comput Animat Virtual Worlds 21:547–559
60. Hendrix C, Barfield W (1995) Presence in virtual environments as a function of visual and auditory cues. In: Proceedings of Virtual Real Annu Int Symp. pp 74–82. doi: 10.1109/VRAIS.1995.512482
61. Lin JJ-W, Duh HBL, Parker DE et al (2002) Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment. In: Proc. IEEE Virtual Real. IEEE Comput. Soc., pp 164–171
62. Freeman J, Avons SE, Pearson DE, IJsselsteijn WA (1999) Effects of sensory information and prior experience on direct subjective ratings of presence. Presence-Teleop Virtual Environ 8:1–13. doi:10.1162/105474699566017
63. Bleumers L, Lievens B, Pierson J (2011) From sensory dream to television format: gathering user feedback on the use and experience of omnidirectional video-based solutions. In: ISPR 2011 Int. Soc. PRESENCE Res. Annu. Conf.
64. Slater M, Khanna P, Mortensen J, Yu I (2009) Visual realism enhances realistic response in an immersive virtual environment. IEEE Comput Graph Appl 29:76–84
65. Welch RB, Blackmon TT, Liu A et al (1996) The effects of pictorial realism, delay of visual feedback, and observer interactivity an the subjective sense of presence. Presence-Teleop Virtual Environ 5:263–273
66. Barfield W, Hendrix C (1995) The effect of update rate on the sense of presence within virtual environments. Virtual Real 1:3–15. doi:10.1007/BF02009709
67. Meehan M (2001) Physiological reaction as an objective measure of presence in virtual environments. University of North Carolina at Chapel Hill
68. Luque FP, Galloso I, Feijoo C, Martín CA, Cisneros G (2014) Integration of multisensorial stimuli and multimodal interaction in a hybrid 3DTV system. ACM Trans Multimedia Comput Commun Appl 11(1s):16:1–16:22. doi:10.1145/2617992
69. Beerends JG, De Caluwe FE (1999) The influence of video quality on perceived audio quality and vice versa. J Audio Eng Soc 47:355–362
70. Slater M, Usoh M, Steed A (1995) Taking steps: the influence of a walking technique on presence in virtual reality. ACM Trans Comput Interact 2:201–219. doi:10.1145/210079.210084
71. De Ruyter B, Aarts E (2004) Ambient intelligence: visualizing the future. In: Proc. Work. Conf. Adv. Vis. interfaces – AVI'04. ACM Press, New York, p 203
72. Waltl M, Timmerer C, Hellwagner H (2010) Increasing the user experience of multimedia presentations with sensory effects. In: 11th Int. Work Image Anal. Multimed. Interact. Serv. (WIAMIS). IEEE, Desenzano del Garda, pp 1–4

73. Ademoye OA, Ghinea G (2009) Synchronization of olfaction-enhanced multimedia. IEEE Trans Multimed 11:561–565. doi:10.1109/TMM.2009.2012927
74. Murray N, Qiao Y, Lee B et al (2013) Subjective evaluation of olfactory and visual media synchronization. In: Proc. 4th ACM Multimed. Syst. Conf. ACM, New York, pp 162–171
75. Ghinea G, Ademoye O (2012) The sweet smell of success: enhancing multimedia applications with olfaction. ACM Trans Multimed Comput Commun Appl 8:1–17. doi:10.1145/2071396.2071398
76. Pallavicini F, Cipresso P, Raspelli S et al (2013) Is virtual reality always an effective stressors for exposure treatments? Some insights from a controlled trial. BMC Psychiatry 13:1–10
77. (1990) ITU-T J.100 Recommendation. Tolerance for transmission time differences between vision and sound components of a television signal. International Telecommunication Union (ITU) – Telecommunication sector.
78. Slutsky DA, Recanzone GH (2001) Temporal and spatial dependency of the ventriloquism effect. Neuroreport 12:7–10
79. IJsselsteijn WA, de Ridder H, Vliegen J (2000) Effects of stereoscopic filming parameters and display duration on the subjective assessment of eye strain. In: Proc. SPIE 3957, Stereosc. Displays Virtual Real. Syst. VII. pp 12–22
80. Kooi FL, Toet A (2004) Visual comfort of binocular and 3D displays. Displays 25:99–108. doi:10.1016/j.displa.2004.07.004
81. Meesters L, IJsselsteijn W (2003) Survey of perceptual quality issues in threedimensional television systems. Proc. SPIE
82. Banos RM, Botella C, Alcaniz M et al (2004) Immersion and emotion: their impact on the sense of presence. Cyberpsychology Behav 7:734–741
83. Timmerer C, Waltl M, Rainer B, Hellwagner H (2012) Assessing the quality of sensory experience for multimedia presentations. Signal Process Image Commun 27:909–916
84. ISO/IEC 23005-3 (2013) Information technology – Media context and control – Part 3: Sensory information. p 104
85. Waltl M, Rainer B, Timmerer C, Hellwagner H (2013) An end-to-end tool chain for Sensory Experience based on MPEG-V. Signal Process Image Commun 28:136–150
86. Yoon K (2013) End-to-end framework for 4-D broadcasting based on MPEG-V standard. Signal Process Image Commun 28:127–135
87. Kim J, Lee C-G, Kim Y, Ryu J (2013) Construction of a haptic-enabled broadcasting system based on the MPEG-V standard. Signal Process Image Commun 28:151–161

Springer
http://www.springer.com/978-1-4939-2025-9

Novel 3D Media Technologies
Kondoz, A.; Dagiuklas, T. (Eds.)
2015, IX, 249 p. 122 illus., 100 illus. in color., Hardcover
ISBN: 978-1-4939-2025-9