Towards a Reference Ontology on Mulsemedia Systems

Estêvão B. Saleme*
Celso A. S. Santos
Ricardo A. Falbo
estevaobissoli@gmail.com
saibel@inf.ufes.br
falbo@inf.ufes.br
Federal University of Espírito Santo
Vitória, ES, Brazil

Gheorghita Ghinea
Brunel University London
London, UK
george.ghinea@brunel.ac.uk

Frederic Andres
National Institute of Informatics
Tokyo, Japan
andres@nii.ac.jp

ABSTRACT

The use of multiple senses in interactive applications has become increasingly feasible due to the upsurge of commercial, off-the-shelf devices to produce sensory effects. Creating Multiple Sensorial Media (MulSeMedia) immersive systems requires understanding their digital ecosystem. Mulsemedia systems encompass a set of applications, and devices of different types assembled to communicate or express feelings from the virtual world to the real world. Despite existing standards, tools, and recent research devoted to them, there is still a lack of formal and explicit representation of what mulsemedia is. Misconceptions could eventually lead to the construction of solutions that might not take into account reuse, integration, standardization, among other design features. In this paper, we propose to establish a common conceptualization about mulsemedia systems through a reference ontology, named MulseOnto, covering their main notions. To evaluate it, we applied ontology verification and validation techniques, including assessment by humans and a data-driven approach. The results showed that MulseOnto can be used as a consensual conceptual model for exploring the knowledge about the whole chain of mulsemedia systems.

CCS CONCEPTS

• **General and reference** → Reference works; • **Information systems** → *Multimedia information systems*; • **Human-centered computing** → *Ubiquitous and mobile computing theory, concepts and paradigms*;

KEYWORDS

Mulsemedia Systems, Reference Ontology, Multimedia

---

*The author is also an Academic Visitor at Brunel University London.

---

1 INTRODUCTION

Continuous advances in technology have enabled the use of mulsemedia to build more immersive experiences for the users. Not only do they engage the senses of sight and hearing in this kind of experience, they can also have involvement with artificially produced sensory effects such as scent, vibration, and flavor [8, 9]. However, building a mulsemedia system entails weaving multiple technologies to connect different entities, distribute the sensory signals, and render sensory effects appropriately such as described in [16, 18, 19, 28–30, 34, 35].

The concept of mulsemedia along with the interactions and inter-relationships between applications, types of sensory effects, and devices in this domain are indeed complex to understand and frequently not so clear-cut because of its heterogeneous digital ecosystem. Despite existing standards, tools, and recent research devoted to mulsemedia, there is still a lack of formal and explicit representation of what mulsemedia is, which may cause a failure to understand it adequately. Misconception might eventually lead to a weak arrangement of how to manage and integrate a plethora of entities. As a result, stakeholders may overlook relevant design aspects such as standardization [36], reuse, and compatibility, to name a few among other design features.

In light of this, capturing the common conceptualization underlying this type of systems is rather important to understand and improve the mulsemedia digital ecosystem by integrating information from varied sources, reducing ambiguity and inaccuracy when interpreting shared information. One way to represent a shared conceptualization is through the use of ontologies. They bring together a shared understanding about a domain that can be communicated between humans and also computers [11]. Ideally, first the structure of the domain conceptualization must be made available to humans through an explicit and formal description of the corresponding portion of reality in terms of a *domain reference ontology*. A reference ontology should be constructed with the sole objective of making the best possible description of the domain in reality, for the purposes of communication, learning and problem-solving. Once a reference ontology is built, an operational version of it (said an *operational ontology*) can be designed and implemented to be processed by machines.

In this paper, we present a reference domain ontology on Mulsemedia Systems, named MulseOnto. It chiefly encompasses the understanding of what mulsemedia systems and their small parts are, which entities they interact with, where the sensory effects come from, and what physical realizations different kinds of media do in this context, therefore providing a big picture of mulsemedia systems. By building this ontology, we propose to establish a common conceptualization about mulsemedia systems to address purposes such as understanding the interrelationship between their entities, describing a common vocabulary for knowledge workers in this domain, certifying that stakeholders are talking over the same concepts when making mulsemedia solutions, and fostering this 'novel' field. To evaluate MulseOnto, we have applied an approach for verifying the concepts and relations, including assessment by humans and a data-driven approach. Additionally, we have done an instantiation of the ontology to find out whether it is able to describe real-world situations. The results showed that this ontology can be used as a consensual conceptual model for exploring the knowledge about the whole chain of mulsemedia systems.

This work is organized as follows. Section 2 brings the main concepts on mulsemedia used to support the ontology. Section 3 presents MulseOnto including the questions it aims to answer and the relationship between its entities. Section 4 discusses how the ontology was evaluated. Section 5 presents related work. Finally, Section 6 concludes the paper and leads to future works.

2 BACKGROUND

Ghinea et al. [9] stated that multimedia applications are usually composed of more than two different media and almost exclusively bisensorial (sight and hearing) by nature, whereas mulsemedia refers to applications that also engages other three primary senses (smell, taste, and touch). Hence, multimedia and also mulsemedia refer to the use of multiple media, but it is necessary to take into account what the meaning of media is. The problem with terminology in mulsemedia starts with the misunderstanding of the concept of media due to the use of this term with different meanings in different contexts.

Heller and Martin [14] proposed a media taxonomy consisting of two dimensions. The first one is related to the well-known existing media types (text, graphics, sound, and motion) and the second one, to the continuum between concrete and abstract expression of each medium. The authors consider that media types can only provide information in aural or visual forms.

Roy and Zeng [26] are particularly interested in multimedia content shared within a social network. They acknowledged that the term media is often used in a broad sense and usually related to the nature of information perceived by humans. Indeed, the authors consider that, in multimedia communication, media are delivered through presentation spaces (e.g. screen, speakers, projector) with presentation attributes (e.g. color, intensity, font) to address the primary human senses. A presentation space includes specialized output devices and can have one or more dimensions (e.g. a computer monitor has two space dimensions, while holography has three ones).

Regarding the time dimension of its presentation space, a medium is classified as discrete or continuous [31]. The first one is composed of time-independent information units, whereas the second one requires a continuous play-out of its information units in time. The time-dependency between the information units of a continuous medium establishes the semantics of this kind of content.

Starting from a conventional view, Roy and Zeng [26] initially defined multimedia as "the use of a variety of communicative media (i.e., information intended for human consumption), including text, audio, visual, and haptic data." Authors also stated that multimedia integrates signals from our primary senses in an attempt to generate a coherent perceptual experience for the users. Afterwards and agreeing with Chang [4], the authors claim that the definition of multimedia has been extended to be much broader and more inclusive, covering a wide spectrum of multimedia applications, going far beyond the conventional audiovisual contents to, for instance, mulsemedia applications.

In an attempt to address the issue of creating a unified terminology, Bordegoni et al. [2] brought not only one but also two definitions of the medium term. The first definition considers a medium as being a certain physical space in which perceptible entities are realized, concentrating on human sensory receptors and derived sensations. As there are different types of perceptible entities (visual, auditory, haptic, gustatory, and olfactory), one may use these terms for making a distinction between media too. In the second definition, medium designates a certain type of information and/or the representation format in which information is stored.

Using the two definitions of medium proposed by Bordegoni et al. [2], mulsemedia could be characterized as: (i) a common physical space in which different perceptible entities can be realized, being two of them, at least, visual and audible entities and the others, haptic, gustative or olfactive ones; or (ii) a composition of basic media types with one or more haptic, gustative or olfactive sensory effects. This composition is characterized by additional properties, such as temporal relationships between the involved media and sensory effect, the expected presentation's behavior, user preferences, etc.

Media content can be directly created by humans (e.g. this paper, pictures, songs), acquired through various sensors (e.g. camera, microphone, motion capture) that capture real-world information, or synthesized using computers (e.g. a virtual 3D space in a game). The sensory effects are usually synthesized. Though haptic effects can be captured, recording taste and smell from the real world remains a great challenge [5].

After media content and sensory effects data have been created, they are usually coded for transmission or storage generally using a standard (e.g. MPEG-4 and AV1 for videos, and MPEG-V [36] or recently NCL [17] for sensory effects). Although the mulsemedia ecosystem should take into consideration issues involving all of these phases, in this paper, we are concentrated only on authoring and presentation phases.

During the so-called authoring phase, media, sensory effects, and additional presentation properties are linked together in a mulsemedia composition. In this phase, if the author is a human, something abstract that exists only in the author's mind must be transformed in a mulsemedia content; if the author is a computer system, an explicit representation of the information content (e.g. video from surveillance cameras, automatic caption from automatic speech recognition, etc.) must be provided.

Mulsemedia contents are represented in a format that is not directly presentable to humans. Hence, for each medium and sensory effects, it must always exist a dedicated physical device that is able to produce perceptible entities. In other words, devices have a function of rendering each mulsemedia content to the human senses. Moreover, the authored content establishes an expected behavior of the presentation through a set of presentation commands and/or constraints, which usually include references to mulsemedia data together with the purposes of engagement, immersion, flow assessment, and so on. Both constraints and commands are assumed to be defined outside the presentation system.

Devices for delivering visual, auditory, and haptic experiences have reached a fairly advanced stage of maturity whereas olfactory and gustatory are relatively newcomers. In this fashion, it is worth underlining the work of Hariri et al. [13], who aimed at stimulating the sense of smell by using electric devices instead of scent diffusers widely used in olfactory [20], and the works of Ranasinghe and Do [24], Ranasinghe et al. [25], and Vi et al. [33], who devised digital interfaces for introducing the sense of taste to digital systems.

Finally, the presentation system is responsible for presenting the mulsemedia content to the user, considering the presentation properties defined in the authoring phase. As a result, this component of the mulsemedia ecosystem must coordinate the integration of outputs emanating from the media and sensory-effect specific devices in all presentation dimensions considered. Moreover, resource limitations, user preferences, and other presentation features can also be taken into account during the mulsemedia content presentation.

The concepts related to mulsemedia digital ecosystems that have been addressed in this section form the basis of our proposed reference ontology.

3 MULSEMEDIA SYSTEMS REFERENCE ONTOLOGY

Our reference ontology on Mulsemedia Systems, called MulseOnto, predominantly incorporates the comprehension of what mulsemedia systems and their small parts are, which entities they interact with, where the sensory effects come from, and what physical realizations different kinds of media do in this context, thus providing a big picture of mulsemedia systems. It does not intend to answer how they affect the Quality of Experience (QoE) of users, how stimuli are perceived by users, how devices produce each sensory effect, how media are loaded, composed and presented in mulsemedia systems, how mulsemedia systems manage timing aspects, how users set up systems and devices within an interactive environment. Moreover, attribute details such as computer media extensions, encoding methods for computer medium, kinds of flavors or scent, the intensity of effects, temporality constraints, colors, and/or in-depth particulars do not take place in this ontology.

To develop MulseOnto, we adopted SABiO, a Systematic Approach for Building Ontologies [7]. We chose SABiO because it has been successfully used to develop domain ontologies, in particular, Software Engineering reference domain ontologies. SABiO's development process comprises five main phases, namely: (i) Purpose Identification and Requirements Elicitation; (ii) Ontology Capture and Formalization; (iii) Design; (iv) Implementation; and (v) Test. These phases are accompanied by supporting processes, such

as knowledge acquisition, reuse, documentation, and evaluation. SABiO aims at developing both reference ontologies (phases i and ii) and operational ontologies (phases iii, iv and v). In this work, we are interested in building only a domain reference ontology, thus we performed only the first two phases, whose the main results produced are discussed next, followed by how we have evaluated MulseOnto.

3.1 MulseOnto Requirements

Analogously to requirements in Requirements Engineering, ontology requirements can be functional and non-functional. Functional requirements refer to the ontology content and can be specified as Competency Questions (CQs). CQs are questions that the ontology should be able to answer [10]. They help to refine the scope of the ontology and are used in the ontology verification process to check whether the ontology elements (concepts, relations, and properties) are able and sufficient to answer the CQs [7]. MulseOnto should be able to answer the following CQs:

- **CQ01**. What is a mulsemedia computer system?
- **CQ02**. What is a mulsemedia software system?
- **CQ03**. What do mulsemedia programs handle?
- **CQ04**. What is a user interface in a mulsemedia computer system?
- **CQ05**. What are the most common classes of output devices in mulsemedia computer systems?
- **CQ06**. What is a medium?
- **CQ07**. What are the most common media classes to be realized in mulsemedia computer systems?
- **CQ08**. What are the most common classes of media in mulsemedia computer systems?
- **CQ09**. What is a sensory effect description?
- **CQ10**. What are the most common classes of sensory effect descriptions in mulsemedia computer systems?

Ontology non-functional requirements, in turn, refer to the characteristics, qualities, and general aspects not related to the ontology content [32]. The following non-functional requirements were defined for MulseOnto:

- **NFR01**. MulseOnto should be integrated to the Software Engineering Ontology Network (SEON) [27], reusing its parts that cover aspects related to the mulsemedia domain.
- **NFR02**. MulseOnto should be developed in a modular way, allowing addressing this complex domain in an iterative fashion, focusing on different concerns.

Concerning **NFR01**, SEON is an ontology network for the Software Engineering domain, which is organized in layers [27]. At its uppermost layer, there is the Unified Foundational Ontology (UFO). UFO is a foundational ontology that is based on a number of theories from Formal Ontology, Philosophical Logics, Philosophy of Language, Linguistics and Cognitive Psychology [12]. At the core layer, there are core ontologies on software (Software Ontology - SwO [6]) and software process (Software Process Ontology - SPO [3]). MulseOnto is a domain ontology included in SEON's domain layer, by extending both SwO and SPO.
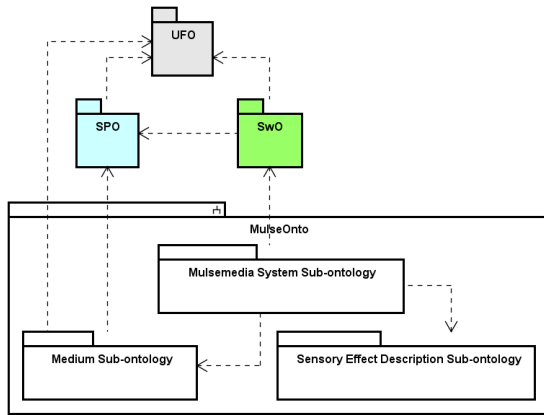
Figure 1: MulseOnto Sub-ontologies.

For addressing **NFR02**, the current version of MulseOnto is divided in three sub-ontologies, namely: (i) Mulsemedia System sub-ontology, which focuses on what a mulsemedia system is, addressing competency questions CQ01 to CQ05; (ii) Medium sub-ontology, which refers to types of media used in mulsemedia systems (CQ06 to CQ08); and (iii) Sensory Effect Description sub-ontology, which regards the types of sensory effects descriptions used in mulsemedia systems (CQ09 and CQ10). Figure 1 presents a UML package diagram, showing the sub-ontologies that comprise MulseOnto and the relationships between them. The dependencies between the sub-ontologies indicate that concepts and relations from a sub-ontology are used by the dependent sub-ontology.

In the following subsections, MulseOnto sub-ontologies are presented. Concepts reused from SPO are shown in blue, preceded by its acronym (SPO::); concepts from SwO are shown in green, preceded by its acronym (SwO::); concepts from UFO are shown in gray, preceded by its acronym (UFO::).
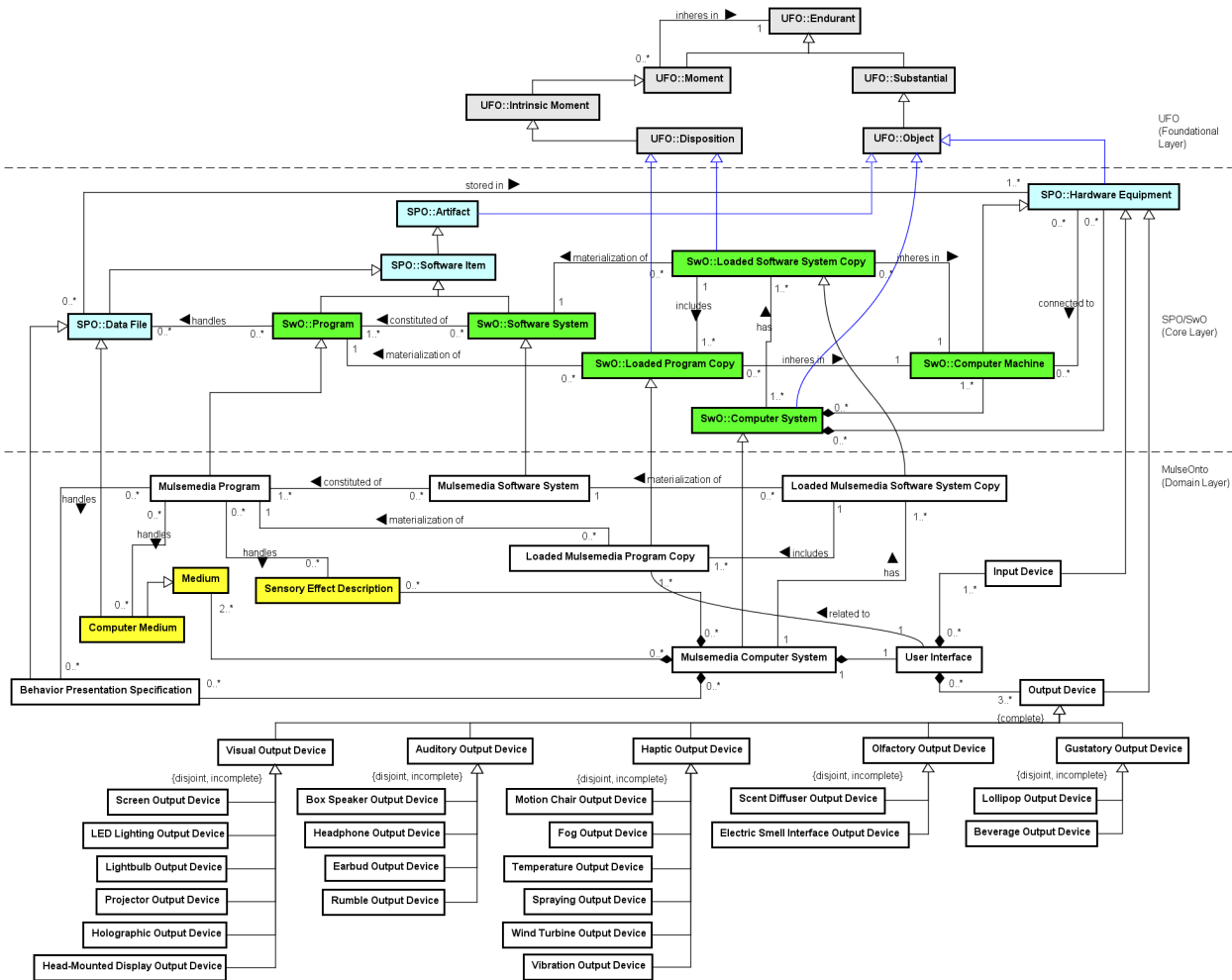

Figure 2: Mulsemedia System Sub-ontology. The yellow and white concepts represent the domain of interest. Above are concepts in which this ontology is grounded.

3.2 Mulsemedia System Sub-ontology

Figure 2 presents the conceptual model of the Mulsemedia System sub-ontology. As discussed above, MulseOnto extends SEON's ontologies from the foundational layer (UFO) and from the core layer (SPO and SwO). The terms used in MulseOnto were derived from Section 2. Concepts from the other two sub-ontologies of MulseOnto are shown in yellow, highlighting the dependency relationships between its sub-ontologies.

The Mulsemedia System sub-ontology is strongly inspired by the Software Ontology (SwO) (green concepts in Figure 2). In SwO, a `Computer System` is a system combining hardware and software. Regarding hardware, a `Computer System` is composed of `Computer Machines` (hardware equipment with processing capacity) and other `Hardware Equipment` connected to them. Regarding software, a `Computer System` has `Loaded Software System Copies` that are installed/loaded in a `Computer Machine` that is part of the `Computer System`. A `Loaded Software System Copy`, in turn, is the materialization of a `Software System`. `Software Systems` are constituted of `Programs` that are also materialized as `Loaded Program Copies` inhering in a `Computer Machine`. `Programs` handle `Data Files` [6].

A `Mulsemedia Computer System` is a `Computer System` that is composed of least two `Medium` and a `User Interface` that is composed of at least one `Input Device` and at least three `Output Devices`. These `Output Devices` should be of different types. Regarding software, a `Mulsemedia Computer System` has at least one `Loaded Mulsemedia Software System Copy` installed/loaded in a `Computer Machine` that is part of the `Mulsemedia Computer System`. `Behavior Presentation Specifications` and `Sensory Effect Descriptions` may also be part of a `Mulsemedia Computer System`, but they are not mandatory, since they can be specified and described directly via code in a `Mulsemedia Program`.

`Loaded Mulsemedia Software System Copy` includes at least a `Loaded Mulsemedia Program Copy`, which is a materialization of a `Mulsemedia Program`. A `Mulsemedia Software System` is constituted of at least one `Mulsemedia Program`, that can handle `Behavior Presentation Specifications`, `Computer Medium`, and `Sensory Effect Descriptions`.

As said before, the `User Interface` of a `Mulsemedia Computer System` should include at least three `Output Devices` of different types among the following: `Visual Output Device`, `Auditory Output Device`, `Haptic Output Device`, `Olfactory Output Device`, and `Gustatory Output Device`, being the first and the second required. MulseOnto depicts some sub-classes of `Output Devices` with the constraint *incomplete*, meaning that other classes of output devices can be considered.

3.3 Medium Sub-ontology

Figure 3 presents the conceptual model of the Medium Sub-ontology. The main concept of this sub-ontology is `Medium`, which is strongly related to mulsemedia systems. Media are a mandatory part of the whole digital ecosystem that permeates this domain.

According to its origin, `Medium` can be classified into `Directly Transmitted Medium`, which is captured by a `Hardware Equipment` or `Computer Medium`, which is stored in a `Hardware Equipment`. `Medium` can also be classified according to its temporal nature
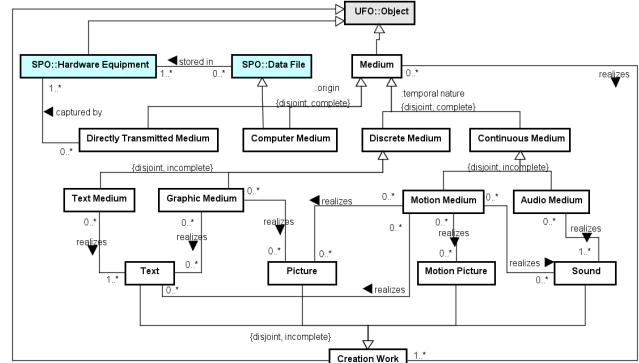

Figure 3: Medium Sub-ontology in the context of mulsemedia systems.

into `Discrete Medium` and `Continuous Medium`. `Text Medium` and `Graphic Medium` are subtypes of `Discrete Medium`, whereas `Motion Medium` and `Audio Medium` are subtypes of `Continuous Medium`. `Medium` realizes at least one `Creation Work`, which can be a `Text`, a `Picture`, a `Motion Picture`, or a `Sound`. The Information Realization Content Pattern was reused from the work of Presutti et al. [23] to represent the difference between abstract and realized (manifested, concrete, etc.) information.

It is worth noting that these specialization criteria (origin and temporal nature) are orthogonal and, thus, they can be combined giving rise to types such as `Directly Transmitted Motion Medium` and `Computer Picture Medium`.

3.4 Sensory Effect Description Sub-ontology

Figure 4 shows a taxonomy of Sensory Effect Descriptions. These concepts are pertinent to this domain because they draw attention to the kinds of sensory effects that can exist in mulsemedia systems.
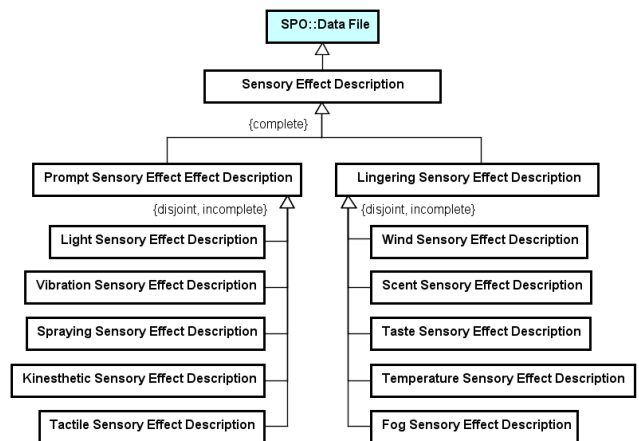

Figure 4: Sensory Effect Description Sub-ontology in the context of mulsemedia systems.

`Sensory Effect Description` is a `Data File` with the description of how the sensory effects will be presented. A significant distinction here is provided by `Prompt Sensory Effect Description` and `Lingering Sensory Effect Description`. The former is related to sensory effects that must be delivered as soon as possible, owing to the fact that delays are more perceptible [30]. The latter is characterized by lingering effects, e.g. the tendency of aromas to linger in the atmosphere and of tastes to linger on the tongue. `Prompt Sensory Effect Description` involves descriptions of light, vibration, spraying, kinesthetic, and tactile sensory effects, whilst `Lingering Sensory Effect Description` encompasses wind, scent, taste, and temperature. Both are constrained by *incomplete*, meaning that other classes can coexist in the model.

4 EVALUATION

In order to evaluate MulseOnto, we performed Ontology Verification & Validation (V&V) activities. First, a verification activity was performed by means of expert judgment to analyze whether concepts and relations defined in the ontology are able to answer their competency questions. Next, a data-driven approach was performed by means of instantiating its concepts and relations from a hypothetical scenario to find out whether it fits its proposal.

4.1 Competency Questions Verification

For verifying MulseOnto, we analyzed if its concepts and relations are able to answer the competency questions. Furthermore, by performing this step, it is also possible to identify whether there are elements besides the point, particularly, those that do not play a part in answering the questions. Table 1 shows which elements of the ontology account for each competency question.

4.2 Ontology Instantiation

In accordance with SABiO [7], we validated whether MulseOnto is capable of representing a real-world case. Given this, we instantiated our ontology using data from the following hypothetical mulsemedia scenario.

Hypothetical Mulsemedia Scenario
A recently created science museum in Tokyo will provide a multi-sensory experience to its visitors. It is elaborating an astonishing exhibition on exploring the earth through places like volcanoes, glaciers, earthquakes, aurora borealis, waterfalls, and locations with exotic food. The goal is that visitors not only see videos and pictures of those places, but also feel immersed in each environment they are experiencing. To do so, the museum will be using a mulsemedia platform called PlaySEM [28], which has a video player and a renderer that support sensory effects. The video-clips are annotated with sensory effect descriptions in MPEG-V (with the help of SEVino [34]), that include light, vibration, spraying, kinesthetic, tactile, wind, scent, taste, fog, and temperature effects. Visual media will be displayed on screens, projectors, head-mounted displays. Some visual media objects will also be read from a behavior presentation specification and be presented in holographic devices. LED strips and lightbulbs will come up with lighting effects. Eventually, a live event from a real place can be directly transmitted to the museum. The behavior presentation specification will cater to mulsemedia program with audio media

Table 1: Competency questions verification.

CQs	Concepts, relations and properties
CQ01	Mulsemedia Computer System *subtype of* Computer System; Mulsemedia Computer System *composed of* Medium; Mulsemedia Computer System *composed of* Sensory Effect Description; Mulsemedia Computer System *composed of* Behavior Presentation Specification; Mulsemedia Computer System *composed of* User Interface; Mulsemedia Computer System *has* Loaded Mulsemedia Software System Copy; Loaded Mulsemedia Software System Copy *materialization of* Mulsemedia Software System.
CQ02	Mulsemedia Software System *subtype of* Software System; Mulsemedia Software System *constituted of* Mulsemedia Program.
CQ03	Mulsemedia Program *handles* Computer Medium; Mulsemedia Program *handles* Sensory Effect Description; Mulsemedia Program *handles* Behavior Presentation Specification.
CQ04	User Interface *composed of* Input Device; User Interface *composed of* Output Device; User Interface *related to* Loaded Mulsemedia Program Copy; Loaded Mulsemedia Program Copy *materialization of* Mulsemedia Program; Loaded Mulsemedia Software System Copy *includes* Loaded Mulsemedia Program Copy.
CQ05	Output Device *subtype of* Hardware Equipment; Visual Output Device *subtype of* Output Device; Auditory Output Device *subtype of* Output Device; Haptic Output Device *subtype of* Output Device; Olfactory Output Device *subtype of* Output Device; Gustatory Output Device *subtype of* Output Device; Screen Output Device *subtype of* Visual Output Device; LED Lighting Output Device *subtype of* Visual Output Device; Lightbulb Output Device *subtype of* Visual Output Device; Projector Output Device *subtype of* Visual Output Device; Holographic Output Device *subtype of* Visual Output Device; Head-Mounted Display Output Device *subtype of* Visual Output Device; Box Speaker Output Device *subtype of* Auditory Output Device; Headphone Output Device *subtype of* Auditory Output Device; Earbud Output Device *subtype of* Auditory Output Device; Rumble Output Device *subtype of* Auditory Output Device; Motion Chair Output Device *subtype of* Haptic Output Device; Temperature Output Device *subtype of* Haptic Output Device; Spraying Output Device *subtype of* Haptic Output Device; Wind Turbine Output Device *subtype of* Haptic Output Device; Vibration Output Device *subtype of* Haptic Output Device; Scent Diffuser Output Device *subtype of* Olfactory Output Device; Electric Smell Interface Output Device *subtype of* Olfactory Output Device; Lollipop Output Device *subtype of* Gustatory Output Device; Beverage Output Device *subtype of* Gustatory Output Device.
CQ06	Medium *subtype of* Object; Medium *realizes* Creation Work.
CQ07	Text Medium *realizes* Text; Graphic Medium *realizes* Text; Motion Medium *realizes* Text; Text *subtype of* Creation Work; Graphic Medium *realizes* Picture; Motion Medium *realizes* Picture; Picture *subtype of* Creation Work; Motion Medium *subtype of* Motion Picture; Motion Picture *subtype of* Creation Work; Audio Medium *realizes* Sound; Motion Medium *realizes* Sound; Sound *subtype of* Creation Work.
CQ08	Computer Medium *subtype of* Medium; Computer Medium *stored in* Hardware Equipment; Directly Transmitted Medium *subtype of* Medium; Directly Transmitted Medium *captured by* Hardware Equipment; Discrete Medium *subtype of* Computer Medium; Continuous Medium *subtype of* Computer Medium.
CQ09	Sensory Effect Description *subtype of* Data File; Prompt Sensory Effect Description *subtype of* Sensory Effect Description; Lingering Sensory Effect Description *subtype of* Sensory Effect Description.
CQ10	Light Sensory Effect Description *subtype of* Prompt Sensory Effect Description; Vibration Sensory Effect Description *subtype of* Prompt Sensory Effect Description; Spraying Sensory Effect Description *subtype of* Prompt Sensory Effect Description; Kinesthetic Sensory Effect Description *subtype of* Prompt Sensory Effect Description; Tactile Sensory Effect Description *subtype of* Prompt Sensory Effect Description; Wind Sensory Effect Description *subtype of* Lingering Sensory Effect Description; Scent Sensory Effect Description *subtype of* Lingering Sensory Effect Description; Taste Sensory Effect Description *subtype of* Lingering Sensory Effect Description; Temperature Sensory Effect Description *subtype of* Lingering Sensory Effect Description; Fog Sensory Effect Description *subtype of* Lingering Sensory Effect Description.

sources. To provide auditory experiences, the museum will arrange box speakers, headphones, earbuds, and rumble devices. To deliver haptic experiences, there will be a motion chair, local air conditioners, spray machines, wind turbines and vibration vests. Olfactory and gustatory effects will be conveyed in two ways. Scent diffusers and electric smell devices to feel the smell of the object media presented to the visitors. A digital lollipop and a beverage electronic device will be set for taste.

Table 2: Ontology instantiation.

Concept	Instances
Mulsemedia Computer System	The mulsemedia computer system of the science museum with software systems, devices, etc.
Mulsemedia Software System	PlaySEM platform software constituted of mulsemedia programs.
Loaded Mulsemedia Software System Copy	The local copy of the PlaySEM platform software system set up in the museum.
Mulsemedia Program	PlaySEM SE Video Player; PlaySEM Sensory Effects Renderer; SEVino.
Loaded Mulsemedia Program Copy	The local copy of the PlaySEM Video Player set up in the museum.
Behavior Presentation Specification	The XML file with the source to the video of Vesuvius volcano.
Screen Output Device	LG 49UJ630V 49 inch 4K Ultra HD HDR Smart LED TV.
LED Lighting Output Device	Individually Addressable LED Strip Light 5050 RGB.
Lightbulb Output Device	Foval smart light Alexa LED Bulb Light with Wi-Fi.
Projector Output Device	Video Projector, MEER 1600 Lumens 130" Wide Screen LED Portable Projector with Built-in Speaker.
Holographic Output Device	Microsoft HoloLens Development Edition Glass 3D Holographic.
Head-Mounted Display Output Device	HTC Vive Pro VR Virtual Reality Headset.
Box Speaker Output Device	Media-Tech Boombox Stereo BT 15 W Black Rectangle.
Headphone Output Device	Sony MDRZX310 Foldable Headphones.
Earbud Output Device	Ultimate Zero-Compromise Total-Wireless Earphones.
Rumble Output Device	Fender Rumble 25 Bass Combo Amp V3.
Motion Chair Output Device	Simuline 4DOF 4D motion chair.
Temperature Output Device	Dyson AM09 Hot + Cool Fan Heater.
Spraying Output Device	Theater Concert Flame Fire Jet Device 3d 3meters.
Wind Turbine Output Device	CITC Hurricane II Wind machine.
Vibration Output Device	KOR-FX gaming vest 4DFX Haptic Feedback System.
Scent Diffuser Output Device	Vortex Activ scent emitter.
Electric Smell Interface Output Device	Digital Smell Interface.
Lollipop Output Device	Digital Lollipop tongue interface.
Beverage Output Device	Vocktail system.
Light Sensory Effect Description	The MPEG-V xml file describing the light effects of the aurora borealis in Iceland.
Vibration Sensory Effect Description	The MPEG-V xml file describing a sequence of vibrations effects of the eruption of Vesuvius volcano.
Spraying Sensory Effect Description	The MPEG-V xml file describing the spraying effects of the eruption of Vesuvius volcano.
Kinesthetic Sensory Effect Description	The MPEG-V xml file describing the kinesthetic effects of an earthquake in Japan.
Tactile Sensory Effect Description	The MPEG-V xml file describing the tactile effects of an earthquake in Japan directly onto the skin of the visitor.
Wind Sensory Effect Description	The MPEG-V xml file describing the wind effects near Iguazu Falls.
Scent Sensory Effect Description	The MPEG-V xml file describing the scent of the durian in Thailand.
Taste Sensory Effect Description	The MPEG-V xml file describing the taste of the durian in Thailand.
Temperature Sensory Effect Description	The MPEG-V xml file describing the temperature effects of the eruption of Vesuvius volcano.
Fog Sensory Effect Description	The MPEG-V xml file describing the fog effects of a mist in Antarctica.
Motion Picture	The pictures in motion of the aurora borealis in Iceland.
Motion Medium	The digital representation of the pictures in motion of the aurora borealis in Iceland.
Sound	The sound of the speaker explaining the eruption of Vesuvius volcano.
Audio Medium	The digital representation of the sound of the speaker explaining the eruption of Vesuvius volcano.
Picture	The picture of the durian in Malaysia.
Graphic Medium	The digital representation of the picture of the durian in Malaysia.
Text	The subtitle of the speaker explaining the eruption of Vesuvius volcano.
Text Medium	The digital representation of the subtitle of the speaker explaining the eruption of Vesuvius volcano.
Computer Medium	The digital representation of the pictures in motion of the aurora borealis in Iceland stored in hard disk.
Directly Transmitted Medium	The digital representation of the pictures in motion of the aurora borealis in Iceland being captured by a camera and broadcasted live on TV.

Table 2 presents the instantiation of the ontology. Bearing in mind that most of the mulsemedia systems presented in [16, 18, 19, 28–30, 34] are smaller, the successful instantiation of the whole ontology with this hypothetical scenario presented in Table 2 gives us a broad sense of what is possible to represent with this reference ontology. Moreover, this instantiation cues for appropriateness of the proposed ontology as a reference model for this domain.

5 RELATED WORK

Oh and Hahn [22] proposed an operational ontology for what they call multi-sensory media service. They represent semantic information about multimedia contents, sensory effects synchronized with the contents and sensory devices. Two ontologies were developed: (i) effect ontology, in which they focus on the attributes of a scene with sensory effects such as time, place, effect, etc; and (ii) device ontology, which represents the sensory effect devices.

To evaluate the work, the authors generated an OWL file about some scenes and then inferred the sensory effects from multimedia content information. Although it is a practical approach, it does not present the mulsemedia ecosystem as a whole, making difficult the understanding of what stakeholders have to deal with.

The core ontology for multimedia stimuli description, created by Horvat et al. [15], is an effort to capture affective multimedia stimuli. The ontology is called STIMONT and represents all relevant knowledge about types of multimedia stimuli such as images, sounds, video, and text. It aims to propose an ontology-based approach to formal descriptions of stimuli metadata in order to improve the processes of multimedia annotation and retrieval from databases. Although it is not a mulsemedia ontology, it is related to this work in the sense that it can be referred to how emotion can be captured. It could be combined with MulseOnto in future works to help to explain the emotions from mulsemedia experiences.

The ontology of Naravane and Lange [21] refers to organoleptic properties to the consumption of food. They argue that organoleptic, as well as biological, chemical, and physical, is a class of phenotypic properties resulting from the intersection of them. Thereby, the authors focus on the details of all possible organoleptic traits of an edible substance. The ontology has several classes of stimuli related to the sensory reaction like appearance, touch, smell, taste, and sound, whereas the sensory ontology presented in this work focus on the knowledge about the detection of the stimulus and subsequent recognition and characterization of it. As the work of Horvat et al. [15], it could be a complementary work towards the perception of the senses produced by mulsemedia systems. However, this is out of the scope of our work in this paper. Furthermore, the perception of the senses might not be the same when combining different sensory effects and is very dependent on the user's QoE.

Following the same way of the aforementioned work, Albert et al. [1] present an ontological representation of sensory perception knowledge. However, they go deeper in detail on specifications for haptic perceptions with the proposition of generic elementary haptic sensations. This focus also does not allow stakeholders to perceive the mulsemedia environment as a whole because of its particular addressing. Moreover, as well as the work of Naravane and Lange [21], the sense of QoE seems to be out of its scope.

As far as mulsemedia systems ontology, this research did not find works dealing with it so far - to the best of our knowledge.

6 CONCLUSION AND FUTURE DIRECTIONS

The concept of mulsemedia is not easy to understand at a glance due to its complex digital ecosystem. By underlining sensory effects, types of different medium, responsibilities to present/produce them, and their physical realizations through a formal and explicit representation, we try to convey the message of what is entailed in mulsemedia systems through MulseOnto, a reference domain ontology on Mulsemedia Systems.

This first effort aimed at establishing a common conceptualization about it including the interrelationship between its entities, describing a common vocabulary for knowledge workers in this domain, certifying the stakeholders are uniform and consistent when discussing mulsemedia solutions, and promoting the development of mulsemedia ecosystems as a whole. We evaluated the reference

ontology against the competency questions and also made an instantiation of the concepts to see if a hypothetical mulsemedia system adheres to it. By doing so, we captured and adapted the ontology to reflect concepts sometimes neglected when reading a plain English text about mulsemedia systems. Nevertheless, further scenarios from real-world environments must be subject of study to strengthen the accuracy of our ontology.

MulseOnto can be considered an initial step to discuss the mulsemedia ecosystem that involves capturing, distributing, rendering and perceiving sensory effects. As future work, we intend to take it a step further by creating a mulsemedia ontology network to cope with those processes in detail. A network of ontologies is an assortment of ontologies connected through a variety of relationships. Therefore, each descendant-ontology can share their relationships with a conceivably large number of other well-established ontologies promoting reuse [32]. Another concern that can be taken into account in the future is the materialization of this conceptual model to an operational ontology to exchange, for instance, sensory effects metadata as an alternative to the current standards such as MPEG-V. Moreover, it can be useful for integrating different mulsemedia standards that might eventually arise enabling interoperability between different solutions by mapping them to this reference ontology.

ACKNOWLEDGMENTS

REFERENCES

[1] B. Albert, C. Zanni-Merk, F. B. de Beuvron, M. Pillet, J-L. Maire, C. Knecht, and J. Charrier. 2017. A Smart System to Standardize the Specifications of Haptic Quality Control. *Procedia Comput. Sci.* 112, C (Sept. 2017), 723–730.

[2] M. Bordegoni, G. Faconti, S. Feiner, M. T. Maybury, T. Rist, S. Ruggieri, P. Trahanias, and M. Wilson. 1997. A Standard Reference Model for Intelligent Multimedia Presentation Systems. *Computer Standards & Interfaces* 18, 6-7 (1997), 477–496.

[3] A. C. O. Bringuente, R. A. Falbo, and G. Guizzardi. 2011. Using a Foundational Ontology for Reengineering a Software Process Ontology. *JIDM* 2, 3 (2011), 511–526.

[4] S-F. Chang. 2013. ACM SIGMM Chair's Message. http://www.sigmm.org/news/chairs_msg_2013. (2013). [Online; accessed 26-June-2018].

[5] F. Danieau, J. Bernon, J. Fleureau, P. Guillotel, N. Mollet, M. Christie, and A. Lécuyer. 2013. H-Studio: An Authoring Tool for Adding Haptic and Motion Effects to Audiovisual Content. In *Proc. 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. https://doi.org/10.1145/2508468.2514721

[6] B. B. Duarte, A. L. C. Leal, R. A. Falbo, G. Guizzardi, R. S. S. Guizzardi, and V. E. S. Souza. 2018. Ontological foundations for software requirements with a focus on requirements at runtime. *Applied Ontology* 13, 2 (2018), 73–105.

[7] R. A. Falbo. 2014. SABiO: Systematic Approach for Building Ontologies. In *1st Joint Workshop Onto.Com/ODISE on Ontologies in Conceptual Modeling and Information Systems Engineering*.

[8] G. Ghinea, F. Andres, and S. R. Gulliver. 2011. *Multiple Sensorial Media Advances and Applications: New Developments in MulSeMedia: New Developments in MulSeMedia*. Information Science Reference.

[9] G. Ghinea, C. Timmerer, and W. Lin. 2014. Mulsemedia: State of the Art, Perspectives, and Challenges. *ACM Trans. Multimedia Comput. Commun. Appl.* 11, 1s, Article 17 (Oct. 2014), 23 pages. https://doi.org/10.1145/2617994

[10] M. Gruninger and M. S. Fox. 1995. Methodology for the Design and Evaluation of Ontologies. In *Workshop on Basic Ontological Issues in Knowledge Sharing*.

[11] N. Guarino. 1995. Formal Ontology, Conceptual Analysis and Knowledge Representation. *Int. J. Hum.-Comput. Stud.* 43, 5-6 (Dec. 1995), 625–640.

[12] G. Guizzardi, G. Wagner, J. P. A. Almeida, and R. S. S. Guizzardi. 2015. Towards ontological foundations for conceptual modeling: The unified foundational ontology (UFO) story. *Applied Ontology* 10, 3-4 (2015), 259–271. https:

//doi.org/10.3233/AO-150157

[13] S. Hariri, N. A. Mustafa, K. Karunanayaka, and A. D. Cheok. 2016. Electrical Stimulation of Olfactory Receptors for Digitizing Smell. In *Proceedings of the 2016 Workshop on Multimodal Virtual and Augmented Reality (MVAR '16)*. ACM, New York, NY, USA, Article 4, 4 pages. https://doi.org/10.1145/3001959.3001964

[14] R. S. Heller and C. D. Martin. 1995. A media taxonomy. *IEEE MultiMedia* 2, 4 (1995), 36–45.

[15] M. Horvat, N. Bogunović, and K. Ćosić. 2014. STIMONT: a core ontology for multimedia stimuli description. *Multimedia Tools and Applications* 73, 3 (01 Dec 2014), 1103–1127. https://doi.org/10.1007/s11042-013-1624-4

[16] L. Jalal, M. Anedda, V. Popescu, and M. Murroni. 2018. QoE Assessment for IoT-Based Multi Sensorial Media Broadcasting. *IEEE Transactions on Broadcasting* 64, 2 (June 2018), 552–560. https://doi.org/10.1109/TBC.2018.2823914

[17] M. Josué, R. Abreu, F. Barreto, D. Mattos, G. Amorim, J. dos Santos, and D. Muchaluat-Saade. 2018. Modeling Sensory Effects As First-class Entities in Multimedia Applications. In *Proc. 9th ACM Multimedia Systems Conf. (MMSys '18)*. ACM, New York, NY, USA, 225–236. https://doi.org/10.1145/3204949.3204967

[18] S-K. Kim and Y. S. Joo. 2014. Sensible Media Simulation in an Automobile Application and Human Responses to Sensory Effects. *ETRI Journal* 35, 6 (Dec. 2014), 1001–1010.

[19] F. P. Luque, I. Galloso, C. Feijoo, C. A. Martín, and G. Cisneros. 2014. Integration of Multisensorial Stimuli and Multimodal Interaction in a Hybrid 3DTV System. *ACM Trans. Multimedia Comp. Comm. Appl.* 11, 1s, Article 16 (Oct. 2014), 22 pages.

[20] N. Murray, O. A. Ademoye, G. Ghinea, and G-M. Muntean. 2017. A Tutorial for Olfaction-Based Multisensorial Media Application Design and Evaluation. *ACM Comp. Surveys* 50, 5, Article 67 (2017), 30 pages. https://doi.org/10.1145/3108243

[21] T. Naravane and M. Lange. 2017. Organoleptic and Sensory Ontology. In *Proceedings of the Joint Ontology Workshops 2017 Episode 3: The Tyrolean Autumn of Ontology, Bozen-Bolzano, Italy, September 21-23, 2017*.

[22] S. Oh and M. Hahn. 2009. The ontology for a multi-sensory media service. In *2009 11th Int. Conf. on Advanced Communication Technology*, Vol. 02. 1025–1030.

[23] V. Presutti, E. Blomqvist, E. Daga, and A. Gangemi. 2012. *Pattern-Based Ontology Design*. Springer Berlin Heidelberg, Berlin, Heidelberg, 35–64.

[24] N. Ranasinghe and E. Y-L. Do. 2016. Digital Lollipop: Studying Electrical Stimulation on the Human Tongue to Simulate Taste Sensations. *ACM Trans. Multimedia Comput. Commun. Appl.* 13, 1, Article 5 (Oct. 2016), 22 pages. https://doi.org/10.1145/2996462

[25] N. Ranasinghe, T. N. T. Nguyen, Y. Liangkun, L-Y. Lin, D. Tolley, and E. Y-L. 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. https://doi.org/10.1145/3123266.3123440

[26] S. D. Roy and W. Zeng. 2015. Media on the Web. In *Social Multimedia Signals*. Springer, 9–18.

[27] F. B. Ruy, R. A. Falbo, M. B. Perini, S. D. Costa, and G. Guizzardi. 2016. SEON: A Software Engineering Ontology Network. In *Knowledge Engineering and Knowledge Management*. Springer International Publishing, Cham, 527–542.

[28] E. B. Saleme and C. A. S. Santos. 2015. PlaySEM: A Platform for Rendering MulSeMedia Compatible with MPEG-V. In *Proceedings of the 21st Brazilian Symposium on Multimedia and the Web (WebMedia '15)*. ACM, New York, NY, USA, 145–148. https://doi.org/10.1145/2820426.2820450

[29] E. B. Saleme, C. A. S. Santos, and G. Ghinea. 2018. Coping with the Challenges of Delivering Multiple Sensorial Media. *IEEE MultiMedia* (2018), 11 pages. In press. https://bura.brunel.ac.uk/handle/2438/16592

[30] E. B. Saleme, C. A. S. Santos, and G. Ghinea. 2018. Improving Response Time Interval in Networked Event-based Mulsemedia Systems. In *Proceedings of the 9th ACM Multimedia Systems Conference (MMSys '18)*. ACM, New York, NY, USA, 216–224. https://doi.org/10.1145/3204949.3204965

[31] R. Steinmetz and K. Nahrstedt. 1995. *Multimedia: Computing, Communications and Applications*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

[32] M. C. Suarez-Figueroa, A. Gomez-Perez, E. Motta, and A. Gangemi. 2012. *Ontology Engineering in a Networked World*. Springer Publishing Company, Inc.

[33] C. T. Vi, A. Marzo, D. Ablart, G. Memoli, S. Subramanian, B. Drinkwater, and M. Obrist. 2017. TastyFloats: A Contactless Food Delivery System. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces (ISS '17)*. ACM, New York, NY, USA, 161–170. https://doi.org/10.1145/3132272.3134123

[34] M. Waltl, B. Rainer, C. Timmerer, and H. Hellwagner. 2013. An End-to-end Tool Chain for Sensory Experience Based on MPEG-V. *Image Commun.* 28, 2 (Feb. 2013), 136–150. https://doi.org/10.1016/j.image.2012.10.009

[35] V. Wanick, G. Xavier, and E. Ekmekcioglu. 2018. Virtual Transcendence Experiences: Exploring Technical and Design Challenges in Multi-Sensory Environments. In *Proceedings of the 10th International Workshop on Immersive Mixed and Virtual Environment Systems (MMVE '18)*. ACM, New York, NY, USA, 7–12.

[36] K. Yoon, S-K. Kim, J. J. Han, S. Han, and M. Preda. 2015. *MPEG-V: Bridging the Virtual and Real World* (1st ed.). Academic Press.