INTERNATIONAL STANDARD

ISO/IEC 23005-1

First edition
2011-04-01

Information technology — Media context and control —

Part 1: Architecture

Technologies de l'information — Contrôle et contexte de supports —

Partie 1: Architecture

Reference number
ISO/IEC 23005-1:2011(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

ISO/IEC 23005-1 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

ISO/IEC 23005 consists of the following parts, under the general title *Information technology — Media context and control*:

— *Part 1: Architecture*

— *Part 2: Control information*

— *Part 3: Sensory information*

— *Part 4: Virtual world object characteristics*

— *Part 5: Data formats for interaction devices*

— *Part 6: Common types and tools*

— *Part 7: Conformance and reference software*

Introduction

The usage of multimedia content is becoming omnipresent in everyday life, in terms of both consumption and production. On the one hand, professional content is provided to the end user in high-definition quality, streamed over heterogeneous networks, and consumed on a variety of different devices. On the other hand, user-generated content overwhelms the Internet with multimedia assets being uploaded to a wide range of available Web sites. That is, the transparent access to multimedia content, also referred to as Universal Multimedia Access (UMA), seems to be technically feasible. However, UMA mainly focuses on the end-user devices and network connectivity issues, but it is the user who ultimately consumes the content. Hence, the concept of UMA has been extended to take the user into account, which is generally referred to as Universal Multimedia Experience (UME).

However, the consumption of multimedia assets can also stimulate senses other than vision or audition, e.g., olfaction, mechanoreception, equilibrioception, or thermoception. That is, in addition to the audio-visual content of, for example, a movie, other senses shall also be stimulated giving the user the sensation of being part of the particular media which shall result in a worthwhile, informative user experience.

This motivates the annotation of the media resources with metadata as defined in this part of ISO/IEC 23005 that steers appropriate devices capable of stimulating these other senses.

ISO/IEC 23005 (MPEG-V) provides an architecture and specifies associated information representations to enable the interoperability between virtual worlds, for example, digital content provider of a virtual world, (serious) gaming, simulation, DVD, and with the real world, for example, sensors, actuators, vision and rendering, robotics (e.g. for revalidation), (support for) independent living, social and welfare systems, banking, insurance, travel, real estate, rights management and many others.

Virtual worlds[1)] (often referred to as 3D3C for 3D visualization & navigation and the 3C's of community, creation and commerce) integrate existing and emerging (media) technologies (e.g. instant messaging, video, 3D, VR, AI, chat, voice, etc.) that allow for the support of existing and the development of new kinds of social networks. The emergence of virtual worlds as platforms for social networking is recognized by businesses as an important issue for at least two reasons:

a) it offers the power to reshape the way companies interact with their environments (markets, customers, suppliers, creators, stakeholders, etc.) in a fashion comparable to the Internet;

b) it allows for the development of new (breakthrough) business models, services, applications and devices.

Each virtual world however has a different culture and audience making use of these specific worlds for a variety of reasons. These differences in existing metaverses permit users to have unique experiences. Resistance to real-world commercial encroachment still exists in many virtual worlds where users primarily seek an escape from real life. Hence, marketers should get to know a virtual world beforehand and the rules that govern each individual universe.

Although realistic experiences have been achieved via devices such as 3-D audio/visual devices, it is hard to realize sensory effects only with presentation of audiovisual contents. The addition of sensory effects leads to even more realistic experiences in the consumption of audiovisual contents. This will lead to the application of new media for enhanced experiences of users in a more realistic sense.

Such new media will benefit from the standardization of a control and sensory information which can include sensory effect metadata, sensory device capabilities/commands, user sensory preferences, and various delivery formats. The MPEG-V architecture can be applicable for various business models for which

1) Some examples of virtual worlds are: *Second Life* (http://secondlife.com/), IMVU (http://www.imvu.com/) and Entropia Universe (http://www.entropiauniverse.com/).

audiovisual contents can be associated with sensory effects that need to be rendered on appropriate sensory devices.

Multi-user online virtual worlds, sometimes called Networked Virtual Environments (NVEs) or massively-multiplayer online games (MMOGs), have reached mainstream popularity. Although most publications tend to focus on well-known virtual worlds like *World of Warcraft*, *Second Life*, and *Lineage*, there are hundreds of popular virtual worlds in active use worldwide, most of which are not known to the general public. These can be quite different from the above-mentioned titles. To understand current trends and developments, it is useful to keep in mind that there is large variety in virtual worlds and that they are not all variations on *Second Life*.

The concept of online virtual worlds started in the late 70s with the creation of the text-based Dungeons & Dragons world MUD. In the eighties, larger-scale graphical virtual worlds followed, and in the late nineties the first 3D virtual worlds appeared. Many virtual worlds are not considered games (MMOGs) since there is no clear objective and/or there are no points to score or levels to achieve. In this report we will use “virtual worlds” as an umbrella term that includes all possible varieties. See the literature for further discussion of the distinction between gaming/non-gaming worlds. Often, a virtual world which is not considered to be an MMOG does contain a wide selection of mini-games or quests, in some way embedded into the world. In this manner a virtual world acts like a combined graphical portal offering games, commerce, social interactions and other forms of entertainment. Another way to see the difference: games contain mostly pre-authored stories; in virtual worlds the users more or less create the stories themselves. The current trend in virtual worlds is to provide a mix of pre-authored and user-generated stories and content, leading to user-modified content.

Current virtual worlds are graphical and rendered in 2D, 2.5 D (isometric view) or 3D, depending on the intended effect and technical capabilities of the platform: web-browser, gaming PC, average PC, game console, mobile phone, and so on.

“Would it not be great if the real world economy could be boosted by the exponential growing economy of the virtual worlds by connecting the virtual - and real world”; in 2007 the Virtual Economy in *Second Life* alone was around 400 MEuro, a factor nine growth from 2006. The connected devices and services in the real world can represent an economy of a multiple of this virtual world economy.

Virtual worlds have entered our lives, our communication patterns, our culture, and our entertainment never to leave again. It's not only the teenager active in *Second Life* and *World of Warcraft*, the average age of a gamer is 35 years by now, and it increases every year. This does not even include role-play in the professional context, also known as serious gaming, inevitable when learning practical skills. Virtual worlds are in use for entertainment, education, training, obtaining information, social interaction, work, virtual tourism, reliving the past and forms of art. They augment and interact with our real world and form an important part of people's lives. Many virtual worlds already exist as games, training systems, social networks and virtual cities and world models. Virtual worlds will change every aspect of our lives: the way we work, interact, play, travel and learn. Games will be everywhere and their societal need is very big and will lead to many new products and require many companies.

Technology improvement, both in hardware and software, forms the basis of this. It is envisaged that the most important developments will occur in the areas of display technology, graphics, animation, (physical) simulation, behavior, and artificial intelligence, loosely distributed systems and network technology.

The figures in this part of ISO/IEC 23005 have been reproduced here with the permission of Samsung, Sharp Electronics, ETRI, University of Klagenfurt, Institute of Science and Technology, Myongji University, Institut national des télécommunications and the partners of the ITEA2 project Metaverse1: Philips, Forthnet S.A., Alcatel-Lucent Bell N.V., Innovalia, Alcatel-Lucent France, Technicolor, Orange Labs, DevLab, CBT, Nextel, Carsa, Avantalia, Ceesa, Virtualware, I&IMS, VicomTECH, E-PYME, CIC Tour Gune, Artefacto, Metaverse Labs, Technical University Eindhoven, Utrecht University, University of Twente, Stg. EPN, VU Economics & BA, VU CAMeRA, Ellinogermaniki Agogi, IBBT-SMIT, UPF-MTG, CEA List and Loria/Inria Lorraine.

Information technology — Media context and control —

Part 1:
Architecture

1 Scope

This part of ISO/IEC 23005 specifies the architecture of MPEG-V (media context and control).

2 MPEG-V System Architecture

A strong connection (defined by an architecture that provides interoperability trough standardization) between the virtual and the real world is needed to reach simultaneous reactions in both worlds to changes in the environment and human behavior. Efficient, effective, intuitive and entertaining interfaces between users and virtual worlds are of crucial importance for their wide acceptance and use. To improve the process of creating virtual worlds a better design methodology and better tools are indispensible. For fast adoption of virtual worlds we need a better understanding of their internal economics, rules and regulations.

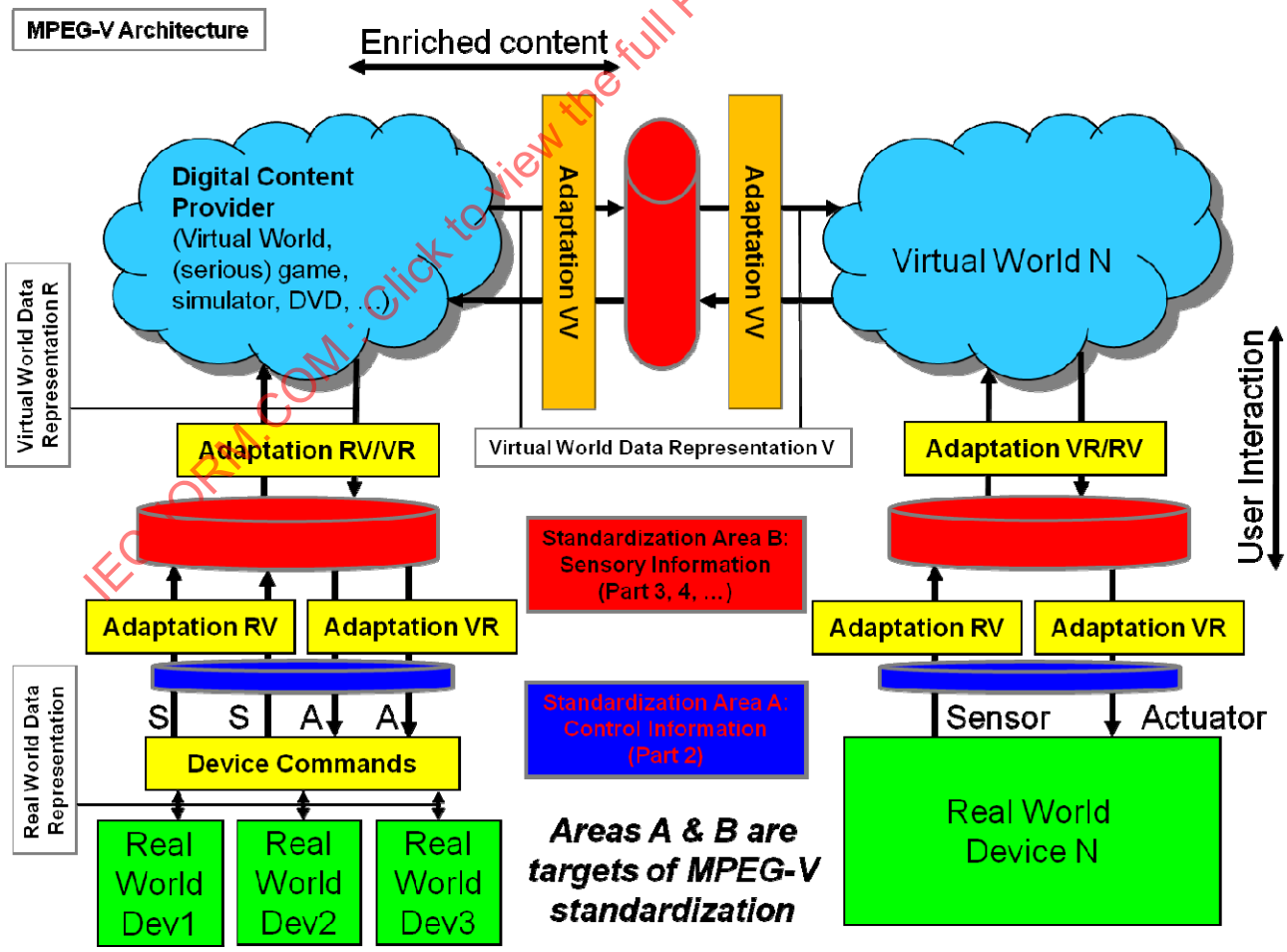

Figure 1 — System Architecture of the MPEG-V Framework

The overall system architecture for the MPEG-V framework is depicted in Figure 1 comprising the standardization areas a: control information and b: sensory information. Please note that standardization area b may be composed of multiple parts of the MPEG-V standard.

The individual elements of the architecture have the following functions:

— **Digital Content Provider**

A provider of digital content, real time or non real time, of various nature ranging from an on-line virtual world, simulation environment, multi user game, a broadcasted multimedia production, a peer-to-peer multimedia production or packaged content like a DVD or game.

— **Virtual World Data Representation R**

The native representation of virtual world related information that is intended to be exchanged with the real world (either exported or imported).

— **Virtual World Data Representation V**

The native representation of virtual world related information that is intended to be exchanged with another virtual world (either exported or imported).

— **Adaptation RV/VR**

The adaptation of the native representation of virtual world related information (that is intended to be exchanged with the real world) to the standardized representation format of MPEG-V in the standardization area B (e.g. sensory information, haptic information, emotion information …) in both directions: that is from the standardized representation into the native representation and vice versa.

— **Adaptation VV**

The adaptation of the native representation of virtual world related information (that is intended to be exchanged with another virtual world) to the standardized representation format of MPEG-V in the standardization area B (e.g. avatar information …) in both directions: that is from the standardized representation into the native representation and vice versa.

— **Sensory Information**

The standardized representation format of MPEG-V in the standardization area B (Sensory Information) (e.g. sensory information, haptic / tactile information, emotion information, avatar information …).

— **Adaptation RV**

The adaptation of the standardized representation of real world related information in the standardized representation format of MPEG-V in the standardization area A to the standardized representation of virtual world related information in the standardized representation format of MPEG-V in the standardization area B.

— **Adaptation VR**

The adaptation of the standardized representation of virtual world related information in the standardized representation format of MPEG-V in the standardization area B to the standardized representation of real world related information in the standardized representation format of MPEG-V in the standardization area A.

— **Control Information**[2)]

The standardized representation format of MPEG-V in the standardization area A (Control Information) (e.g. bi-directional control information, preference information, capability information …) related to the following elements of the architecture:

— Virtual World Data Representation R

— Virtual World Data Representation V

— Real World Data Representation

— **Real World Data Representation**

The native representation of real world related information that is intended to be exchanged with the virtual world (either exported or imported).

— **Device Commands**

Device commands is responsible for the adaptation of the native representation of real world related information (that is intended to be exchanged with the virtual world) to the standardized representation format of MPEG-V in the standardization area A (control information) (e.g. bi-directional control information, preference information, capability information …) in both directions: that is from the native representation into the standardized representation and vice versa.

— **Real World Device S**

A real world device containing a sensor (e.g. a temperature, light intensity, blood pressure, heartbeat …)

— **Real World Device A**

A real world device containing an actuator (e.g. a display, speaker, light speaker, fan, robot, implant …).

NOTE Real world devices can contain any combination of sensors and actuators in one device.

In the MPEG-V standard the following areas are addressed:

— **Standardization Area A: Control Information**

This area covers the information representation of the control information to and from devices in the physical world and into and from the virtual world. Examples of these representations are the representation of sensory input devices like smart vision systems, environmental and body sensors and the like and sensory output rendering devices like lights, heaters, fans, displays, speakers and the like.

— **Standardization Area B: Sensory Information**

This area covers the (bidirectional) information representation of information exchanged between the physical world and the virtual world as well as the information exchange between virtual worlds. Examples of these representations are the representation of haptic, emotion and avatar information.

2) In general, control information is strongly related to de-facto industry solutions for e.g. sensors, actuators and virtual worlds.

3 Instantiations

3.1 Instantiation 1: Exchanges within the real world

3.1.1 Instantiation 1.1: Representation of Sensory Effects (RoSE)

3.1.1.1 Introduction and Motivation

Traditional multimedia with audio/visual contents have been presented to users via display devices and audio speakers as depicted in Figure 2. In practice, however, users are becoming excited about more advanced experiences of consuming multimedia contents with high fidelity. For example, stereoscopic video, virtual reality, 3-dimensional television, multi-channel audio, etc. are typical types of media increasing the user experience but are still limited to audio/visual contents.

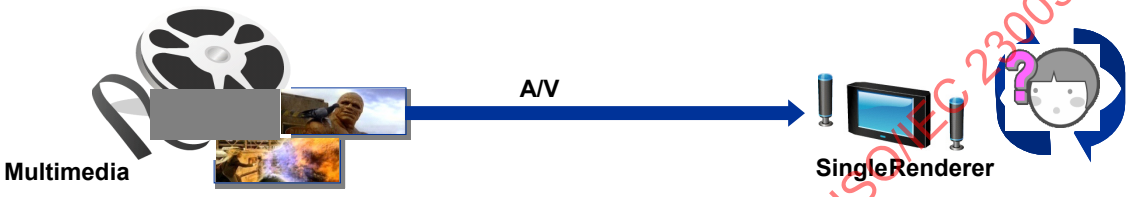

Figure 2 — Traditional Multimedia Consumption

From a rich multimedia perspective, an advanced user experience would also include special effects such as opening/closing window curtains for a sensation of fear effect, turning on a flashbulb for lightning flash effects as well as fragrance, flame, fog, and scare effects can be made by scent devices, flame-throwers, fog generators, and shaking chairs respectively. Such scenarios would require enriching multimedia contents with information enabling consumer devices to render them appropriately in order to create the advanced user experience such as described above. Figure 3 shows an example configuration adopting a multimedia multiple device (MMMD) approach for an advance user experience compared to the multimedia single device (MMMD) approach as illustrated in Figure 2. In this configuration, the multimedia contents are not rendered by a single device but with multiple devices in a synchronized manner.

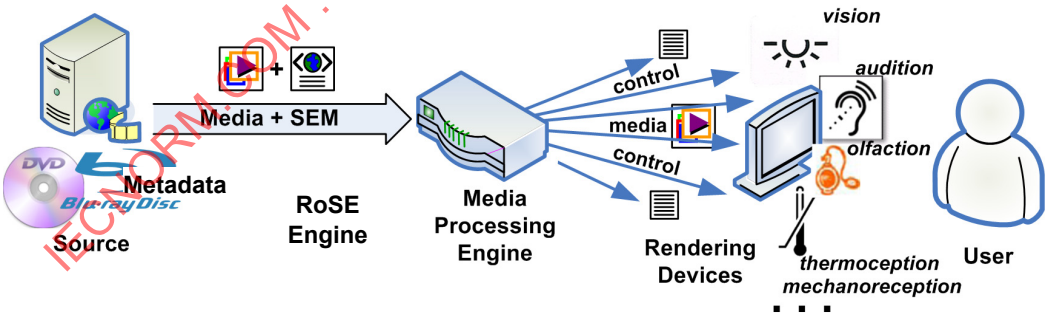

Figure 3 — RoSE-enabled Multimedia Consumption for Advanced User Experience

From a technical perspective, this requires a framework for the *Representation of Sensory Effects (RoSE)* information which may define metadata about special or sensory effects, characteristics of target devices, synchronizations, etc. The actual presentation of the RoSE information and associated audio/visual contents allows for an advanced, worthwhile user experience.

The next Subclause provides a brief overview of the RoSE system architecture.

3.1.1.2 RoSE System Architecture

The overall system architecture for the RoSE framework is depicted in Figure 4 comprising Sensory Effect Metadata, Sensory Device Capabilities, Sensory Device Commands, User Sensory Preferences, and a so-called RoSE Engine which generates output data based on its input data.

It is important to note that the *Sensory Effect Metadata*, *Sensory Device Capabilities*, *Sensory Device Commands*, and *User Sensory Preferences* are within the scope of standardization and, thus shall be normatively specified. On the other side, the *RoSE Engine* as well as *Provider* entities and *Consumer Devices* are informative and are left open for industry competition.

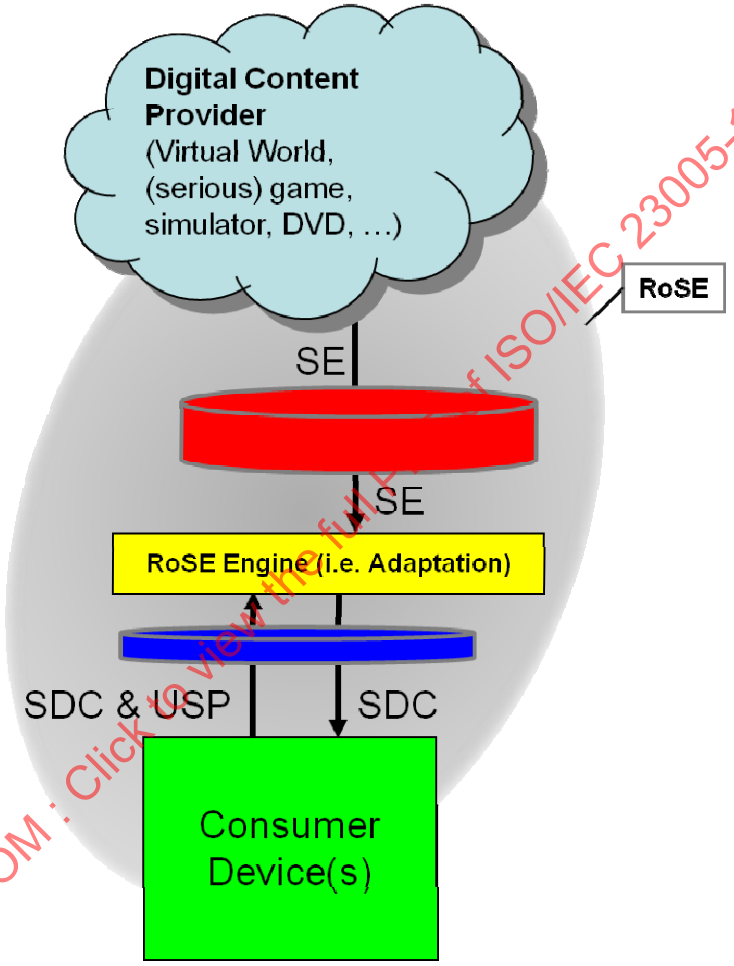

Figure 4 — RoSE System Architecture

A *provider* within the RoSE framework is referred to as an entity that acts as the source of the *sensory effect metadata* such as a broadcaster, content creator/distributor, or even a service provider. The *RoSE Engine* is an entity that takes the *sensory effect metadata*, the *sensory device capabilities* and the *user sensory preferences* as inputs and generates *sensory device commands* based those in order to control the consumer devices enabling a worthwhile, informative experience to the user.

Consumer devices are entities that act as the sink of the *sensory commands* and act as the source of *sensory device capabilities*. Additionally, entities that provide user sensory preferences towards the RoSE engine are also collectively referred to as consumer devices. Note that *sensory devices* (see below) are sub-set of consumer devices including fans, lights, scent devices, human input devices such as a TV set with a remote control (e.g., for preferences).

The actual *sensory effect metadata* provides means for representing so-called *sensory effects*, i.e., an effect to augment feeling by stimulating human sensory organs in a particular scene of a multimedia application. Examples of sensory effects are scent, wind, light, etc. The means for transporting this kind of metadata is referred to as sensory effect delivery format which, of course, could be combined with an audio/visual delivery format, e.g., MPEG-2 transport stream, a file format, or Real-time Transport Protocol (RTP) payload format, etc.

The *sensory device capabilities* define description formats to represent the characteristics of sensory devices in terms of which sensory effects they are capable to perform and how. A sensory device is a consumer device by which the corresponding sensory effect can be made (e.g., lights, fans, heater, fan, etc.). *Sensory device commands* are used to control the sensory devices. As for sensory effect metadata, also for sensory device capabilities and commands corresponding means for transporting this assets are referred to as sensory device capability/commands delivery format respectively.

Finally, the user sensory preferences allow for describing preference of the actual (end) users with respect to rendering of sensory effects for also a delivery format is provided.

3.2 Instantiation 2: Exchanges between real world and virtual world

3.2.1 Instantiation 2.1: Full motion control and navigation of avatar/object with multi-input sources

Full motion control and navigation of avatar/object with multi-input sources allows for the full motion control and navigation of 3D objects and avatars in a Virtual World. Recently, user interest in human-computer interaction has grown considerably based on large volumes of recent research. Through the development of VR technology, it has applied to various fields. Especially, the Entertainment area is commercialized such as 3D virtual online communities like *Second Life* and 3D Game station. Nintendo Wii provides new game experience using 3D input device. Especially the control of objects and avatars in 3D virtual space requires more complex methods than conventional input devices such as mouse, keyboard, joystick and etc. The Figure below shows the example picture of these systems and like this style, it is applied to home, school or other place for various purposes such as entertainment or education including digital contents of 3D virtual world.

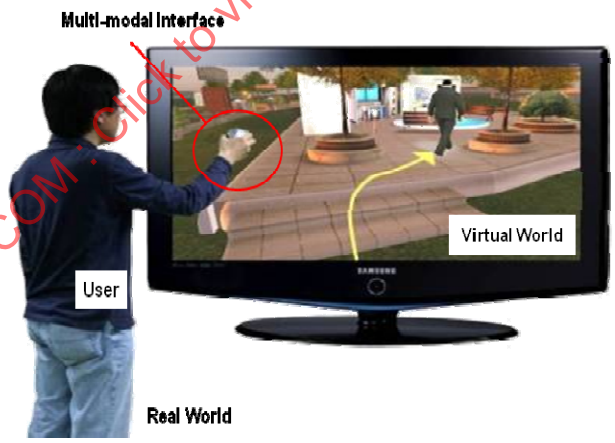

Figure 5 — Full motion control and navigation of avatar/object with multi-input sources

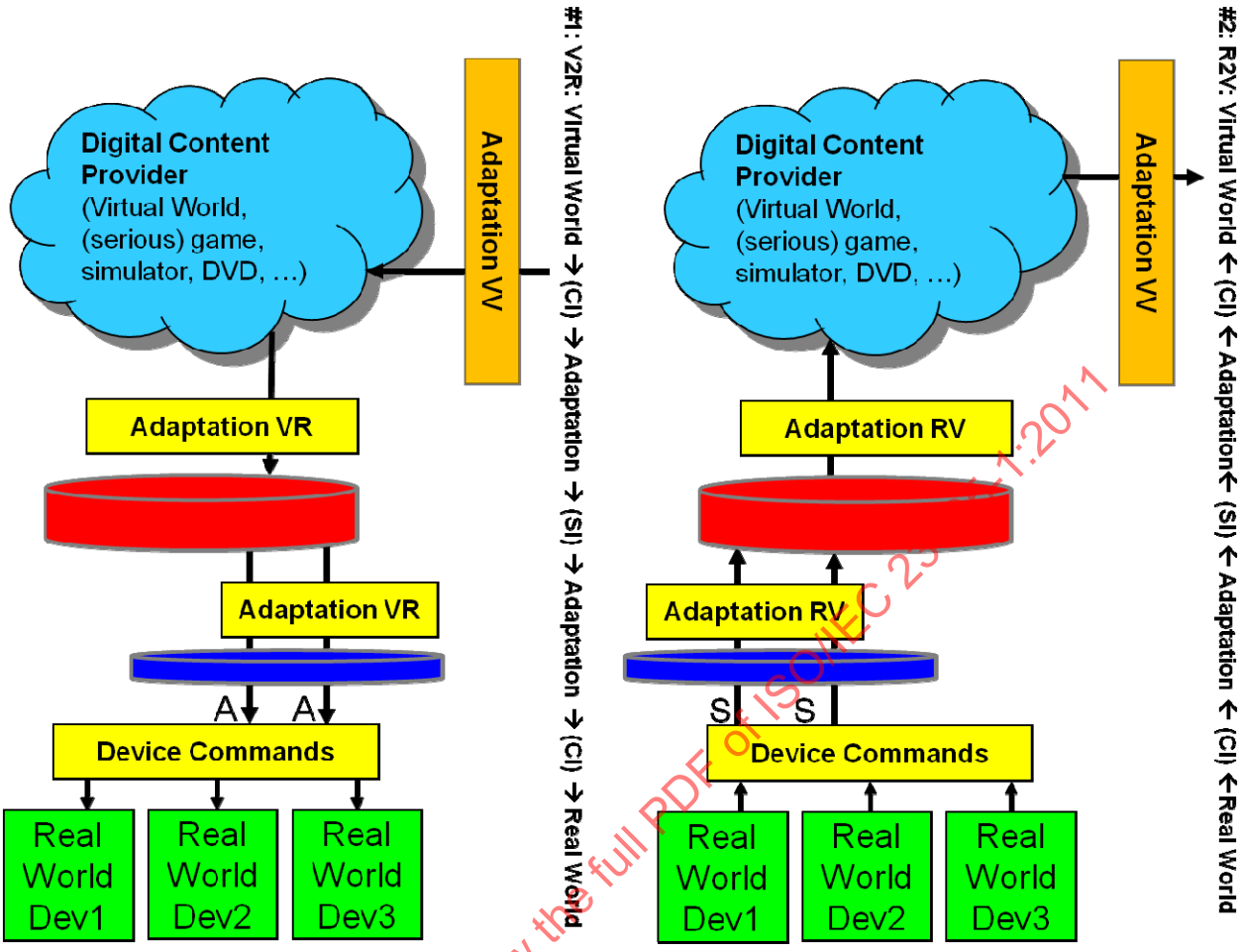

Figure 6 — (Possible) System Architecture for (bidirectional) exchange of information between real world and virtual world

3.2.2 Instantiation 2.2: Virtual Travel

Tourism has become a popular global leisure activity. It is defined as people who travel to and stay in places outside their usual environment for not more than one consecutive year, for leisure, business and other purposes. With this use case, we are contributing to change a little bit this concept, as the main goal of the virtual travel use case is to offer to the people, the possibility of visit a tourist destination, in this case Las Palmas de Gran Canaria, with no ticket required, no money spent and no need to leave their seat, only whit the help of elaborated 3D images and pictures of this place.

So the virtual tourist will be able to arrive at the airport, to take a taxi, tram or bus, to eat traditional food or to visit the most interesting tourist places, such as virtual museums, where the user can interact with the objects, which are placed and exhibited within or virtual guides around the city.

One of the objectives of the use case will be the travel motivation to this destination, as he has a lot of information about it, being easier than other tourist places. The Virtual Travel will be the part before the travel.

However if the user decide to visit this destination, it will be able to use the virtual world, due to the Virtual Traces and Real Places use case, where he can share all experiences, impressions and feelings, he gained during the stay, with family and friends. It will be the part after the virtual travel.

Figure 7 — Visualization of virtual travel

The “Virtual Travel” uses case, as proposed by the Spanish partners however supported by a larger subset of the Metaverse1 consortium, can rely on the support of the body that is in Charge of Tourism Activities in the Canary Island (Patronato de Turismo de Gran Canaria, Ignacio Mol (Managing Director), Las Palmas de Gran Canaria, Spain) has agreed to provide the consortium with the necessary requirements in the area of Tourism and Metaverse in order to assess the *appropriateness of the technology and standards to their real system and activities.*

3.2.3 Instantiation 2.3: Serious gaming for Ambient Assisted Living

The “Serious gaming for Ambient Assisted Living”, subtitle “The example of physical exercise” use case is proposed by the Dutch partners however supported by a larger subset of the Metaverse1consortium. Today, in an environment where people have the option to be increasingly inactive in their daily lives, the requirement for physical exercise as an important factor for a healthy lifestyle is critical, particularly for the elderly part of the society. An individual's physical as well as psychological wellbeing is dependent on daily exercise giving one a sense of self-efficacy and independence. As European societies become increasingly aged, the necessity to exercise is of utmost importance for the physical and mental fitness as well as the self-respect of their people.

However, people in general tend to be less motivated to maintain their level of fitness as they become older. While people may recognize the need for behavioral change the same individuals are very creative in justifying excuses. Thus, one of the greatest challenges for behavioral change is bridging good intentions with factual behavior. The majority of elderly people do not exercise much or at least not often and intensively enough. Increasing the frequency and duration of exercise remains one of the most important and difficult endeavors that challenge health care professionals.

Unfortunately, making individuals aware of the benefits of physical exercise is likely to only change their behavioral intentions, but rather unlikely to affect their established lifestyle habits. In fact, a widening of the discrepancy between desired intentions and undesired behavior contributes to the emergence of guilt, and consequently may reduce self-esteem and the experience of self-efficacy.

It is a necessity to provide the tools to guide a person through a behavioral goal to change his/her behavior by offering a sustainable positive and reinforcing experience. This can best be accomplished through the implementation of a virtual agent that takes a leadership role in such intended behavioral changes. In this project, elderly people will be provided with a solution through serious gaming that will impact their habits and successfully transform them into new, desired behavior patterns.

3.2.4 Instantiation 2.4: Virtual Traces of Real Places

When we travel, we gain new impressions and experiences to which we associate emotions and feelings. Once we arrive home we want to share these experiences, impressions and feelings with our friends and families. With the widespread availability of digital cameras we can capture images and video of a remote place and bring it home. However, this way of sharing an experience has two main drawbacks:

a) Only a part of the experience can be shared with only a limited impact on the audience

b) There is a significant time gap between the moment we gain the experience (at some remote place) and when we share it with our friends and family (at home)

Wouldn't it be nice to stay in touch with our friends and family while travelling? Wouldn't it be nice to instantly share an experience with those who stay home and to make them a part of that experience? The goal of this use case is to incorporate travel experiences into a virtual world that can be shared with friends and family. Digital content (images, video, sounds, light atmosphere, position, activity, motion etc) captured by someone in a remote place is associated to a virtual world. Family and friends that stay at home can experience these virtual traces as they move around in the virtual world. Various sensory effect devices are used to enhance the recreated experience at home, providing more of the feeling of “really being there” to users.

Imagine, for example, Sue and John travelling to Gran Canaria for a short break. Their two teenage children Dan and Max stay at home. The father of Sue, although physically getting weaker through his old age, is very interested in the places his daughter visits.

Sue and John both carry a mobile device to capture digital content (such as pictures, videos, GPS coordinates, sounds, etc). One morning Sue and John decide to visit the big crater of Caldera de Tejada. In the afternoon they make a boat-trip along the coastline and in the evening they enjoy a beautiful sunset on the beach, close to a sea-side cocktail bar.

The pictures, videos and sounds they capture with their mobile devices are automatically geo-tagged and published in the family's virtual world. When pictures are uploaded, they are automatically annotated with information about the light atmosphere, sound atmosphere and motion at the time and place the picture was taken.

At home, the children of John and Sue just got in to have dinner. But first, they switch on the IPTV in the living room. The IPTV indicates that new content is available on the family's virtual world channel. Dan uses the remote control pointer to select the channel, curious what his parents have been up to. The 3D Metaverse1 browser on the TV gently glides to a virtual representation of Gran Canaria. The photos created today by John and Sue are all in there, placed at the appropriate positions in the virtual world. Dan can easily navigate this world using his pointing device.

He selects a replay of his parents' experiences of today. First, the browser shows impressive rock formations and vistas in the Caldera de Tejada. The sun was shining bright there, and the lights connected to the TV turn on bright to mimic the experience that Sue and John had there. Then, Dan lets his avatar fly to the coastline where his parents made photos during their afternoon boat trip. While Dan lets his avatar follow the route taken by his parents, the boat trip is recreated in the living room: the lights are adapted, impressions of the sound from the trip can be heard, photos can be seen, and Dan and Max even feel the humming of the engine and the rocking of the waves through the Touch Experience technology integrated into the family's couch.

Finally, they end up at the virtual version of the beach where their parents are right now, enjoying the sunset. The lights in the living room get a deep orange color. Dan and Max can even feel on their couch the rolling and crashing of the strong Atlantic waves, as Dan moves his avatar into the water. “Not bad!”, says Max, “Maybe we should've joined mum and dad after all. Hey, Dan, check out that place over there.” Max points his brother to the nearby cocktail bar, where good music can be heard. Dan navigates his avatar into the bar and, in almost no time, he and Max are busy complimenting the girls through the voice-chat system. The bass of the music gently drums on their bodies, amplified through the Touch Experience couch. Dan thinks the couch is great for places like this - even the occasional hug or tap on the shoulder from others can be felt on your body!

Unfortunately, not everyone is as technology-savvy as Dan and Max. Sue's father for example is not able to handle a mouse, pointer or keyboard because of his old age. However, Sue gave him a digital picture frame that is connected to the family virtual world. This picture frame currently shows a map of Gran Canaria and on that map are some thumbnails of pictures that Sue and John have taken today. A simple touch on a thumbnail allows him to see the full picture and to share in the experiences his daughter had today.

The “Virtual Traces of Real Places” use case is proposed by the Dutch partners however supported by a larger subset of the Metaverse1 consortium and has a strong link to the “Virtual Travel” use case.

3.3 Instantiation 3: Exchanges between virtual worlds

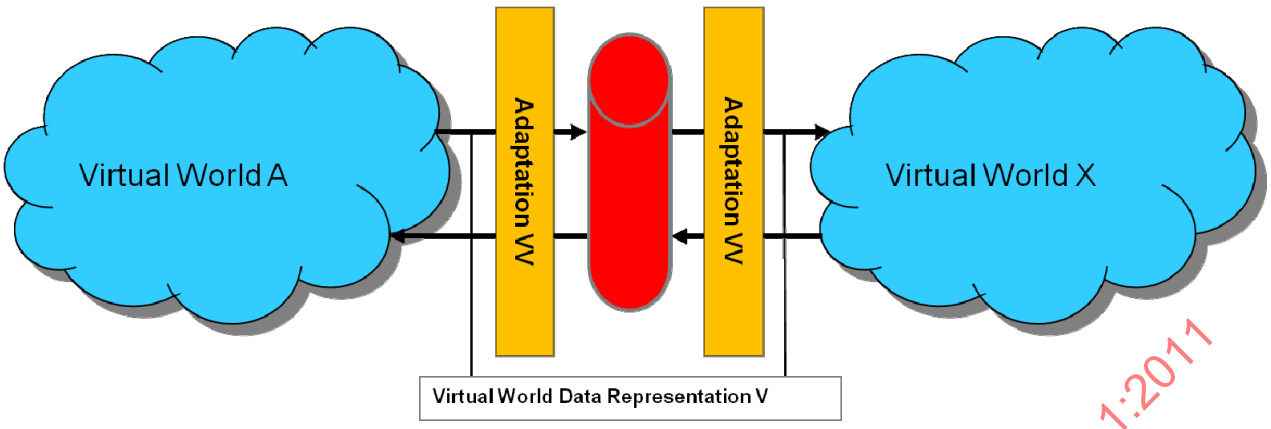

Figure 8 — (Possible) System Architecture for (bidirectional) exchange of information between virtual worlds

3.3.1 Instantiation 3.1: Avatar Appearance

Avatar Appearance allows specifying the appearance of an avatar in a common way, allowing for the exchange of its characteristics between different virtual worlds and, in this way, allowing a user having a virtual anonymous but common identity in any virtual world.

3.3.2 Instantiation 3.2: Social presence

There is a universe where people live in groups that occupy certain places of it, almost exclusively on the planet earth. People create these groups that we call societies, because of the need to communicate and collaborate with mankind. Societies, scant of number, accumulate into numerous and more numerous groups that occupy towns, cities and states, eventually spreading throughout the planet.

During their life, people may travel to a lot of different places for various reasons and explore things to collect more and more information to reason with. They produce record and transfer knowledge, while their activities point out a necessity for the creation of aiding tools to assist them. Countless procedures and tools, which constitute people's Technology, were designed, produced and improved due to people's innovative thoughts, discoveries and inventions.

People now design the Metaverse. The virtual universe that will overcome the physical distance between two remote places of the universe by exchanging information between them in what we call “real time”. Metaverse technology will be able to process this information in real time, in order to distribute a representation of each user's activity to many remote places and also provide it's users with the impression of being in another place, giving them the opportunity to operate applications and devices that run or are located to remote places. Metaverse will also provide a connection with numerous services available to the real world.

Hence in the future, Paul living in Paris will use his 3D visualization system to deceive himself of having a conversation in the same room with his daughter Penny, who lives in Stockholm. Together, they may watch TV, participate in a quiz, play a game, or visit an exhibition that is represented in the Metaverse. Sometimes they call Anne, Paul's wife and mother of Penny.

Penny is a theatre actor and she likes to comment on things by making a show for her dad. For that she uses two robots located in Paul's office, Ego and Alter-ego. Using her Metaverse system, Pennys wear a number of different kinds of wearable sensor-devices and records some of the movements for each robot while watching the outcome in her 3D visualization system in real time. She also records their “words”. After that she chooses some existing automated movements for the robots in order to increase the complexity of the show and with her order, her system will execute the show by operating the two robots. When Penny has prepared

something, she likes to take over one of the robots and operate it in real time to go and tease her dad in order to get his attention.

Her mother, Anne, is a physics teacher. She likes to build virtual worlds and try to visualize the invisible for her lessons. With a “crowd managing” service, she builds virtual environments full of automated and adapting characters and structures for her students to use in order to observe, interact and learn through their journeys in her 3D virtual worlds.

John, the son of the family, usually sends an invitation to all family members to join him in various events. He particularly enjoys visiting museums. All members of the family can see at any time the declared status of each other. One day, John planned a meeting in the Virtual Museum of the Science and the Industry of “La Villette” located in Paris (see Figure 1). Anne, who is very sociable, started a chat with a lot of real visitors located in the real place of the museum. They all have a lot of different kind of devices to use with their Metaverse systems such as “wiimote-like” devices, multi-touch table, cameras and different kind of sensors and actuators.

Figure 9 — Geode of industry and sciences museum overview

3.3.3 Instantiation 3.3: Group Decision-making in the context of Spatial Planning

Virtual Worlds offer the opportunity to communicate with different people from all over the world, regardless of time and geographical location. Therefore, for organizations in general, and international operating parties such as multinationals and NGO's more specific, Virtual Worlds can function as an important future *communication platform.* This use case will focus on the potential of Virtual Worlds as a medium to support social interaction and group decision making processes regarding spatial planning issues.

In comparison to 2D worlds such as *Facebook*, the added value of 3D worlds lies particularly in the possibility to control visual cues. These visual cues in Virtual Worlds can have an effect on factors such as eInclusion, the involvement of participants, the development of shared mental models, and the equality of group member's participation in the decision making process. For instance, a more equal participation of group members might be achieved in Virtual Worlds since the presence of visual cues, which reflect group member's physical appearance (e.g. gender, ethnicity physical disability or even emotional state of being), can be controlled. Moreover, Virtual Worlds might facilitate decision making tasks in the context of spatial planning, since different parties are able to visually experience the final result of the zoning plan. For instance, group members are able to negotiate with each other while virtually walking in and around the to-be-built venue of a sushi bar, a hotel lobby or a city area.

A community council has to decide on the destination of an untitled piece of land located in their city area. A specific urban area, including the untitled piece of land, will be reconstructed on a virtual estate in *Second Life*. The members of the city council can walk though this virtual reconstruction of their city area and as such develop a zoning plan and decide on the future use of the untitled piece of land.

Figure 10 — Piece of land in city area

Keywords: zoning plan, spatial planning, interaction support, process support, shared mental model social acceptance, experience economy…

This case seeks to examine how Virtual Worlds can contribute to group decision-making in general and public decision making in spatial contexts in particular. Relevant issues in this context are: the effects of visual cues on eInclusion, involvement and equal participation of group members (e.g. to what extent does the ability to control visual cues influence decision making strategies?); what are the effects of a 3D environment on the process of group decision making; to what extent do Virtual Worlds support group interaction; and to what extent does the opportunity to virtually experience the final result of a zoning plan, facilitate more effective and efficient decision making?

Figure 11 — Group Decision Making in Virtual World

The product obtained of this case is a theoretical framework (conceptual model) of the determinants of effective group decision making in Virtual Worlds. This established framework can be applied to spatial planning in particular, and to a certain extent it can be generalized to other contexts [e.g. (strategic) business meetings].

3.3.4 Instantiation 3.4: Consumer Collaboration in Product Design Processes along the Supply Chain

During the past decade, the relationship between firms and consumers has transformed from a singular top down perspective towards a more equal, dialogue kind of approach. On the one hand, this shift has been stimulated by active consumers who seek interaction with firms, as they demand products that have been

customized to suit their personal preferences. For instance, the sneaker brand PUMA anticipated this consumer need by offering consumers the possibility to customize their favorite sneaker on an online platform called *The Mongolian Shoe BBQ.* On the other hand, businesses themselves seek the interaction and creative input of their end-users, as they are becoming more aware of the benefits of consumer interaction in the early stages of the product design process. Virtual Worlds can function as an interesting platform that easily brings together the different players in a supply chain, offering them a playing field for co-creating and virtually testing new products before launching them in real life.

Different members of a specific supply chain (e.g. a manufacturer and an architect) will virtually meet with consumers in a challenging 3D environment in order to engage in a “consumer driven” co-creation process.

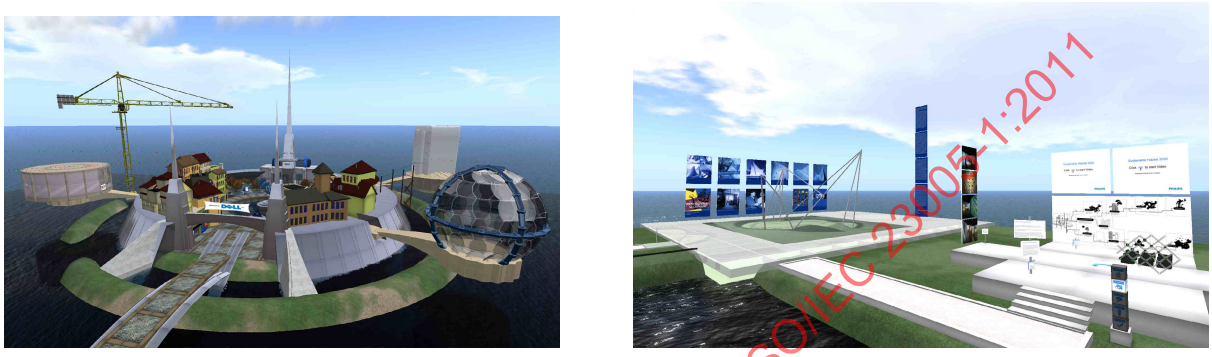

Figure 12 — Examples of manufacturers present in *Second Life* (Dell and Philips)

Keywords: innovation, crowd-sourcing, co-creation, audience participation, customized communication, social acceptance.

This use case seeks to explore the added value of Virtual Worlds as platforms for supporting processes in the supply chain such as co-creation and testing of new products. Interesting issues in this context are: how can firms anticipate specific underlying mechanisms of *involvement* in order to establish a successful co-creation relationship with consumers in Virtual Worlds; what determines the success of consumer participation in the product design process in a Virtual World; to what extent does electronic “word-of-mouth” communication by opinion leaders in Virtual Worlds influence 'the adaptation process' of new product ideas; and how can these co-creation relationships successfully be deployed for marketing purposes.

Figure 13 — The Decision Dome: Challenging group decision site in Second Life

The product of this case is obtained will be a theoretical framework (conceptual model) of the determinants that can used to achieve effective collaboration within the context of product innovation.

3.3.5 Instantiation 3.5: Virtual Objects

When users have information for a target virtual object, the information should be used in different virtual worlds such as 3D virtual world community and 3D game console that are illustrated in Figure 13. For example, the basic format for the visual information that is applied to a target avatar's clothes in the 3D virtual world community (e.g., Second Life) can also be used for the avatar, which has the same kind of avatar in the 'Second Life', in the 3D game console.

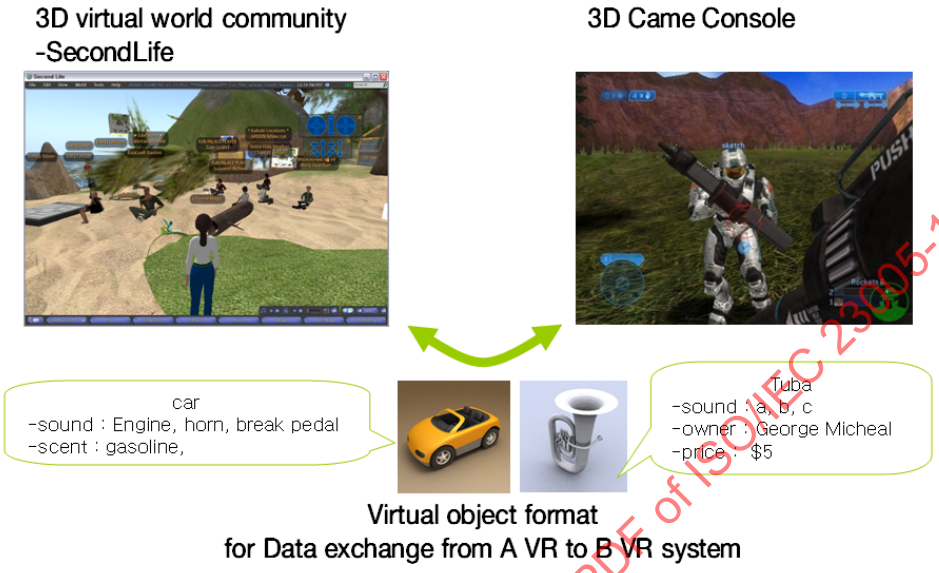

Figure 14 — Virtual object format for Data exchange

Consider the example of virtual shopping in Figure 14. The use case provides that virtual objects are controlled by input devices in real world. Virtual shop in 3D virtual world provides a realistic experience better than the existing web shopping malls which are based on product images. Instead of browsing some photos, a user manipulates a virtual product in 3 dimensions. It would provide a powerful realistic experience as novel online shopping services.

In order to provide virtual shopping services, the virtual products require a data structure which includes product classification form, manipulation/control method of a product, etc. The metadata of virtual objects can be used to information for a virtual shopping.

Virtual object meta data	Information for a virtual shopping mall
Object family	product classification form
Price/ owner	product base information
Control	manipulation/control method (translation, rotation, scaling)
Animation, sound, scent, haptic properties	product usage information

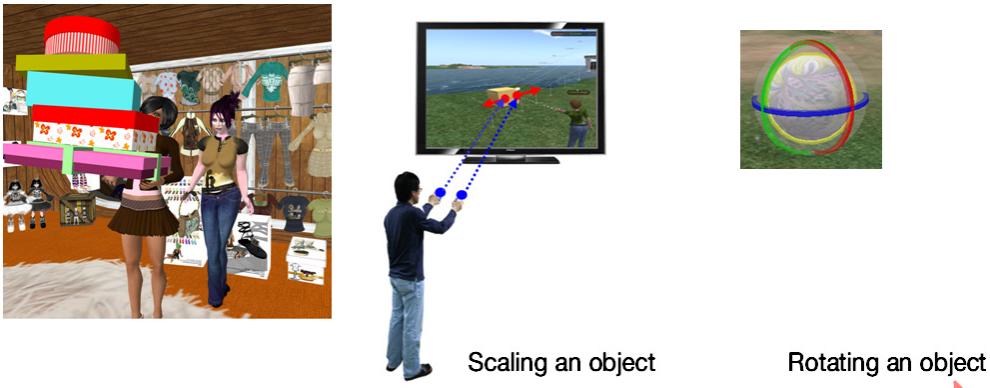

Figure 15 — Object manipulations for virtual shopping

For Virtual learning (v-learning) the virtual objects can be created as approximately as possible to imitate their counterparts in the real world because the newly proposed metadata is formed by considering the real objects' properties.

An example of virtual learning is a (virtual) sport lesson, (e.g., golf). Virtual golf club has the same physical property as the real one which a user owns. A user plays golf in the real world, and then virtual golf club reflects the same physics in virtual world as illustrated in Figure 15.

Figure 16 — Object manipulations for virtual learning (Golf)

3.4 Instantiation 4: Control of avatars and other virtual objects by real world signals

In this Subclause, several use case scenarios with haptic information are presented in order to describe how haptic contents can be applied in MPEG-V system through kinesthetic and tactile devices.

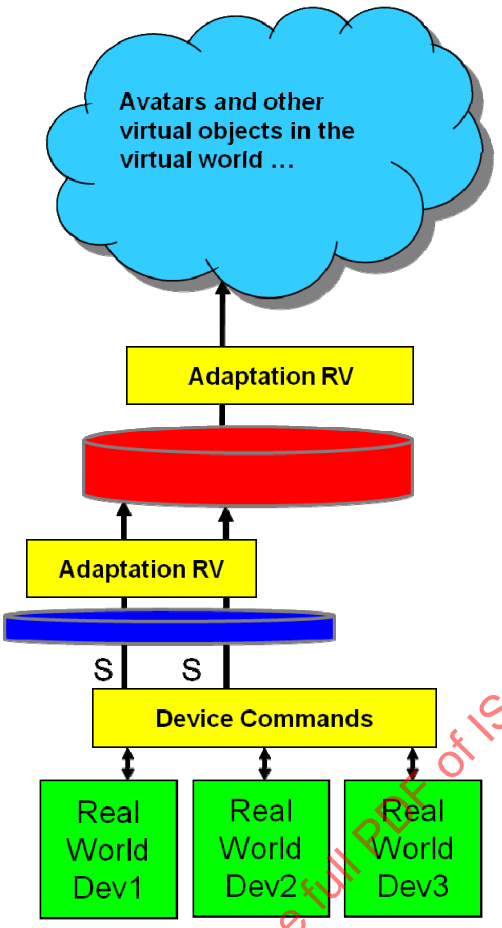

Figure 17 — (Possible) System Architecture for the control of avatars and other virtual objects by real world signals

3.4.1 Instantiation 4.1: Internet Haptic Service - YouTube, Online Chatting

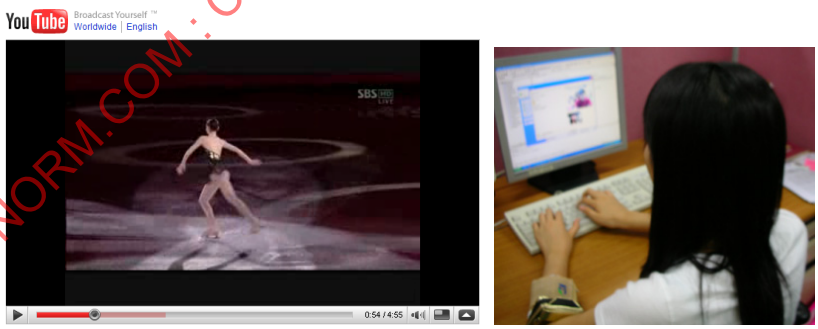

Figure 18 — Use Case Scenario for Internet Service

Description: A user is wearing a haptic wristband and browsing Internet.

- The user connects to www.youtube.com and searches for the term 'figure skating' and then selects one of the video which is Yuna Kim's performance played last week with haptic information authored by the user's favourite haptic director.

- Once the video is selected, the auto haptic display software checks how many arrays of tactile display the user is wearing and then provides the authored haptic information of Yuna Kim's performance video.
- After clicking the play button, the user can enjoy the music and skater's performance with haptic displays.
- At the time of spinning scenes of the skater, the arrays of haptic display circle patterns as if the skater spins around on the user's wrist.
- After the user finishes watching the video for YouTube, the user starts to chat with her boyfriend. She can send texts and look at the boyfriend's face and then send a haptic icon as well.
- After she sends the tickle haptic icon to him, she receives a heart haptic icon which has heartbeat sensation by haptic display.

3.4.2 Instantiation 4.2: Next Generation Classroom – Sensation Book

Description: The classroom is equipped with an immersive book, so called a sensation book, speakers, kinesthetic and tactile devices. During the class, students listen to a teacher's lecture and then they explore the sensation book.

Figure 19 — Use Case Scenario for Next Generation Classroom

- In the history class, students learn famous temples of Korea's Chosun Dynasty in Korea.
- While students read a part related to the bell in the book, a virtual bell is visually augmented on the image.
- Some students who want to hit the bell may hold a haptic device and touch the bell. In addition, the sound of bell may be out from a speaker. The harder students hit the bell, the louder sound comes.
- In the science class, students learn about lithology.
- As soon as students press 'feel the stones' button, several virtual sample stones (sandstone, shale, conglomerate, granite, whinstone) are loaded in front of them.
- By touching them with haptic device they can distinguish the difference among stones since each virtual stone provides its own hardness, roughness, friction to users.

3.4.3 Instantiation 4.3: Immersive Broadcasting – Home Shopping, Fishing Channels

Description: In digital home environment, a haptic device is connected to TV set-top box.

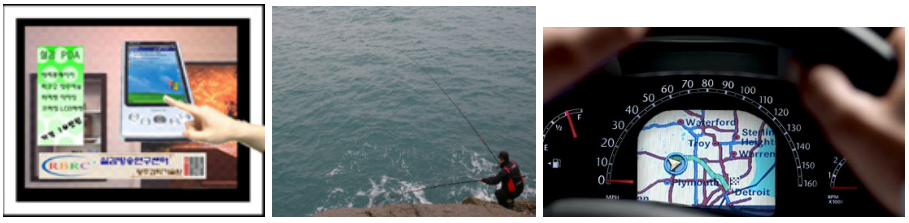

Figure 20 — Use Case Scenario for Immersive Broadcasting

- A viewer turns on TV and changes the channel to home shopping.
- The product, which is being sold, is a PDA and the 3D virtual object of PDA is displayed on TV screen.
- A shopping host explains functions of PDA and asks viewer to click the buttons.
- When the viewer who holds a haptic device clicks a button of the PDA, he or she can feel a button click sensation with the glassy body of the PDA.
- In addition, the viewer who holds the PAD can measure the weight and (s)he thinks it is portable.
- Finally, the viewer decides to purchase the PDA.
- After shopping, the user changes to a fishing channel.
- When a fish takes bait, the fishing pole starts to vibrate in the screen. Then the haptic device the viewer is holding also vibrates.
- In addition, the viewer feels force when the fisherman in the screen catches the fish.

3.4.4 Instantiation 4.4: Entertainment – Game (Second Life, Star Craft), Movie Theater

Description: In the digital entertainment room, a haptic movie system and an immersive game system are installed.

- A user comes to the entertainment room to watch a movie “Spiderman”.
- Video, Audio systems are turned on and the user wears tactile glove.
- Haptic information is authored synchronously with audio video information and is displayed together with the synchronized audiovisual media.
- The scene that a web is coming out of Spiderman, tactile device makes a straight pattern based on the authored tactile information in the pre-stage. Therefore, the user can feel a web coming out from his/her hand.
- After the movie ends, the user connects to the Second Life application.
- When the user shakes hands with a business partner in the Second Life the haptic device approaches to the users and provides shaking-hands motion and force feedback.
- After spending an hour, the user starts the game “Starcraft”
- While the user is attacking the enemy, (s)he feels right-bottom (5 O'clock) of back of the tactile vest is hot since the base of the user is firing and attacked by other enemies.