INTERNATIONAL STANDARD

ISO/IEC 23005-1

Third edition
2016-07-15

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 23005-1:2016(E)

ISO copyright office
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

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.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

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. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL: www.iso.org/iso/foreword.html.

The committee responsible for this document is ISO/IEC JTC 1, *Information technology*, SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

This third edition cancels and replaces the second edition (ISO/IEC 23005-1:2014), which has been technically revised.

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 (actuator) capabilities/commands, user’s sensory preferences, and

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

various delivery formats. The MPEG-V architecture can be applicable for various business models for which audiovisual contents can be associated with sensory effects that need to be rendered on appropriate sensory devices (actuators).

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

Part 1: **Architecture**

1 Scope

This part of ISO/IEC 23005 specifies the architecture of MPEG-V (media context and control), its three associated use cases of information adaptation from virtual world to real world, information adaptation from real world to virtual world, and Information exchange between virtual worlds.

2 Terms and definitions

2.1 Device Command

description of controlling actuators used to generate **Sensory Effects**

2.2 R ➔ V Adaptation

procedure that processes the Sensed Information from the real world in order to be consumed within the virtual world's context; takes the Sensed Information with/without the Sensor Capabilities from Sensors, the Sensor Adaptation Preferences from Users, and/or the Virtual World Object Characteristics from a Virtual world; controls the Virtual World Object Characteristics or adapts the Sensed Information by adapting the Sensed Information based on the Sensor Capabilities and/or the Sensor Adaptation Preferences

2.3 Sensed Information

information acquired by sensors

2.4 Sensor

device by which user input or environmental information can be gathered

EXAMPLES Temperature sensor, distance sensor, motion sensor, etc.

2.5 Sensor Adaptation Preferences

description schemes and descriptors to represent (user's) preferences with respect to adapting sensed information

2.6 Sensor Capability

description of representing the characteristics of sensors in terms of the capability of the given sensor such as accuracy, or sensing range

2.7 Sensory Device

consumer device by which the corresponding sensory effect can be made

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

2.8 Sensory Device Capability

description of representing the characteristics of actuators used to generate sensory effects in terms of the capability of the given device

2.9 Sensory Effects

effects to augment perception by stimulating human senses in a particular scene

EXAMPLES Scent, wind, light, haptic [kinesthetic-force, stiffness, weight, friction, texture, widget (button, slider, joystick, etc.), tactile: air-jet, suction pressure, thermal, current, vibration, etc. Note that combinations of tactile display can also provide directional, shape information].

2.10 Sensory Effect Metadata

metadata that defines the description schemes and descriptors to represent sensory effects

2.11 User's Sensory Preferences

description schemes and descriptors to represent (user's) preferences with respect to rendering of sensory effect

2.12 User

the end user of the system.

2.13 Virtual World

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

2.14 V ➔ R Adaptation

procedure that processes the Sensory Effects from the Virtual World in order to be consumed within the real world's context; takes Sensory Effect Metadata from a Virtual World, Sensory Device (Actuator) Capabilities from the Sensory Devices (Actuators), the User's Sensory Preferences from users, and/or the Sensed Information as well as the Sensor Capabilities from Sensors as inputs; generates the Device Commands by adapting the Sensory Effects based on the Sensed Information, the Capabilities and/or the Preferences

2.15 VW Object Characteristics

description schemes and descriptors to represent and describe virtual world objects (from the real world into the virtual world and vice versa)

3 MPEG-V System Architecture

Figure 1 — 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 world technologies we need a better understanding of their internal economics, rules and regulations. The overall system architecture for the MPEG-V framework is depicted in エラー! 参照元が見つかりません.

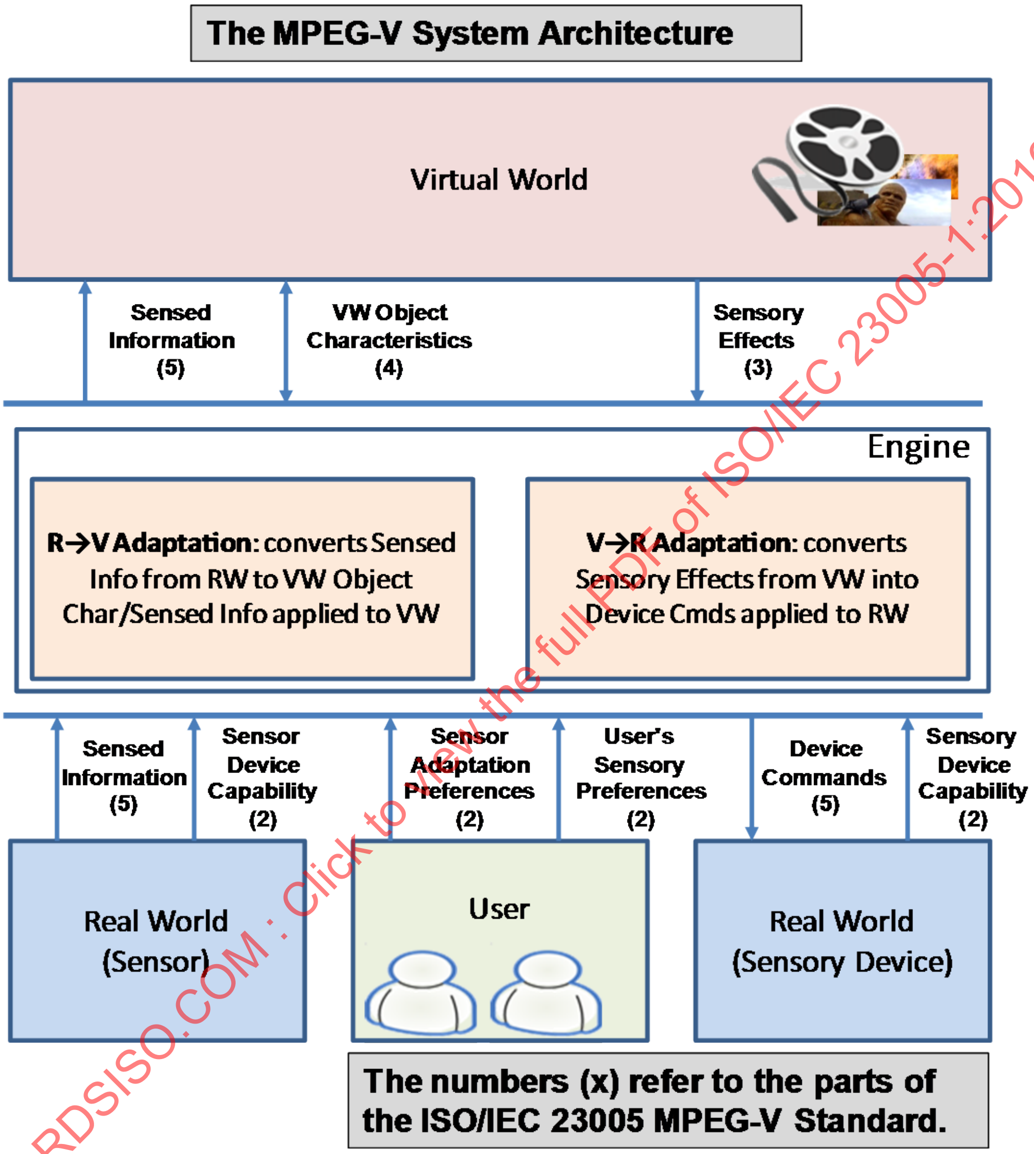


Figure 1 — System Architecture of the MPEG-V Framework

The MPEG-V System Architecture can be used to serve three different media exchanges. There are two types of media exchanges occurring between real world and virtual world, i.e., the information exchange from real world to virtual world and the information exchange from virtual world to real world. An additional type of media exchanges is the information exchange between virtual worlds. The three media exchanges are defined as use cases in Clause 4.

It is important to note that *Sensory Effect Metadata*, *Sensory Device Capability*, User's *Sensory Preferences, Device Commands*, *Sensed Information, Sensor Device Capability, Sensor Adaptation Preferences, and Virtual World Object Characteristics* are within the scope of standardization and, thus shall be normatively specified. On the other side, the *V→R Adaptation Engine, V→R Adaptation Engine, Virtual World,* as well as *Devices (Sensors and Sensory devices)* are informative and are left open for industry competition.

Metadata within the scope is formed other parts of the ISO/IEC 23005. *Sensor Device Capability, Sensory Device Capability, Sensor Adaptation Preferences*, and User's *Sensory Preferences* are specified in Part 2: Control information. *Sensory Effect Metadata* is specified in Part 3: Sensor information. *Virtual World Object Characteristics* is specified in Part 4: Virtual world object characteristics. *Device Commands* and *Sensed Information* are specified in Part 5: Formats for interaction devices.

4 Use cases

The three media exchanges require information adaptations in order for a targeting world to adapt information based on capabilities and preferences: information adaptation from virtual world to real world, information adaptation from real world to virtual world, and information adaptation between virtual worlds.

4.1 Information adaptation from virtual world to real world

- **System Architecture for information adaptation from virtual world to real world**

The system architecture for the information adaptation from virtual world to real world is depicted in **Figure 2**. It represents V→R adaptation comprising *Sensory Effect Metadata*, *VW Object Characteristics*, *Sensory Device Capability (actuator capability)*, *Device Commands*, User's *Sensory Preferences*, and a *V→R Adaptation* Engine which generates output data based on its input data.

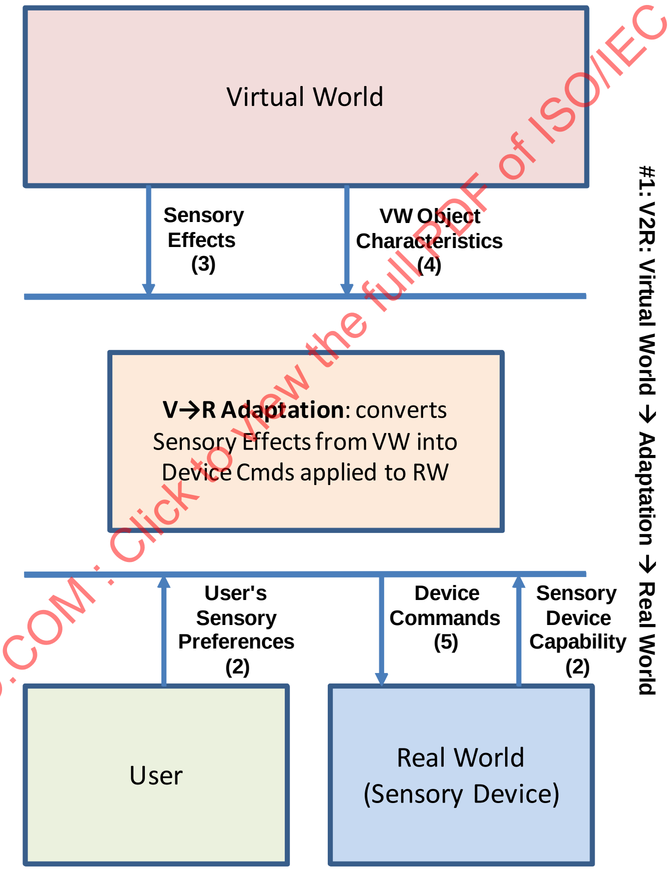


Figure 2 — (Possible) System Architecture for information adaptation from virtual world to real world

A *virtual world* within the framework is referred to as an entity that acts as the source of the *sensory effect metadata* and *VW Object Characteristics* such as a broadcaster, content creator/distributor, or even a service provider. The *V→R Adaptation Engine* is an entity that takes the *sensory effect metadata*, the *sensory device (actuator) capability* and the user's *sensory preferences* as inputs and generates the *device commands* based those in order to control the consumer devices enabling a worthwhile, informative experience to the user.

Real world devices (sensory devices) are entities that act as the sink of the *device commands* and act as the source of *sensory device (actuator) capability*. Additionally, entities that provide user's *sensory preferences* towards the *RoSE* engine are also collectively referred to as *real world devices*. Note that *sensory devices* (actuators) are sub-set of *real world 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., effects 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 capability* defines description formats to represent the characteristics of sensory devices (actuators) in terms of which sensory effects they are capable to perform and how. A *sensory device* (actuator) is a consumer device by which the corresponding *sensory effect* can be made (e.g., lights, fans, heater, fan, etc.). *Device commands* are used to control the *sensory devices (actuators)*. As for *sensory effect metadata*, also for *sensory device (actuator) capability* and *device commands* corresponding means for transporting this assets are referred to as sensory device capability/commands delivery format respectively.

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

4.2 Information adaptation from real world to virtual world

- **System Architecture for information adaptation from real world to virtual world**

The system architecture for information adaptation from real world to virtual world is depicted in Figure 3. It represents R2V adaptation comprising *VW Object Characteristics*, *Sensed Information*, *Sensor Capability*, *Sensor Adaptation Preferences*, and an *R→V Adaptation Engine* which generates output data based on its input data.

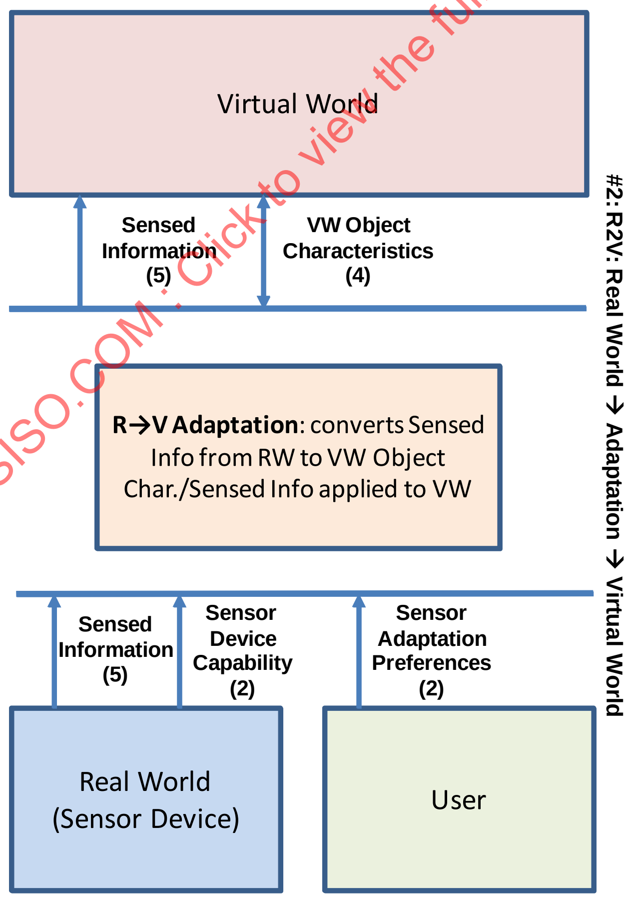


Figure 3 — (Possible) System Architecture for information adaptation from real world to virtual world

Entity that processes the Sensed Information from the real world in order to be consumed within the virtual world's context; takes the Sensed Information with/without the Sensor Capabilities from Sensors, the Sensor Adaptation Preferences from Users, and/or the Virtual World Object Characteristics from a Virtual world; controls the Virtual World Object Characteristics or adapts the Sensed Information by adapting the Sensed Information based on the Sensor Capabilities and/or the Sensor Adaptation Preferences.

There are two possible implementations to adapt information from real world to virtual world. In the first system implementation, *R→V adaptation* takes the *Sensor Capabilities* as inputs, the *Sensed Information* from *Sensors*, and *Sensor Adaptation Preferences* from *Users*; adapts the *Sensed Information* based on the *Sensor Capabilities* and/or *Sensor Adaptation Preferences*.

In the second system implementation, R→V adaptation takes the *Sensed Information* with/without the *Sensor Capabilities* from *Sensors*, the *Sensor Adaptation Preferences* from *Users*, and/or the *Virtual World Object Characteristics* from a *Virtual world*; controls the *Virtual World Object Characteristics* adapting the *Sensed Information* based on the *Sensor Capabilities* and/or the *Sensor Adaptation Preferences*.

4.3 Information exchange between virtual worlds

- **System Architecture for exchanges between virtual worlds**

The system architecture for information exchange between virtual worlds is depicted in Figure 4. It represents information exchange comprising *VW Object Characteristics* which generates exchangeable information within virtual worlds.

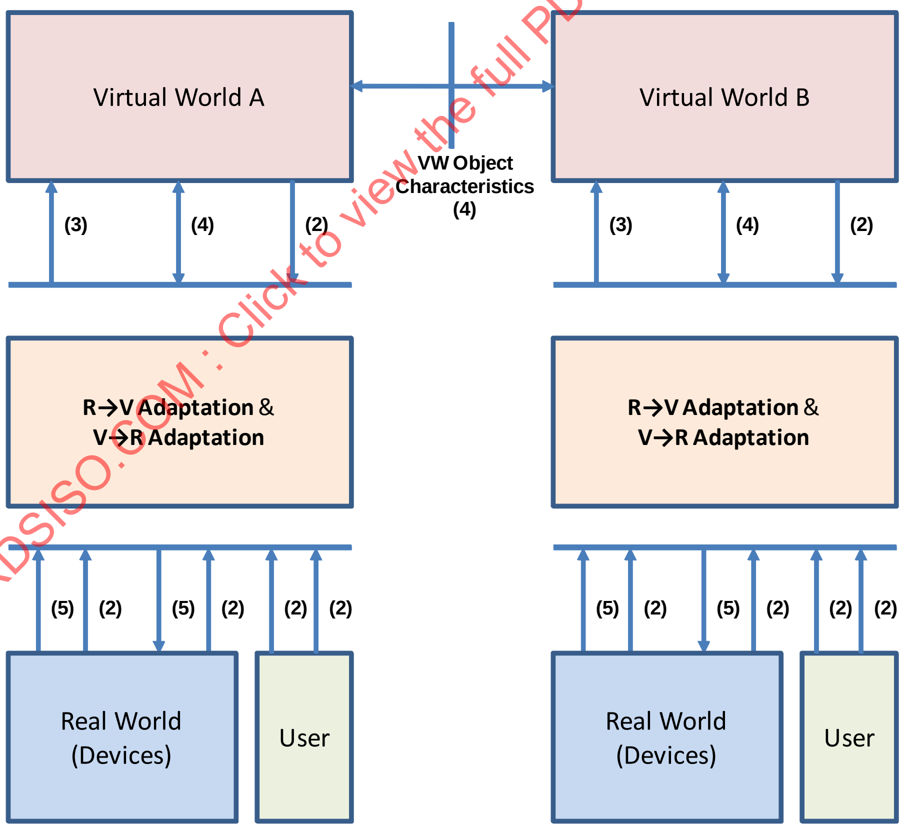


Figure 4 — (Possible) System Architecture for (bidirectional) exchange of information between virtual worlds

V→V adaptation adapts proprietary virtual world object characteristics from a Virtual World to *VW Object Characteristics* and sends the *VW Object Characteristics* from the *Virtual World* to another *Virtual World* to support interoperability. Based on the data provided in *Virtual World Object Characteristics*, the *Virtual World* will internally adapt its own representation for virtual object/avatar.

5 Instantiations

5.1 Instantiation A: Representation of Sensory Effects (RoSE)

- **System Architecture for Representation of Sensory Effects**

The system for representation of sensory effects is partly instantiated from the system architecture of information adaption from virtual world to real world. The overall system architecture for Representation of Sensory Effects (RoSE) is depicted in Figure 5 comprising Sensory Effect Metadata, Sensory Device (actuator) Capability, Device Commands, User's Sensory Preferences, and a so-called RoSE Engine which generates output data based on its input data.

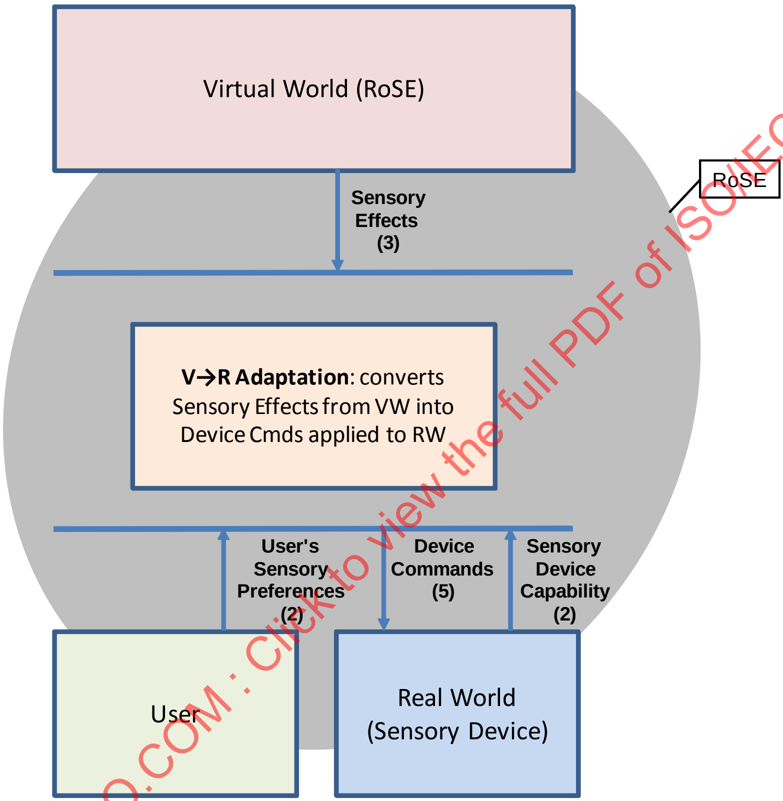


Figure 5 — RoSE System Architecture

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 (actuator) capabilit*y and the user's *effect preferences* as inputs and generates the *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 *device commands* and act as the source of *sensory device (actuator) capability*. Additionally, entities that provide user's *sensory preferences* towards the *RoSE* engine are also collectively referred to as *consumer devices*. Note that *sensory devices (actuators)* 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., effects 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 (actuator) capability* defines description formats to represent the characteristics of sensory devices (actuators) in terms of which sensory effects they are capable to perform and how. A *sensory device (actuator)* is a consumer device by which the corresponding *sensory effect* can be made (e.g., lights, fans, heater, fan, etc.). *Device commands* are used to control the *sensory devices (actuators)*. As for *sensory effect metadata*, also for *sensory device (actuator) capability* and *device commands* corresponding means for transporting this assets are referred to as sensory device (actuator) capability/commands delivery format respectively.

Finally, the user's *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.

- **Instantiation A.1: Multi-sensorial Effects**

Traditional multimedia with audio/visual contents have been presented to users via display devices and audio speakers as depicted in Figure 6. 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.

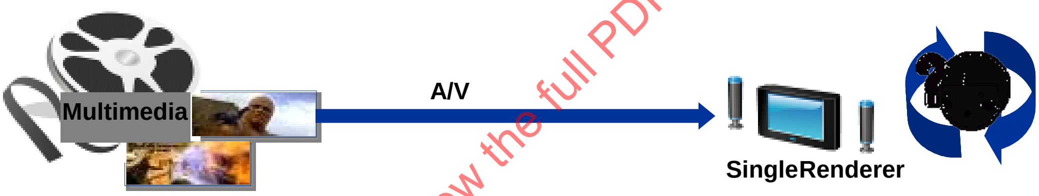


Figure 6 — Traditional Multimedia Consumption

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 7 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 7. In this configuration, the multimedia contents are not rendered by a single device but with multiple devices in a synchronized manner.

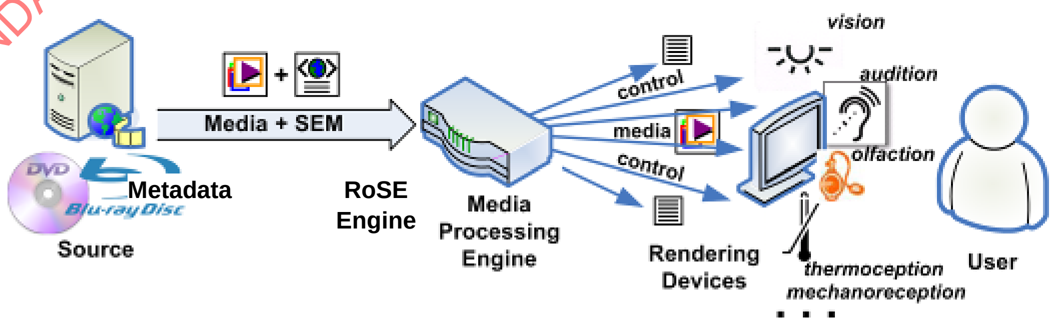


Figure 7 — RoSE-enabled Multimedia Consumption for Advanced User Experience

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.

- **Instantiation A.2: Motion effects**

One of the important sensory effects that we should not ignore is the effect related to the motion. The motion effect gives a user a similar feeling on the movement like the actor/actress feels in the movie. The motion effect is popular sensory effects commonly used in such a places like theme park, game room, and the movie theater now a day. The motion effect is usually provided by the motion chair. The motion chair usually has motor(s) and axis underneath or above the chair. The number of motor and the length of axes determine the range and depth of the movement of the chair. There are a lot of manufacturers of motion chair in the world and each of them has its own mechanical characteristics as shown in Figure 8 below.

Figure 8 — Various motion chairs and their characteristics

For example, the motion effect chair of the Simnoa Technologies provides several types of motions including tilt motion (pitch) with various speed and accelerations, fast falling, continuous wave motion at variable speeds, swaying sideways in variable speeds and accelerations, vibration, and combination of wave and sway motions. The 4D chair from Changjin also supports falling down, rolling, and pitching motion with speed, but not yawing or forward/backward move motion. On the other hand, the 4D chairs from Fantawild or Acouve only supports vibration and falling effect.

Therefore, when designing the sensory effect of motion, our thoughts shall not be limited by the capabilities of a specific chair. We have to consider that the author produces sensory effect metadata based on the audio visual data, and he does not know the mechanical characteristics of the motion chair that the sensory effect will be rendered. As a matter of fact, the author does not know in which device the motion sensory metadata will be rendered. This means that the motion sensory metadata should not restrict the actual movement of motion chair. It is proper to describe the conceptual motion in the scene. Figure 9 shows an example to explain why the motion sensory effect should be conceptual. For example, the author wants to express “Turn left” motion effect. There are two kinds of motion chair. First chair supports rolling, yawing, and surging. On the other hand, second chair only supports rolling. If motion effect (SI) is expressed with physical term like “Yawing 90 degrees”, it depends on the capabilities of chair whether the chair can render it or not. However, If motion effect (SI) is expressed with conceptual term like “Left turn”, both of chair can render is with its own capability.

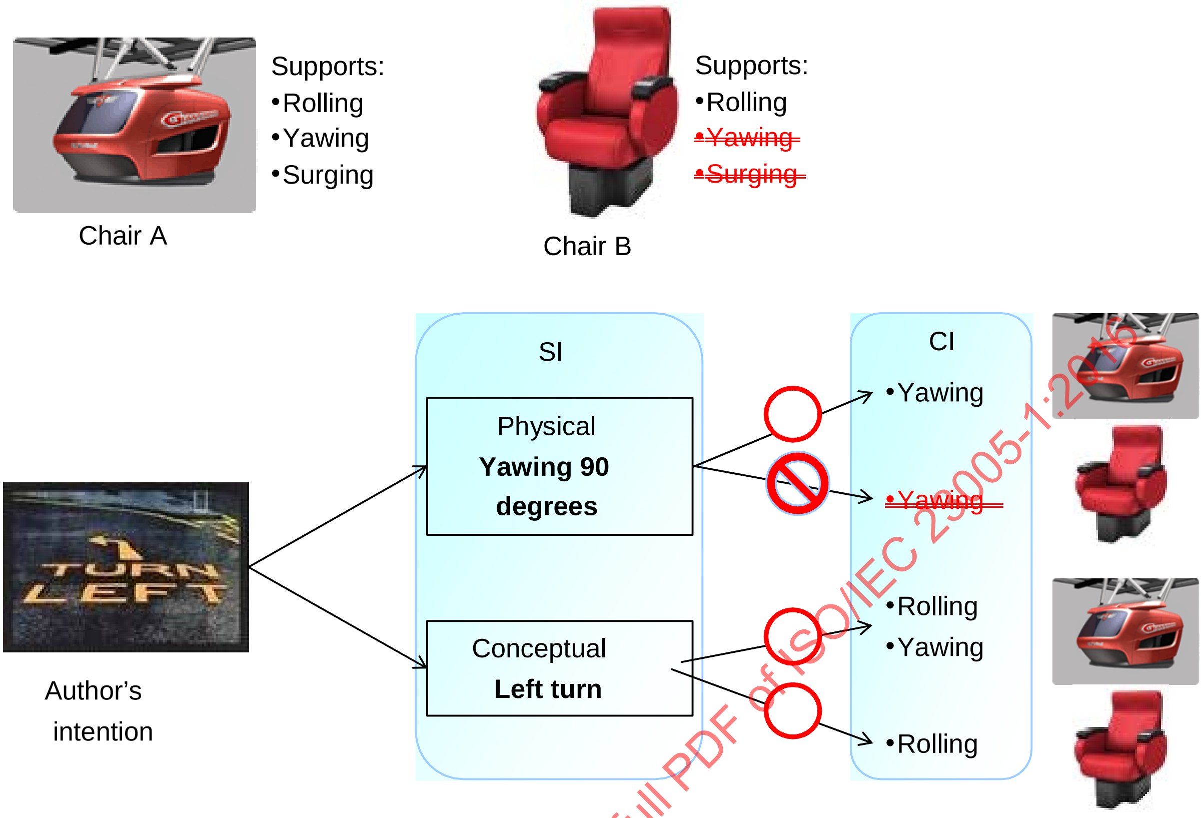
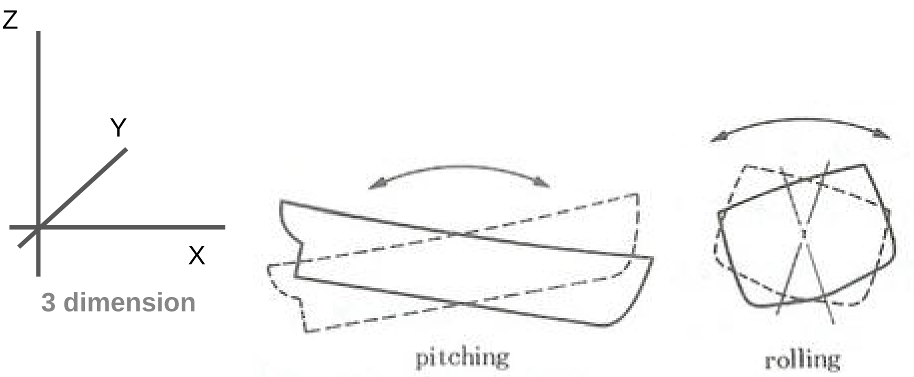


Figure 9 — Conceptual Vs. Physical SI

In other words, considering the process of adaptation by the engine, it is recommended for the SI to carry the semantics of the motion effect or the intention of the author, so that the adaptation engine can find the best combination of command information to satisfy the author's intention under the given restrictions of the specific motion effect chair.

Proposed schema for motion effect is based on "**Six Degrees of Freedom (6DoF)**" which is commonly used for motion description in robotics and engineering. 6DoF is composed of 3 dimensional axes, pitch, yaw, and roll as shown in Figure 10. It is a well known fact that any motion of a rigid body can be described by the 6 DoF motion. We abstracted the 6 DoF motions into several **basic motion patterns** and also added more **combinational motion patterns,** which are based on the repetition or a combination of basic patterns and have specific semantics.



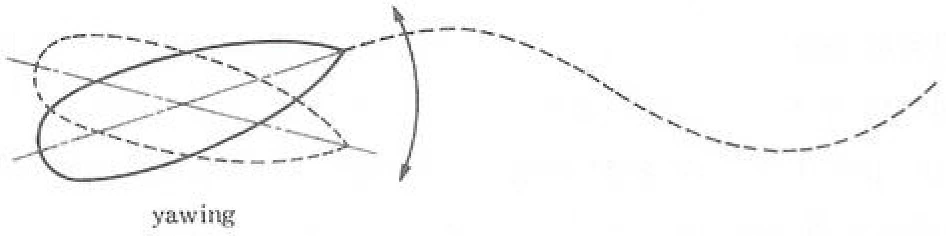


Figure 10 — Six degrees of freedom

5.2 Instantiation B: Natural user interaction with virtual world

- **System Architecture for Natural user interaction with virtual world**

The system for natural user interaction with virtual world is instantiated from the system architecture of information adaption from real world to virtual world. The sensors for such interaction include gaze tracking sensors, multi-point sensors, smart cameras, motion sensors, gesture recognition sensors, intelligent cameras, etc. *R→V adaptation engine* analyzes the interaction intention from the information from the *sensors* and adapts *VW Object Characteristics* and/or the *sensed information* to send the intention to *a virtual world.*

- **Examples of sensors**
  - **Gaze tracking sensors**

Eye tracking is the process of measuring either the point of gaze ("where we are looking") or the motion of an eye relative to the head. An eye tracker is a device for measuring eye positions and eye movement. Eye trackers are used in research on the visual system, in psychology, in cognitive linguistics and in product design. There are a number of methods for measuring eye movement. The most popular variant uses video images from which the eye position is extracted. Other methods use search coils or are based on the electro-oculogram.

In the video-based eye trackers, a camera focuses on one or both eyes and records their movement as the viewer looks at some kind of stimulus. Modern eye-trackers use contrast to locate the center of the pupil and use infrared and near-infrared non-collimated light to create a corneal reflection (CR). The vector between these two features can be used to compute gaze intersection with a surface after a simple calibration for an individual.

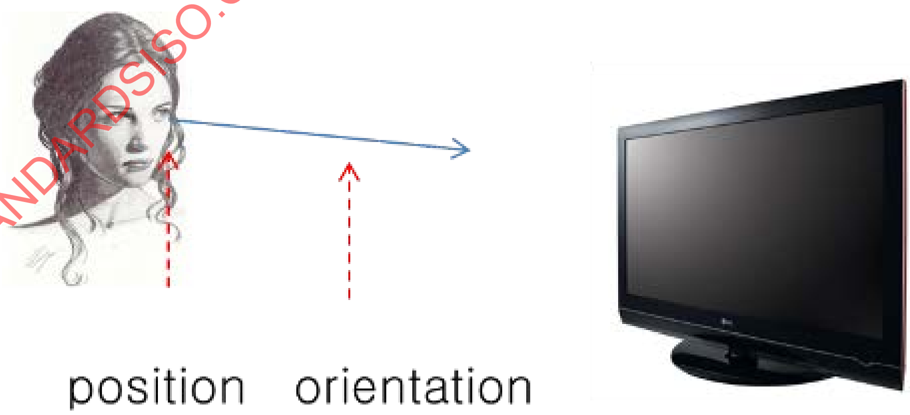


Figure 11 — Main components sensed by gaze tracking sensors

In recent years, the increased sophistication and accessibility of eye tracking technologies have generated a great deal of interest in the commercial sector. Applications include web usability, advertising, sponsorship, package design and automotive engineering.

- **Multi-pointing sensors**

One of the most useful devices to interact with display devices is a kind of remote controllers. In general, people can operate a display device through a remote controller to indicate a position or to order a command. As the recognition technology evolves, new kinds of remote controllers are being introduced. For instance, smart phones and intelligent tables already support multi touching features. Moreover, the gesture recognition technology can be used to handle the objects on a screen. Multi-pointing sensors are the sensors to detect the positions of feature points and the status of buttons. These types of sensors are a mouse, gesture recognition device, multi-touch device and so on.

Figure 12 — Examples of multi-pointing sensors, mouse, touch screen, and gesture recognition device

- **Smart camera**

Smart cameras are vision systems that are, next to the base line positioning sensors like GPS, Gyro scope, and Magneto scope, are equipped with sensors that are capable of doing imagery, sensing human facial expression and body motion and the like.

- **Instantiation B.1: Full motion control and navigation of avatar/object with multi-input sources**

Figure 13 — 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. Figure 13 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.

- **Instantiation B.2: 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. An individual's physical as well as psychological well being is dependent on daily exercise giving one a sense of self-efficacy and independence. 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 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.

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. The agent takes on the role of a Virtual Exercise Coach (VEC), and takes a leadership role in prompting exercise, guiding, correcting and reinforcing the right behavior on the basis of user history and actual user needs. The appearance of the VEC can be selected and shaped by the user to fit his/her profile (sex, age, ethnicity language, personality). What is obviously very important is that the behavior of the agent is convincing, natural, helpful, and suited for the coaching task.

Coaching agents do exist already, although often the interaction between agent and human is limited. It is clear that we will see much improved agents in the near future that show more natural behavior, better facial expressions, better and more natural bodily behavior including non-verbal behavior. The simple trainers that we see today might very well evolve into full fledges games, where one exercises within an interesting virtual landscape, possibly together with other users.

Figure 14 — VR and Avatars in ambient assisted living

Avatars representing other members of a community join in the virtual environment for exercise. Social interaction and community building is facilitated in meeting each other and sharing experiences. It is of utmost importance that the users can communicate real with each other to harvest the motivational peer resources.

Clearly, there is the need for detailed multi-modal control of virtual humans, along the lines of the Saiba framework. Virtual Humans playing the role of a training coach should be able to perform detailed and complex body animation, proper synchronized speech and facial expressions and appropriate body poses. For uses communities, possibly employing different virtual worlds as a platform, this type of multi-modal behavior should be communicated from one virtual world to another. And for realization of behavior by means of some animation engine, there should be proper communication between the animation engine and system components that plan such behavior.

- **Instantiation B.3: Gesture recognition using multipoint interaction devices**

In the information technology dominated era, new interaction devices are introduced virtually every year. For instance, multi-point mechanism is not only applicable to special devices for a special application but also easily confronts in everyday use consumer electronics such as touchpad, advanced remote controllers. Therefore, it is necessary to have a common data format to represent its characteristics and beyond multi-point device, lots of interaction devices are waiting for being standardized which are considered as interfaces on existing scene description standards to utilize this already popular interaction mechanism. These new classes of devices require a set of additional data formats for advanced user interaction such as a gaze sensor representing the position of the eye and the direction of the gaze.

Figure 15 — Intelligent table with multi-touch screen

Figure 15 shows a commercial device which interacts with users through a touch-screen. Users simply operate this device with their fingers and the device detects the position of the touched screen and specific toggle motions such as tapping which are used as buttons.

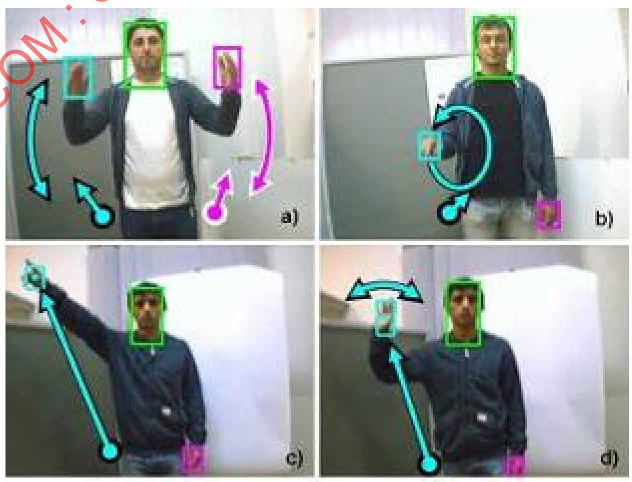


Figure 16 — An example of gesture recognition (ref.: http://paloma.isr.uc.pt/gesture-reco/)

One of other application of multi-pointing sensor is gesture recognition. Figure 16 shows some gesture examples. User may represent their intention using their hand or finger and the sensor detects the position of hands and fingers and specific motion such as holding.

- **Instantiation B.4: Avatar facial expression retargeting using smart camera**

The cameras in this use case are sensors that are capable of doing imagery, sensing human facial expression and body motion and the like. These cameras may also provide other types of sensed information from integrated sensing modalities such as global positioning sensor (GPS) and / or orientation (Gyro, Movement or Magnetic) sensors. These intelligent cameras have a great potential to use it as a mean of interaction devices. Field of applications includes facial expression retargeting and cloning, tele-presence, and augmented-reality.

Controlling a facial expression of an avatar by the user's facial movement has important interaction implications in virtual reality. When source data (features obtained from a camera) and a target facial model of an avatar do not have the same facial configuration, retargeting a facial expression is necessary to create appropriate facial animations to the target facial model. It defines facial expression basis which controls a distinct part of facial morphology such as eye opening and mouse opening, instead of using each feature point of face.

Specifically, there are two steps for the facial expression retargeting shown in Figure 17 (a). The first step is an initialization to normalize a facial expression of an individual human face. Initialization shall include the distances between the key feature-points (a facial morphology), the default value of each expression unit when there is no expression (neutral face) on face and the range of each basis in order to map the ranges of the facial expression of the individual user and the one of a target avatar. Second, once the initialization is finished, the sensed value of each expression basis from the intelligent camera can directly control each corresponding facial expression basis of an arbitrary avatar face. Unfortunately, the range of each expression basis cannot be easily obtained without user's various facial expressions and it requires time consuming efforts in the initialization stage. In order to reduce the painful initialization stage shown in Figure 17 (b), the intelligent camera may send its facial expression characteristics to the database in order to reuse the users' facial characteristics and download the initialization data onto the other intelligent cameras to avoid the initialization phase.

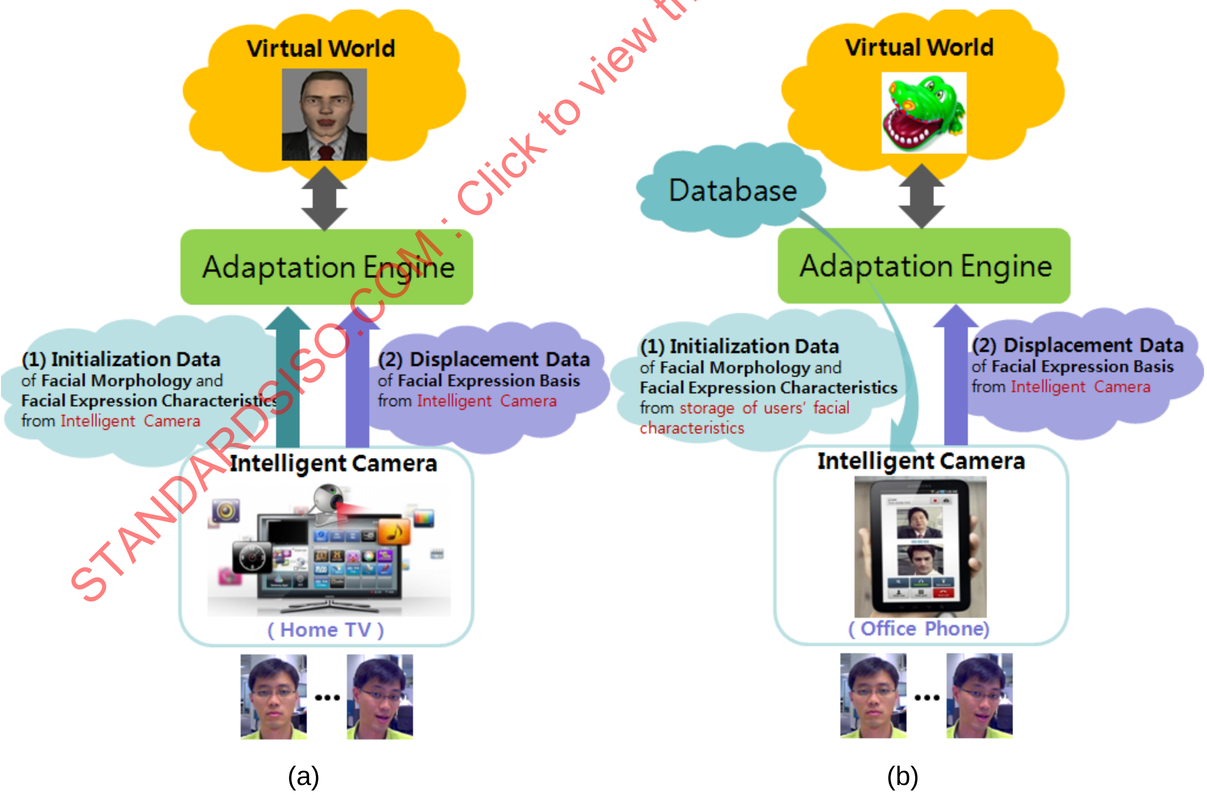


Figure 17 — Use case of intelligent camera for facial expression retargeting

- **Instantiation B.5: Motion tracking and facial animation with multimodal interaction**

  - **Motion tracking**

Motion tracking is now possible with various devices. The result of motions tracking is a detailed dataset, specifying motion data at a rather low level of abstraction. In order to use such data in multimodal context, where it needs to be fused with other data streams, and where synchronization is important, one needs an intermediate level of detail where meaningful phases and meaningful synchronization points can be annotated and communicated.

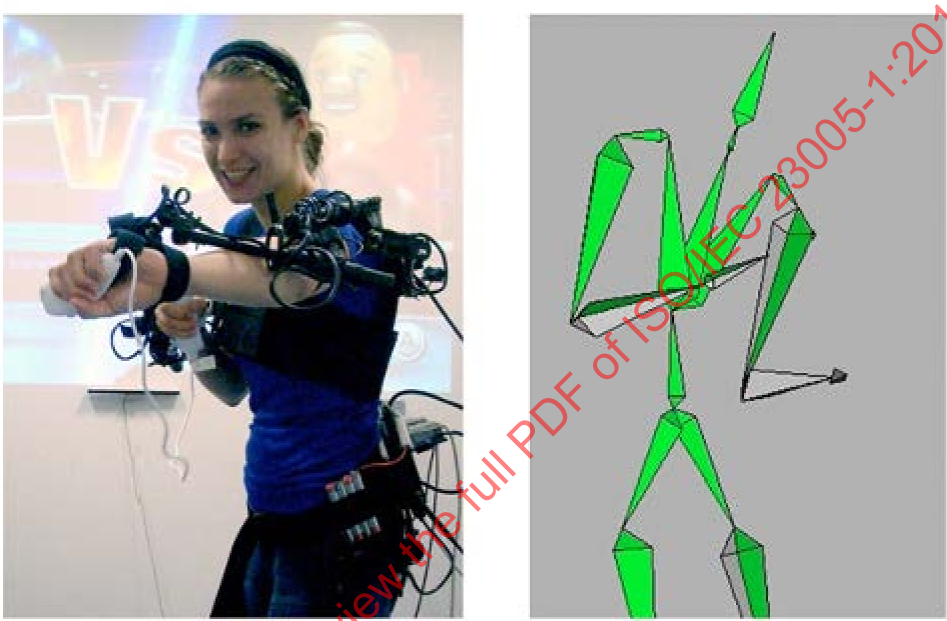


Figure 18 — Detailed motion tracking for VR

  - **Facial animation**

Facial animation is addresses within MPEG-4, for instance, by defining Face Definition Parameters and face Animation Parameters.

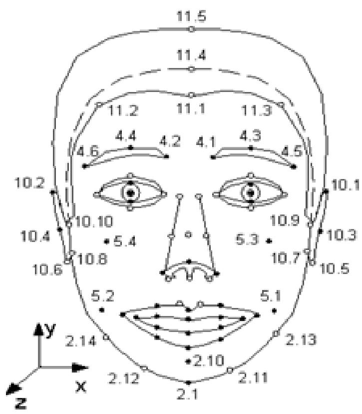


Figure 19 — MPEG-4 FDP and FAP

- **Instantiation B.6: Serious gaming and training with multimodal interaction**

A well known example of serious gaming is the ``Mission Rehearsal Exercise System'' created at the University of Southern California. (Hill, Gratch, Marsella, Rickel, Swartout, Traum) The goal was to create an experience learning system, where believable embodied conversational agents can talk and interact with trainees. The typical scenario that is simulated by the system is that of a lieutenant has just arrived in a Balkan town. One of his soldiers has been involved in an accident with a civilian vehicle.

The challenging situation is meant to become more familiar with the local culture, how to handle intense situations with civilians and crowds and the media, and how to make decisions in a wide range of non-standard (in military terms) situations. From a more technical point of view, multi-modal interaction, detailed timing and synchronization between modalities are highly important in order to create believable avatars and scenarios.

Figure 20: Scene from a serious game

- **Instantiation B.7: Virtual museum guide with embodied conversational agents**

Virtual guides within exhibitions or museums are typical candidates for Virtual Humans where great demands are put on the quality and naturalness of interaction. The final goal is to create systems that do not rely on simple pre-planned scripted multimedia, but rather to have more spontaneous interaction between visitor and guide. It is clear that detailed and high quality control of speech, gesturing, and facial animation is required.

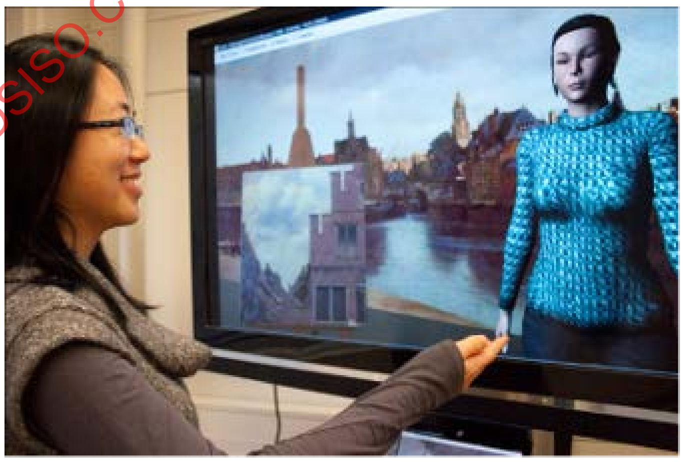


Figure 21 — Interaction with a virtual museum guide

5.3 Instantiation C: Traveling and navigating real and virtual worlds

- **System Architecture for traveling and navigating real and virtual worlds**

The system for traveling and navigating real and virtual worlds can be instantiated from the two use cases such as the system architecture of information adaptation from real world to virtual world and the system architecture of information adaptation from virtual world to real world.

- **Examples of sensors and path finding mechanisms**

  - **Attributed coordinate sensors**

A position sensor is any means that permits position measurement. It can either be an absolute position sensor or a relative one (displacement sensor). Position sensors can be either linear or angular. Some position sensors that available today GPS, GSM based triangulation, Wi-Fi based triangulation, Inductive Non-Contact Position Sensors, Proximity sensor and the like.

The attributed coordinate sensors combine the position information with additional (vital) information about an event taking place at that particular location and time like an e-call for an automotive vehicle (including vital information (vehicle conditions and / or passenger conditions) with respect to the accident next to the position information) or a heart failure related emergency call (including vital body signs next to the positioning information).

  - **Path finding**

Navigating within Virtual Worlds and or of unmanned vehicles in real the world Life can be hard, especially for novice users. A mechanism is proposed in which the user can easily navigate through the virtual world (i.e. Second Life) or real world without directly controlling the motion of the avatar or unmanned vehicle. He/she should be led to and along (interesting for various reasons) places in a smooth way. The user can travel through the virtual world by using a *flying device*, such as a *virtual air-scooter* or *magic carpet*. This device contains a dashboard with a map of the world. By clicking on a place on the map, the device flies to the indicated position, using a smooth motion, avoiding collisions and leading the user along interesting locations on the route. Together with the scooter's path, a smooth camera motion is provided which gives the user a clear view of the environment. Such a motion consists of camera placements and camera aim positions. Hence, three (linked) paths are specified for a tour. Besides this personal tour, the user can also choose for a more general guided tour.

  - **General technical details**[2]

An algorithm for camera planning needs to have a simplified geometrical representation of the world. Initially, this will be the 2D footprint of the world, which is made up of a collection of geometric primitives, such as points, lines, polygons, and disks, all lying in the ground plane. While the planning is performed in 2D, the motions can be projected on an elevated terrain. Next, based on the footprint, a data structure (i.e. the *Explicit Corridor Map* (ECM)), needs to be build which represents the collision-free space in the world. In summary, the ECM is the *medial axis* G = (V,E), where V and E are its Voronoi vertices and Voronoi edges. The edges are annotated with event points (placed at the intersections of the obstacle normals and Voronoi edges) together with their closest points to obstacles (i.e. obstacle normals). This data structure is generated automatically using the graphics card, allowing for real-time computations. Its size is optimal, i.e. linear in the number of vertices describing the obstacles.

Given an ECM, the radius of a character, and a query (consisting of a start and goal position), a two-dimensional corridor is extracted which provides the global route for the character (or flying device device). Next, a shortest path is extracted that has a user-specified amount of minimum clearance to the obstacles. These calculations are *optimal*, i.e. linear in the number of cells describing the corridor. This path acts as an *Indicative route* for the character, providing the input for the *Indicative route method* (IRM). The IRM uses the

---

[2] More technical details on 'Planning Short Paths with Clearance using Explicit Corridors' as described above can be found at: http://people.cs.uu.nl/roland/motion_planning/ecm.html

corridor, indicative route, and query to generate a smooth path. This path is the final path of the character, or flying device depending on the use case. The next phase is the extraction of the camera path given the ECM and the character's final path.

- **Automatically obtaining of the height map (extension)**

A camera planning system has been implemented in Second Life (SL) to realize a virtual tour guide. A '*scanner'* has been developed that gives a height map of an area in SL, which is needed as basis for the path finding algorithm. The scanner collects data by flying over an island and shooting probes. The data is collected and transferred to a web server running a PHP script. The scanner makes can be used with any virtual environment (such as SL), making it a powerful tool. After transferring the height map, it is converted into a bitmap. From this bitmap, its footprint, consisting of polygons, is extracted. After this footprint is fed into the ECM software (see above), which resides on the server, it is ready to be queried from SL.

- **Instantiation C.1: 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 22 — 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.

- **Instantiation C.2: 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:

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

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.

- **Instantiation C.3: Virtual tour guides**

Virtual tours are mechanisms which support the user to easily navigate through the virtual world (i.e. Second Life) without directly controlling the motion of the avatar. He/she should be led to and along interesting places in a smooth way.

The user can travel through the virtual world by using a *virtual air-scooter*. This scooter contains a dashboard with a map of the world. By clicking on a place on the map, the scooter flies to the indicated position, using a smooth motion, avoiding collisions and leading the user along interesting locations on the route. Together with the scooter's path, a smooth camera motion is provided which gives the user a clear view of the environment. Such a motion consists of camera placements and camera aim positions. Hence, three (linked) paths are specified for a tour.

Besides this personal tour, the user can also choose for a more general guided tour. Later, other team members could build in technology to get information about the interesting spots. In addition, technology for multi-lingual or text input could be could be integrated.

A virtual tour brings the user from its current location to a goal location in the virtual world. Besides these two locations, a tour may additionally consist of a series of intermediate locations. To construct a virtual tour, we need user input for retrieving the name of the selected tour or a specific goal location. A tour consists of three smooth paths (and each path consists of a series of 3D coordinates):

- The character path. This is a path which is traversed by the scooter.
- The camera path. This is a path which is traversed by the camera. The camera must keep some minimal distance to the scooter and obstacles in the virtual world so that the user has a clear view of the environment.
- The camera's aim path. This is the series of locations which are being looked at during the tour. To reduce the risk of getting motion-sick, the camera- and aim combination must move smoothly. In addition, the camera must look at future positions (about 1 second ahead) such that the user can anticipate about what is going to happen.

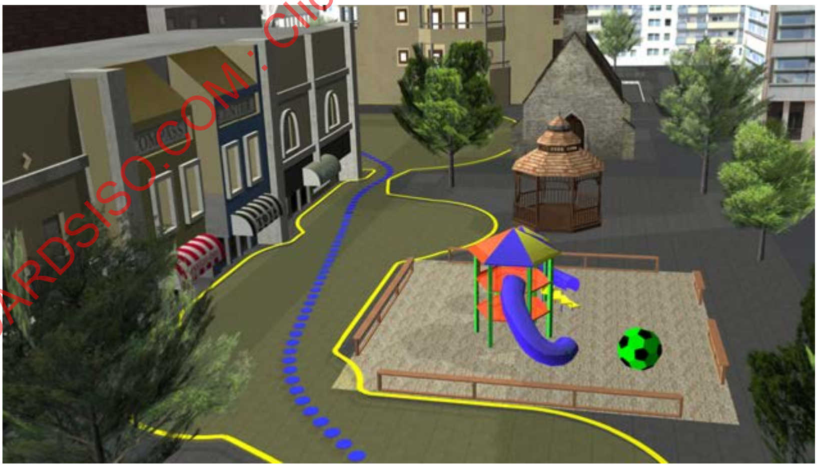


Figure 23 — An example of a smooth, short path with clearance inside an Explicit Corridor

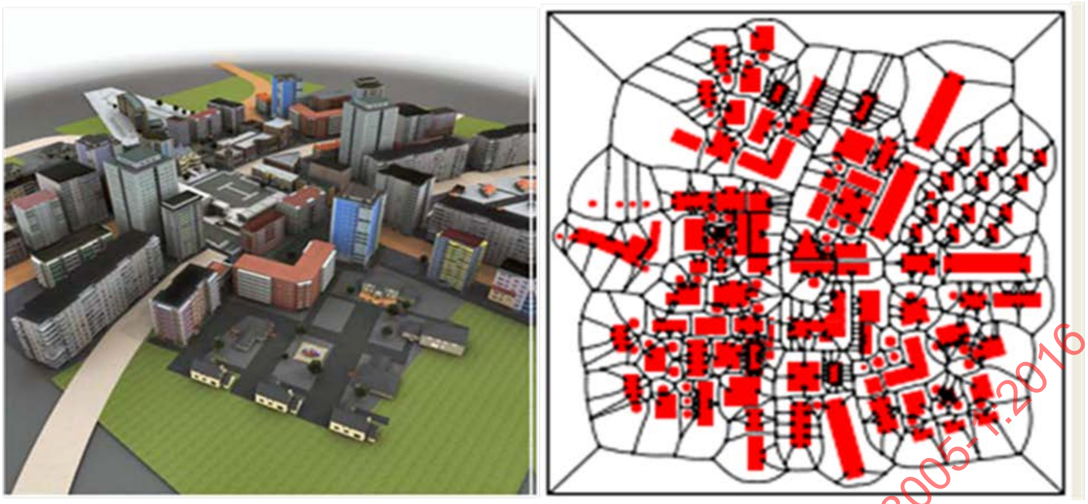


Figure 24 — Example of a virtual test city environment and its footprints and corridor map

- **Instantiation C.4: Unmanned aerial vehicle scenario**

This use case is presenting a scenario of controlling an unmanned aerial vehicle through a map interface at the land station. The vehicle at real-world and the map interface, which can be considered as the enhanced virtual world, shall be synchronized in the sense that the actual position of the vehicle shall be represented in the virtual world and the control of the vehicle in the virtual world shall be reflected through the vehicle at real-world.

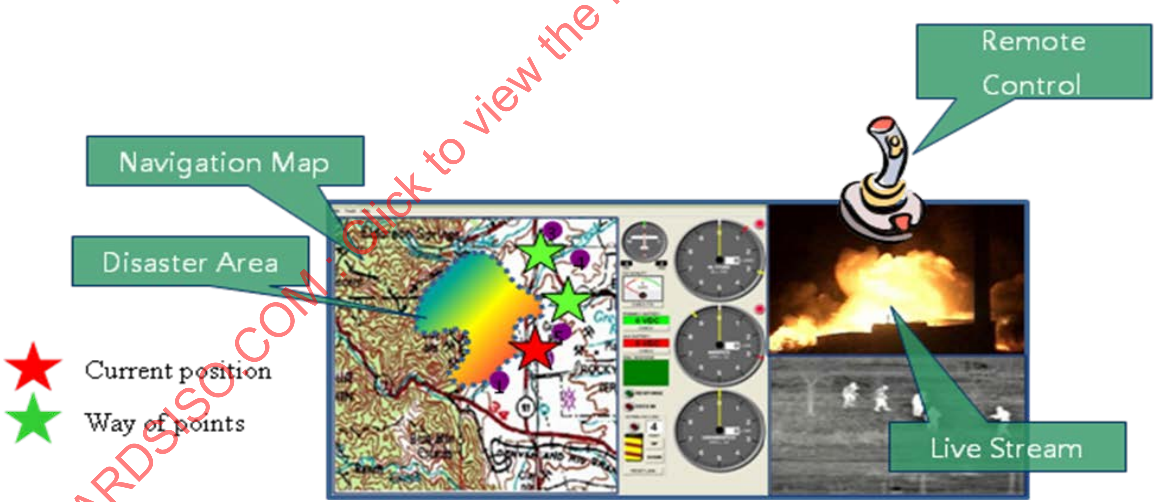


Figure 25 — Usage scenario of enhanced interface and the unmanned aerial vehicle (UAV)

Figure 25 shows one usage scenario of enhanced interface and the unmanned aerial vehicle (UAV). In this scenario, a point of mountain fire is reported and displayed on the map of the land station called HeliStation. The controller of the HeliStation commands a nearby stationed UAV to fly over the fire area and to stream the real-time scene of the site. There are two ways of commanding the UAV. One is by simply giving the destination position in terms of global coordinate and the altitude and letting the intelligent navigation system of the UAV do the rest. The other is by manipulating the movement of the UAV remotely, by commanding the movement of the UAV using the control with description of 6 degrees of freedom, watching the navigation map and the real-time streamed video. Figure 26 shows the current version of the actual HeliNavi interface, through which the controller can control the UAV either manually or intelligently.

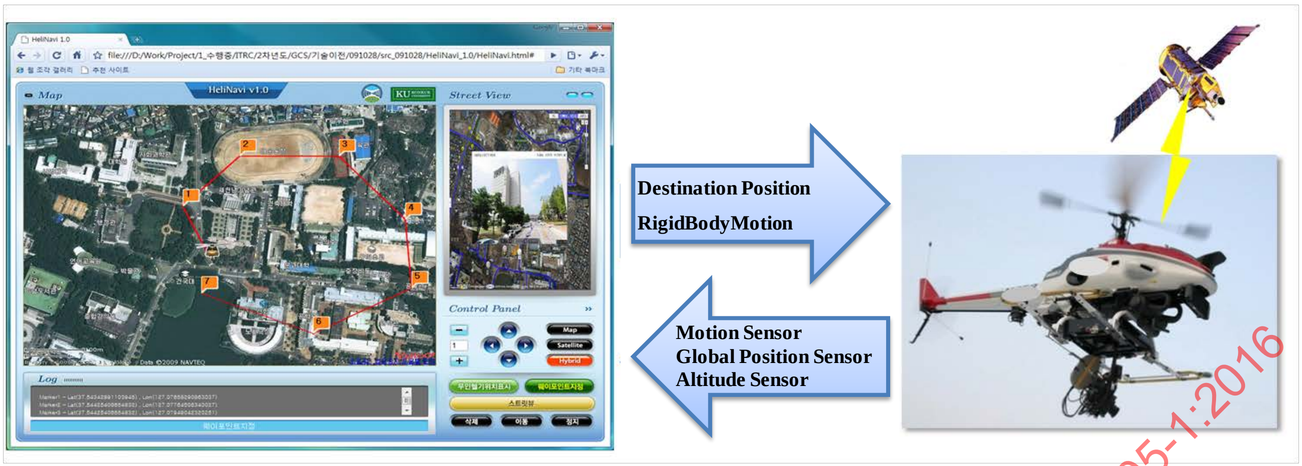


Figure 26 — HeliNavi interface

5.4 Instantiation D: Interoperable virtual worlds

- **System Architecture for interoperable virtual worlds**

The system for interoperable virtual worlds can be instantiated from the system architecture of information exchange between virtual worlds.

- **Instantiation D.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.

- **Instantiation D.2: 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 27. 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.

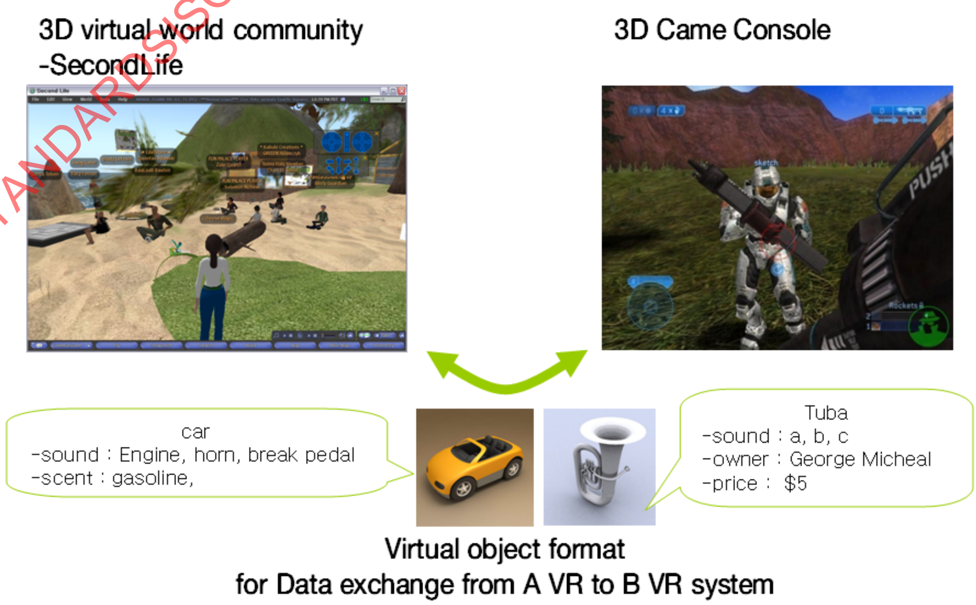


Figure 27 — Virtual object format for Data exchange

Consider the example of virtual shopping in Figure 28. 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

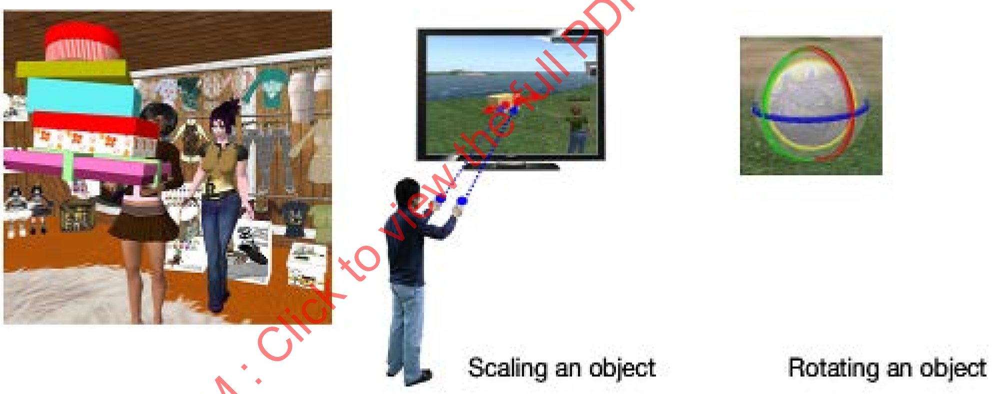


Figure 28 — Object manipulations for virtual shopping

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 29.

Figure 29 — Object manipulations for virtual learning (Golf)

5.5 Instantiation E: Social presence, group decision-making and collaboration within virtual worlds

- **System architecture**

The system for social presence, group decision-making and collaboration within virtual worlds can be instantiated from the system architecture of information exchange between virtual worlds.

- **Instantiation E.1: 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 30). 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 30 — Geode of industry and sciences museum overview

- **Instantiation E.2: 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 31 — 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 inclusion, 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 32 — 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].

- **Instantiation E.3: 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.

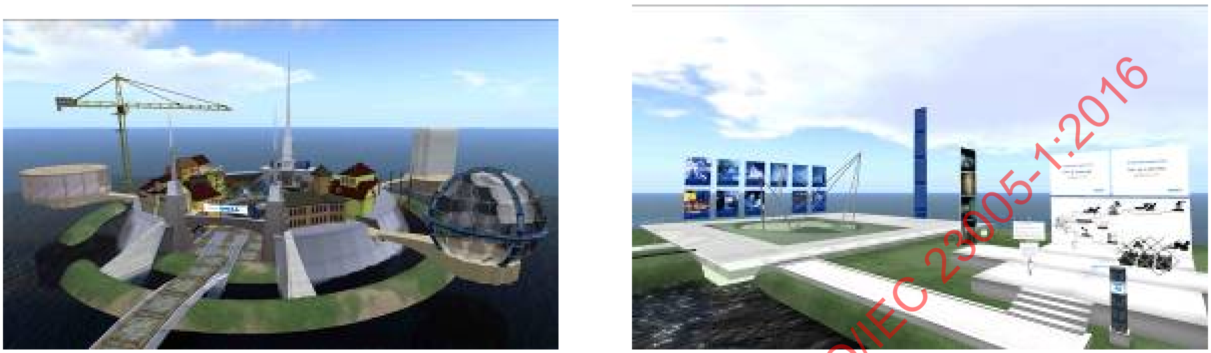


Figure 33 — Examples of manufacturers present in *Second Life* (Dell and Philips)

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 34 — 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.

5.6 Instantiation F: Interactive haptic sensible media

- **System architecture for interactive haptic sensible media**

The system for interactive haptic sensible media combines the two use cases such as information adaptation from real world to virtual world and the one from virtual world to real world. Specifically, Figure 35 shows the exact system architecture for the instantiation. The instantiated system architecture can be adopted by any interactive sensory effect media.

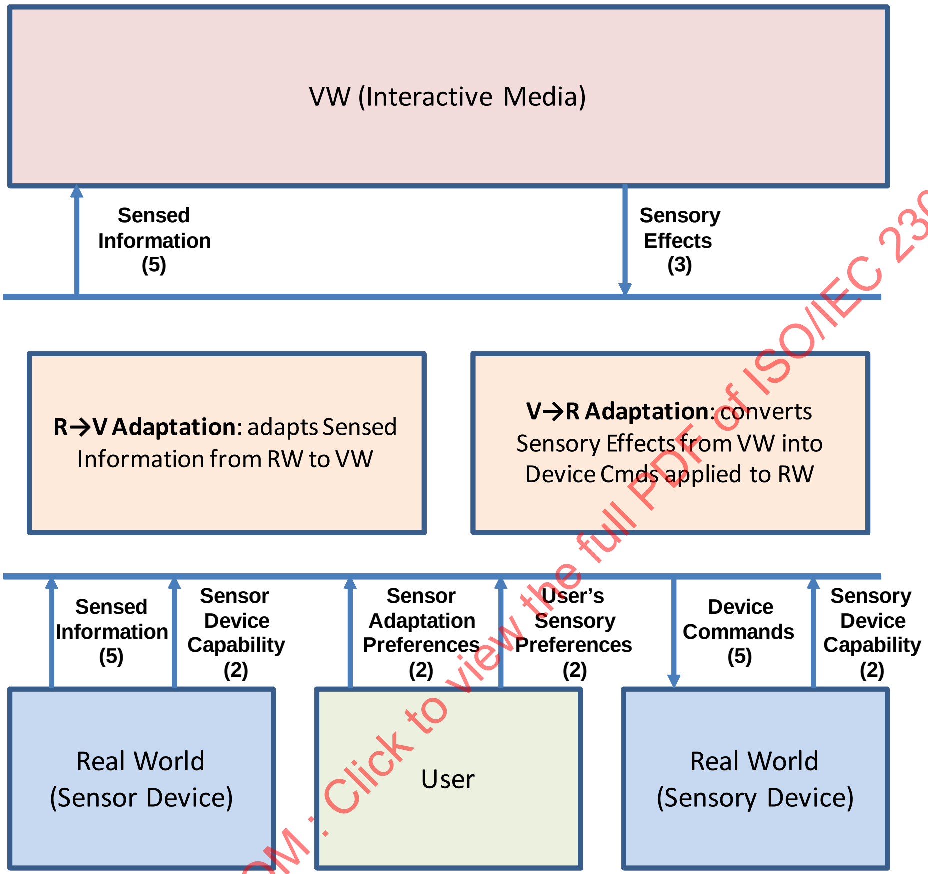


Figure 35 — (Possible) System Architecture for Interactive haptic sensible media

The sensors for interactive haptic sensible media include position, velocity, acceleration, orientation, angular velocity, force, torque, and pressure sensors. The actuators for this instantiation include haptic sensory effects such as rigid body motion, tactile, and kinesthetic effects. 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.

- **Instantiation F.1: Internet haptic service - YouTube, online chatting**

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

Figure 36 — Use Case Scenario for Internet Service

- 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.

- **Instantiation F.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 37 — 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 (e.g. sandstone, shale, conglomerate, granite, and whinstone) are loaded in front of them.
- By touching them with a haptic device they can distinguish the difference among stones since each virtual stone provides its own hardness, roughness, friction to users.

- **Instantiation F.3: Immersive broadcasting – home shopping, fishing channels**

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

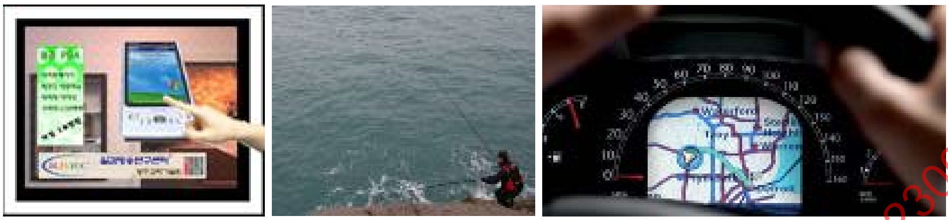


Figure 38 — Use Case Scenario for Immersive Broadcasting

- 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.

- **Instantiation F.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.
- Tactile vest (including thermal and vibration display) provides heating mode when the base is burning and vibration mode when soldiers are attacked on the back of user according to the map of the game.

- **Instantiation F.5: Virtual simulation for training – military task, medical simulations**

Military Training Task

Description: In the virtual simulation room, soldiers are performing building-clearance training. They are wearing tactile vests that let the soldiers know which parts of the building are cleared by providing vibration on different parts of their backs.

Medical Simulations

Description: In a medical school, a medical student has the first simulated MIS (Minimally Invasive Surgery). A virtual patient is anesthetized and two assistant nurses are standing next to the patient.

- The medical student put endoscope into the virtual patient's abdomen and then the inside scene of abdomen is displayed on the screen according to the position and angle of the endoscope.
- The student touches a gall bladder with the tool for diagnosis and (s)he feels it is harder than a healthy gall bladder.
- The gall bladder is being cut by laser tool and during the resection of the gall bladder the student is provided the feeling of the movement of the gall bladder and deformable force feedback.

5.7 Instantiation G: Bio-sensed information in virtual world

- **System architecture for bio-sensed information in virtual world**

System for bio-sensed information in virtual world can be instantiated from the system architecture of information adaptation from real world to virtual world. Biosensors are sensors that detect vital body signals like nerve/muscle activity, heartbeat, blood pressure, and the like. Biosensors have the potential to affect many areas. Field application areas including medicine, physical therapy, music, and the video game industry, can all benefit from the introduction of biosensors.

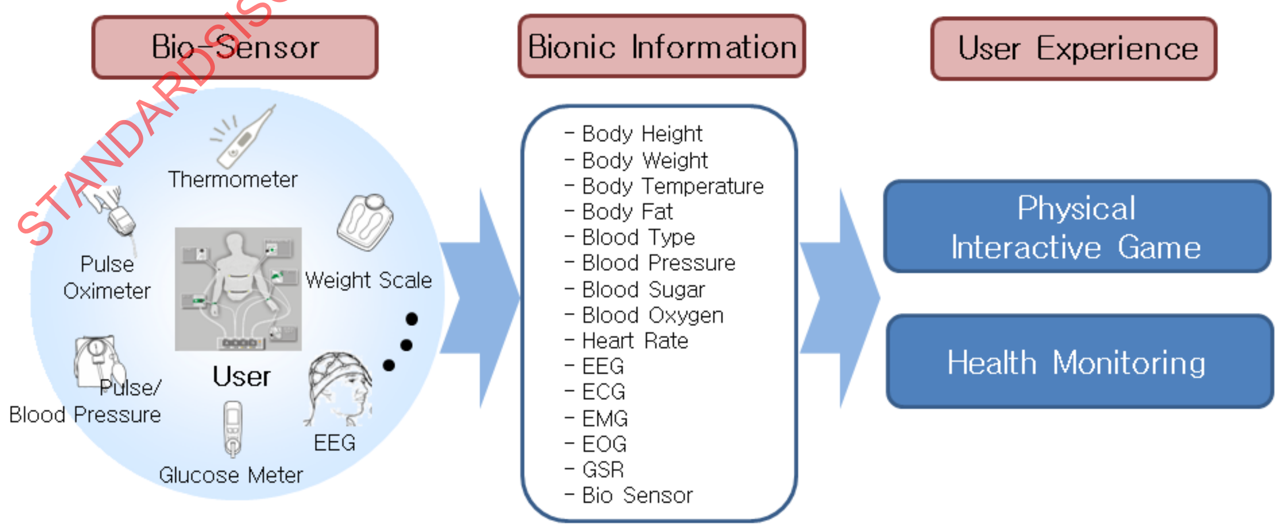


Figure 39 — Examples of bio-sensors and use cases

- **Instantiation G.1: Interactive games sensitive to user's conditions**

Biosensors potentially have a number of uses in the emerging field of Virtual Reality, particularly in the areas of user interaction and the development of interaction devices. For example, eye movement also has important interaction implications in virtual reality. Eye movement is determined from biosensors (ex. EOG sensor) placed strategically on the forehead and under the eyes. Wherever the user's eyes look, the virtual environment could be displayed appropriately. Furthermore, convergence of the eyes on a specific object in the environment could be detected. This could be used to select and object in the virtual environment. Figure 40 illustrates the potential application of biosensors in physically interactive games.

Figure 40 — Examples of using biosensors in interactive games

- **Instantiation G.2: Virtual hospital and health monitoring**

Another possible application area of biosensors is "e-Health", a globally distributed process, in which communication and collaboration of geographically dispersed users (patients, older people, or therapists…) play a key role. Within this process, an important role will be played by intelligent environments for health care in which multimedia contents integrate and enrich the real space. Figure 41 illustrates the potential application of biosensors in e-Health.

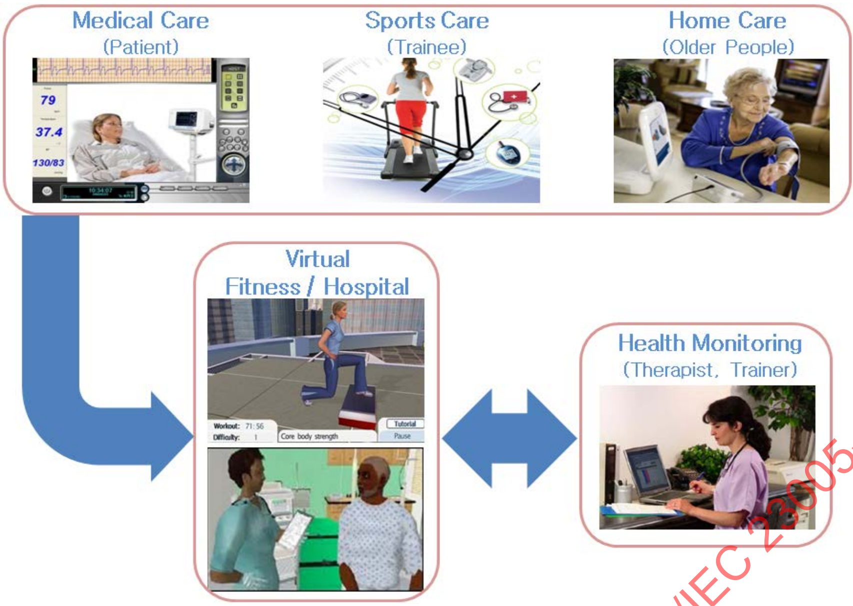


Figure 41 — Examples of using biosensors for health monitoring.

- **Instantiation G.3: Mental health for lifestyle management**

The WHO (World Health Organization) predicts depressive syndromes will be higher than any other health problem within 20 years (450 million people). Anti-depressant market in EU in 2010 is 4,5B€.

Figure 42 — Illustration of the Mental Health for Lifestyle Management use case

The purpose of the mental health application is to monitor the mental wellbeing of the user unobtrusively by utilizing sensors embedded in the everyday life of the user. The application collects unobtrusively information about the user's behavior by using a large variety of information sources such as:

- Wearable sensors (accelerometers, ECG, EEG);
- Home monitoring (bed sensors)
- Electronic information (amount of email, electronic calendar, phone usage)
- User input (diary, questionnaires)

These data are fused to assess the mental strain of the subject. Personalized computer-aided therapy and intervention applications are developed which help the users to balance their mental load and resources.

The user will benefit from the improved quality of care and access to care provided by the developed computer-aided therapy tools.

The economical impact of mental health problems equals 3-4% of the GNP in the EU states and mental health problems are the most common reason for early retirement in OECD countries. However, the mental health problems are currently under-diagnosed and under-treated because of the limited healthcare resources.

- **Instantiation G.4: Food intake for lifestyle management**

The World Health Organization (WHO) has identified current issues and predicted significant future health problems stemming from inappropriate eating habits in Europe (WHO, 2004).

Figure 43 — Illustration of the Food Intake for Lifestyle Management use case

The health consequences arising from poor nutrition are serious, in particular cancer and cardiovascular disease, but also obesity which carries a related set of health risks in the form of cardiovascular disease, type 2 diabetes, musculoskeletal disorders like osteoarthritis, and some cancers. The cost of obesity to society is enormous: approaching 1% of the gross domestic product in some countries in the WHO European Region (WHO/Europe, 2006). The reasons for poor diet and nutrition are complex and require multi-sectorial interventions. In this respect, the tools of assessment of nutritional interventions form an important problem. These currently require complex and extremely time consuming participant involvement (e.g. the maintenance of food diaries). The lack of easy-to-use tools to assess food intake, limits the ability to create personalized interventions.

Using rich multi-sensor data, the aim is to provide people with insights in:

- eating patterns (moments of eating),
- the type of food,
- the quantity of food eaten (portion sizes), and
- the consumed nutrients and calories.

The latter can be estimated based on the measured data and food composition databases. Using a device/service combination, awareness of one's food intake is created by displaying the measured sensor data in an easy-to-understand manner. Moreover, personal habits and food preferences can be derived using data mining algorithms applied to the multi-sensor data. Based on the needs of a user (e.g. lose weight, consume a sufficient amount of vegetables and vitamins), these habits and preferences are used to provide actionable, tailored food advice at the appropriate time. The advice is combined with behavior change strategies to persuade the users to improve their food intake.

- **Instantiation G.5: Cardiovascular rehabilitation for health management**

The main objective of this application domain is to provide added safety and an improved quality of care for cardiac patients during exercise training, in particular while still in cardiovascular rehabilitation.

Figure 44 — Illustration of the Cardiovascular Rehabilitation for Health Management use case

After a cardiac infarction and treatment in the cardiac cath lab or heart surgery, patients should participate in rehabilitation activities. These activities normally include rehabilitation sports such as indoor bicycle ergometer training and outdoor training activities (e. g. walking). While advanced systems supporting individual training parameters and a supervision of vital signs have been developed for indoor ergometer training, e.g. by the SAPHIRE and OSAMI projects, support for outdoor activities is limited so far. Available products are either intended for athletic use and unsuitable for cardiac patients, or limited to simple ECG acquisition systems with the ability to forward vital parameters through a mobile phone if the patient does not feel well or, e. g. a tachycardia is detected. Furthermore there are no “smart shirts” based on textile sensors available yet. Therefore, detecting vital signs requires separately applicable sensors that deteriorate users’ compliance due to their complexity.

A system permitting a continuous acquisition of ECG during outdoor training using an unobtrusive ECG measurement (e. g. ECG shirt) and a review of the vital parameters by a medical supervisor accompanying a group of patients during outdoor training, as well as forwarding the data to a rehabilitation centre for second opinion could for instance significantly improve the (lifestyle) conditions for the patient while lowering the cost of the procedure.

Techniques for automatically detecting of adverse events during training, for improving training outcome e. g. by an automatic adjustment of training duration and “load”, and for predicting trends in the patient’s health status based on individual patient models, a larger set of sensory input (including non-medical data such as GPS location information) and the capabilities of the sensor broker component developed in the project. This includes a system for outdoor exercises that is individually adapted to the patient by stress tests to allow for a predication of heart frequency changes. This individual body model will also be used to indicate health problems during exercises and instruct the device to send medical data to the hospital or call for help on severe problems. As an example for an optimization of training conditions, a track may be planned by a routing algorithm based on maps with GPS information including altitude to predict load (and, therefore, heart frequency) on track for every patient individually. Furthermore, vital signs and other health related data of a patient will be stored in a medical document called “training report” during a training session, based on the HL7 Clinical Document Architecture. Data in this document will be annotated with timestamps or time intervals if this information is available during acquisition. Temporal reasoning fusing different types of data and taking into account the uncertainty of the information recorded will allow for a better forecasting of trends in the patient’s health status.

- **Instantiation G.6: Glucose level / diabetes management for health management**

The main objective of this application domain is enabling person-centric diabetes management processes and efficient and sustainable hospitalization infrastructures through collaborative platforms, embedded technologies and interoperable multi-tiered architectures.

Figure 45 — Illustration of the Glucose level / Diabetes management use case

The purpose is to extent the concept of diabetes management to spread the “Active Prevention” to all the phases involved in the care cycle and to improve the quality of life of the patients and citizens by moving the care services to their regular living environment. Making the concepts Ambient Intelligence, Independent Living and all related technologies possible is a key issue in achieving an Active Prevention in all the Care Cycle phases, allowing active prevention provision by means of enabling the ambient to interact with personal devices and personal health or social information.

Patient telemonitoring will play a key role in improving the interaction between patients and doctors. On one hand, this will help the patient to feel more comfortable and safe, and this system offers to the physician valuable information for a better analysis of the disease evolution and prediction of risks complications. The educational effort in the project will serve to promote adherence, and thus the patient benefit of the technology. Remote monitoring technologies can transmit data on a regular, real time basis and prevent hospitalizations by identifying and treating problems by triggering adjustments in care before negative trends reach crisis stage.

Different kind of sensors will be needed to get the information from the patients in each scenario, these sensors will follow the “Continua health alliance” standards or, on the contrary, any type of bridge will be developed to use these sensors without these standards.

Beyond the remote monitoring and information provision to the patients and citizens for prevention, the project proposes to advance in the state of the art of all those technologies allowing the ambient to participate in a personalized health management, mainly by focusing in all kinds of prevention.

5.8 Instantiation H: Environmental monitoring with sensors

The system for environmental monitoring with sensors can be instantiated from the system architecture of information adaption from real world to virtual world. The associated sensors include gas sensors and a wind sensor.

- **System architecture for environmental monitoring**

System for environmental monitoring can be instantiated from the system architecture of information adaptation from real world to virtual world.

Naturally, environmental pollution problems have been local and minor because of the Earth's own ability to absorb and purify minor quantities of pollutants. However, the industrialization of society, the introduction of motorized vehicles, and the explosion of the population are factors contributing toward the growing air pollution problem. Therefore, introducing effective methods to measure air pollutants and clean up the air become urgent and critical.

The primary air pollutants found in most urban areas are carbon monoxide, nitrogen oxides, sulfur oxides, hydrocarbons, and particulate matter (both solid and liquid). These pollutants are dispersed throughout the world's atmosphere in concentrations high enough to gradually cause serious health problems along with global warming problems. Serious health problems can occur quickly when air pollutants are concentrated, such as when massive injections of sulfur dioxide and suspended particulate matter are emitted by a large volcanic eruption or sandy dust clouds. Figure 46 illustrates major air pollutants with possible application areas.

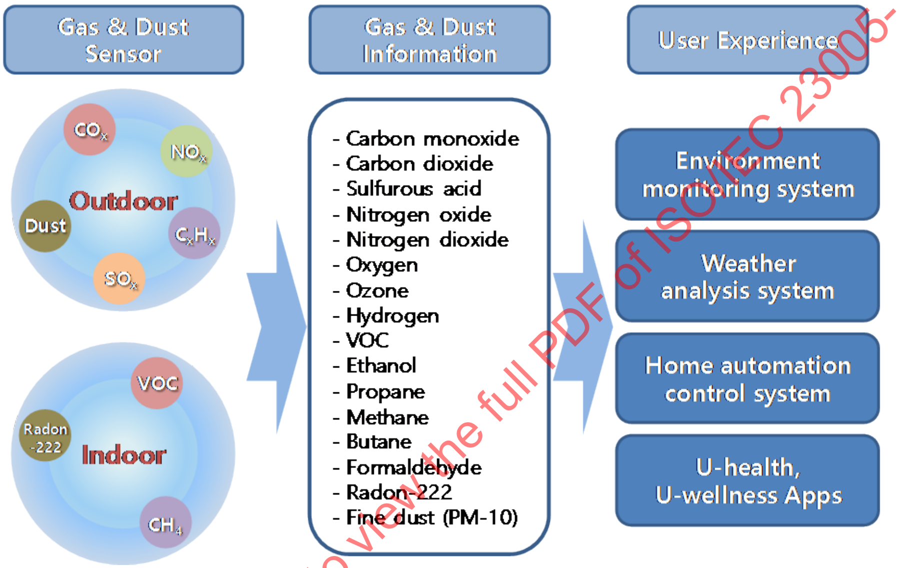


Figure 46 — Gas & Dust sensor information and use case examples

Wind sensor, which is also known as anemometer, is the sensor which detects the velocity of the wind and the position of the sensor. Wind speed and its orientation obtained by wind sensors is one of important information to describe the room condition where people consume some contents.

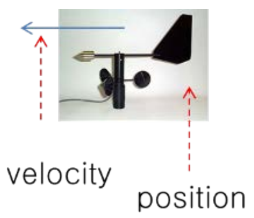


Figure 47 — Main elements sensed by wind sensors

- **Instantiation H.1: Environmental monitoring system**

You cannot escape air pollution, not even in your own home. The Environmental Protection Agency (EPA) reported that toxic chemicals found in the air of almost every American home are three times more likely to cause some type of cancer than outdoor air pollutants. The health problems in these buildings are called "sick building syndrome". The EPA has found that the air in some office buildings is 100 times more polluted than the air outside. Poor ventilation causes about half of the indoor air pollution problems. The rest come from specific sources such as copying machines, electrical and telephone cables, mold and microbe-harboring air conditioning systems and ducts, cleaning fluids, cigarette smoke, carpet, latex caulk and paint, vinyl molding, linoleum tile, and building materials and furniture that emit air pollutants such as formaldehyde. Another major indoor air pollutant is radon-222, a colorless, odorless, tasteless, naturally occurring radioactive gas produced by the radioactive decay of uranium-238. Radon is considered to be the second leading cause of lung cancer, and is an underrated threat. Also, there can be liquid natural gas (LNG) leaking caused by a careless mistake. Sensing those indoor gases can be critical to the home automation control system, for instance.

The gas and dust sensors can easily be adopted as mobile applications like “environmental monitoring system”, “weather analysis system”, or “home automation system”. The “u-Health” or “u-wellness” is another closely related area that can adopt gas and dust sensors as well. Figure 48 illustrates three instances of an air pollutant detect and warning system embedded in a mobile device.

Figure 48 — AR application scenario using gas & dust sensors

5.9 Instantiation I: Virtual world interfacing with TV platforms

- **System architecture for virtual world interfacing with TV platform**

System architecture for virtual world interfacing with TV platforms can be instantiated from the system architecture of information adaptation from real world to virtual world. The objective is to control virtual world applications (like Second Life) from an (IP) TV Set and render the output of the virtual world applications (like Second Life) on the (IP) TV platform. jointSPACE is an Open Source project that will allow every user/supplier to develop applications for TV displays. It is based on the SPACE architecture which was developed by Philips to ease internal development. At a certain point of time, Philips decided to open its architecture to allow everyone developing code for the TV target. jointSPACE addresses this by opening and extending the current TV architecture. Obviously this is beneficial for the development of new TV applications, but also very interesting as application itself by using remote applications. These applications are running on external devices, making use of TV capabilities to render GFX and media.

Figure 49 — Remote application example

External devices can be any computing device (including iPod/iPhone, game consoles, PDA, lightweight PC, PC servers, MAC and the like) and:

- Remote applications are using the IP network (wired or wireless) to communicate with the TV.
- Remote applications can be used to extend TV functionality with customizable features integrated together with "standard" TV applications.
- Remote applications can be controlled with the TV remote controller.
- Remote applications are making use of DirectFB/VooDoo technologies and follow (joint) SPACE architecture rules.

To capture these screenshots, a remote application is developed that runs simultaneously with Second Life on the user's home pc. It automatically selects the Second Life window, captures what is inside and pushes it via the local network to the Philips TV, where it is rendered. The remote control key handling is integrated with the remote application and controls the avatar.

- **Instantiation I.1: The TV platform as a virtual worlds I/O device**

Via the users' home network the Second Life game content is transferred towards the Philips TV set, while the user can control the avatar with the TVs' remote control. With current settings a frame rate of 10 fps with good image quality is achieved. By altering settings, a higher frame rate is possible, however the image quality decreases, this because the bandwidth of the network in the limiting factor. Therefore, for now it is all about finding a balance between image quality and frame rate. In the future, as (jpeg) compression and/or other techniques to decrease bandwidth for this application, this is not a limiting factor anymore.

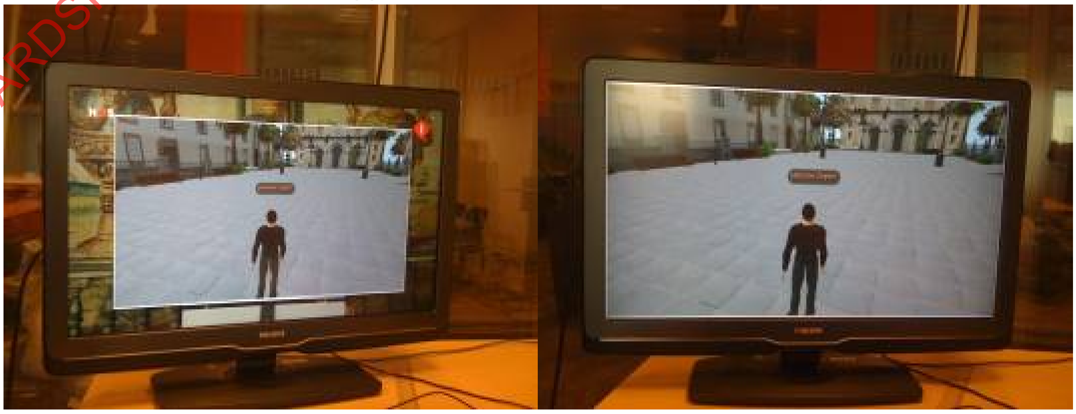


Figure 50 — Virtual World (Second life) content displayed and controlled on a (IP) TV set3.

[3] The window size can be controlled via the remote control; on the left image a small window with SL is shown with on the background a regular broadcast, while on the right image the full screen mode is shown.

5.10 Instantiation J: Seamless integration between real and virtual worlds

- **System architecture for seamless integration between real and virtual worlds**

System for seamless integration between real and virtual worlds can be instantiated from the two use cases of the information adaptation from real world to virtual world and the information adaptation from virtual world to real world. Real and virtual sensors and robots are the means of integrating the two worlds.

Virtual worlds are becoming ubiquitous as an alternative form of communication, commerce, relationships and interaction in general. Nowadays it is common to see humans of all ages subscribing to, and using virtual worlds where they form communities and establish bonds with both avatars and other real people. Even more, the interaction is reaching levels where the real and virtual worlds merge, showing a synergy that allow virtual and real agents to become essential parts of our lives.

However, the seamless integration of virtual and real worlds needed to ensure the embrace of these new technologies, still faces challenges that must be overcome before a real -and useful- interaction between smart networked devices, virtual environments and, most importantly, humans can occur. A key aspect for the integration of the real and virtual worlds is the standardization of the communication between and description of real and virtual devices.

Due to the growth of commercial available low cost sensors and actuators the need for interfacing with these devices to be able to monitor and control events in everyday life rises. Because of this growth, the variation among devices, and a lacking communication standard, software needs to be tailored to specific hardware which slows down the adoption of sensor and actuator usage in virtual world applications.

However, to be able to visualize complex real world activities also virtual sensors with higher abstractions are needed. In order to visualize human behavior from the real world into the virtual world, there must be a mapping between the identity of the person in the real and virtual worlds, together with a sensor that detects the presence of a person.

Because the real world is not static and a good mapping between a real world and a virtual world location is hard to standardize, there exists a need to have relative device position information. With this information sensor data from specific devices can be spatially mapped and for instance be used for tracking and tracing. This relative device position information should be obtained via a device proximity sensor which obtains this information in the real world for instance via the wireless RSSI.

- **Instantiation J.1: Seamless interaction between real and virtual worlds with integrating virtual and real sensors and actuators**

Integration between real and virtual worlds can be achieved by means of integrating virtual and real sensors and actuators. Sensors can offer perceptions on different levels of abstraction that can range from raw video and audio data to location tracking, presence, light, temperature, etc. Actuators, on the other hand also offer the possibility of performing actions with different levels of granularity that range from commands to move to a specific coordinate in 3D space, to more general instructions like “call my daughter”.

Because the structure of sensor and actuator data varies among different virtual and real devices, to be able to visualize sensor information in the virtual world and control actuators in the real world (and vice versa), a standard format for communication between devices needs to be defined. For this, new technology must be developed to allow for the correct handling of sensor and actuator information and its transformation into such standard representation.

The access to sensors and actuators (both virtual and real) must be abstracted from the peculiarities of each particular device by means of Device Services. A Device Service provides all the necessary information about the device and access methods that conforms to the standard defined for that service type, and that allows other components in the system to interact, process and/or visualize abstract data and events.

The exchange of information between virtual and real devices can only be achieved if the different services available for the devices in the system “speak” the same language. Therefore, we propose a message format called Device Service Message (DSM). DSM is an “envelope” format that allows for dynamic construction of a message, with parameters that are dependent on the required device service.

The advantage of such system design is that the particularities of a device are hidden from the system users. For example, consider the case of a Second Life avatar trying to access real video from the living room of its parent's house and display it into a virtual TV (inside the virtual world), which would go more or less like this:

- The virtual TV searches for "live video" in the list of available device services
- The service directory has registered a service that offers video in three different resolutions and 4 different formats
- The virtual TV selects the service, and calls the service's generic method to obtain a video stream, indicating the desired resolution and video format
- It is up then to the device service to talk to the device that registered the service, in this case a Phillips webcam, and transmit the appropriate configuration to obtain the desired video stream, which will be sent back to the TV in the virtual world.

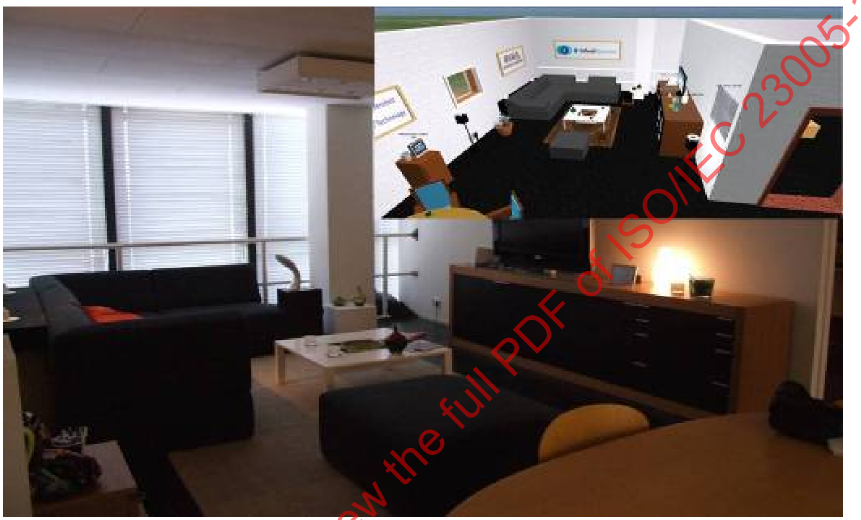


Figure 51 — Example actual living environment and it's representation in the virtual world

While device services provide a way to hide the complexity of communicating with and obtaining data from different devices, a final component of the system (implicitly included in the example) is still missing: There is a need for a software layer that performs the translation of a request received from a service (in DSM format) into device dependent data formats and function calls that will execute the desired action and/or obtain the desired data. This layer is called Device Message Layer (DML).

In the previous section we outlined the proposed platform architecture for communication between real and virtual devices. Figure 52 shows a use case diagram of the interaction between devices in our system: either they send/receive sensor data (e.g. a camera sending video), send/receive an action/command to be executed (e.g. a request for a switch to turn lights off), or send/receive other type of information which we call "general purpose" (e.g. request the virtual or real device status, debugging, etc).