INTERNATIONAL STANDARD

ISO/IEC 23005-1

Fourth edition
2020-08

Information technology — Media context and control —

Part 1: Architecture

Technologies de l'information — Contexte et contrôle des medias —

Partie 1: Architecture

Reference number
ISO/IEC 23005-1:2020(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword

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

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 ISO documents 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 of the voluntary nature of standards, 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 www.iso.org/iso/foreword.html.

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

This fourth edition cancels and replaces the third edition (ISO/IEC 23005-1:2016), which has been technically revised.

The main changes compared to the previous edition are as follows:

— added a new use case for 3D printing;

— added six new use cases for olfactory information in virtual world;

— added two new use cases for virtual panoramic vision in car;

— added a new use case for adaptive sound handling.

A list of all parts in the ISO/IEC 23005 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

The ISO/IEC 23005 series provides an architecture and specifies information representation of data flowing in and out of the real world and virtual worlds.

The data for the real world are communicated through sensors and actuators. The data for virtual worlds consist of properties of virtual objects and multi-sensorial data embedded in audio-visual content. The ISO/IEC 23005 series specifies data formats for sensors, actuators, virtual objects and audio-visual content.

Data captured from the real world can need to be adapted for use in a virtual world and data from virtual worlds can also need to be adapted for use in the real world. This document does not specify how the adaptation is carried out but only specifies the interfaces.

Data for sensors are sensor capabilities, sensed data and sensor adaptation preferences.

Data for actuators are sensory device capabilities, sensory device commands and sensory effect preferences.

Data for virtual objects are characteristics of avatars and virtual world objects.

Data for audio-visual content are sensory effects.

This document contains the tools for exchanging information for interaction devices. To be specific, it specifies command formats for controlling actuators (e.g. actuator commands for sensory devices) and data formats for receiving information from sensors (e.g. sensed information from sensors) as illustrated as the yellow boxes in Figure 1. It also specifies some examples. The adaptation engine is not within the scope.

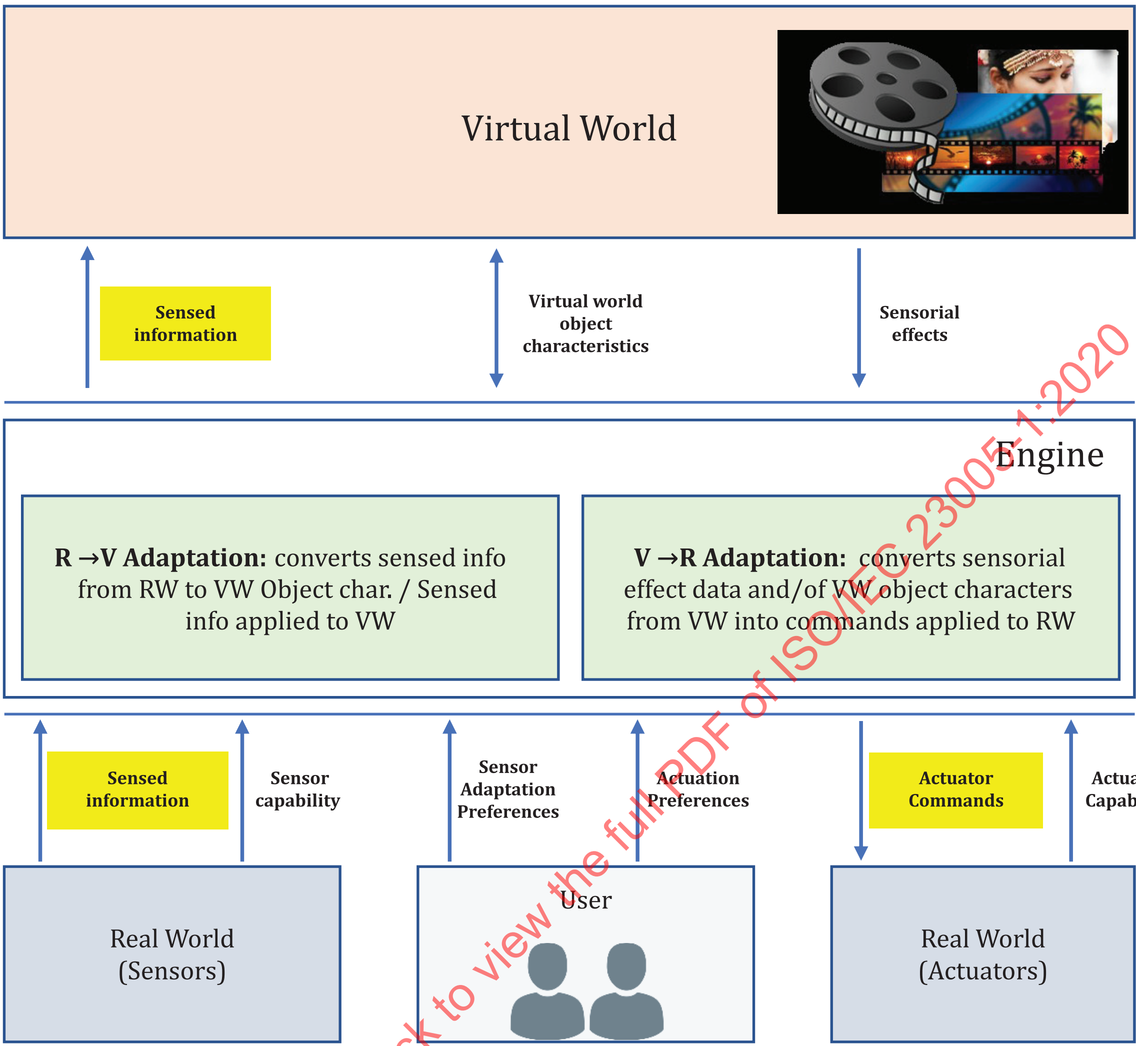


Figure 1 — Scope of the data formats for interaction devices

When this document is used, the adaptation engine (RV or VR engine), which is not within the scope of standardization, performs bi-directional communications using data formats specified in this document. The adaptation engine can also utilize other tools defined in ISO/IEC 23005-2, which are user's sensory preferences (USP), sensory device capabilities (SDC), sensor capabilities (SC) and sensor adaptation preferences (SAP) for fine controlling devices in both real and virtual worlds.

The International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC) draw attention to the fact that it is claimed that compliance with this document may involve the use of a patent.

ISO and IEC take no position concerning the evidence, validity and scope of this patent right.

The holder of this patent right has assured ISO and IEC that he/she is willing to negotiate licences under reasonable and non-discriminatory terms and conditions with applicants throughout the world. In this respect, the statement of the holder of this patent right is registered with ISO and IEC. Information may be obtained from the patent database available at www.iso.org/patents.

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

Information technology — Media context and control —

Part 1: Architecture

1 Scope

This document specifies the architecture of MPEG-V (media context and control) and its three types of associated use cases:

— information adaptation from virtual world to real world;

— information adaptation from real world to virtual world;

— information exchange between virtual worlds.

2 Normative references

There are no normative references in this document.

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at http://www.electropedia.org/

3.1
device command
description of controlling actuators used to generate *sensory effects* (3.9)

3.2
R→V adaptation
procedure that:

— processes the *sensed information* (3.3) 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* (3.4), the *sensor adaptation preferences* (3.5) from *users* (3.12) and/or the virtual world object characteristics from a virtual world;

— controls the *virtual world* (3.13) object characteristics or adapts the sensed information by adapting the sensed information based on the sensor capabilities and/or the sensor adaptation preferences

3.3
sensed information
information acquired by *sensors* (3.4)

3.4
sensor
device by which *user* (3.12) input or environmental information can be gathered

EXAMPLE Temperature sensor, distance sensor, motion sensor, etc.

3.5
sensor adaptation preferences
description schemes and descriptors to represent (user's) preferences with respect to adapting *sensed information* (3.3)

3.6
sensor capability
representation of the characteristics of sensors in terms of the capability of the given *sensor* (3.4) such as accuracy, or sensing range

3.7
sensory device
consumer device by which the corresponding *sensory effect* (3.9) can be made

Note 1 to entry: Real-world devices can contain any combination of *sensors* (3.4) and actuators in one device.

3.8
sensory device capability
representation of the characteristics of actuators used to generate *sensory effects* (3.9) in terms of the capability of the given actuator

3.9
sensory effect
effect to augment perception by stimulating human senses in a particular scene

EXAMPLE 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.]

3.10
sensory effect metadata
metadata that defines the description schemes and descriptors to represent *sensory effects* (3.9)

3.11
user's sensory preferences
description schemes and descriptors to represent (user's) preferences with respect to rendering of *sensory effect* (3.9)

3.12
user
end user of the system

3.13
virtual world
digital content, real time or non-real time, of various nature

EXAMPLE On-line virtual world, simulation environment, multi-user game, broadcast multimedia production, peer-to-peer multimedia production or packaged content like a DVD or game.

3.14
V→R adaptation
procedure that:

— processes the *sensory effects* (3.9) from the *virtual world* (3.13) in order to be consumed within the real world's context;

— takes *sensory effect metadata* (3.10) from a virtual world, sensory device (actuator) capabilities from the sensory devices (actuators), the *user's sensory preferences* (3.11) from *users* (3.12) and/or the *sensed information* (3.3) as well as the sensor capabilities from *sensors* (3.4) as inputs;

— generates the *device commands* (3.1) by adapting the sensory effects based on the sensed information, the capabilities and/or the preferences

3.15
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)

4 MPEG-V system architecture

A strong connection (defined by an architecture that provides interoperability through standardization) between the virtual and the real world is needed to reach simultaneous reactions in both worlds to changes in the environment and human behaviour. 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 indispensable. For fast adoption of virtual world technologies, a better understanding of their internal economics, rules and regulations is needed. The overall system architecture for the MPEG-V framework is depicted in Figure 2.

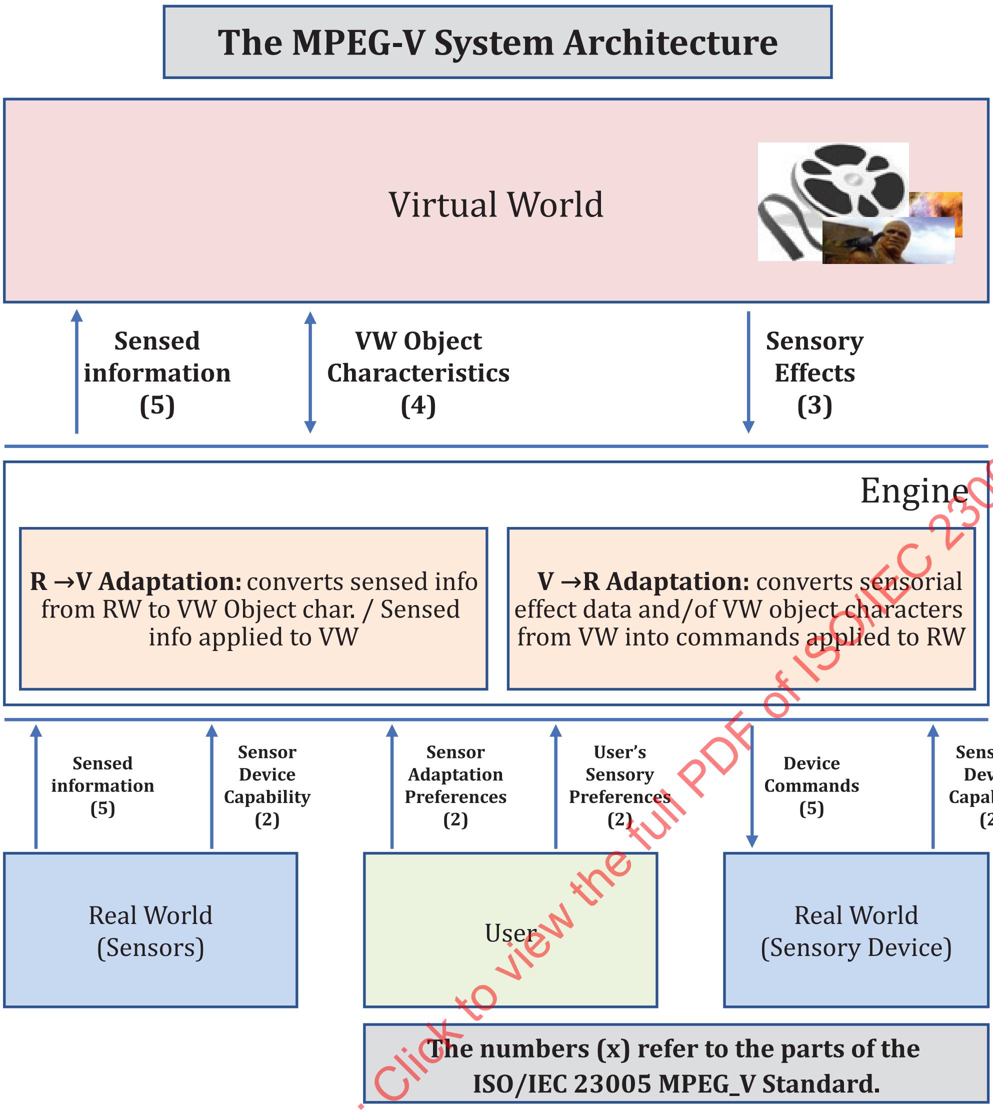


Figure 2 — System architecture of the MPEG-V framework

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 5.

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 are specified in other parts of the ISO/IEC 23005 series.

On the other side, the V→R adaptation engine, R→V adaptation engine, virtual world as well as devices (sensors and sensory devices) are left open for industry competition.

Metadata is specified in other parts of the ISO/IEC 23005 series. Sensor device capability, sensory device capability, sensor adaptation preferences and user's sensory preferences are specified in ISO/IEC 23005-2. Sensory effect metadata is specified in ISO/IEC 23005-3. Virtual world object characteristics is specified in ISO/IEC 23005-4. Device commands and sensed information are specified in ISO/IEC 23005-5.

5 Use cases

5.1 General

The three types of media exchanges require information adaptations 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.

5.2 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 3. 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.

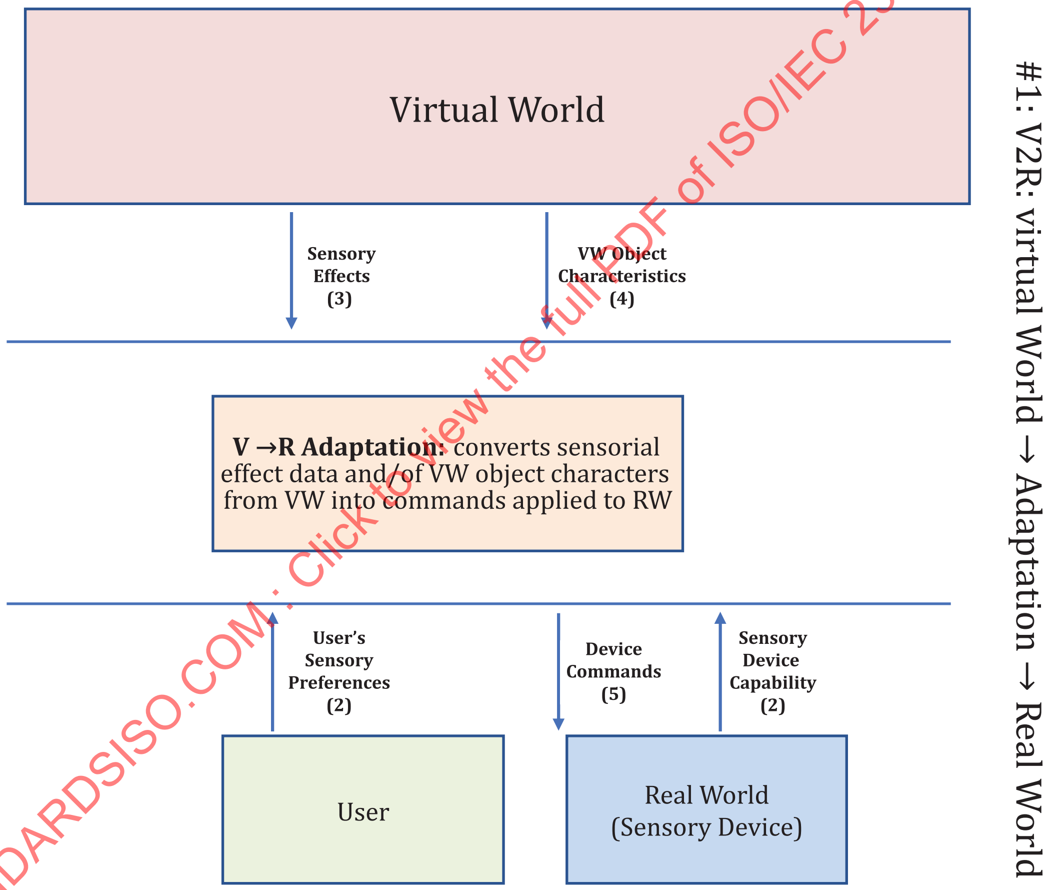


Figure 3 — Example of system architecture for information adaptation from virtual world to real world

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 on 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 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 a 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, can be combined with an audio/visual delivery format, e.g. MPEG-2 transport stream, file format, 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 of performing 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 these assets are referred to as sensory device capability/commands delivery format respectively.

Finally, the user's sensory preferences allow end users to describe their preferences with respect to rendering of sensory effects.

5.3 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 4. 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.

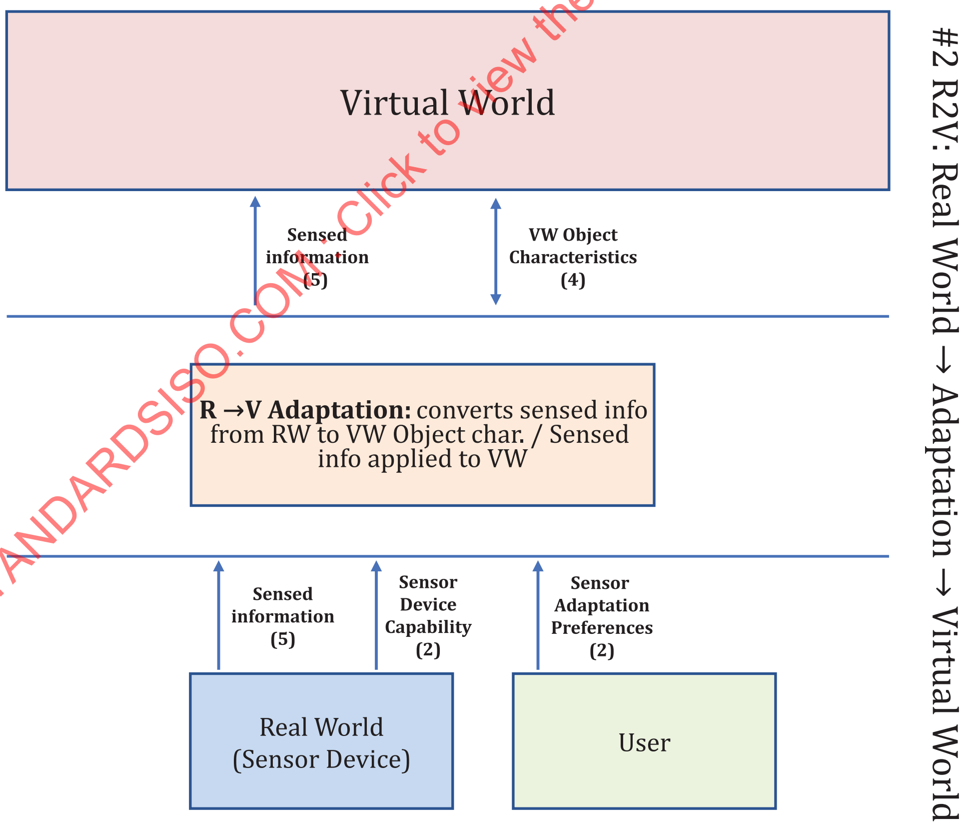


Figure 4 — Example of system architecture for information adaptation from real world to virtual world

R→V adaptation engine is an 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, and 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, and controls the virtual world object characteristics adapting the sensed information based on the sensor capabilities and/or the sensor adaptation preferences.

5.4 System architecture for exchanges between virtual worlds

The system architecture for information exchange between virtual worlds is depicted in Figure 5. It represents information exchange comprising VW object characteristics which generates exchangeable information within virtual worlds.

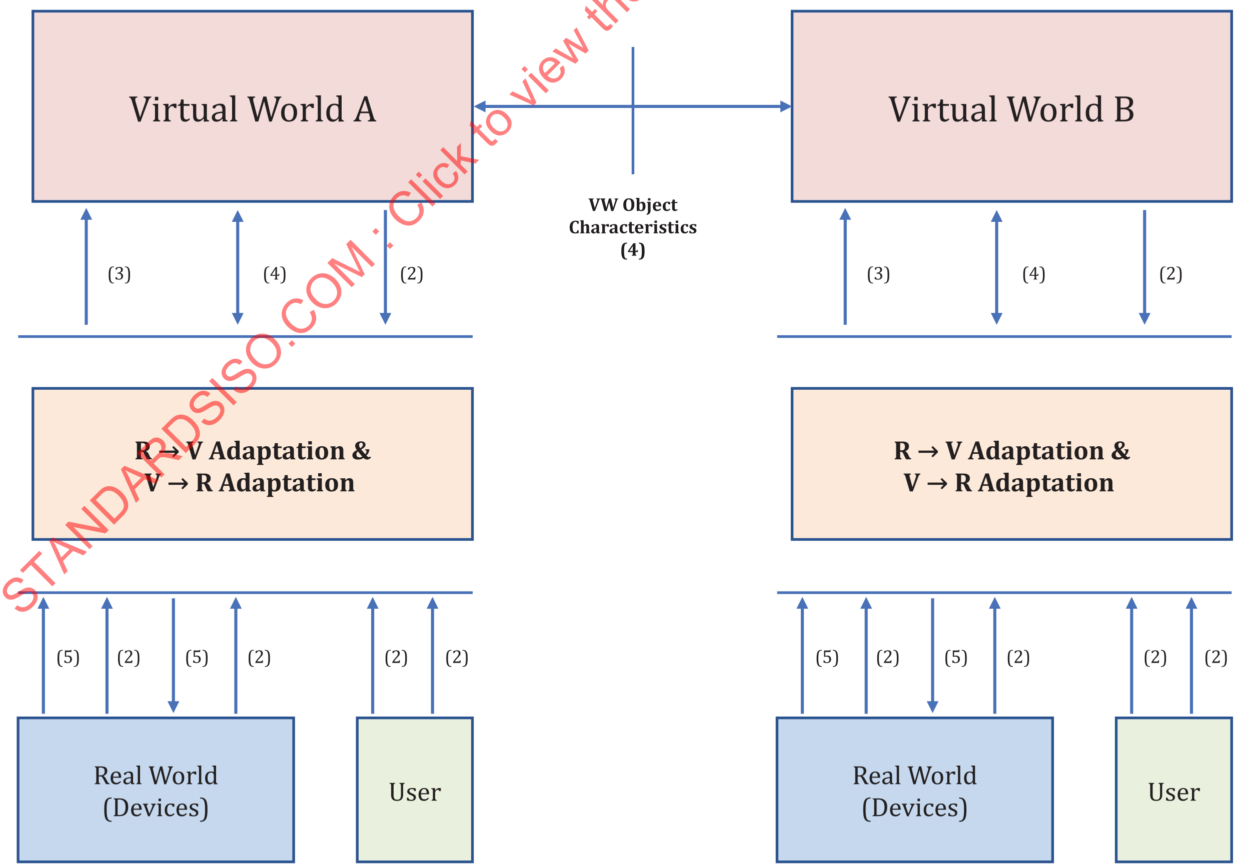


Figure 5 — Example of system architecture for (bidirectional) exchange of information between virtual worlds

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 internally adapts its own representation for virtual object/avatar.

6 Instantiations

6.1 Instantiation A: representation of sensory effects (RoSE)

6.1.1 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 6 comprising sensory effect metadata, sensory device (actuator) capability, device commands, user's sensory preferences and a V→R adaptation engine which generates output data based on its input data.

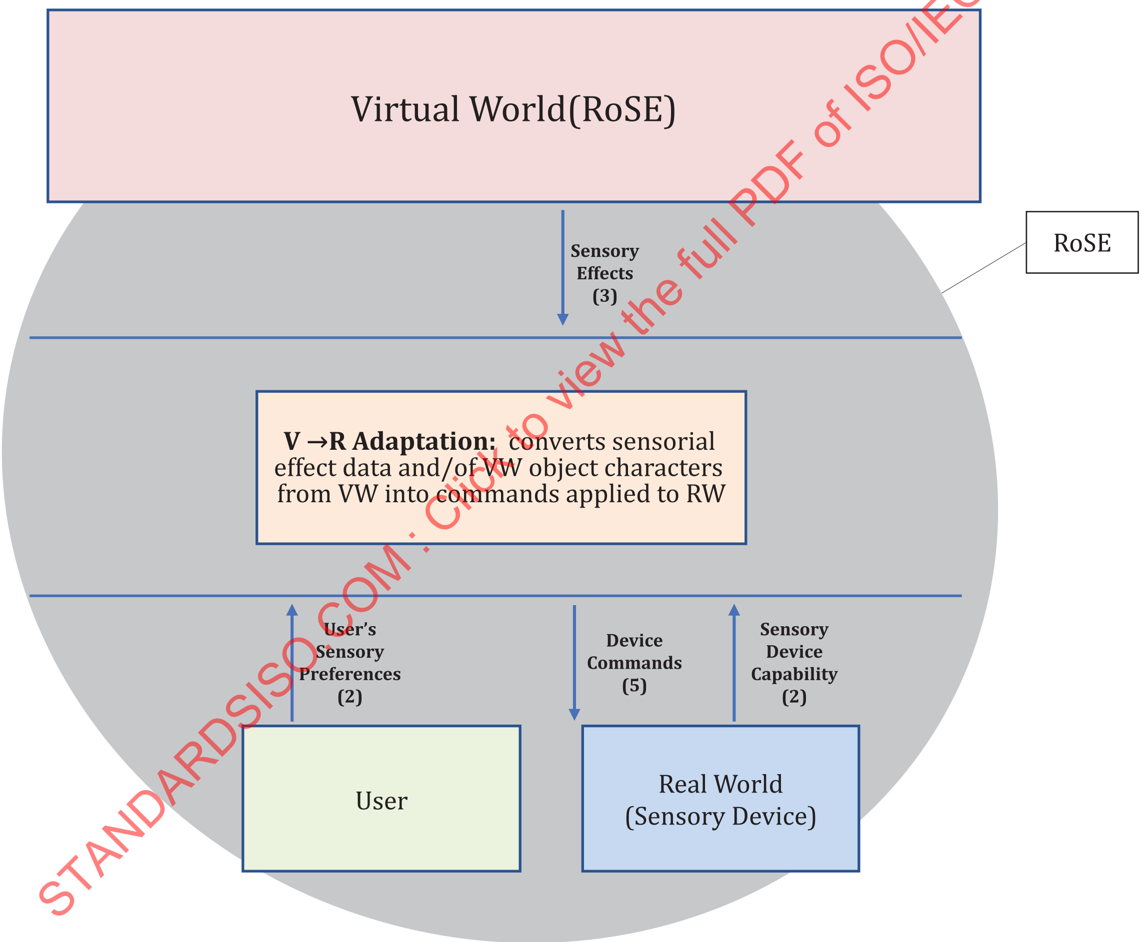


Figure 6 — RoSE system architecture

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 V→R adaptation engine is an entity that takes the sensory effect metadata, the sensory device (actuator) capability and the user's effect preferences as inputs, and generates the device commands 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 V→R adaptation engine are also collectively referred to as consumer devices. Note that sensory devices (actuators) are a 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).

6.1.2 Instantiation A.1: multi-sensorial effects

Traditional multimedia with audio/visual contents have been presented to users via display devices and audio speakers. 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.

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, which 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 as described above.

From a technical perspective, this requires a framework for the representation of sensory effects (RoSE) information which can 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.

6.1.3 Instantiation A.2: motion effects

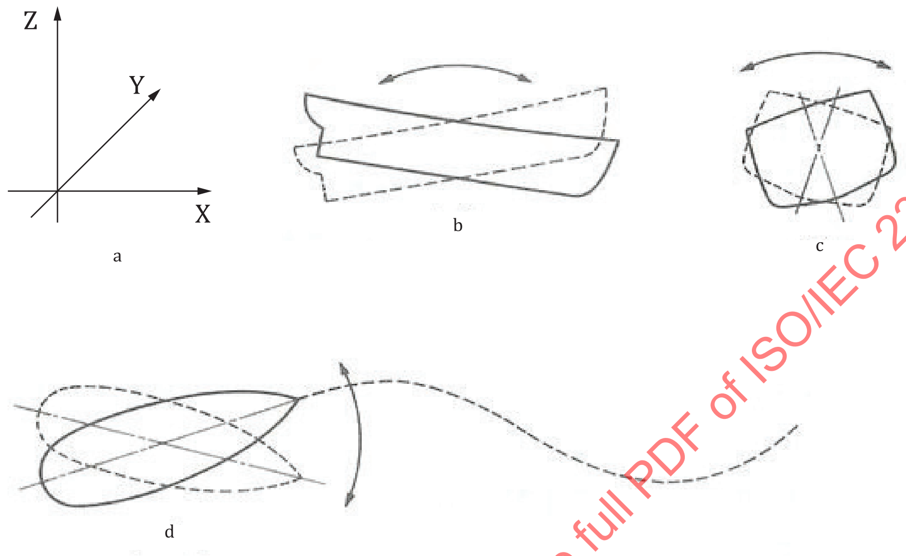
One of the important sensory effects that should not be ignored is the effect related to the motion. The motion effect gives a user a similar feeling on the movement like the actor/actress feels in the movie. The motion effect is a popular sensory effect commonly used in places like theme parks, game rooms and movie theatres nowadays. 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 motors and the length of axes determine the range and depth of the movement of the chair. There are a lot of manufacturers of motion chairs in the world and each of them has its own mechanical characteristics.

For example, the motion effect chair of a manufacturer 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 another manufacturer also supports falling down, rolling and pitching motion with speed, but not yawing or forward/backward move motion. On the other hand, there are 4D chair manufactures which only support vibration and falling effect.

Therefore, designing the sensory effect of motion shall not be limited by the capabilities of a specific chair. We need to consider that authors produce sensory effect metadata based on the audio-visual data and they do not know beforehand the mechanical characteristics of the motion chair by which the sensory effects 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 appropriate to describe the conceptual motion in the scene. Figure 10 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. The first chair supports rolling, yawing and surging. On the other hand, the second chair only supports rolling. If motion effect (SI) is expressed in physical terms like "Yawing 90 degrees", whether the chair can render it or not depends on the capabilities of chair. However, if motion effect (SI) is expressed in conceptual terms like "Left turn", both chairs can render this with their own capability.

In other words, considering the process of adaptation by the engine, the SI should 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.

The 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 7. It is a well-known fact that any motion of a rigid body can be described by the 6 DoF motion. The 6 DoF motions have been abstracted into several basic motion patterns and more combinational motion patterns, which are based on the repetition or a combination of basic patterns and have specific semantics, were added.



a 3-dimension.

b Pitching.

c Rolling.

d Yawing.

Figure 7 — Six degrees of freedom

6.1.4 Instantiation A.3: arrayed light effects

From a rich media perspective, realistic media coupled/assisted with their target devices are very beneficial to users because their experiences on media consumption can be greatly enhanced. Business market such as 4D cinema can be enriched by coupling media and devices together. Multiple devices are used to enhance realistic experiences, which can simulate the effect of light, wind, flash and more. When creating a 4D movie, a light effect can be used to indicate the mood or warmth of the space, but an arrayed light effect can be used to augment the events in a story with a much higher resolution by expressing complex light sources with an array of multiple light display actuators.

An example of arrayed light effect use case is a simulation of fireworks. Fireworks visual effects can be simulated in sync with the blasting event occurring in the movie with two arrayed light actuators in 4D theatre. This 4D theatre provides a big main screen in front of 4D chairs and two light display actuators are installed on the left and the right side. While the video of fireworks is playing on the big screen, the arrayed light effect represented with a sequence of *m* by *n* matrices is simulated by the light display actuators according to the timeline of the movie contents. The source of arrayed light effect can be light sequences composed of an *m* by *n* matrix mapped to a light display actuator or a file of video or image composed of *m* by *n* matrix.

Another use case of the arrayed light effect is given by visually augmenting cars that are passing through a tunnel with coloured lights in 4D theatre. When cars are passing through a tunnel with tunnel lights of red, blue and green colours in the movie, effect editors can add an arrayed light effect in order to make viewers to experience the same light effect as the passengers in the cars. The arrayed light effect metadata can be authored by editors who are unaware of the capability or the number of light display actuators installed in the theatre. Then, the adaptation engine generates device commands for actuating proper light actuators by analysing and extending the arrayed light effect metadata.

6.2 Instantiation B: natural user interaction with virtual world

6.2.1 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. The R→V adaptation engine analyses 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.

6.2.2 Examples of sensors

6.2.2.1 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 centre 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.

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.

6.2.2.2 Multi-pointing sensors

One of the most useful devices to interact with display devices is a remote controller. 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 sensors that detect the positions of feature points and the status of buttons. These types of sensors include mouses, gesture recognition devices, multi-touch devices, etc.

6.2.2.3 Smart camera

Smart cameras are vision systems with integrated cameras. They usually include a base line positioning sensor like GPS, gyroscope and magnetoscope in addition to the sensors that are capable of doing imagery, sensing human facial expression and body motion, etc.

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

Full motion control and navigation of avatar/object with multi-input sources allows for the full motion control and navigation of 3D objects and avatars in a virtual world. Recently, user interest in human-computer interaction has grown considerably based on large volumes of recent research. Through the development of VR technology, it has applied to various fields. Especially, the entertainment area is commercialized such as 3D virtual online communities like Second Life®[1)] and 3D game station. Nintendo Wii®[1)] 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, etc.

6.2.4 Instantiation B.2: serious gaming for ambient assisted living

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 can recognize the need for behavioural change, the same individuals are very creative in finding excuses. Thus, one of the greatest challenges for behavioural change is bridging good intentions with factual behaviour. Most 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 endeavours that challenge health care professionals.

It is a necessity to provide the tools to guide a person through a behavioural goal to change his/her behaviour 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 behavioural 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 behaviour based on user history and actual user needs. The appearance of the VEC can be selected and shaped by the user to fit their profile (sex, age, ethnicity language, personality). What is obviously very important is that the behaviour of the agent is convincing, natural, helpful and suited to the coaching task.

Coaching agents do exist already, although the interaction between agent and human is often limited. It is clear that we will see much improved agents in the near future that show more natural behaviour, better facial expressions, better and more natural bodily behaviour including non-verbal behaviour. 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.

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

Clearly, there is the need for detailed multi-modal control of virtual humans. 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 user communities, possibly employing different virtual worlds as a platform, this type of multi-modal behaviour should be communicated from one virtual world to another. And for realization of behaviour by means of some animation engine, there should be proper communication between the animation engine and system components that plan such behaviour.

1) Nintendo Wii® and Second Life® are trademarks of Nintendo Co., Ltd and Linden Lab, Inc., respectively. They are given as examples for the convenience of users of this document and do not constitute an endorsement by ISO or IEC.

6.2.5 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.

There is a commercial device which interacts with users through a touchscreen. 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.

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

6.2.6 Instantiation B.4: avatar facial expression retargeting using smart camera

The cameras in this use case are sensors that can perform imagery, sensing human facial expression and body motion, etc. These cameras can 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 8 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 8 b), the intelligent camera can 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.

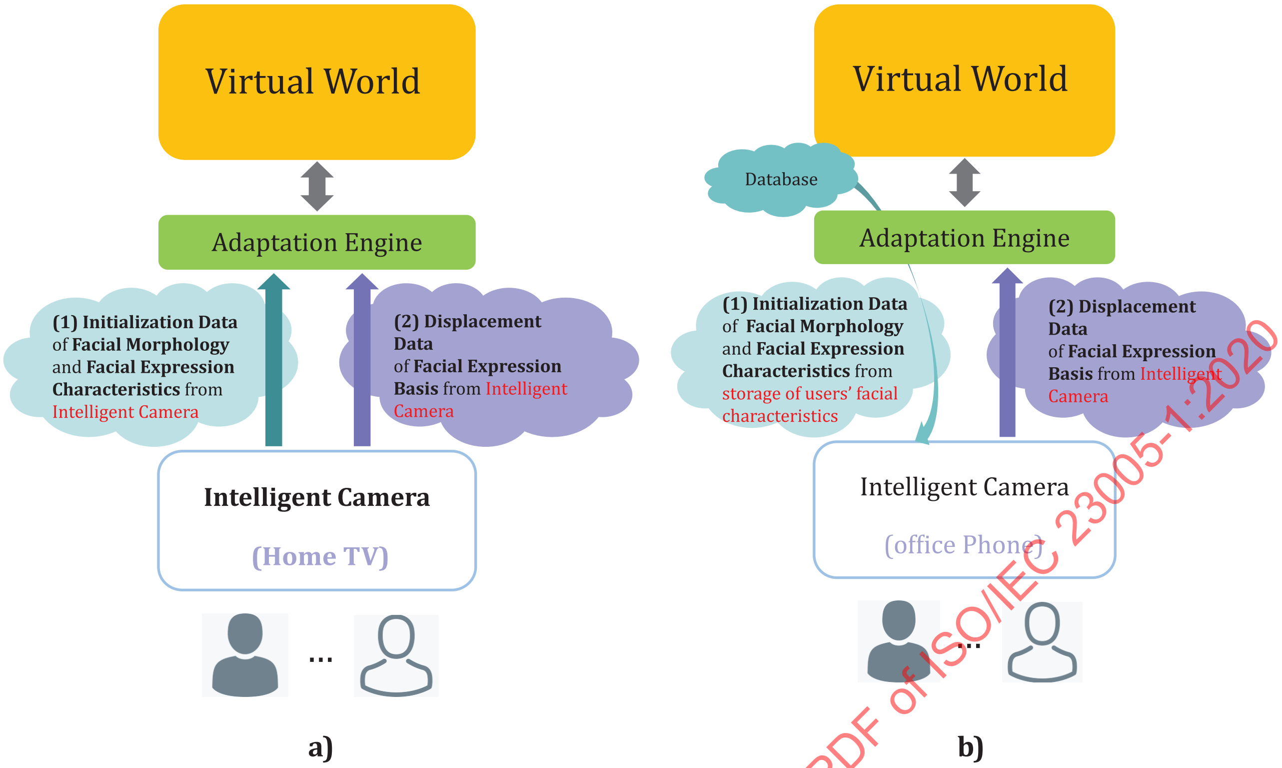


Figure 8 — Use case of intelligent camera for facial expression retargeting

6.2.7 Instantiation B.5: motion tracking and facial animation with multimodal interaction

6.2.7.1 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.

6.2.7.2 Facial animation

Facial animation is addresses within MPEG-4, e.g. by defining face definition parameters and face animation parameters.

6.2.8 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 Californi. (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 who has just arrived in a town in the Balkans. 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.

6.2.9 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.

6.3 Instantiation C: traveling and navigating real and virtual worlds

6.3.1 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.

6.3.2 Examples of sensors and path finding mechanisms

6.3.2.1 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, etc.

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).

6.3.2.2 Path finding

Navigating within virtual worlds and or unmanned vehicles in real 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. The user 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.

6.3.2.3 General technical details

NOTE More technical details on planning short paths with clearance using explicit corridors, as described in 6.3.2.2, can be found in Reference [2].

An algorithm for camera planning needs to have a simplified geometrical representation of the world. Initially, this is 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 normal vectors of obstacle surfaces and Voronoi edges) together with their closest points to obstacles (i.e.

normal vectors of obstacle surfaces). 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 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.

6.3.2.4 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 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.

6.3.3 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. This use case aims at contributing to changing a little bit this concept, as the main goal of the virtual travel use case is to offer 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 with the help of elaborated 3D images and pictures of this place.

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

One of the objectives of the use case is the travel motivation to this destination, as the user has a lot of information about it, more easily available than other tourist places. The virtual travel is the part before the travel.

However, if users decide to visit this destination, they are able to use the virtual world, due to the virtual traces and real places use case, where they can share all experiences, impressions and feelings they gained during the stay with family and friends. It is the part after the virtual travel.

6.3.4 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).

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, giving users the feeling of really being there.

Imagine, for example, Sue and John travelling to Gran Canaria for a short break. Their two teenage children Dan and Mary stay at home. Sue's father, 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 John and Sue's children 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 about what his parents have been up to. The 3D 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 Sue and John had. Then, Dan lets his avatar fly to the coastline where his parents took 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 Mary 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 colour. Dan and Mary 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 Mary, “Maybe we should have joined mum and dad after all. Hey, Dan, check out that place over there.” Mary 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 Mary share their thoughts on the bar 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 Mary. 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 Sue had today.

6.3.5 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 opt for a more general guided tour. Later, other team members can build in technology to get information about the interesting spots. In addition, technology for multi-lingual or text input can 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 can 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 is a path which is traversed by the scooter;

— the camera path is a path which is traversed by the camera. The camera keeps 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 is the series of locations which are being looked at during the tour. To reduce the risk of motion sickness, the camera and aim combination moves smoothly. In addition, the camera looks at future positions (about 1 s ahead) such that the user can anticipate what is going to happen.

6.3.6 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.

There is a usage scenario of enhanced interface and an 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.

6.4 Instantiation D: interoperable virtual worlds

6.4.1 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.

6.4.2 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, this way, allowing a user having a virtual anonymous but common identity in any virtual world.

6.4.3 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 the 3D virtual world community and 3D game console. 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.

Consider an example of virtual shopping. 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 a new online shopping service.

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.

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 a virtual golf club reflects the same physics in the virtual world by identifying physical characteristics of the real golf club such as sound, weight, texture and stiffness.

6.5 Instantiation E: social presence, group decision making and collaboration within virtual worlds

6.5.1 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.

6.5.2 Instantiation E.1: social presence

There is a universe where people live in groups that occupy certain places of it, almost exclusively on planet Earth. People create these groups that we call societies, because of the need to communicate and collaborate with others. 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 can 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 its users with the impression of being in another place, giving them the opportunity to operate applications and devices that run or are located in remote places. Metaverse will also provide a connection with numerous services available to the real world.

Hence, in the future, Paul, who lives 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 can 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, Penny wears 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 executes 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 science and industry of "La Villette" located in Paris. 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.

6.5.3 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 NGOs more specifically, virtual worlds can function as an important future communication platform. This use case focuses 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 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 can 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 can facilitate decision making tasks in the context of spatial planning, since different parties are able to visually experience the 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.

If a community council is 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, can 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.

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?);

— the effects of a 3D environment on the process of group decision making;

— the extent to which virtual worlds support group interaction;

— the extent to which the opportunity to virtually experience the final result of a zoning plan facilitates more effective and efficient decision making.

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).

6.5.4 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 favourite 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) virtually meet with consumers in a challenging 3D environment in order to engage in a consumer driven co-creation process.

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 firms can 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;

— the extent to which electronic word-of-mouth communication by opinion leaders in virtual worlds influences the adaptation process of new product ideas;

— how these co-creation relationships can be successfully deployed for marketing purposes.

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

6.6 Instantiation F: interactive haptic sensible media

6.6.1 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 9 shows the exact system architecture for the instantiation. The instantiated system architecture can be adopted by any interactive sensory effect media.

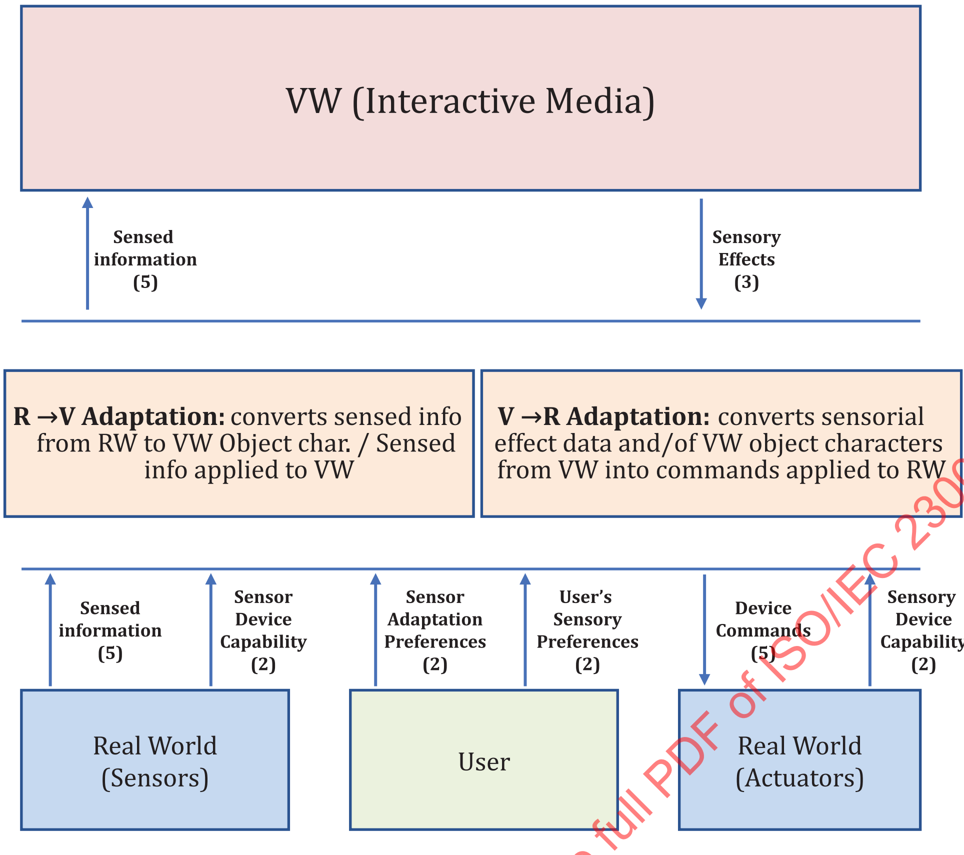


Figure 9 — Example of system architecture for interactive haptic sensible media

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.

6.6.2 Instantiation F.1: Internet haptic service — YouTube, online chatting

Assume a user is wearing a haptic wristband and browsing the Internet.

— The user connects to www.youtube.com, searches for the term "figure skating" and then selects one of the videos 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 on 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.

6.6.3 Instantiation F.2: next-generation classroom — sensation book

The classroom is equipped with an immersive book, such as 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.

— 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 can hold a haptic device and touch the bell. In addition, the sound of bell can be out from a speaker. The harder students hit the bell, the louder the sound.

— In the science class, students learn about lithology.

— As soon as students press the 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.

6.6.4 Instantiation F.3: immersive broadcasting — home shopping, fishing channels

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

— A viewer turns on TV and changes the channel to home shopping.

— The product which is being sold is a PDA and the 3D corresponding virtual object is displayed on the TV screen.

— A shopping host explains functions of the PDA and asks viewers to click the buttons.

— When the viewer who holds a haptic device clicks a button of the PDA, he/she can feel a button click sensation with the glassy body of the PDA.

— In addition, the viewer who holds the PDA can measure the weight and see that 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 that the viewer is holding also vibrates.

— In addition, the viewer feels force when the fisherman on the screen catches the fish.

6.6.5 Instantiation F.4: entertainment — game (Second Life®, StarCraft®[2]), movie theatre

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 gloves.

— Haptic information is authored synchronously with audio video information and is displayed together with the synchronized audio-visual media.

2) StarCraft® is a trademark of Blizzard Entertainment, Inc. It is given as an example for the users of this document and does not constitute an endorsement by ISO or IEC.

— During the scene when a web is coming out of Spiderman, the 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 Second Life®, the haptic device approaches the users and provides shaking-hands motion and force feedback.

— After an hour, the user starts the game StarCraft®.

— While the user is attacking the enemy, he/she feels the bottom right corner on the back of the tactile vest (5 o'clock) is hot since the bottom of the user is firing and attacked by other enemies.

— The tactile vest (including thermal and vibration display) triggers heating mode when the bottom is burning, and vibration mode when soldiers are attacked from the back of the user according to the map of the game.

6.6.6 Instantiation F.5: virtual simulation for training — military task, medical simulations

6.6.6.1 Military training task

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.

6.6.6.2 Medical simulations

In a medical school, a medical student has the first simulated minimally invasive surgery (MIS). A virtual patient is anesthetized and two assistant nurses are standing next to the patient.

— The medical student puts an endoscope into the virtual patient's abdomen. Then, the inside of the abdomen is displayed on the screen according to the position and angle of the endoscope.

— The student touches the gall bladder with the tool for diagnosis and feels it is harder than a healthy gall bladder.

— The gall bladder is cut by a 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.

6.7 Instantiation G: bio-sensed information in the virtual world

6.7.1 System architecture for bio-sensed information in the virtual world

System for bio-sensed information in the 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, etc. 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.

6.7.2 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 (for example, EOG sensor) placed strategically on the forehead and under the eyes. Wherever the user's eyes look, the virtual environment can be displayed appropriately. Furthermore, convergence of the eyes on a specific object in the environment can be detected. This can be used to select and object in the virtual environment.

6.7.3 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, therapists, etc.) play a key role. In this process, intelligent environments for health care play an important role in deciding which multimedia contents integrate and enrich the real space.

6.7.4 Instantiation G.3: mental health for lifestyle management

The World Health Organization (WHO) predicts depressive syndromes will be higher than any other health problem within 20 years (450 million people). The anti-depressant market in the European Union (EU) was 4 500 000 000 EUR in 2010.

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

— wearable sensors (accelerometers, ECG, EEG);

— home monitoring (bed sensors);

— electronic information (amount of e-mail, 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 economic impact of mental health problems equals 3 % to 4 % of the GNP of EU states and mental health problems are the most common reason for early retirement in OECD countries. However, mental health problems are currently under-diagnosed and under-treated because of the limited healthcare resources.

6.7.5 Instantiation G.4: food intake for lifestyle management

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

The health consequences arising from poor nutrition are serious, in particular with cancer and cardiovascular diseases, but also obesity which carries a related set of health risks in the form of cardiovascular diseases, type 2 diabetes, musculoskeletal disorders like osteoarthritis and some cancers. The cost of obesity to society is enormous: approaching 1 % of the GDP in some countries of 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 issue. These currently require complex and extremely time-consuming participant involvement (e.g. for 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 behaviour change strategies to persuade the users to improve their food intake.

6.7.6 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.

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 if, for example, 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.

For instance, a system permitting 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, can significantly improve the (lifestyle) conditions for the patient while lowering the cost of the procedure.

Techniques were developed in the project for:

— automatically detecting adverse events during training;

— improving training outcome e.g. by an automatic adjustment of training duration and "load"; and

— 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.

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 is also 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 can 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 are stored in a medical document called "training report" during a training session, based on the HL7 clinical document architecture. Data in this document is 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 allows for a better forecasting of trends in the patient's health status.

6.7.7 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.

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 of 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 the one hand, this helps the patient to feel more comfortable and safer. On the other hand, this system offers valuable information to the physician for a better analysis of the disease evolution and prediction of risks complications. The educational effort in the project serve to promote adherence and, thus, the benefit of the technology for patients. Remote monitoring technologies can transmit data on a regular, real-time basis and prevent hospitalizations by identifying and treating problems as well as triggering adjustments in care before negative trends reach crisis stage.

Different kind of sensors are needed to get the information from the patients in each scenario. These sensors follow the Continua Health Alliance standards or, on the contrary, any type of bridge 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.

6.8 Instantiation H: environmental monitoring with sensors

6.8.1 General

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.

6.8.2 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 growth 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 becomes urgent and critical.

The primary air pollutants found in most urban areas are carbon monoxide, nitrogen oxides, sulfuric oxides, hydrocarbons and particulate matter (both solid and liquid). These pollutants are dispersed throughout the 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 sulfuric dioxide and suspended particulate matter are emitted by a large volcanic eruption or sandy dust clouds. Figure 10 illustrates major air pollutants with possible application areas.

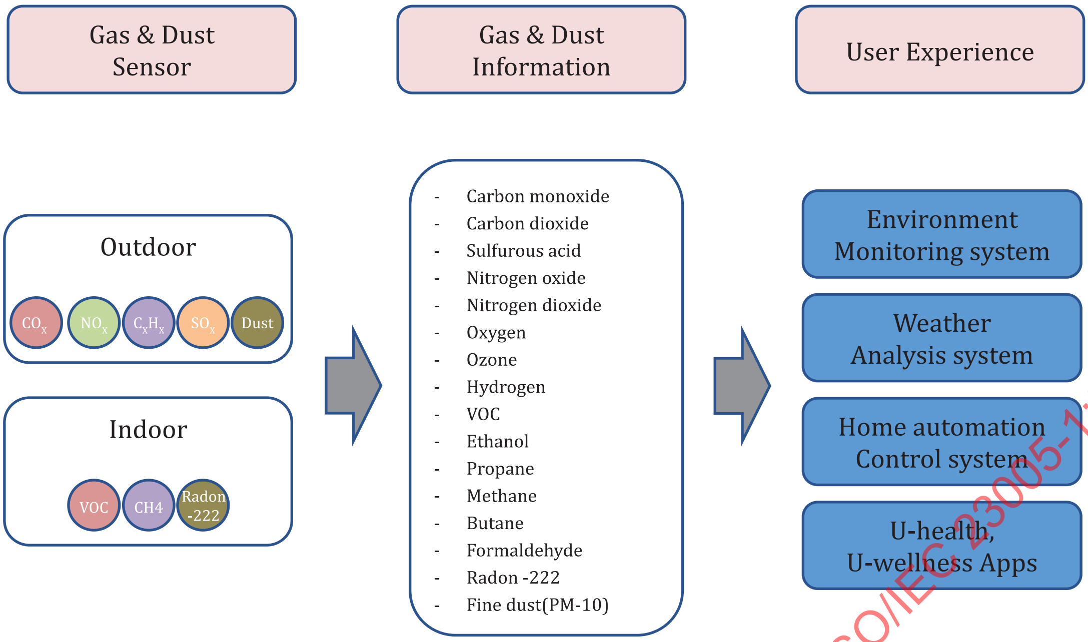


Figure 10 — Gas and dust sensor information and use case examples

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 are important information to describe the room condition where people consume some contents.

6.8.3 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 comes from specific sources such as copying machines, electrical and telephone cables, mould and microbe-harbouring air conditioning systems and ducts, cleaning fluids, cigarette smoke, carpet, latex caulk and paint, vinyl moulding, linoleum tile and building materials and furniture that emit air pollutants such as formaldehyde. Another major indoor air pollutant is radon-222, a colourless, odourless, 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.

6.9 Instantiation I: virtual world interfacing with TV platforms

Via the users' home network, the Second Life® game content is transferred to the 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 decrease as the bandwidth of the network is reduced.

6.10 Instantiation J: seamless integration between real and virtual worlds

6.10.1 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 commercially 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, virtual sensors with higher abstractions are also needed. In order to visualize human behaviour from the real world into the virtual world, there needs to 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 is a need to have relative device position information. With this information, sensor data from specific devices can be spatially mapped and, for instance, 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.

6.10.2 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 needs to 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) needs to 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.

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 as follows.

— 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 four 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 to the device service to talk to the device that registered the service (in this case, a webcam) and transmit the appropriate configuration to obtain the desired video stream, which is sent back to the TV 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 into device-dependent data formats, and function calls that execute the desired action and/or obtain the desired data.

In 6.10.1, we outlined the proposed platform architecture for communication between real and virtual devices. Figure 11 shows a use case diagram of the interaction between devices in our system. They can:

— 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.).

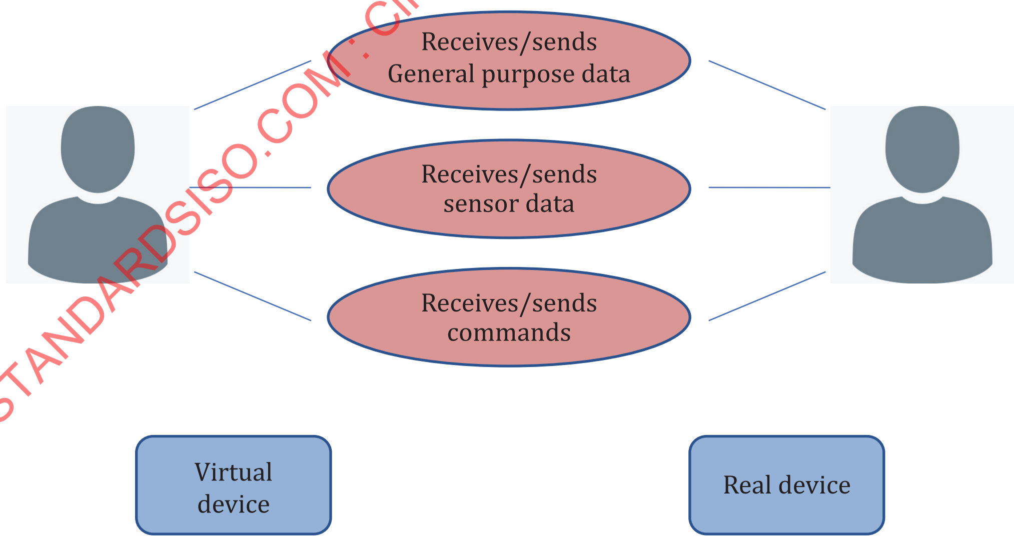


Figure 11 — Interaction between virtual and real devices

The server sensor node is the node that gathers all the information from the wireless sensors for processing. On the server side, there is a service selector which is a management tool for wireless devices. All the services of the discovered wireless devices are listed by the service selector. Via this manager, individual and groups of sensors can be selected to route their data to a particular service.

Sensor services are modules with specific sensor logic. A sensor service can be localization. Based on the selected sensors (nodes), in-network localization can be performed using values from those sensors (nodes). Each sensor service conforms to a service interface which makes them pluggable to different outputs. The followings are the basic services that is assumed for enabling given scenario.

— **Presence notification**: This service broadcasts a message to the system when one (or more) person enters the room. If a request is made with no parameters, the service returns a list with all the known persons IDs and their current presence status. The service can receive as a parameter the unique ID of a person and return their presence status.

— **Location tracking**: This service provides the current location of a person in the room. The service receives as a parameter the unique ID of the person whose location is needed and returns the corresponding pose in 3D space (Cartesian system).

— **Chair usage**: This service indicates if the chair is being used at the given moment or not. If the type of chair allows for the use of more than one person, the service provides the amount of people currently sitting as well as their IDs. A request with no parameters obtains a response with a list of all the chairs available in the system and their status of occupancy. If the parameter is a chair's unique ID, it returns its occupancy status.

— **Thermometer**: This service provides information on the ambient temperature in the room.

— **Illuminometer**: This service provides information on the brightness of the lighting in the room.

— **Audio**: This service has two modalities:

a) a request without parameters produces a list of all audio capture devices available in the system and their properties;

b) a request with a specific device ID tries to establish a connection with it and initiate audio capture.

— **Speech**: This service has two modalities:

a) a request without parameters produces a list of all text to speech devices available in the system;

b) a request with a specific device ID tries to establish a connection with it, opening a port to transmit text that is translated into speech.

— **Video**: This service has two modalities:

a) a request without parameters produces a list of all video capture devices available in the system and their properties;

b) a request with a specific device ID tries to establish a connection with it and initiate video capture.

— **Robot commands:** This service allows sending commands to be executed by the robot. Upon a request without parameters, the service provides a list of the available commands and their parameters. If the request contains a command, this is sent to the robot for execution.

6.11 Instantiation K: hybrid communication

With the success of virtual worlds (800 million registered accounts in early 2010, with about 20 % quarter-on-quarter increase in 2009), more and more people are experiencing the reality of parallel universes in almost every sector: (serious) gaming, socializing, gambling, dating, educational, events/promotional, sports, music, fashion, TV, etc.

But human beings need to share, exchange, communicate, and the borders of existing worlds defined by their inner rules (currency, appearance, roles, grades, physics, politics, communities and communications, etc.) are (pragmatically) not acceptable regarding the end-user experience.

If we consider the telecommunication facet of virtual worlds (VW), players naturally need to communicate with the “outside”, friends or connections that are not necessarily connected or existing in the current world. This need is partially fulfilled thanks to workarounds such as communication applications simultaneously launched with the VW client (e.g. Skype, MSN, GTalk). Other solutions, either hardware or software, were also released to address this issue (mostly for players), such as TeamSpeak, Ventrilo or Roger Wilco. Finally, system-wide solutions like Steam[3)] (digital game distributor) integrated community and communication solutions in their offer as they identified the users’ wishes and the market potential in term of free advertising.

The existence of such workarounds, dedicated software or emerging game system features proves that a need actually exists and is, somehow, partially addressed. But the number of different approaches also illustrates the lack of a unified solution for virtual–real worlds inter-human communication whereas a generic service would benefit to all actors.

— **Customers**: Providing intra-world communication features is just obvious but users also need to connect with the “outside” (e.g. friends). But connecting worlds implies connecting identities which need to be handled very carefully, e.g. SupaNoobKilla is possibly not the most appropriate ID for use with your boss. So, privacy and reachability but also the communications mean exposure and its control should be simultaneously ensured.

— **Service providers**: Integrating generic and interoperable communication functions would be a great advantage for developers, accelerating the deployments, easing the maintenance and most of all connecting their users to an unlimited number of potential new customers (free advertising). In return, third parties (such as telecom operators) can then easily benefit from the openness of the model, leveraging new revenue opportunities.

Considering the previous works of ALU Bell Labs Hybrid Communication department on telecommunication technologies between heterogeneous environments (web, fixed and mobile networks, etc), we are in position to propose similar solutions adapted to the intrinsic characteristics of virtual worlds.

The previously developed technologies (see live experimentation at Reference [3]) were realized generically enough to be easily extended to different use cases (here the virtual environments). The concept of Communication Hyperlinks suits perfectly with virtual worlds’ communication requirements.

— **Privacy protected**: Personal details are not disclosed (no phone number, no identity, the communication hyperlink is interpreted by the server which plays the role of secured proxy).

— **Pure-web approach**: The service can be easily embedded into any application.

— **Full control**: The publisher can modify at any time the properties and the configuration of his COM hyperlinks. So, a hyperlink can be linked to any communication mean (phone, SIP, e-mail, IM, web-based COM app, etc.).

— **Open model**: A set of API allowing third-party applications (communication mash-ups) to handle the COM hyperlinks.

In the framework of Metaverse1 project, several prototypes have been developed to demonstrate this technology in virtual world conditions. The technology was implemented in a specific container in Dundal.com and tested with the massively multiplayer online video game World of Warcraft® and the Orange Labs’ open source platform Solipsis[4)].

The solution is generic, i.e. it would be implemented the same way in different virtual worlds as far as the underlying rendering engine is able to handle web texture (quite common nowadays, e.g. Second

3) STEAM is a trademark of Valve Corporation and World of Warcraft is trademark of Blizzard Entertainment, Inc. They are given as examples for the users of this document and do not constitute an endorsement by ISO or IEC.

4) Solipsis is a platform by Orange Labs. It is given as an example for the users of this document and does not constitute an endorsement by ISO or IEC.