INTERNATIONAL ORGANIZATION FOR STANDARDIZATION
ORGANISATION INTERNATIONALE NORMALISATION
ISO/IEC JTC 1/SC 29/WG 11
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC 1/SC 29/WG 11/N9902
Archamps, France, May 2008

Title: **Draft requirements for MPEG-V**
Source: **Requirements**
Status: **Approved**

Table of Content:

Abstract

The ‘Information exchange with Virtual Worlds’ project intends to provide a standardized global framework and associated interfaces, metadata definitions and the like, to enable the interoperability between virtual worlds (as for example Active Worlds, Second Life, IMVU, Google Earth, Virtual Earth and many others) and with the real world (sensors, actuators, vision and rendering, robotics (e.g. for revalidation), (support for) independent living, social and welfare systems, banking, insurance, travel, real estate, rights management and many others).

Virtual worlds (often referred to as 3D3C for 3D visualization & navigation and the 3C's of Community, Creation and Commerce) integrate existing and emerging (media) technologies (e.g. instant messaging, video, 3D, VR, AI, chat, voice, etc.) that allow for the support of existing and the development of new kinds of social networks.

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

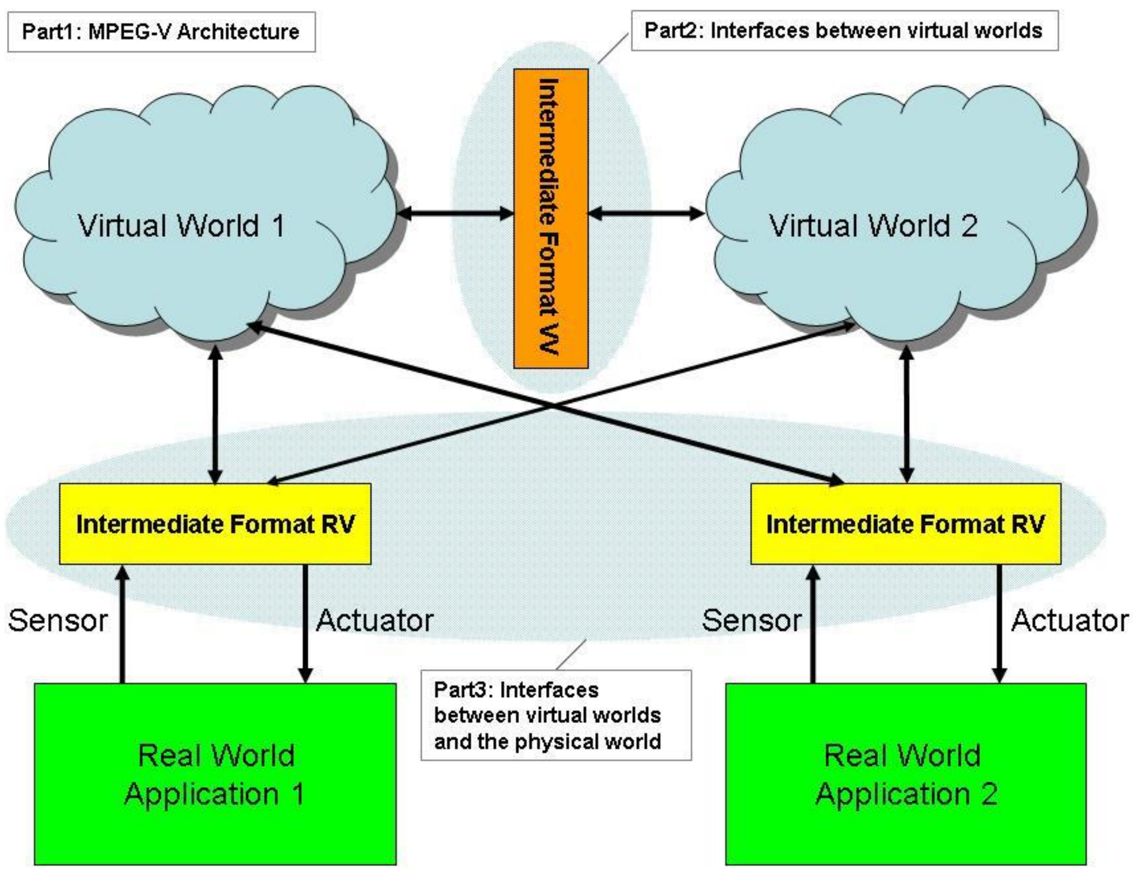
Virtual worlds have entered our lives, our communication patterns, our culture, and our entertainment never to leave again. It's not only the teenager active in Second life and World of Warcraft, the average age of a gamer is 35 years by now, and it increases every year. This does not even include role-play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

Technology improvement, both in hardware and software, forms the basis. It is envisaged that the most important developments will occur in the areas of display technology, graphics, animation, (physical) simulation, behavior and artificial intelligence, loosely distributed systems and network technology. Furthermore, a strong connection between the virtual and the physical world is needed to reach simultaneous reactions in both worlds to changes in the environment and human behavior. Efficient, effective, intuitive and entertaining interfaces between users and virtual worlds are of crucial importance for their wide acceptance and use. To improve the process of creating virtual worlds a better design methodology and better tools are indispensible. For fast adoption of virtual worlds we need a better understanding of their internal economics, rules and regulations. And finally interoperability achieved trough standardization.

As stated in the w9424 "Information exchange with Virtual Worlds 'Context, Objectives and Use Cases'" output document of the 82nd MPEG meeting (Shenzhen, China, October 22 – 26, 2007) the intended standard for 'Information exchange with Virtual Worlds' will consist of at least three parts (see also the illustration below):

- Part1: Architecture
- Part2: Interfaces between virtual worlds
- Part3: Interfaces between virtual worlds and the physical world

The first part (Part1) will describe an overall architecture that can be instantiated for all the foreseeable combinations of virtual worlds and real world deployment. This architecture will be derived from the requirements defined by a set of use cases (a number of these use cases are presented in the appendix of this document) that will be chosen to cover the entire foreseeable virtual world / real world combination application domains (see the illustration below for an example of the foreseen architecture and related interfaces for interoperability). The architecture positions all the related technologies, endorsed technologies and the MPEG technologies defined in this project. (This includes the interfaces between virtual worlds (Part2) and the interfaces between virtual worlds and the physical world (Part3)).



As an example for Part2 (Interfaces between virtual worlds), imagine that a user want to change his activities from one virtual word to another virtual world for instance due to a change of interest. With current technologies this means that the user loses all the effort and value invested to create his avatar and environment and has to start from scratch in the other virtual world. If a format can be defined that allows the export of specific characteristics from one world and import of this format to another world a lot of the user invested effort and value is preserved.

An example for Part3 (Interfaces between virtual worlds and the physical world) would be the interfaces with peripherals that allow for the extension of the experience created in the virtual world into the physical world and vice versa.

Requirements Areas

Initially the following (not complete list of) requirements areas are foreseen for the ‘Information exchange with Virtual Worlds’ standard:

Requirements related to Part2: ‘interfaces between virtual worlds’:

- MPEG-V shall define an intermediate format that that can be used to move characteristics natively represented in virtual world 1 (export) to the native format of any other virtual world (import)

- The intermediate format shall be able to represent characteristics of assets (avatars, objects, …), e.g.:
    - for an avatar: body size, relative sizes of body elements, color and type of hair, color and shape of eyes, …
    - clothing for the avatar: type of clothing, shape description, color, texture, …
    - accessories carried by the avatar: jewelry, watches, …
    - personal attributes the avatar can make use of: weapons, wallets, animations, …
    - other objects: cars, house, furniture, …

- The intermediate format shall be able to represent dynamic attributes, e.g.:
    - movements of avatars
    - behaviors of objects

- The intermediate format shall have scalability features e.g.: the ability to move characteristics form a complex format to a simple format and the other way around

- MPEG-V shall include communication protocols to
    - export and import assets in the intermediate format
    - ensure security, trust and privacy

- MPEG-V shall support metadata related to assets e.g.:

    - Location

- MPEG-V shall support an intermediate format for
    - user identity
    - user profile
    - ownership of virtual objects
    - rights and obligations associated to virtual goods

- MPEG-V shall support an intermediate format for
    - the transfer of (virtual) currencies

- Additional requirements from platform diversification ?

Requirements related to Part3: ‘Interfaces between virtual worlds and the physical world’:

- MPEG-V shall support a representation format for sensor & actuator data
    - blood sensor, smart camera, temperature, humidity, smell, …
    - robotic devices, illumination, …

- MPEG-V shall support audio / video communication between users from real and virtual worlds

- Additional requirements for interfaces between real and virtual worlds platform?

Appendix 1: Examples of requirements areas

This remainder of this document presents a number of inputs for the definition of the areas of requirements that have to be allocated to the applicable parts of the intended standard at a later stage. These inputs have been received as input for the Call for Requirements for 'Information exchange with Virtual Worlds' at the 84th MPEG meeting in Archamps, France from April 28 to May 02, 2008. These inputs have been prepared by a subset of the members of the Metaverse1 consortium before the actual start of the Metaverse1 initiative and are for that reason incomplete. The first phase of the Metaverse1 initiative, when started, is targeted at an extensive use case based requirements analysis in parallel with an extensive state of the art overview, the results of this exercise will be made available to the 'Information exchange with Virtual Worlds' standardization process. These provided preliminary inputs don't have the intention to be complete but serve as a starting point for the requirements gathering process for the specific area.

1 Avatar 'Mobility'

Since some years ago first virtual characters –so called *avatars*- with the function of human-device communication or human-human communication through a device appeared, either acting in an autonomous way or as direct representation of the user, their use has gone growing. Nowadays, avatars can be found in different fields of application, among which virtual communities can be point out.

Although generally, the role of the avatar in these environments is the user representation, being them driven by the user all the time, the likeness that is expected to be obtained between these Metaverses and the real world, we can difference four roles to fulfill by an avatar:

- **User puppet.** As it happens in current virtual communities, these avatars are driven by the user, taking a role of middleman in human-virtual world communication, whether it is with the goal of communicating with other virtual world inhabitants or interacting with it.
- **Author puppet.** It is the one that executes the orders defined by the author of the application. It does not have a decision taking capacity, as it is a passive avatar.
- **Interactive avatar.** The interactive avatar is capable of interacting with the user, with the application and with other avatars, by interpreting each event produced in the interaction and generating the relevant response.
- **Autonomous assistants/inhabitants.** The role of autonomous assistants is to offer services and assists the inhabitants in their necessities without mediation of a human. Examples of this role can be a shop assistant or a virtual teacher. In these cases, the avatars are no directly driven by user, but they acts in an autonomous or semi-autonomous way. In the case of virtual world inhabitants, they are autonomous avatars not driven by users. Their role is not necessarily assists users, but they cohabit with them in Metaverses.

In the case of representation of the user and author puppets, the functional features of the avatars more and more allow different expressions, gestures and motion ways. In short,

they allow features that expect users can express their thought in a more natural way or show their emotions in a more clear way than with only text messages.
In the case of autonomous or semi-autonomous avatars, they usually have artificial intelligence modules that provide them with a behavior and level of response nearby to the humans.
Interactive avatars use a reduced mix of these features, its behavior is driven for some rules and it is important that they shows the response to interactions in a like-human way.

However, as all novel technology, it presents some lacks yet, over all referred to the definition of the **appearance and behavior** of an avatar. The definition of the whole avatar is a closed format, specified for each application, and without possibilities of transfer between them.

Up to now, some standardization, which can cover the avatar appearance and behavior in all of its possible roles, might be created. This specification will allow to specifying in an easy way all the aspects related with the virtual character in the virtual environment.

The description of the (movement) behavior of the avatars and possible other (movable) objects in the virtual world will allow for the export of these motions for computer-controlled characters, that are characteristics of for instance an avatar, and the import into another / new virtual world. In the other area, the interfacing between virtual worlds and the real world these motions can be developed in the virtual world for an avatar resembling a device (for instance a robot) and can be mapped on a real robot for the use of robots in dangerous areas, for maintenance tasks or the support for disabled of elderly people and the like. These descriptions should include:

- A description of the places where the crowd and vehicles can come.
- Shape of the objects, including textures and animations and their variety.
- Density of characters and vehicles, probably location dependent, and variation thereof.
- A high-level description of the type of behavior (wandering, shopping, …)
- A description of the required type of character interaction (meeting, joining, avoiding, ...)

A (formal) language should be developed for behavior descriptions like for instance (or an adaptation of) the Behavior Markup Language (BML) and the Emotion Markup Language (EML) to model the behavior and the appearance & emotions of avatars in order to transfer the behavior and / or appearance & emotions between virtual worlds. This formalization can however also be used between the virtual world and the real word and vise versa. In the case of the transferring behavior and appearance & emotions from the virtual world to the real world implies the ‘rendering of the behavior and appearance & emotions on physical devices like robots like devices and / or devices that create sensory experiences. In the case of transferring between the physical and virtual world captured behavior and appearance & emotions in the real world can and be used as input for applications and services in the virtual world.

In this section of the proposal, we try to give a step forward, allowing define in a common and transferable way the appearance and movements, emotions and gestures of the avatar by means of the specification of a easily understandable markup language. This language will be XML-compliant and will allow independence of the virtual world in which it is used.

Proposed language elements and exclusions

The goal of this avatar standardization proposal section is to obtain **a descriptive format** specifying the avatar features and not the way of render them, it is said, the 'what' and not the 'how'.
So, the goal is not to define a fixed format for the avatar. Each virtual world could be rendered by different animation engines, animation techniques, geometric formats, etc.
In fact, a powerful animation engine will be able to show the same avatar configuration with more complexity –it is said better- than a less powerful engine.
For obtaining this independent specification, the avatar file format will fulfill the next requisites:

- **Independence of the geometrical format.** There exist a lot of geometrical file formats –obj, 3ds, fbx, vrml...-, some of them are open and others are proprietary, and each virtual world uses one or some of them or it uses any own format. A format that allows transfer an avatar between virtual worlds should be independent of these formats.
- **Independence of the animation techniques.** The language will define movements and gestures of the avatar but not the way they are rendered. The format will specify 'my avatar can jump', but not 'the jump is done by means of inverse kinemactics'
- **Independence of the animation engine.** In the same way, the file format will not contain information about specific issues of a concrete animation engine.
- **Independence of parsers/loaders.** This section proposes a file specification, not the way of loading its content in a virtual world. It is left in the virtual world designer hands, in order to obtain the more efficient load way for each case.

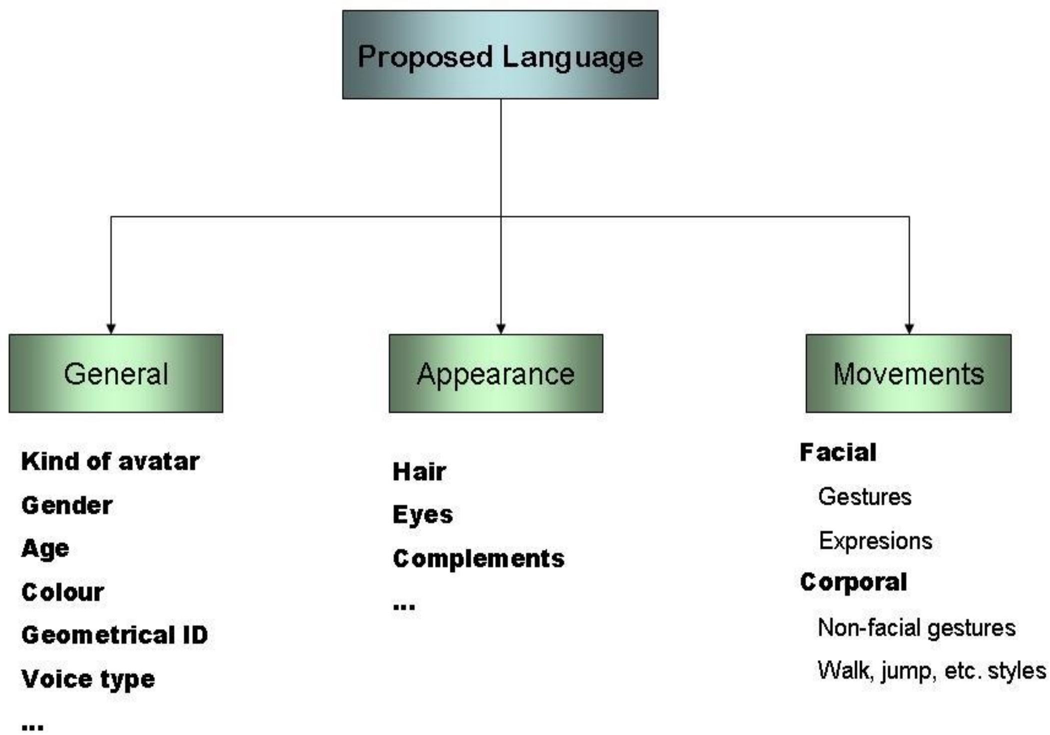
The language is defined as a XML-compliant language specifying the attributes and appearance of the avatar. The decision of using a XML-compliant language has been taken due to the simplicity and legibility of files defined in this way. They are plain text files based on tags with attributes, and it is very easy to develop authoring tools that allows create the file without necessity of understand the markup language.
Moreover, a lot of programming libraries and editors that gives support to languages defined according to this markup standard exist. Prove of its efficacy is that there exist XML-compliant markup languages for defining practically all the things that can be imagined.

Brief Description of Language Contents

The language contains information about three kinds of avatar features:

- **General features.** In the highest level tag –*'person'* tag- there is a set of attributes specifying general features of the avatar: Human or not, gender, age, skin color and a identifier of the geometrical model of the avatar.
  This identifier will tell virtual world engine what kind of geometrical model load for the avatar.
- **Appearance features.** There are tags specifying the appearance features of the avatar: Hair, clothes, complements. All of them have an identifier attribute that tells the virtual world what kind of geometrical model load for each feature.
- **Movement features.** A set of tags specifies the movement capabilities of an avatar. Each movement has an associated identifier. The virtual world animation engine will use its own animation techniques in order to obtain the avatar motion.
  The features are divided in facial and corporal features depending on if they consist in facial expressions or gestures or if the movement implies an spatial translation of the avatar or of part of it.

The figure below shows the language structure:



Language structure

So, a file using this format will contain a set of nest tags specifying these attributes and using it, the engine of each virtual world will be able of select the most similar appearance and movements for an avatar when it enters in the virtual world.

Advantages of this approach

The main idea of avatar standardization is to have appearances that identify the users in all Metaverses they enter. A current simile with this is the use of cookies for web sites. These cookies contain information about user preferences about a web site and are stored in the local computer of the user.
When the user enters the web site again, the information is retrieved and the web site is modified in order to reflect these preferences.

Advantages from user point of view

The users will be able to use the appearance/s that they prefer for their avatar in all Metaverses. Due to the format of the language and its independence of loaders/parsers, authoring tools can be created that allows the user configure his/her avatar in his local machine, without necessity of enter in a virtual world for configure it.
The user can use the authoring tool that s/he prefers for creating its avatar appearance and the results will be the same that with other one.
The resulting configuration would be stored in the machine and used when s/he want to enter in a virtual world.

Advantages from a Virtual World designer's point of view

The proposed language is independent of technical aspects. A virtual world will have its own geometrical models in its own format and will create the animations with techniques and engines that the designers have developed.
The language tags are a guide for the avatar appearance and movements, and the model identifiers only tells virtual world to apply the most similar model that it has to the avatar.
If a virtual world has not a model in its database, it could replace it with other similar that is in the database.
In the other hand, if a virtual world can provide extra models –as selling packages with extra complements-, neither there is a problem. These models will be translated to their correct tags in the final XML file. In the case the user enters in this concrete virtual world, the models will be shown exactly, and in other virtual worlds they will be replaced with other models.

2 Exchange of virtual objects

Next to the movement, behavior and appearance & emotions of avatars and other objects the (descriptions of) avatars and / or other objects themselves need to be exchangeable between virtual worlds, the exchange of 3D scripted models (avatars, synthetic objects) between virtual worlds:

Interchange of assets and synchronization of events between real and virtual worlds:

- Exchange of 3D scripted models (avatars, synthetic objects) between virtual world and real world

- Exchange of (2D) video snippets (streaming, file based) based on position information between virtual world and real world
- Exchange of (3D) video snippets (streaming, file based) based on depth and/or location information between virtual world and real world
- Exchange of (3D) interactions (triggers) from real world (sensors) to virtual world (3D models like avatars, synthetic objects)
- Exchange of (3D) interactions (events) from virtual world to real world (activators)

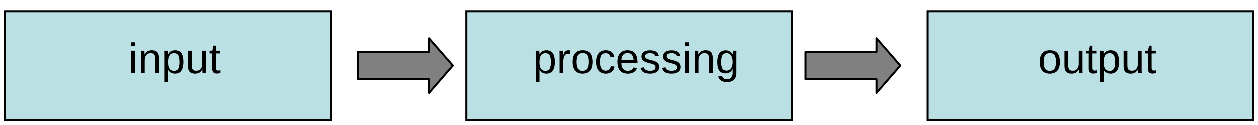
Note: that sensors and actuators in this context can be either intimate (directly connected to human/animal/plant/...) or ambient (HVAC, domotica, smart objects/...).

3 Bidirectional immersive sensory effects

In the recent years networked virtual worlds have become popular and further growth is predicted. There is a trend towards a more immersive and realistic experience. Up to now conventional media with audiovisual contents has been presented via display devices and speakers. However, virtual experiences can be made more immersive by adding other multi-sensory effects, like light, force feedback, motion etc. as well as more natural input methods.

Model

Interaction with virtual worlds can be viewed on a high level as:



- Inputs: human sensors (bio-signals), commands, voice, and gestures as well as traditional key/mouse or controller input. Besides human sensors, environmental/context sensors could also be used as inputs, e.g. movement, activity sensing and GPS positioning.
- Processing: translates the inputs to state and actions. It maps inputs to outputs and checks for constraints (collision detection, Artificial Intelligence, world simulation, etc).
- Output: display devices, audio output systems, or other output effect devices like amBX, force feedback, motion, light, robotics, etc.

Abstraction from the ‘raw’ input

To avoid that each virtual world has to process the raw input data in a proprietary way, a kind of middleware or standards would be useful to map raw input to a semantically meaningful, machine-readable representation. (Of course exceptions will be there: e.g. a virtual world *server* that parses the raw samples of spoken commands by the user, and maps it to meaningful AI actions that are only valid in the context of this specific world.)

Often, inputs can be straightforwardly mapped to output and there is no need for abstract representation of the raw input. E.g. if I voice chat, my voice is sent out into the virtual world and propagates according to simulated physics. Other people hear me chat. However, an extra layer of **interpretation** can be added for some cases. Example: If I speak to an AI character (not a real player), the character can determine from the tone of my voice and what I say whether I'm friendly or hostile. This can be used in an RPG to determine the course of the story and development of your character and in-world reputation. The same holds for many other input types: if interpretation is added (i.e. an abstract machine-readable representation of my inputs) it allows interesting interaction with simulated world entities, making future virtual worlds much more immersive and life-like. The question remains where this interpretation would occur: client-side in processor, on-device or server-side or spread over several places.

Considerations for the MPEG-V requirements

Input

- Standard should define how sensor input signals are represented in a standardized way
  - For a broad range of current and foreseen sensor (input) devices
  - In a way that allows fast, real-time processing of signals taking into account timing requirements of the applications
- Immersive virtual-real interaction requires natural interaction e.g. via human sensors, not only traditional “chat / text” oriented interaction
  - Implies security, privacy issues (cameras, human observation, …) to be considered
- Standard should define timing requirements for latency (=delay), and throughput, for sensor input signals
  - Values will be different for different categories of applications, eg gaming worlds, mirror worlds, social worlds etc. This has to be investigated.
  - Latency could be defined in a number of ways. For example, latency measured from the moment of user input to the moment the virtual experience is rendered (both on screen, and using effect output devices.)
- Plug and play (of input and output devices) ➔ see also MPEG-RoSE
- Specific sensor domains to consider:
  - Human poses -- standard could define a single representation for human body position and gestures. This representation can be decoupled from the actual sensors used to obtain this data. For example, it could be generated by a webcam, by an array of cameras, or body-worn motion sensors. Still the representation is the same.
  - Robotic poses – user manipulates a toy robot or digital pet to provide input into a virtual world this way

Processing

- The standard considers (where appropriate) interpolation of sensor input / sensory effects output. To be investigated ... interpolation and extrapolation are useful e.g. in case of delayed or lost server updates.
    - E.g. in sensor input data there could be 'hints' added to the server on how to extrapolate/interpolate in case of delayed or lost packets of sensor data.
- Potentially a streaming format for 3D content? (Includes 2D overlays, textures, etc.).

Output

- MPEG-V standard should offer similar features as the MPEG standard sensory effect output devices (RoSE). Perhaps, it could re-use this standard (by inclusion) and build on top of it.
- Plug and play (of input and output devices) → see also MPEG-RoSE
- Could define standards on how to drive robotics output? (avatar-to-robot)

4 Generic requirements

Next to the above area specific categories more general requirements related to the interfacing of the virtual world to existing systems ((displays, control devices, sensors, robotics, …), infrastructures, services and the like) in the real world exist that support optimal interfacing the virtual world to the vast amount of usable and useful (legacy) systems, infrastructures and services.

Appendix: Use Cases

1 Virtual Travel

The sun is shining and the waves are lapping at the sand. Just beyond the lounge chairs and palm trees, a group of people with perfect tans listen to music coming from a pair of seaside speakers. Between them, there are two girls having a great time: Anne and Carol. For them it is, in short words, a dream holiday, they are immersed in a place that they have only dreamed of (the beautiful island of Gran Canaria) from the comfort of their own home. They are enjoying incredible vistas and architectural wonders at anytime, without traveling outside their own doors.

The tour for Anne and Maria was exciting: sunbathing at the delicious beach of Maspalomas, having a walk close to the big crater of Caldera de Tejada and dancing at a beachside disco in Las Canteras. Time elapsed? Less than two hours.

With no tickets required, no money spent and no need to leave their seat. With the help of elaborate 3D images tourists can watch their avatars lay at the beach, have dinner at a romantic restaurant, or go out dancing at a crowded nightclub.
As they travel to the Gran Canaria virtual world, visitors can interact with other participants from all over the (real) world. Visitors can even capture a few photos or home videos to remind them of their trip and to hold them on image-sharing sites around the web. These vacations also promise things that real travel can't: no jet lag, no crowds, perfect companions and the chance to visit distant places.

2 Virtual Traces of Real Places

When we travel, either for business purposes or for leisure, 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 camera’s we can capture images and video footage of a remote place and bring it home. However, his way of sharing an experience has two main drawbacks:

1) Only a part of the experience can be shared with only a limited impact on the audience
2) There is a significant time gap between the moment we gain the experience (at some remote place) and when we share it with our friends and family (at home)

Wouldn't it be nice stay in touch with our friends and family while traveling? – Wouldn't it be nice to instantaneously share an experience with those who stay at home and to make them part of that experience? The goal of this use case is to incorporate travel experiences into a virtual world that can be instantaneously shared with friends and family. Digital content (images, video, sound, light atmosphere, 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.

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

Sue and John both carry a Metaverse compliant mobile device that is able to capture digital content (such as pictures, GPS coordinates, sound intensity, etc). One morning Sue and John decide to visit the big crater of Caldera de Tejada. In the afternoon they do some shopping and in the evening they enjoy a beautiful sunset on the beach.

The pictures they take with their mobile device are automatically geo-tagged and published in the virtual world. Some of these pictures they want to keep private, others look so great that they want to share them with everybody. When pictures are uploaded to the virtual world, they are automatically annotated with information about the light atmosphere and the sound intensity at the time and place the picture was taken.

At home the children of John and Sue are playing a computer game. Suddenly, there is a slight, temporary change in the color of the light that sits next to the computer. They know that this indicates that their parents have published new content in the virtual world. They stop their computer game and open their 3D Metaverse browser and gently glide to a virtual representation of the beach where their parents are.

Virtual Trace of a real place

The meta-data from the picture is used to re-create the appropriate light settings (see Figure 1) and then to play some music from the families' media centre in accordance to the atmosphere on the beach.

Sue's father is not able to handle a mouse and keyboard because of his old age. However, his daughter gave him a digital picture frame that is connected to the virtual world. This picture frame shows a map of Gran Canaria and on that map are small thumbnails of pictures that Sue and John have taken. A simple touch of a thumbnail allows him to see the full picture and to share the beach experience with his daughter.

The "Virtual Traces of Real Places" has a strong link to the "Virtual Travel" use case.

3 Language Learning

The Metaverse is an amazing opportunity for various types of learning; a few examples are illustrated below:

Object examination

A student can examine a skeleton; look closely at each of the bones.
It can take it a part; examine the physics of the skeleton.
A teacher can use a pointer to point at specific places.
A student can position various parts of the body (heart) inside.

This is a generic case of complex machines that can be viewed, and explained with a 3D environment. Similar representations can be done for cars, engines, homes, bridges, molecules etc.

Standardization opportunities: how to see a copy able version of this object; how to re-arrange it after the student have move the bones; how to update the skeleton;

Hard Core simulation.

In this example, taken from an experiment in the Tel Aviv university robotic lab we see a robot that solves a maze.
Technically the robot examines the 3D maze; sends the data to a remote "Artificial Intelligence –like program" and gain the instructions to solve the maze.
For the first time student can examine the actual work of motion planning algorithms in a simulated; joint; almost real life environment.

Standardization opportunities: how to allow fast and reliable communication from the virtual world into the real world; how to connect multiple robots in the same environments; how to connect the robot to real robots, and real sensors.

Language Learning (picture taken from the Aloft project)

The Metaverse is immersive; social; and kind.
Each object can have a written name on it; it can sense the language of the visitor and give its name in other (the target language); real mentors – with or without voice can add a human dimension.
This can be done across nations. So a Mexican kid can be learning English from an Indian teacher.

Standardization opportunities: scaling; cost; language marking; auto translation of text in chat etc.

These are just three of the thousands of examples of learning experiences and environments that can be carried in the Metaverse.
What is unique about the Metaverse in terms of a learning environment?

1. Ability to duplicate any learning object; track its usage; allow for specific times to be used.
2. Universal access to complex environments (measured).
3. Social factors (class mates; teachers, Teaching assistants)
4. Ability to connect your teaching with money. Brining the human touch into teaching for real money.
5. Ability to share learning and action with others (users can see how some one build and does experiments).
6. Ability to create; in multiple teams; across nations with real collaborative/competitive environment.
7. Ability to connect engineering; artistic; social considerations in learning.
8. Ability to capture processes in the Metaverse for later examination (log; movies; etc)
9. Ability to keep learning objects in a virtual learning personal space. (augmented intelligence)
10. Cost is 1/1000 of similar systems. All that you need is an internet connection and a computer (current hardware cost is a bit high but is assumed to get cheaper later).

4 Interplay of Virtual and Real World Social Behavior

It all started with Graham Bell who enabled people to talk to each other also when they were in geographically dispersed locations. While a tremendous progression for mankind, the rich context that is available when people are in the same location is partly or even completely lost. The new car in front of the door is not there to talk about, the painting on the wall neither, there are no kids running around and dinner is never served.

The progression from 2D to 3D interaction, from 2D to 3D visualization, and from 2D or static to 3D dynamic recording provides people with intuitive means to enhance their communication. Using smart objects, multi-touch tables or gesture control, they will be able to manipulate a virtual cup of tea on a 3D screen. Real-time replacement of a person’s live image by an intuitively controlled avatar will allow people to anonymously participate in discussions with a committee on environmental topics or with the society at large. Real-time video scanning of real world objects will accelerate the creation of immediately available new 3D content. These enhancements will accelerate the re-introduction of rich context that will serve as a source of inspiration to talk about. Likewise, public or semi-public "Metaverse Booths" that can act as 2D or 3D looking glasses between the real and the virtual world can take up the role of the familiar but disappearing public phone booths again. People in the street will be able to seamlessly interact with creatures appearing in a virtual world by approaching these billboard-alike terminals.

Combining the rich communication terminal at the home with the Metaverse Booth in public or semi-public places, people will be able to participate in immersive city games as real or virtual players, enjoying seamless cross-over of real-world and virtual world interaction in solving quests simultaneously. Augmented reality will allow adding artifacts to the real world for virtual world spectators and visa versa. Such artifacts can be anything from flying textual information, virtual 3D models or avatars to filtered real-time video streams. In a city context, the mobile phone can play an essential role in multimedia experience. Nevertheless, the use of a mobile phone will initially be limit to its use as a "remote control and browsing" device.

An example of such city game would be the quest of finding the in 1934 stolen panel "The Just Judges", part of the "The Lamb of God" multi-panel painting (see the illustration below). People in the city and people at home are encouraged to collaborate in finding the stolen panel. They can become private investigator, art collector, museum broker, thief or any other figure that might be related to the theft of the panel or in finding it. For this, they will have to tackle challenges in the real as well as in the virtual world. When a player in the real world found a clue, he can use a Metaverse Booth to show and pass it on to the virtual participants. While communicating, the video image of the real world may show some additional artifacts that may help in solving the quest.

The Ghent Altarpiece: The Adoration of the Mystic Lamb
(http://en.wikipedia.org/wiki/Ghent_Altarpiece)

The Ghent Altarpiece or Adoration of the Mystic Lamb (Dutch: Het Lam Gods or The Lamb of God; completed 1432) is a very large and complex Early Netherlandish polyptych panel painting which used to be in the Joost Vijdt chapel at Saint Bavo Cathedral, Ghent, Belgium, but was later moved for security reasons to another part of the cathedral. Commissioned by the wealthy merchant and financier Joost Vijdt, it was painted by Hubert van Eyck, who died in 1426 whilst work was underway, and his younger brother Jan van Eyck. The lower panel at the far left, The Just Judges, was stolen in 1934. Although several people have claimed to know its whereabouts, it has never been recovered and is now believed to be destroyed. It was replaced by a copy of 1945 by Jef Vanderveken.

5 Serious Gaming for Ambient Assisted Living

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

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

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

It is a necessity to provide the tools to guide a person through a behavioral goal to change his/her behavior by offering a sustainable positive and reinforcing experience.

This can best be accomplished through the implementation of a virtual agent that takes a leadership role in such intended behavioral changes. In this project, elderly people will be provided with a solution through serious gaming that will impact their habits and successfully transform them into new, desired behavior patterns.

6 Serious Games for Disaster management

Serious game are simulations which have the look-and-feel of a game, but these games are played with a serious purpose, such as training staff members for business operations,

military operations or rescue operations. For example, in New York serious games are used to quickly train new firemen as the understaffed New York fire brigade is in need of new firemen because of the World Trade Centre catastrophe.
Serious games are intended to provide an engaging, self-reinforcing context in which players motivate and educate themselves. The serious game paradigm is born on the inability of firms to provide educational software with an effective return on investment due to the cost of dedicated software and hardware. Traditional simulators cost millions of euros not only to develop, but also to deploy, due to the specialized hardware involved. Instead of dedicated media or computers for high-end simulators, serious games only require a traditional computer to play these games. Multiple deployments can easily be realized when the serious game is made available through a website. Finally, video and computer game developers are accustomed to develop games quickly and are used to create games that simulate functional entities such as persons, vehicles, building, etc... Using existing games library, developers can develop simulation games at a fraction of the cost and time of traditional simulator developers.

The purpose of the serious game use case is to go further than the current state of the art by leveraging on a virtual world. Doing so should bring a lot of advantages to the existing situation. In the current situation simulation games are not planned to be massively multi players, this is often only an add-in that enables the game to become multi-player but limits itself to only a few players. On the other hand virtual worlds are built on an infrastructure enabling a massive number of gamers, and all the rest of the game is built upon that. In the frame of the use case virtual worlds will be used to simulate equipment and people:

- Real equipment simulation, ie: simulation of equipment used in the real life
  - Rescue equipments (fire truck, , …)
  - Communication equipments (radio, mobile phone, …
  - Sensor equipments (RFID, GPS, …)
- Population simulation
  - People
  - Rescue team

By linking real world equipment and virtual word simulated ones, it would be possible to create a near real experimentation field to teach and observe a rescue team. This way it would also be possible to simulate most of the expensive equipment and only have real equipment when needed. For instance we may link real video from a survey camera into the virtual world and linking virtual world communication tools with the real world, this way it would be possible to use real equipment such as fire truck, PMR radio, mobile phone to communicate with other rescue team members or victims immersed into the virtual world. This will reduce the amount of persons needed to be involved and material into a real exercise the way this is currently done in big cities to train rescue team in case of massive terrorist attacks.

The use case for serious games relates to the exercise of crisis management after a nuclear accident through a 3D simulation in a virtual world. Various participants concerned (owner of the nuclear installation, public authorities) playing their part via an avatar in the virtual world and this through terminals with various capabilities. According to the capabilities of the terminals and roles' of the actors, 3D information transmitted will be suitably rich and realistic. Example:

1. In the Operational Headquarters (located near the place of the accident) and in the Central of Crisis (with the Prefecture or Ministry) the persons in charge for the crisis management has synthetic information.
2. On the place of the accident, the intervention services must have a 3D representation to control the deployment of technical means (gendarmerie, fireman, technical team of the installation, first-aid services).
3. The people in charge of an intervention (piloting a vehicle, a robot of intervention to take a sample or to put in safety the installation) must have a very realistic 3D representation with physical phenomena simulation (management of the collisions, realistic mannequin posture and aerosol dispersion). These people will have to manage material or people contaminated.

The simulation environment of the Metaverse will have to manage the access to documentation (technical documentation on the installation, emergency procedures (Plan of Prevention), reporting of the events) - (right of access).

- The virtual participants in the crisis management will have to interact with other avatars "population civil" which are not actors of the scenario. One will be able to thus observe their behaviors.
- Each participant will have to register himself and according to the role, it will have some rights and will have certain interaction possibilities with the virtual world.
- The interoperability of the various simulation modules and of the multi-model interaction modes (MMI) will be used to create virtual world standards.
- People who do not take part in the scenario, can "see" certain aspect of the interventions (displacement of the vehicles on the roads and in the streets of a city), but the situations on site will have "to be masked" with the visitor lambda (virtual world area protected).

The benefits of this serious game approach are as follows:

- The realization of crisis exercise through a network simulation (serious game) makes it possible to simulate accident situations difficult to realize in reality. With a realistic physical phenomena simulation, one confronts the actors with more realistic situations. One can thus much better sensitize the actors and better form them with the "good practices".
- One can initialize simulation with certain phases of the accidental situation and concentrate the exercise on certain aspects and thus shorten the duration of exercise.
- It is possible to tracking the actions of each participant to make a very detailed experience feedback.
- Simulation makes it possible to validate the prevention plans (scenario). It can also contribute to their design.
- The cost of the virtual exercises is much lower because they do not require an implementation of real technical means.

7 Shopping

Mary wake ups Saturday morning. She is very lazy but she has to buy bread and some cakes for a lunch with her family. Last week, she installed a special plugin for the Operative System of her computer. This plugin allows configuring a virtual character that represents the user in any virtual world that s/he desires. She configured her virtual character called Samantha selecting her favorite appearance from the multiple options and additional complements that allows the plugin.

This morning she has decided to do the shopping virtually and so, it is not necessary go out. When she does the shopping, a courier will bring it home in some minutes. She turns on her computer and enters in the FIVICOM virtual world transforming herself in Samantha. The most of complements that she selected last week for her virtual character are loaded correctly: skirt, a sweater, and big hoop earrings. The exact kind of Samantha's hair is not in the virtual world database, but it is replaced by other very similar.

In the virtual world, he goes to nearest "Bread & Cakes" shop centre. It is a cake shop chain that has shops in the main virtual worlds existing. Between the shop options of the chain, they provide a tool that allows configure the desired appearance of the virtual dependent and, this appearance will be the appearance of the dependent in all the shops that the chain has spread by the virtual worlds.

So, she enters in "Bread & Cakes" and her custom dependent attends her. The dependent is an autonomous character not controlled by any user. Both, it and Samantha will speak in natural language and the dependent automatically transmit the order to the "Bread & Cakes" orders database. Samantha asks for a baguette and some cakes. The dependent shows her the wide selection of cakes that have the shop and she inflates her chubby cheeks, trying to express indecision. It is one of the gestures that Mary could customize for Samantha in the Operative System plugin and that makes Samantha be different of the other avatars. The majority of virtual worlds are able to reproduce it. After some minutes of indecision, Samantha selects a dozen of cakes and asks the dependent for send them and the baguette to her habitual address. In some minutes, a courier brings the order to Mary's home and she can go back bed.

8 Go for a Walk

Sunday afternoon, it is raining again. Charles and his friends thought to meet in the –real-street for spending the afternoon, but due the weather, they decide meet at the main square of FIVICOM virtual world. Charles turns on his computer and enters the world with his usual virtual personality, Xenic. In general, it uses this personality and appearance in all virtual world that he enters. Xenic takes the main street in direction of the square. At the middle of way, an autonomous inhabitant of FIVICOM asks him about a pub. Xenic explains it how to go and continues his way. In the square, he meets with his friends at once. Their appearance is always the same when they meet. A friend has brought a present for Xenic, a new virtual watch. He accepts the present and put on it in his wrist. Automatically, this new complement is updated in the database of Charles avatar, and next time he enters in any virtual world, Xenic will wear the watch. They expend the rest of the afternoon going for a walk in FIVICOM virtual world.

9 Immersive active TV gaming / sporting

PHILIPS

Immersive active TV gaming/sporting (1)

- "take-over" from PC to TV and peripherals for rendering the experience
- physical workout by playing
- TV, amBX, LivingColor lamp body input sensors, etc. render experience

CONFIDENTIAL 1

PHILIPS

Immersive active 3D TV gaming/sporting (2)

- physical workout by playing
- Free movement, no wires
- Movement capturing via webcam or via remote control
- No extra game console needed
- IPTV Game (provided by internet service)
- No extra device needed
- No Game CD needed

CONFIDENTIAL March 14, 2008 19

10 Real emotions in virtual worlds

Online communities are more and more becoming important parts of many people's social lives. Although these communities are a great way to communicate with and meet people with different backgrounds from across the globe, the interaction between people that is offered is still quite limited by the computers in between the people. To add the spontaneity of real-life social interactions, it would be desirable to add non-verbal communication to the online worlds, for instance by giving avatars the facial expressions recorded from the 'real' user during online sessions.

The first thing Jane did when entering the online world Aurora Universe was defining her personal avatar. Most of her friends joined Second Life but Jane likes Aurora Universe because it has a very advanced system to sense Jane's emotions and express these through her avatar. Jane does not want to show her real identity and appearance but she wishes to share her true emotions and feelings with her friends online. The emotional avatar instantly shows her expression on-line, improving her online communication because Kathryn can now show and interpret the non-verbal communication as well. It feels like a real face-to-face chat.