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)

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.

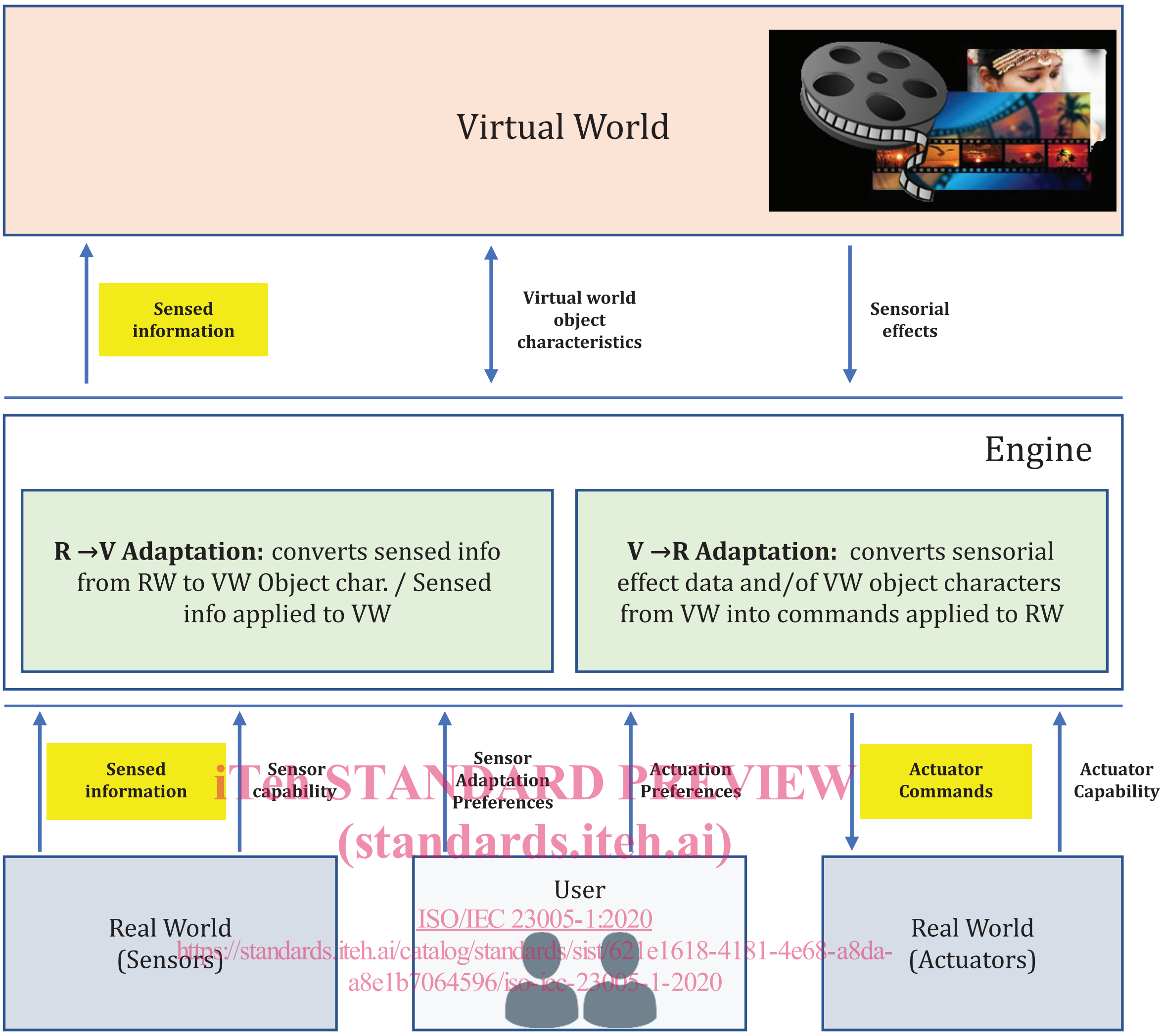

Figure 1 — Scope of the data formats for interaction devices

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.

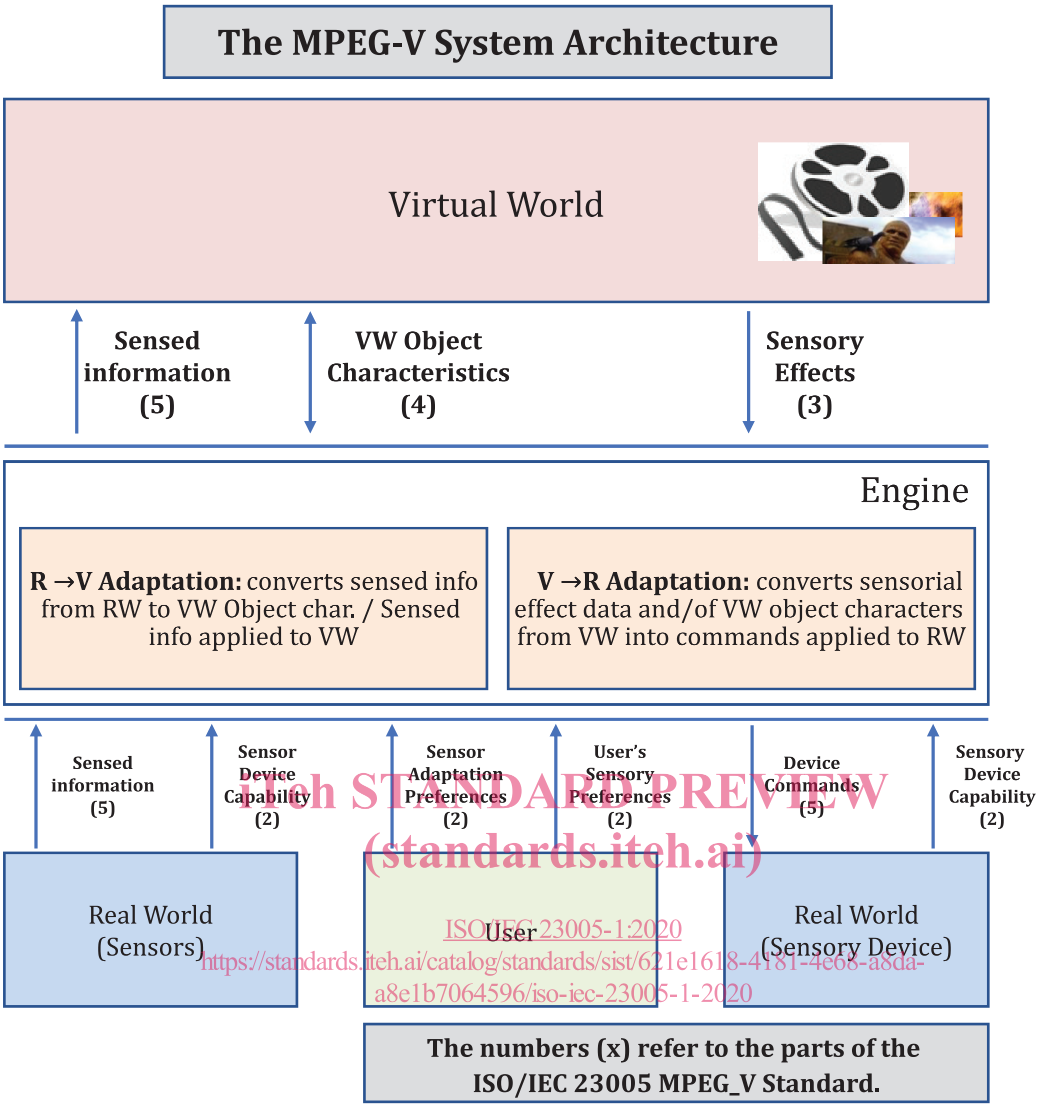

Figure 2 — System architecture of the MPEG-V framework

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.

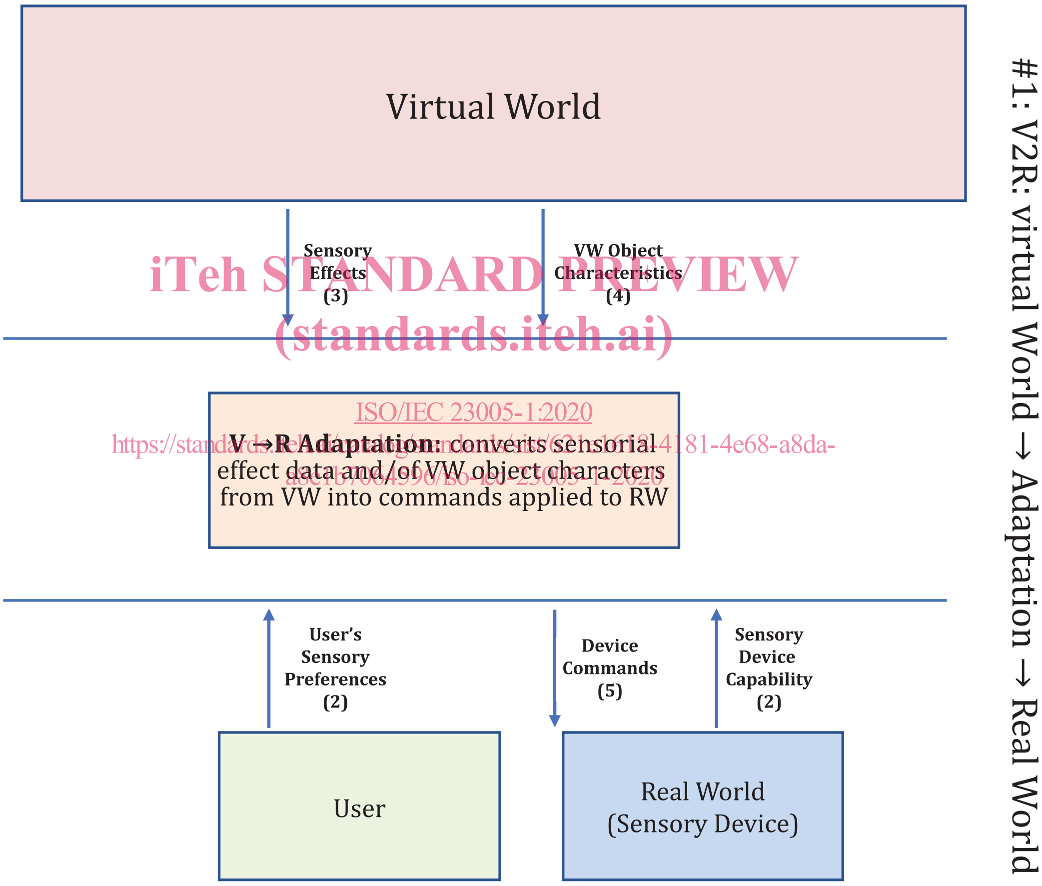

Figure 3 — Example of system architecture for information adaptation from virtual world to real world

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.

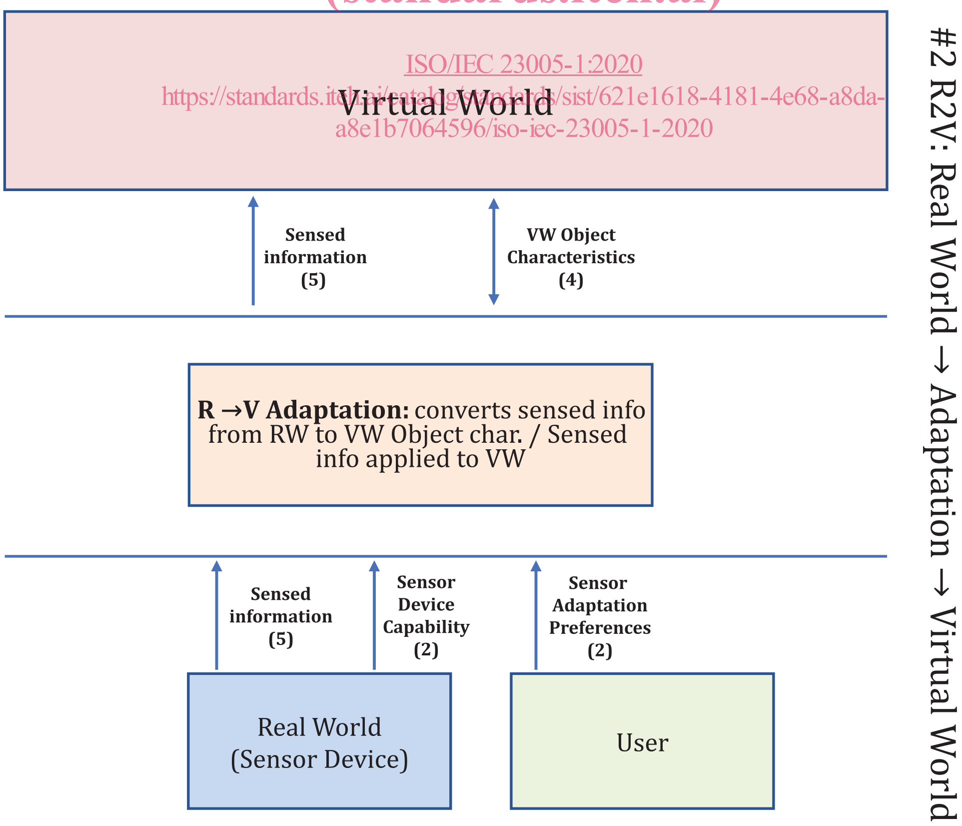

Figure 4 — Example of system architecture for information adaptation from real world to virtual world

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.

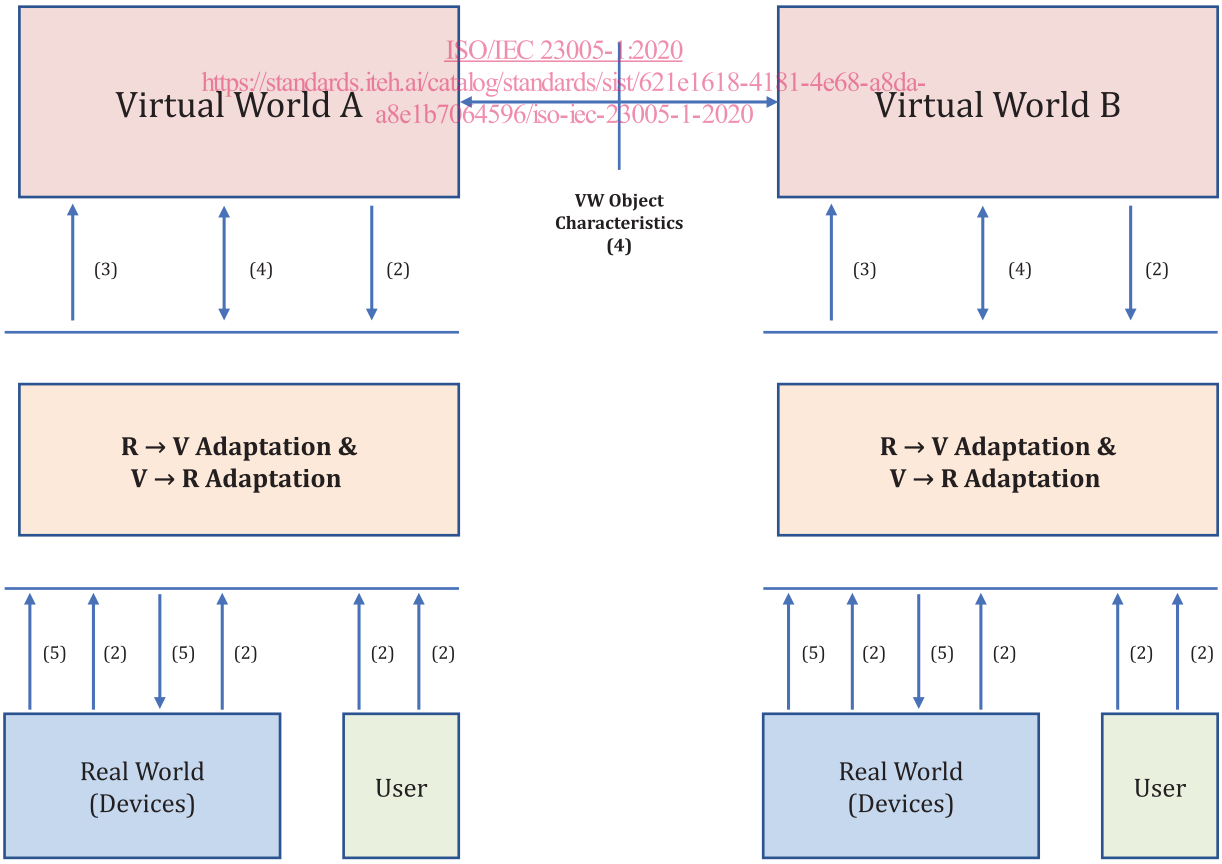

Figure 5 — Example of system architecture for (bidirectional) exchange of information between virtual worlds

