INTERNATIONAL ORGANISATION FOR STANDARDISATION

ORGANISATION INTERNATIONALE DE NORMALISATION

**ISO/IEC JTC 1/SC 29/WG 11**

**CODING OF MOVING PICTURES AND AUDIO**

**ISO/IEC JTC 1/SC 29/WG 11/ w15200**

**Feb 2015, Geneva, Switzerland**

|  |  |
| --- | --- |
| **Source:** | **Requirement** |
| **Title:** | **Exploration on Wearable MPEG** |
| **Author:** | **Mihai Mitrea (IMT), Sungmoon Chun (Insignal), Danilo Pau (STMicroelectronics), Miran Choi (ETRI)** |

1. **Introduction**

An important factor contributing to the growing adoption of IoE (Internet of Everything) is the emergence of wearable devices, a category with high growth potential. Wearable devices, as understood commonly, are devices that can be worn by or embedded in a person and have the capability to connect and communicate to the network either directly through embedded cellular connectivity or through another device (primarily a smartphone) using Wi-Fi, Bluetooth, or another technology. Examples are shown in figure 1.

Wearable devices, for example, allow the user to track his time, distance, pace and calories via a set of sensors in a T-shirt or on smart shoes. Another example can be smart glasses which combine innovative displays with some novel gestural movements for interaction. A classic example of implanted wearable device is a pacemaker or a heart rate monitor intelligent band aid.

Wearable devices are developed for and applied with sensor integration, to measure physical phenomena and e.g. to help people see better (whether in task-specific applications like camera-based welding helmets or for everyday use like computerized "digital eyeglass") or to help people to live and understand the world better or for health care monitoring systems.

Wearable device generated data are of great interests for the cloud and therefore it is expected that future personalized cloud services will elaborate them to produce advanced personalized feedbacks and therefore interoperability between wearable and those future services it’s anticipated as a need.

1. **Measuring the Trend of the Emerging Wearable**

This section focuses on the continued growth of the emerging trend of wearable devices and let the reader to understand their impact in the market. Wearable devices are making computing and connectivity very pervasive in our day-to-day lives.

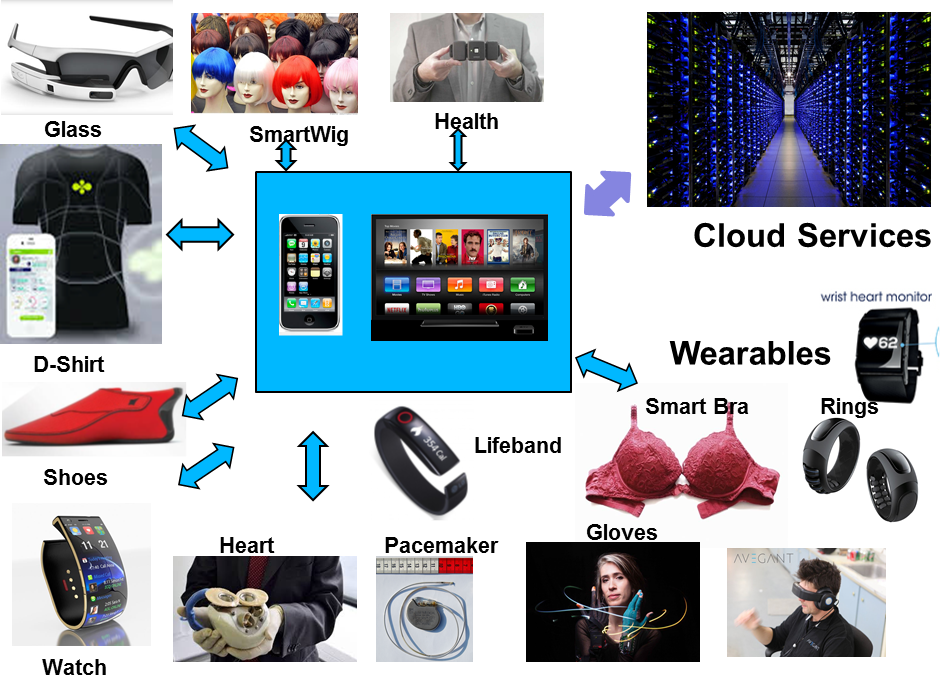


Figure 1 Typically wearable communicate to Cloud Services via Smartphones and Smart TVs

These devices come in various shapes and forms, ranging from smart watches, smart glasses, heads-up displays (HUDs), health and fitness trackers, health monitors, wearable scanners and navigation devices, smart clothing, et al. The growth of these devices has been fuelled by technology advances that have supported compression of computing and other electronics into devices are now both lightweight and autonomous enough to be worn.

These advances are being combined with fashion to match personal styles, especially in the consumer electronics segment, along with network improvements and the growth of applications, such as location-based services and augmented reality. Although there have been vast technological improvements to make wearable possible as a significant device category, wide-scale availability of embedded cellular connectivity still has some barriers to overcome for some applications—such as technology limitations, regulatory constraints, and health concerns.

Different market forecast studies issued recently by analyst show positive evolution of wearables market growth. These studies differ in their estimates figures, taking different users’ adoption scenarios and categories for wearables. A recent study from CCS Insight, gives the Global Wearables Forecast up to 2018 (Figure 2 – ref. <http://www.ccsinsight.com/press/company-news/2137-wearables-market-2015-is-year-that-will-make-or-break-the-smartwatch>).

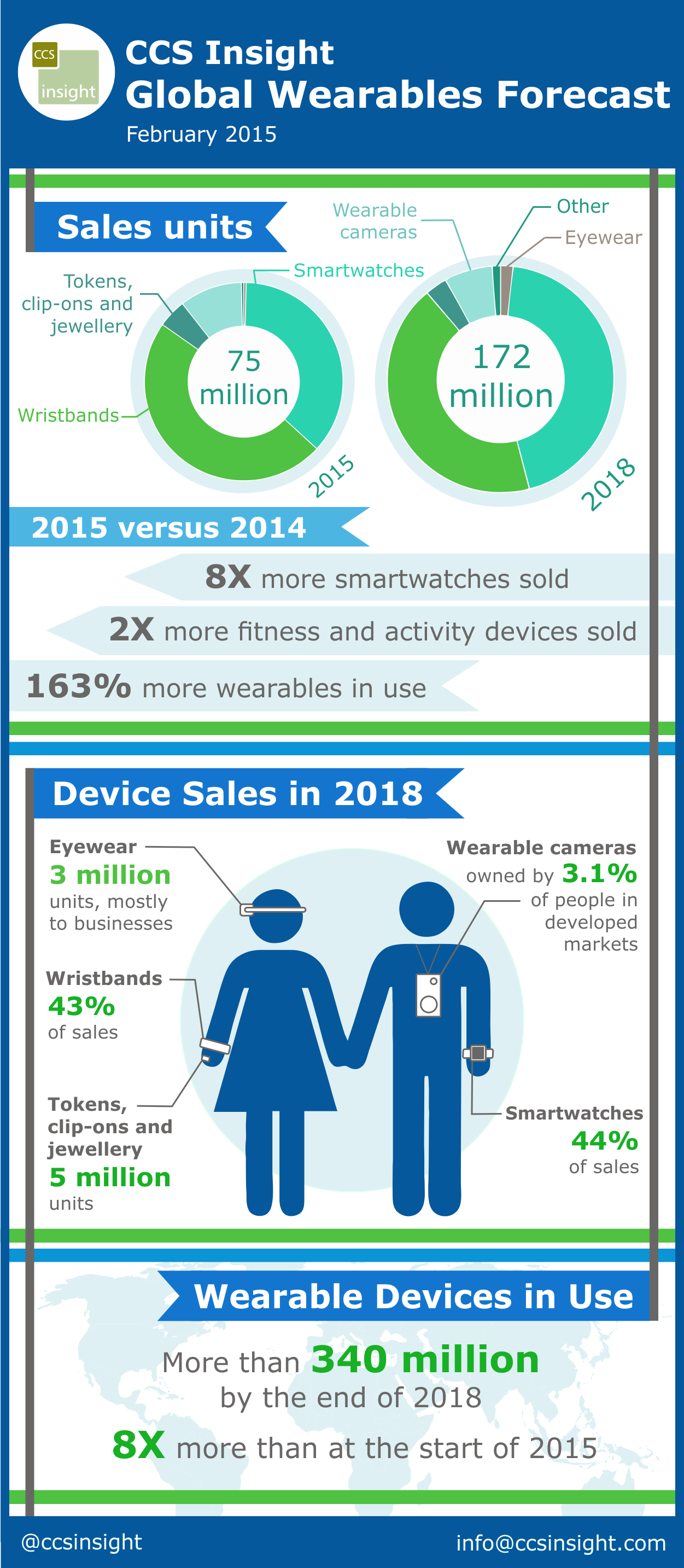


Figure 2 - Wearables Market Forecast; Source: CCS INSIGHT, Feb. 2015

By 2019, CISCO estimate that there will be 578 million wearable devices globally, growing fivefold from 109 million in 2014 at a CAGR (Compound Aggregated Growth Rate) of 40% (Figure 3).

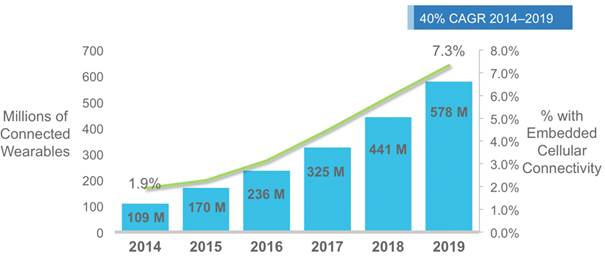


Figure 3 Global Connected Wearable Devices; Source: Cisco VNI Mobile, 2015

As mentioned earlier, there will be limited embedded cellular connectivity in wearables through the forecast period. Only 7% will have embedded cellular connectivity by 2019, up from 2% in 2014.

Regionally, North America will have the largest regional share of wearables, with 33% share by 2019. Other regions with significant share include Asia Pacific with 34% share in 2014, declining slightly to 32% by 2019.

The wearables category will have a tangible impact on mobile traffic, because even without embedded cellular connectivity they can connect to mobile networks through smartphones. Globally, traffic from wearables will account for 1.4% of smartphone traffic by 2019 (Figure 4).

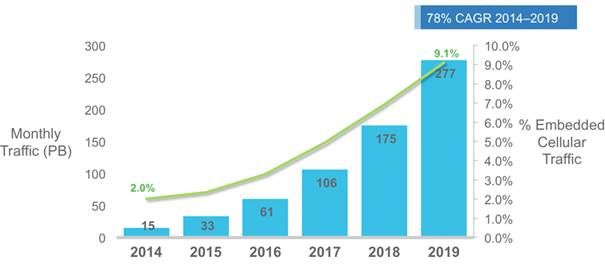


Figure 4 Global Wearable Devices Traffic Impact; Source: Cisco VNI Mobile, 2015

Globally, traffic from wearable devices will grow 18-fold from 2014 to 277 petabytes per month by 2019 (CAGR 78 percent). Globally, traffic from wearable devices will account for 1.1% of total mobile data traffic by 2019, compared to 0.6% at the end of 2014.

While the technology is in its infancy stage, analysts at Morgan Stanley believe it will become a $1.6 trillion business in less than 5 years. "*Wearable devices will far surpass market expectations, and become the fastest ramping consumer technology device to date, in our view*" accordingly to Morgan Stanley analysts. The analysts considers that wearable devices will have "far-reaching" impacts by creating a new category and disrupting or even accelerating change within industries outside technology. The analysts’ project sales of wearable devices will grow at a 154% annual compound rate through 2017, when 248 million devices will be sold. The figure will grow even further when sales of wearable technologies will reach one billion in 2020.

The analysts identify six sectors where wearable technologies could prove to be disruptive:

1. Traditional watches: wearable technology will change how consumers view traditional watches.
2. Apparel: wearable could accelerate an already strong health and wellness trend.
3. Payments: new Apple Pay payment system, if adopted to its upcoming watch, could make Pay even easier to use, generating intense competition for others to follow suit.
4. Chinese Retail: Alibaba is working with Intime to equip stores and malls with technology to improve customer experience and data analytics.
5. Health care: wearable devices could address health care system inefficiencies.
6. Industrials: wearables could be a catalyst offering users something unique as many companies are building smart home products.
7. **Terminology**

Within the wearable MPEG, the following terminology is defined.

**Wearer:** any living organism that is sensed by a Wearable.

**Wearable:** any thing that senses the Wearer; it may have control, communication, storage and actuation capabilities, and sense the Wearer environment.

**Mwearable:** a Wearable having at least one of media communication or storage capabilities.

**User:** any living organism, physical object or software interacting with and / or acted on by a Wearable; it may also communicate with the Processing Unit. In some applications, the Wearer is also the User.

**Processing Unit:** a unit or a set of units, some of which can reside in a local client and/or in a remote server that intelligently processes the information received from and provides the results to the Wearable and/or the Wearer.

1. **Scope**

As shown in figure 5, a Wearable may contain several types of sensors, integrated as a system on a module, which can be media (audio and visual) or non-media (heterogeneous sensors to enable body activity monitoring) oriented. The sensors sense physiological phenomena from the Wearer or from surrounding environment.

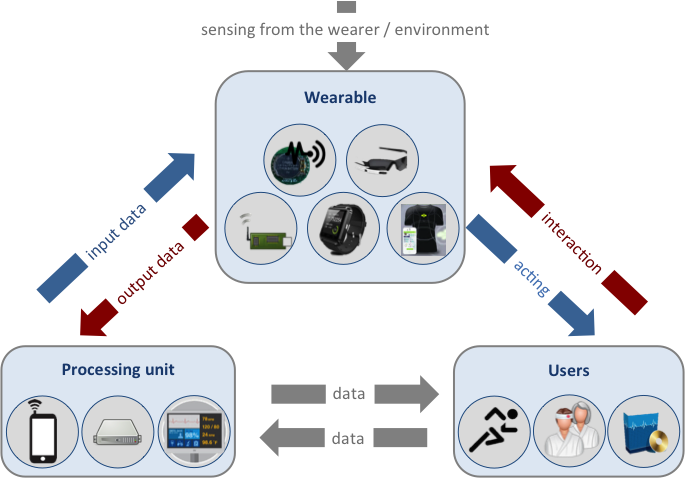


Figure 5 Conceptual model for WearableMPEG

Some of the sensors are already enumerated and specified in term of output bistream by ISO/IEC 23005-5 latest edition. The sensor of interests for the Wearable are listed in Appendix A. A note is included if a sensor is already specified or is missing considering the use cases submitted at MPEG 111th.

The Wearable processes this information and acts (e.g. display some visual content, play some audio content, send an alert, provide force feedbacks) on the User/Wearer. The User/Wearer can interact with the Wearable by a set of commands and may do so via the processing unit. The Wearable can also communicate an aggregated set of sensed information and receive data to/from the Processing unit.

**The scope of the WearableMPEG is to standardise**

1. **The interaction commands from User to Wearable**
2. **The format of the aggregated and synchronized data sent from the Wearable to the Processing unit (represented as red arrows in Figure 5)**
3. **A focused list of sensors that the Wearable may integrate.**

**The blue arrows in Figure 5 (acting from the Wearable to the User and the data input from the Processing Unit to the Wearable) correspond to MPEG technologies listed in Appendix B.**

**Grey arrows indicate entities that are out of scope of the standard**

1. **The sensing**
2. **The data exchanged between the User and the Processing Unit.**
3. **Use cases**
   1. **Multimedia communication of Wearable device (smart glasses)**

Monoscopic and Stereoscopic content can be created from a mono and stereoscopic camera on smart glasses. This content will contain mono-stereoscopic still images and/or motion pictures. The user will be able to Upload/Download this content via a wireless or wired network. Existing networks allow real-time transmission of content.

The Stereoscopic content player on smart glasses will be able to play transmitted real-time files or stored files. The player will be able to play monoscopic content as well.

The main function of smart glasses between smart phones **can be included** as video call, voice call and multimedia communication. This multimedia communication on smart glasses can be compatible with different OSs (Androidwear/Google, iOS/Apple,TizenOS/Samsung)

* 1. **Gesture recognition of Wearable device (smart glasses)**

Gestures can originate from any motion of the body or state but commonly originate from the face or hand. Current focuses in the field include emotion recognition from the face and hand gesture recognition. Many approaches have used cameras and computer vision algorithms to interpret sign language. However, the identification and recognition of posture, gait, proxemics, and human behaviours is also the subject of gesture recognition techniques. Gesture recognition can be seen as a way for computers to begin to understand human body language,

There are three essentially different input devices of gesture recognition: single camera, stereo cameras, and depth cameras. Depending on the type of the input data, the approach for interpreting a gesture could be done in different ways. However, most of the techniques rely on key pointers represented in a 3D coordinate system

* 1. **Speech translation of Wearable device (smart watch)**

Speech translation for people of different languages is a very convenient service in the multi-cultural, multi-lingual society and in a global environment. Starting from being mounted on PC through laptop and tablets to smart-phone, speech translation systems are getting more usable with wearable devices. When a user speaks to the microphone embedded in the smart watch in one language in a conversation with another user with different languages, it will be translated to the target language.

The result of the translation can be heard to the user of the target language through the smart watch. The translation engine is either in the remote server (remote translation system) or in the smart-phone (stand-alone translation system) which is connected to the smart watch. With the wearable translation service, the user is able to use his hands freely while the conversation is translated. The smart watch is also used for finding someone automatically who can speak one of the languages the embedded translation system handle in a travelling situation.

* 1. **Natural language communication of Wearable device (smart watch)**

Natural language communication is a very convenient interface between a human user and the computer. The user can control the machine with a speech command, find information, and ask questions to the intelligent agent inside the smart-phone or a server. With the wearable devices and the natural language communication, the user can have the freedom to participate in an activity without using a keyboard or any other terminal devices.

* 1. **Accessibility/protection functionality of Wearable device (smart watch)**

Wearable devices can provide accessibility functionalities for people with disabilities. Smart watch/glasses can guide people with low or no vision for the direction using location and visual information and speech interface. Smart devices can provide sound information to the people with hearing problems using vibration of the watch or some light signals from the smart-glasses when some emergency situation occurs. People of low intelligence or reading problems can also benefit from wearable devices through reading software embedded in or connected to the devices.

* 1. **Wearable device in cloud**

Nowadays, the use of wearable devices in cloud-based environments is restricted because of their hardware/software heterogeneity and of their very limited communication/computing/storage capabilities.

A potential solution to this issue is the virtualisation of the wearable devices as a cloud resource.

In order to obtain simple and effective systems, this virtualisation is expected to be unitary, independent with respect to the wearable device particularities.

Consequently, a standard able to turn a wearable device into a virtual resource ready to be integrated into cloud would be an appropriate solution. Such a standard is expected to deal with at least three complementary aspects:

* The cloud/device interfaces for handling the various types of content (be it multimedia or not) exchanged between the cloud and device;
* The interfaces for the operations which can be applied to the above mentioned content (e.g. hooks towards HEVC compression/APIs/etc);
* The interfaces required by the user in order to control the exchanged data and their operations.
  1. **Visual Communication via Wearable Device**

* + 1. **Concept of Visual Communication**

Text and voice are the main means of communication. But people use visual message such as rough map and emoticon to deliver intention which is hard to express using text and voice. Visual Message can used as follows.

* Case 1. Using visual message, without text and voice.
* Case 2. Using visual message assisted with text or voice.

* + 1. **Visual Communication for Wearable Devices**

Currently, there are a variety of wearable devices that are commercially available: Google Glass, Samsung Galaxy gear, Apple watch, LG Watch, Moto 360 and so on. Visual messages can improve the efficiency of the interaction between the user and the wearable device when the display resources are very restricted.

* + 1. **Benefit of Visual Communication**

In a visual communication everyone can understand intuitively a pictogram. So people can easily express an implicit meaning or an ambiguous emotion. When people can't express implicit meaning or ambiguous emotion, it's possible to visually communicate, e.g.via simple emoticons.

* 1. **Digital shirt**

A digital t-shirt is also called d-shirt. It is made of intelligent fabrics integrating sensors in its design for monitoring and gathering data (audio/video + semantic description about them) from the person wearing it and the environment where the person is. d-shirts also send/apply some information/feedback to the wearer by vibration and enables haptic interaction [<http://www.cityzensciences.fr/en/>].



Figure 6 D-shirt, an intelligent digital t-shirt

* + 1. **Use case**

A person is making smart sports by wearing a smart digital shirt (d-shirt) for optimizing his activity. While running, the d-shirt is capable of recognizing the type of activity, his state, and provides a feedback on the persons skin. If the person is running in a new environment and needs guidance for the route, by following the feedback from his d-shirt he is able to make his course as planned, with a tempo well adapted for his age, weight and condition. While running he can also record the environment (audio/video).

By allowing the d-shirt to communicate with the processing unit (an intelligent and more powerful device), the user is capable to see his results at the end of his activity, and let the processing unit make a recommendation for the next activity.

* + 1. **Parameters from and to the wearable device**

**Input**

*Information read from the body*:

* temperature
* pulse
* movements
* position
* pressure
* etc…

*Information read from the environment*:

* video stream
* audio stream
* temperature
* etc…

*Information read from the user gesture*:

* single touch
* multi touch
* surface touch
* hands gesture
* arms gestures

**Output**

*Information sent to the body*:

* tempo
* alarm
* directions
* haptic interaction
* etc…

*Information sent to the processing unit*:

* audio
* video
* temperature
* pulse
  1. **Artificial heart**

Current day artificial harts are generally managed by some parameters set before the patient leaves the hospital. However, in order to increase the live autonomy of the patient, solutions for ensuring the possibility of bi-direction communication between the heart and the physician, in the sense of distant control and monitoring of its state and the state of its host (the human) are searched for.

* + 1. **Use case**

A patient with an artificial heart is having a breathing problems, and an ambulance car is taking him to the hospital. While the patient is transported, he is connected with a range of sensors: blood pressure, body temperature, breathing, etc. and monitored with a real-time camera. All these sensors (sources of precious information) are connected to a processing engine, which gathers all data and transmits them to the hospital where the patient is heading. At the same time the physician remotely accesses the sensors, reads the data and actually interact with them, especially with the patient artificial heart, by controlling the tempo, the level of compression etc.

* + 1. **Parameters from and to the wearable device**

**Input**

*Information read from the patient*:

* core temperature
* pulse
* breathing tempo
* facial expression

*Information read from the processing engine*:

* heart rate
* compression type

**Output**

*Information sent to the processing unit*:

* heart rate
* core temperature
* timing
  1. **In-body**

The possibility of deploying reliable services around in-body devices [<http://www.holstcentre.com/en/FutureVisions.aspx>] is currently under investigation. These devices are battery-free wireless intelligent sensor modules implanted in the body to enable continuous health monitoring in the future. Tiny sensor-devices gather physiological information and communicate this to the outside world if deviations in physiological properties are detected.

* + 1. **Use case**

John Smith, 55 years old, visits his doctor for implanting the latest health sensor, which is the size of a rice grain. He loves sports and he goes for jogging every day. The sensor is connected to a processing engine, which sends the relevant data of his heart rate and core temperature to his wrist device. His training schedule automatically adjusts to his physical performance, through an intelligence medical service.

Two years later, the sensor registers abnormal patterns in the heart rate, and via an internet connection it makes the diagnosis while he is calling his wife.

A visit to his doctor is scheduled automatically because of the serious heart disorder. The doctor refers him to hospital for a lifesaving operation. The intelligence medical service (cloud based) takes care of John health wherever/whenever he is.

* + 1. **Parameters from and to the wearable device**

**Input**

*Information read from the body*:

* heart rate
* core temperature

**Output**

*Information sent to the processing unit*:

* heart rate
* core temperature
* timing
* person performance
  1. **QA through wearable device**

Question answering (QA) is an advanced function to generate answers for the user’s question in a natural language. More systems in the future are expected to be equipped with QA functions for advanced user experience. The QA architecture in Figure 7 is the general one which takes the user question as input and produces an answer as output after following each step of the QA process. In this contribution, we introduce a use case of applying QA through wearable devices.

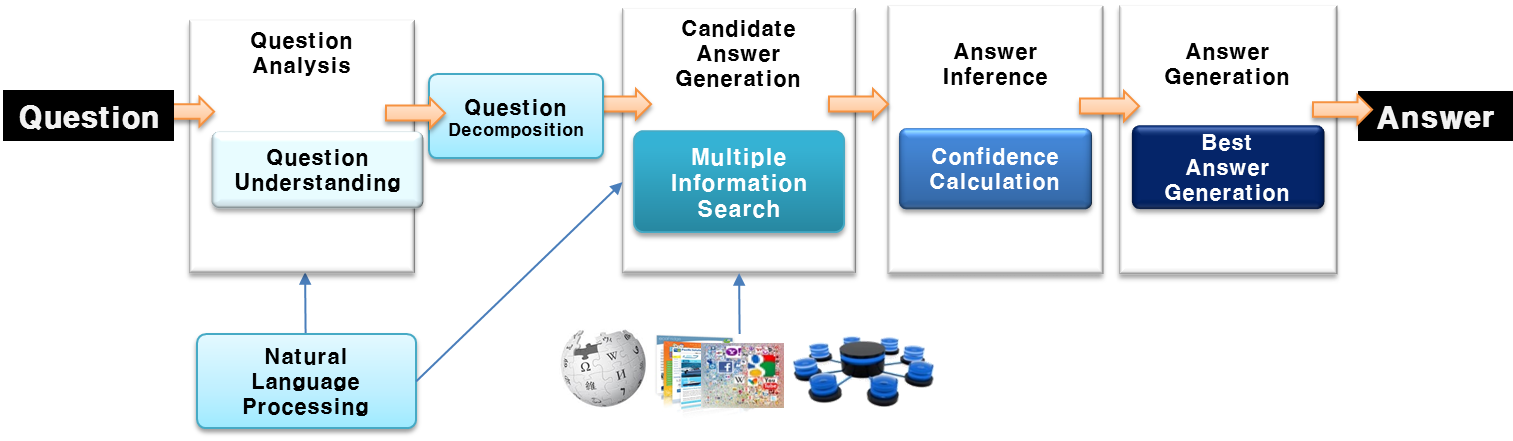


Figure 7 Example of Question Answering Service

* + 1. **Use case**

Michael visits Milano, Italy for his vacation. It is his first time in Italy and he does not have much knowledge on history or location of the various attractions. Using his wearable device, a smart headphone, he asks all the questions in natural language via speech interface and receives answers conveniently through the intelligent QA service. He can travel around in Milano easily without help of a tour guide

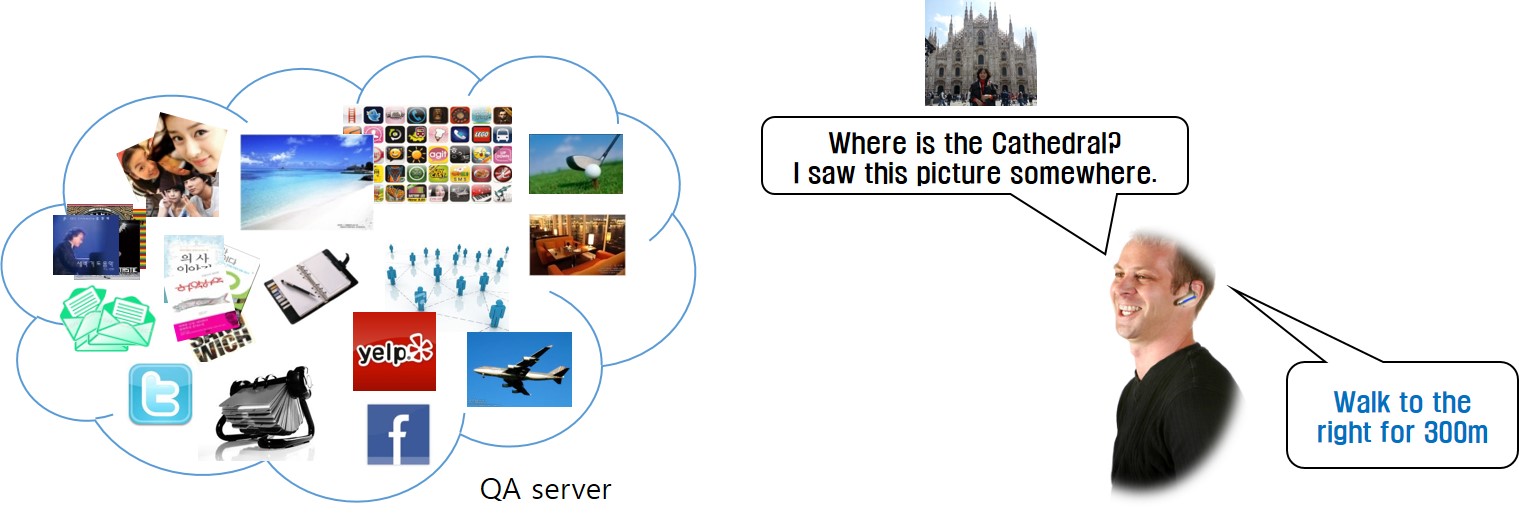


Figure 8 QA through Wearable device

* + 1. **Parameters from and to the wearable device**

**Input**

*Question in natural language*:

* Speech signal
* User information (UD)
* Multimedia contents

**Output**

*Answer from the QA server in natural language or pictures, maps*:

* Speech signal
* Multimedia contents
  1. **Conclusion**

By investigating the use cases listed above, the following formats of data / commands are identified for standardisation and are reported in Table 1.

Table 1

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Use case** | | **Exchange types of data (formats) / commands** | | | |
| **Wearable 🡪**  **User** | **User 🡪 Wearable** | **Wearable 🡪 ProcessingUnit** | **ProcessingUnit 🡪 Wearable** |
| **1** | Multimedia communication | * display mono video and still image * display various formats of stereoscopic video and still image   (side by side, top and bottom, frame sequential and so on) | * gesture command * voice command * touch command | * three(gesture or voice or touch) command * mixed(gesture, voice and touch) command   modified three command by wearable | * mono video and still image * various formats of stereoscopic video and still image * control data of multimedia contents |
| **2** | Gesture recognition | not defined | User:   * hand or body * skin keyboard * head motion * various touch   Wearable : Input devices of gesture command:   * one camera * two camera   9-axis sensor | * modified gesture command by wearable | control data for multimedia contents using gesture recognition engine |
| **3** | Speech translation | * audio * text | * audio * text * gender * location * speech style | * audio * text * gender * location * speech style | * audio * text |
| **4** | Natural language communication | * audio * text * interaction | * audio * video * text * interaction * location * gender | * audio * video * text * interaction * location * gender | * audio * video * text |
| **5** | Accessibility/protection functionality | * audio * video * text * interaction * vibration * light | * audio * video * text * interaction * location * accessibility features | * audio * video * text * interaction * location * accessibility features | * audio * video * text * interaction * signal for vibration and light |
| **6** | Wearable device in cloud | * render visual * notifications | * single touch * multi touch * voice control | * audio * video * interaction | * audio * video * graphics |
| **7** | Visual Communication | * visual message | * gesture command * voice command   touch command | * edited visual message   (if wearable can processing) or  edit information for visual message | * edited visual message   (if wearable can processing) or  edit information for visual message |
| **8** | Digital shirt | * audio * video * temperature * vibration | * single touch * multi touch * surface touch * hands gesture * arms gestures * voice control | * audio * video * temperature * pulse | * run * jump * stand * lay down * walk |
| **9** | Artificial heart | - - | - - | * heart rate * core temperature * timing | * heart rate * compression type |
| **10** | In-body | - - | - - | * heart rate * body temperature * blood pressure * heart disorder * blood pressure too high | - - |
| **11** | QA | * audio * text * interaction | * audio * video through second device * text * interaction * UD | * audio * video through second device * text * UD | * audio * video through second device * text |

1. **Relation between Wearable MPEG and MPEG V**

The relationship between wearable MPEG and MPEG V is represented in figure 9 by synoptically comparing side by side their conceptual models. For terminology about Wearable, please refer to chapter 3 of this documen and ISO/IEC 23005-5 latest edition for MPEG V.

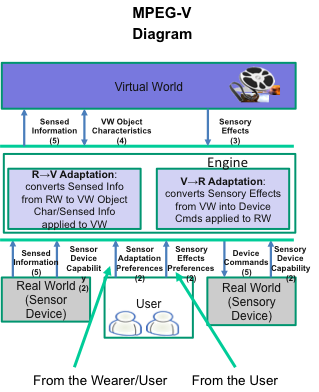
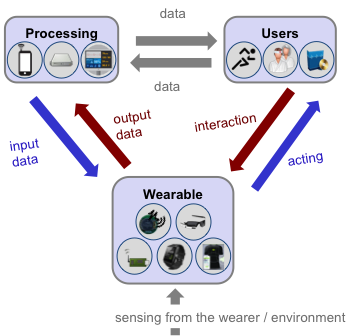


Figure 9 On the left the Wearable system and on the right the MPEG V

The similitudes and the differences are summarized in the following table 2:

Table 2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Function** | **MPEG V provides** | **Wearable Needs** | **To be standardized in Wearable** |
| **Sensors** to Engine (MPEGV)/ Processing Unit(Wearable) | Sensing | Sensor definition and capability | Aggregation of Sensors per device | Multiplexed and synchronized data stream from aggregated sensors |
| Engine(MPEGV)/ Processing Unit(Wearable) to **Sensory** | Acting | Device Commands | Commands and data | Only Vibration is included while other ones should be defined  Master clock to synchronize multiple wearable devices (D- shirt, shoes) |
| Engine(MPEGV)/Processing Unit(Wearable) to **Sensors** | Sensing | Single Sensor capability | Aggregation of Sensors per device capabilities’ query | Discover what type of sensors are available in the wearable as a set |

Investigation between Wearable and MPEG V about user interaction is necessary.

From figure 9 following types on MPEG V system diagrams (figure 10, 11, 12) capturing Wearable can be derived:

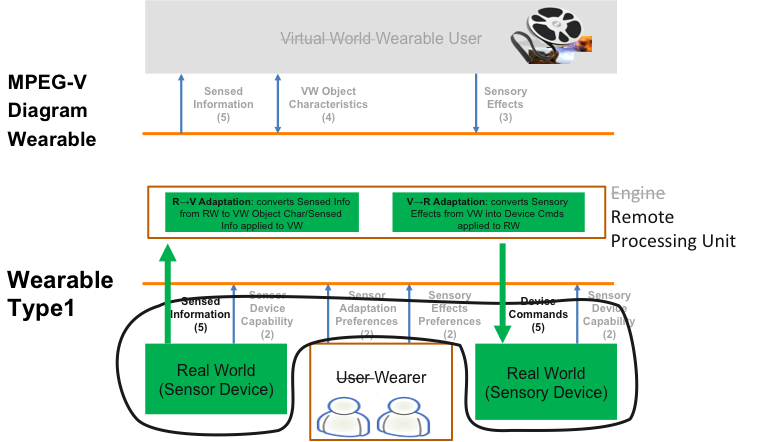


Figure 10 Wearable Type 1

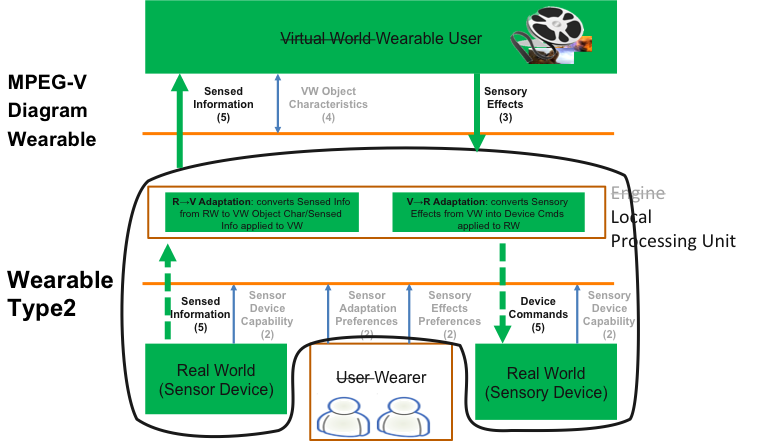


Figure 11 Wearable Type 2

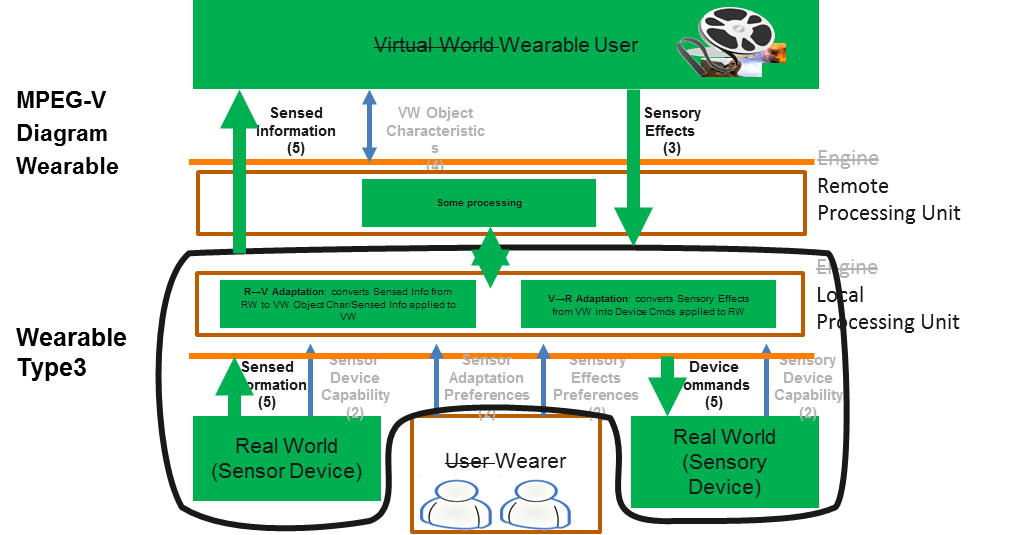


Figure 12 Wearable Type 3

1. **Relation between wearable MPEG and MIoT**

Within the IoT MPEG, the following terminology was defined at MPEG 111th.

**Thing**: any thing that can communicate with other Things; additionally, it may sense and/or act on Entities.

**Entity**: any physical or virtual object that is sensed by and/or acted on by Things.

**Mthing**: a Thing with at least one of audio/visual sensing and actuating capabilities.

1. **Conclusion**

The exploratory work carried out by the wearable MPEG and MPEG V during the MPEG 111th meeting brought to light the synergies and complementarities among wearable MPEG, MioT and MPEG V. This will be further jointly investigated and assessed from the standardization point of view during the next meeting (MPEG 112th).

**Appendix A**

The sensors of interests for the Wearable are listed here including a notice if they are already specified or missed considering the use cases submitted at MPEG 111th.

From ISO/IEC 23005-5 latest edition, these are the sensors of interest for Wearable as per a preliminary check. An in depth check has to be performed during the next Ad Hoc Period.

1. Light sensor
2. Ambient noise sensor
3. Temperature sensor
4. Humidity sensor
5. Distance sensor
6. Atmospheric pressure sensor
7. Position sensor
8. Velocity sensor
9. Acceleration sensor
10. Orientation sensor
11. Angular velocity sensor
12. Angular acceleration sensor
13. Force sensor
14. Torque sensor
15. Pressure sensor
16. Motion sensor
17. Intelligent camera
18. Multi Interaction point sensor
19. Gaze tracking sensor
20. Wind sensor
21. Global position sensor
22. Altitude sensor type
23. Bend sensor type
24. Gas sensor type
25. Dust sensor type
26. Body height sensor
27. Body weight sensor
28. Body temperature sensor
29. Body fat sensor
30. Blood type sensor
31. Blood pressure sensor type
32. Blood sugar sensor
33. Blood oxygen sensor
34. Heart rate sensor
35. Electrograph sensor
36. EEG sensor type
37. ECG sensor type
38. EMG sensor
39. EOG sensor
40. GSR sensor
41. Bio sensor
42. Weather sensor
43. Facial expression sensor
44. Facial morphology sensor
45. Facial expression characteristics
46. Geomagnetic sensor
47. Proximity sensor
48. Camera Sensor
49. Color Camera Sensor
50. Depth Camera Sensor
51. Stereo Camera Sensor
52. Infrared Camera Sensor
53. Thermographic camera sensor

**Appendix B**

The actions from the Wearable to the User and the data input to the Wearable from the Processing Unit) correspond to technologies already existing in MPEG listed here.

1. ISO Base Media File Format
2. Video : MPEG4 Simple profile based stereoscopic codec
3. Video: MPEG4 AVC based stereoscopic codec
4. HEVC
5. Still Image : JPEG based stereoscopic still image
6. Audio: MPEG-1 Layer-2, MPEG-2 AAC+, MPEG-4 ER-BSAC MPEG-4 HE-AAC
7. System: IOD/OD, BIFS, MPEG-4 File format, AVC File format, MMT.
8. User metadata: MPEG-7, TV-Anytime, etc
9. Stereoscopic VAF
10. MPEG-UD