Can MPEG cope with new media technologies?

Leonardo Chiariglione

CEDEO.net
Convenor, ISO/IEC JTC1/SC29/WG11

Abstract
This paper scans the almost 20 year of MPEG life to discover the movements of the industry, as represented by the standards produced, and analyses the current status.

1. INTRODUCTION

Formally the Moving Pictures Experts Group (MPEG) is Working Group 11 (WG 11) Coding of moving pictures and audio" of Subcommittee 29 (SC 29) "Coding of audio, picture, multimedia and hypermedia information" of the Joint Technical Committee 1 (JTC 1) "Information Technologies" of the International Organisation for Standardisation (ISO) and the International Electrotechnical Commission (IEC).

Working Groups are typically established by a Subcommittee to deal with some specific technical issue. Once the issue has been resolved working groups are typically disbanded.

This has not been the case of MPEG. Born in 1988 as an experts group of the then WG 8 of SC 2 MPEG was driven by a vision whose main points can be summarised as below:

- *Standards in a competitive environment are either rubber stamping or anticipatory.* Against all traditions of the CE industry MPEG has chosen to be the latter. Successful standards are the result of careful planning of the technology and market evolution.
- *Standards as enablers of mass phenomena.* Many people think of standards as the little thing that enables a business. It is more productive to see standards as the enablers of mass phenomena where people – millions of them – discover that they can do things they could not do before.
- *The converging force of digital technologies.* Analogue technologies were invariably the result of efforts designed to make audio and video systems with the constraints of or in symbiosis with physical entities – be they tapes or wires. Digital technologies can instead handle the audio and video content per se, leaving out the interfacing with the physical systems – the core of the "convergence" of digital technologies.
- *The value of a standard not just is in what the standard enables.* As so many technologies are required to make digital standards, the old model of a standard as the enabler of business for all concerned must leave room to those for which the value of the standard is in the IP of the standard.
- *Standards must provide the means to remunerate R&D investments.* A vibrant standards-based multimedia industry must provide the means to remunerate those who invest in technology that becomes part of a standard.
- *The unintended consequences of developing a standard.* When dealing with mass phenomena it is hard to forecast exactly how the standard will end up being utilised. This means that it is often dangerous to develop a standard for a specific use, as the use may not be what the market wanted.
- *Profiles as the consequence of the generic approach.* With digital technologies the size of a video frame can easily become a parametre. But to make a practically usable standard specific values are needed, i.e. profiles.
- *Expressing a digital standard in software.* When everything is expressed in bits it is more effective to represent the operations performed on the bits by means of computer code.
- *Conformance testing as the enabler of horizontal markets.* This is a traditional approach that MPEG has inherited and enhanced.
- *Assessing the value of what a standard offers.* A standard, particularly when there are many of them, is best demonstrated by its performance (what MPEG calls "verification tests").

2. THE BREAD AND BUTTER

2.1. MPEG-1

This standard was developed in the early days when it was not really clear what the impact of digital would be. Indeed MPEG-1 meant having the technology to handle audio and video bits for the first time with the magic word of "interactivity".

In those days anything that implied digital audio and video had to confront with issues of hardware – i.e. integrated circuit (IC) – implementation. So all decisions were weighed against possible impacts on IC design.

- Without calling them as such MPEG-1 Audio had Profiles – the 3 MPEG-1 Audio layers are an examples of 3 hierarchical profiles
- Without call them as such MPEG-1 Video had Levels – the Constrained Parameter Set of MPEG-1 Video defines a level for all practical purposes
- The reference software was already available, probably the first example of software attached to an ISO/IEC standard
- Conformance was already available, using the reference software to text video and audio bitstreams
- Verification tests were carried out for audio providing the expected results of transparency of stereo sound coded at 384 kbit/s (Layer I), 256 kbit/s (Layer II) and 192 kbit/s (Layer III).

The most serious problem faced by MPEG-1 was one of the actual use of digital technologies based on the standard. The Digital Compact Cassette (DCC), meant to be a user of MPEG-1 Audio Layer I was a flop, the Digital Audio Broadcasting (DAB) standard, meant to be a user of MPEG-1 Audio Layer II, is not really flying even now and Copact Disc Interactive (CD-i), meant to be a user of the complete set of MPEG-1 technologies, is long dead.
Still MPEG-1 is well alive and kicking: Video CD uses the complete set of MPEG-1 technologies and MP3 uses MPEG-1 Audio Layer III. Both need no introduction.

2.2. MPEG-2

Once it was proved that the machine could work it was possible to make true the long-held dream of television made digital. What MPEG added was the definition of a truly universal digital video format (audio, as it well known, is not part of the truly universal digital television format).
The development of MPEG-2 Video was obsessed by the issues of hardware implementation. In the early 1990s designing an MPEG-1 Video chip was becoming possible, but an MPEG-2 chip was another story. Hardware problems existed not only for the signal processing part but also for the Random Access Memory (RAM) that was required to store decoded pictures. A special profile (Simple) was designed to avoid use of too much RAM, although with some picture quality loss.
The issue of patents came to the fore with all its importance. The Consumer Electronics (CE) industry had its ways to handle the issue of patents in its standards: the IPR of a successful device belonged to the company that had originally conceived it, but with MPEG quite a few companies held IPR on MPEG-2 Video. The first patent pool after many years was established to handle the over 100 patents that the patent pool administrator has found to be relevant to implementing the MPEG-2 standard. The royalty model followed the established CE practice: a fixed amount per device and Digital Versatile Disc (DVD) disc carrying MPEG-2 Video encoded content.

2.3. MPEG-4

MPEG-4 was conceived as a unified multimedia coding standard. Indeed MPEG-4 Video has been designed to cover all bitrates and all video formats and MPEG-4 Audio is capable of handling music and speech.
MPEG-4, however, is also capable of handling synthetic content, such as 2D and 3D graphics, human faces and bodies, synthetic music etc. Moreover it has a powerful technology to handle the composition of natural and synthetic information.
MPEG-4 has been extremely prolific in Visual extensions and profiles. The basic Advanced Audio Coding (AAC) technology inherited and extended in MPEG-4 has undergone many extensions.
MPEG-4 did not develop a transport mechanism for MPEG content but only a generic interface to transport. It did develop, however, a Real Time Protocol (RPT) payload – and after untold vicissitudes got it adopted by the Internet Engineering Task Force (IETF) – for generic content.
MPEG-4 has been an innovator in several respects. The first is the fact that MPEG-4 has provided the first example to manage Intellectual Property Rights (IPR) in a complete fashion through its Intellectual Property Management and Protection (IPMP).
The second form of innovation was brought by MPEG-4 as the convergence point of the traditional hardware-based and the new software-based audio-visual industry. This prompted the transformation of the reference software from a nice-to-have add on to the formal status of standard on par with the textual form of the standard. The development process of the reference software was also codified – independently – along the lines of the Open Source Software tradition.
Another innovation was made by MPEG-4 in patent licensing. While MPEG-2 Video has a licensing model that is based on physical objects, be they set top boxes (STB) or DVD discs, MPEG-4 Visual also has a licensing form that charges the use of content per unit time.

3. NEW LAND

3.1. MPEG-7

The MPEG-1/-2/-4 standards have each played a role in shaping the digital audio and video industry as we know it today. MPEG-7 has the potential to reshape the industry in even more substantial ways.
The natural observation – a tautology today when we use a search engine to find information on the web – is that content is important but more so is the *description* of the content so that it can be found and organised.
The idea of MPEG-7 was triggered by the realisation that as much as a standard form of audio and video coding had created horizontal markets of content, a standard form of audio and video descriptions would create horizontal markets of descriptions.
The needed standard technology has all been developed. There are visual descriptions, audio descriptions and multimedia descriptions. It is also possible to define Multimedia Description Schemes (MDS) using the MPEG-7 Description Definition Language (DDL).
Today the MPEG-7 vision has still some way to go. The main reason is that digital technologies have removed value from "free" content. Premium content, for which value is still high, is offered as proprietary package using Digital Rights Management (DRM) technologies.
We are not exactly in a horizontal market and this will remain so for as long as there will be no interoperable DRM (iDRM) solution broadly adopted by the industry. iDRM is the only way for content to retain its value thereby providing an incentive for companies to add more features to their offers.

3.2. MPEG-21

As much as MPEG-7 is not an audio and video coding standard, MPEG-21 is also not a "signal processing" standard. Its need was triggered by the realisation that compression and description are two important components of "Multimedia" but to enable a complete multimedia market where content can be traded at the atomic level there are more disparate technologies to integrate.
To achieve this goal a number of new technologies are required:

- To bundle together in an organised way all the data that refer to one or more resources that is the object of trading. This bundle is called Digital Item by MPEG
- Once a Digital Item has been created there must be a means to identify it
- In the digital space a Digital Item is typically not just "given away" possibly against a transaction, but permissions to carry out only certain actions are granted to users. There is a need for a technology capable of describing such permissions
- The meaning of actions may well be understood and agreed by humans, but machines are a different story. There is a need to define the meaning of the verbs that are employed in the permissions
- In order to use the resources in the Digital Item it is necessary to describe the DRM Tools required.
- In order to perform certain actions on the resources in the Digital Item (e.g. subsampling, extraction of a layer in a scalable video stream etc.) it is required to describe the features of the Tools that needs to be used to operate on the resources in a Digital Item
- A Digital Item may need to be moved via a file or in a stream
- Etc.

4. NEW HORIZONS

4.1. MPEG-A

Some 19 years after MPEG began its existence the situation today is that there is a lot of multimedia technologies out there both from MPEG and from other sources, but making successful multimedia applications is as hard as ever – just too many choices.
On the other hand MPEG has an impressive repository of multimedia tools and has the know how to define the appropriate combination of technologies that enable an application.
This is the rationale for the MPEG-A (A as Application) standard for which the following has been done or is under way

- Music Player: a file format for MP3 music and MPEG-7 wrapped ID3 metadata. The unprotected version has been completed and the protected music version is under development
- Photo Player: a file format with JPEG pictures and MPEG-7 Visual metadata to be used for automatic organisation of a photo album
- Musical Slide Show Player: a file format for displaying a time-synchronised sequence of slides. It is a superset of the Music Player and the Photo Player
- Media Streaming Player: a set of formats (mostly drawn from MPEG-21) and protocols to govern streamed content
- Audio Archival: a format for storage of high resolution audio material.

4.2. MPEG-E

The MPEG Multimedia Middleware (M3W) is another non-coding and non-signal processing but very much multimedia oriented standard.
The standard will be eventually made up of the following parts

1. Architecture
2. Multimedia API
3. Component Model

4. Resource and Quality Management
5. Component Download
6. Fault Management
7. System Integrity Management

The M3W API is intended to be a generic API suitable for use in multiple products with different capabilities that have significant overlap in the functionality they require from the platform, but there are also differences.

5. THE SITUATION TODAY

While MPEG is running towards the end of its second decade it is possible to make a few assessments regarding its field of work.

- *There is a continuous need for more efficient audio and video compression.* Advanced Video Coding (AVC) has provided a factor-of-2 improvement in compression and, in spite of its increased complexity, several application fields are deploying the new standard. The same applies to the audio field where AAC has been repeatedly extended with High-Efficiency AAC (HE-AAC).
- *More applications of video and audio compression are popping up.* MPEG-Surround, a technology to add a very reduced-budget multichannel information to a mono or stereo coded signal is likely to provide innovation in several digital audio application fields. Scalable Video Coding (SVC) is at last coming of age with efficient support of scalability features. Multiview Video Coding (MVC) is the first step towards the creation and navigation of 3D spaces of video information
- *New flexibility in the definition of video and audio codecs.* So far each new audio and video codec has been a tightly designed bundle of video and audio coding tools. This may well continue to happen but it will soon be possible via the Reconfigurable Video Coding (RVC) standard under development to employ new video codecs that utilise the existing set of video coding tools and to add new tools to those existing that produce demonstrable improvement.
- *MPEG continues to be the place for multimedia technology integration.* The MPEG-4 standard continues to grow. It collects the most recent audio and video coding tools, but also a variety of other technologies that are needed to make a complete multimedia experience such as the last addition Open Font Format.
- *Gone are the days of integrated Systems/Video/Audio standards.* MPEG-1/-2/-4 have been packages of systems, video and audio coding technologies but the maturity and complexity of the industry does not justify this integrated approach any longer. MPEG-B, -C and –D have been introduced to cope with the fact that MPEG is now adding technologies as need arises, e.g.
  - A binary format for XML structures used by MPEG
  - Fragment Request Unit
  - Fixed point implementation of DCT/IDCT
  - Auxiliary Video Data Representation
  - MPEG Surround.
- *More Information Technology with Digital Signal Processing (DSP).* MPEG has started as a DSP standards group but its co-petition with IT has started from early on, namely from the moment it decided to use a pseudo C code to represent the working of its standards. The DSP roots of MPEG are not in question but the role of IT is on the increase.
- *More flexible licensing of MPEG technologies.* Significant innovations in the way IP is licensed has happened but after 20 years it is clear that some technologies can no longer be licensed. It is necessary to design a way to preserve the value of IPR in standards while recognising that a mature industry has competition from within itself.

6. CONCLUSIONS

In its life span close to 2 decades MPEG has played a key role in moving the incumbent (20 years ago) audio and video industry to the digital shore and creating the new converged digital audio and video industry that we see around us.

Many things have changed in these years and the industry is far from reaching a stable configuration. As the state of technology continues to progress the need for standards will continue.

MPEG will continue its role of anticipating the standardisation needs of the industry by continuing to stick to its original vision, relying on the considerable technology assets developed so far and designing new standards in a less structured world.