MPEG-4 – The Media Standard

The landscape of advanced multimedia coding

m4-out-20027-R3.pdf

This document presents an overview of the technology and business landscape, the applications, advantages and benefits of MPEG-4 - The Media Standard. Comments are welcome at: papers@m4if.org

Table of Contents

WHAT IS MPEG-4?

The media standard

MPEG-4 is an open standard, representing thousands of man-years of work shared by hundreds of companies. No one company can hope to match the technical and intellectual resources of an entire competitive market. No other technology has the potential to become as deeply developed and widely supported by multiple industries, vendors and service providers, and to be trusted by end users with their video and multimedia needs.

It is the only open standard that can address the opportunities enabled by the digital revolution: easily deploy multimedia content for any and all platforms.

MPEG-4 dramatically advances audio and video compression, enabling the distribution of content and services from low bandwidths to high-definition quality across broadcast, broadband, wireless and packaged media.

MPEG-4 is an open toolbox to build bitstreams and decoders for all multimedia content. MPEG-4 provides a standardized framework for many other forms of media — including text, pictures, animation, 2D and 3D objects – which can be presented in interactive and personalized media experiences. To support the diversity of the future content market MPEG-4 offers a variety of so-called “profiles,” tool sets from the toolbox, useful for specific applications, e.g. in audio-video coding, simple visual or advanced simple visual profile. Users need only implement the profiles that support the functionality required.

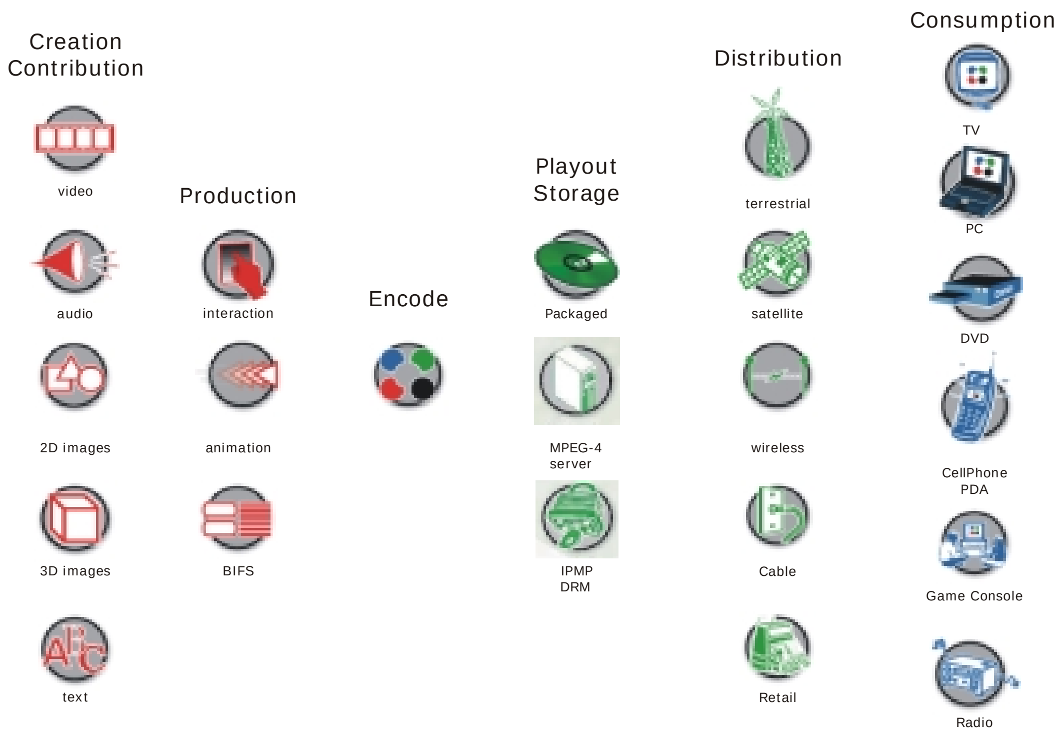

Figure 1: The MPEG-4 ecosystem liberates multimedia for delivery across any network to any user of any device

MPEG-4 is a transport-agnostic multimedia representation technology, suited for interactive multimedia delivery systems at any bandwidth. Software that implements the MPEG-4 standard is free for download, and IP owners are committed to license their patents on terms that are fair, reasonable and non-discriminatory. Comprehensive interoperability programs are up and running at ISMA (www.isma.tv[i]), M4IF (www.m4if.org[ii]) and other organizations.

MPEG-4 is developed by the Moving Picture Experts Group (MPEG), a workgroup of the International Organization for Standardization (ISO) (www.iso.org[iii]) and the International Electro-technical Committee (IEC) (www.iec.org[iv]) – the group that designed MPEG-1, which includes MP3 digital audio and MPEG-2 (the standard for the digital television, both DVB and DVD).

A revolution in functionality

MPEG-4 is a fundamental revolution in multimedia experience and functionality. At the same time, it is a progressive evolution, as MPEG-4 works in existing infrastructures, including MPEG-2 environments.

MPEG-4 goes beyond delivering traditional sequences of 2D video & audio to end-users as efficiently as possible. MPEG-4 also delivers interactive multimedia scenes with mixed media in 2D and 3D. In MPEG-4, while video and audio are coded with state-of-the-art compression techniques, graphics, text and synthetic objects have their own coding, rather than forcing them into pixels or waveforms. This makes representation more efficient and handling them much more flexible.

MPEG-4 offers revolutionary synthetic content tools, like structured audio (a language for describing sound generation), animated faces and bodies, 2D and 3D meshes, and vector-based graphics. This way, a minimum of generic information is transmitted, and rendering is done locally. This allows separate interaction with each object and sharp rendering at all bandwidths, e.g. subtitles that can be turned on and off and remain sharp even at high compression rates..

Interactive scenes, objects, elementary streams

Because MPEG-4 is object-based, it is possible to construct multimedia scenes which revolutionize the possibilities of interactive media. Authors can allow end-users to interact with objects in the scene: to change the color of a car to see how it will look, to tag a player on the field and watch all their moves, or to personalize enhancements to an enhanced video program. MPEG-4 also opens up new revenue opportunities through the integration of back-office systems including transaction/ e-commerce systems.

With MPEG-4, advanced interactive programming can be authored seamlessly integrating audio/video with 2D, 3D objects, animation and interactivity. For example, a viewer can navigate a sporting event's course from a 3D map, select information about aspects of the program, listen to commentary within a picture-in-picture window, and watch sponsored advertising – all within a single MPEG-4 stream supporting multiple media objects. MPEG-4 allows the same interactive programming to be used across different delivery channels. The same interactive program can be used on a DVD or delivered across a broadband network, something that was so far impossible.

An evolution in infrastructure

MPEG-4 is not an “in with the new - out with the old” revolution, since it is designed to smoothly integrate into existing, hard-working MPEG environments. MPEG-4 is in itself transport-agnostic. For some environments, transport mappings have been defined to allow smooth integration and interoperable solutions. The MPEG committee has specified how to carry interactive MPEG-4 content in an MPEG-2 Transport Stream, and has worked with IETF on a mapping of MPEG-4 onto RTP transport.

Saves money, makes money

MPEG-4 saves money for content businesses by eliminating the need for costly duplicate production in a number of different multimedia formats. MPEG-4 makes money for content businesses by offering multiple new broadband and narrowband platforms for the distribution of their content, such as wireless networks, digital television and the Internet. Together with content description (specified in MPEG-7) and Digital Rights Management tools from MPEG-4 and MPEG-21, MPEG-4 assists companies to find out where their content has been distributed while helping users to find content and to acquire it.

MPEG-4 is the standard that new entrants to the market are using to gain competitive advantage, e.g. by delivering content to multiple appliances across multiple networks. Multi-platform deployment of content is of key importance to building global Information and Entertainment brands. While it is not possible with MPEG-2 to create interactive content and then deploy it on both DVDs and interactive broadcast networks, this is easy with MPEG-4, and a feature which is attractive to major movie studios.

Risks of proprietary technology

Some vendors may claim their solution is cheaper since they are a “one-stop shop.” Using the open MPEG-4 standard avoids the dangerous hidden costs of proprietary technology, such as becoming hostage to third party business and pricing models; the level of risk in depending on proprietary and confidential third party technology road maps; conflicting agendas when the licensing entity is both a supplier and a competitor to the licensee; exploitative licensing terms, such as when the license includes many more bundled features than required; single sourcing problems with respect to pricing, competition, product-sourcing, new product versions, and bug fixes.

MPEG-4 is being widely adopted

MPEG-4 is being broadly and progressively adopted, across traditional industry barriers.

It is already the established standard for low bandwidth multimedia on 3G mobile terminals.

MPEG-4 is currently being discussed in some of the most active groups within DVB, the world's leading digital television standards drafting body.

Several streaming providers have adopted MPEG-4 including Apple, who adopted MPEG-4 Simple Visual Profile and Advanced Audio Coding for its QuickTime platform, RealNetworks, who supports decoding of MPEG-4 content, and the popular DivX codec is also MPEG-4 compliant.

In fact, most – if not all - of the major streaming players support, either natively or through plug-ins, the MPEG-4 standard in their currently deployed infrastructure and products.

RealNetworks supports MPEG-4 on its Helix servers, which makes it quite easy to utilize MPEG-4 today. Apple supports MPEG-4 natively in their Darwin Streaming Servers, which are available for free, and in its new QuickTime 6. At the time of writing, QuickTime 6 has been downloaded 25 million times since its introduction in summer 2002. Last but not least, a number of MPEG-4 vendors offer plug-ins for Microsoft's Windows Media Player that enable users to watch MPEG-4 content in this player.

In terms of hardware vendor support, Sony is shipping its newer PDAs and DVCams with MPEG-4. Several commercially available digital still cameras allow the capture of MPEG-4 onto flash memory media. Snell & Wilcox's Ingest Station features a browsing system based on MPEG-4. Both this browser and the company's SmartFile Transcoder were developed in conjunction with codec vendor Dicas.

Consumer electronic device manufacturers and semiconductor companies are building MPEG-4 devices for home entertainment systems including set-top boxes and DVD players and handheld devices with MPEG-4, as well as MPEG-1 and MPEG-2 playback capabilities. The DVD Forum is evaluating MPEG-4 for the next generation of storage media.

Finally, MPEG-4 is being deployed by many network and service operators around the world where new services are being added to take advantage of the broadband infrastructure build out.

A STANDARDS LANDSCAPE

This Section briefly describes the standards environment of MPEG-4: the other MPEG standards and some of the technology developed outside of MPEG.

ISO/IEC's Moving Pictures Experts Group (MPEG)

Since the end of the 1980s, the Emmy Award-winning MPEG committee has built the foundations of digital content delivery with its highly successful standards.

MPEG-1

Today, for the home video enthusiast, MPEG-1 video and audio (among which is MP3) are possibly the best formats to choose if you want to share your content with others, because they are universally understood. Combining a good compression ratio with high quality and universal deployment on PCs, these formats have been very popular for many years. MPEG-1 is the basis for the multimedia revolution on PCs. MPEG-1 was originally designed for CD-interactive and video-on-demand-like applications, and is still hugely popular in the form of Video CDs.

MPEG-2

MPEG-2 is the audiovisual standard most widely used for entertainment video applications. MPEG-2 enables digital television and DVDs, and there are several hundred million MPEG-2 decoders deployed in Digital Satellite and Cable set-top boxes, DVD players and PCs. It is a more powerful format than MPEG-1, capable of achieving higher compression ratios and supporting interlaced video. MPEG–2 decoding and encoding are more CPU-intensive than for MPEG-1, especially for video.

Virtually every image you see on television today, even on an analog receiver, has at some point been coded and decoded in MPEG-2.

MPEG-2 delivery rates have come down dramatically since the early adopters first went on air, without any changes to the standard. When the standard was just released, it took over 6Mbps to get a picture quality acceptable for broadcast. Today, many service providers successfully deliver quality images with bitrates between 2-2.5 Mbps – using the same, unchanged decoders. MPEG-4 has started its ascent on the same quality/bitrate curve, driven by a force only present in standards that are open for anyone to implement and improve: free competition on a level playing field. And note that the mechanism works for audio just as it does for video.

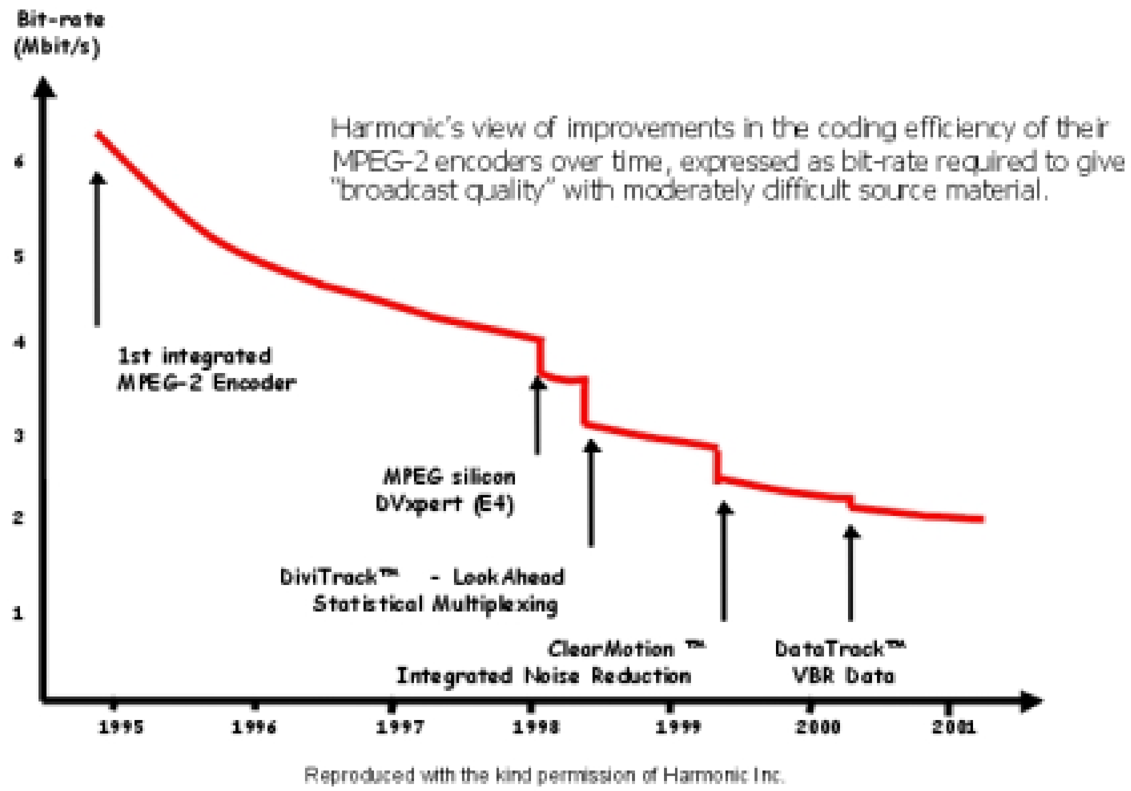

Figure 2: Improvements in Coding efficiency are facilitated by MPEG-based Competition [source: Harmonic[v]]

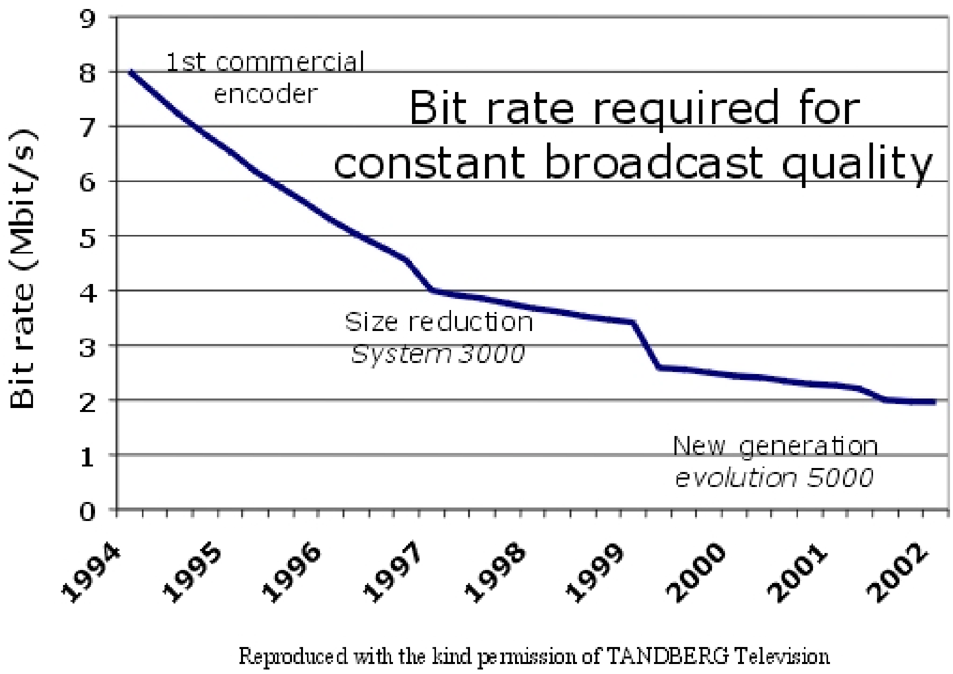

Figure 3: Improvements in coding efficiency are facilitated by MPEG-based Competition [source: TandbergTV[vi]]

Both MPEG-7 and MPEG-21 are additional toolsets which extend the functionality of MPEG and interface tightly with MPEG-4 to create new content management features. MPEG has taken care that MPEG-4 integrates well with MPEG-7 and MPEG-21. MPEG-7 descriptions and metadata can be carried as MPEG-4 streams, and MPEG-21's specifications are being written to complement MPEG-4's content representation.

MPEG-7

MPEG-7 is a recently finalized standard for description of multimedia content. It will be used for indexing, cataloging, advanced search tools, program selection, smart reasoning about content and more. The standard comprises syntax and semantics of multimedia descriptors and descriptor schemes. MPEG-7 is an important standard because it allows the management, search and retrieval of ever-growing amounts of content locally stored, on-line and in broadcasts.

For example, a tune can be whistled into a wireless device, the tune is converted locally to an MPEG-7 description, the description is transmitted to a search engine which returns the searched piece. Another example is facilitating complex editing tasks based on rich, hierarchical descriptions of the raw footage. A third example is broadcasting MPEG-7 Descriptions along with TV content, allowing TVs and Personal Video Recorders (PVRs) to autonomously choose programs based on user preference. MPEG-4 has a built-in MPEG-7 data type, allowing the close integration of MPEG-7 descriptions and MPEG-4 content.

MPEG-21

MPEG-21 is an emerging standard with the goal of describing a “big picture” of how different elements to build an infrastructure for the delivery and consumption of multimedia content – existing or under development – work together. The MPEG-21 world consists of Users that interact with Digital Items. A Digital Item can be anything from an elemental piece of content (a single picture, a sound track) to a complete collection of audiovisual works. A User can be anyone who deals with a Digital Item, from producers to vendors to end-users.

Interestingly, all Users are “equal” in MPEG-21, in the sense that they all have their rights and interests in Digital Items, and they all need to be able to express those. For example: usage information is valuable content in itself; an end-user will want control over its utilization. A driving force behind MPEG-21 is the notion that the digital revolution gives every consumer the chance to play new roles in the multimedia food chain. While MPEG-21 has lofty goals, it has very practical implementations.

MPEG-21 includes a universal declaration of multimedia content, a language facilitating the dynamic adaptation of content to delivery network and consumption devices, and various tools for making Digital Rights Management more interoperable.

MPEG-21 is about managing content and access to content. Even if you have fully interoperable coding there are still things you have to do to guarantee that all the features of various different networks work. MPEG-21 is a framework which allows interoperability and portability of content.

Internet Engineering Task Force (IETF)

The Internet Engineering Task Force (IETF)[vii] is a large, open international community of network designers, operators, vendors, and researchers concerned with the evolution of the Internet architecture and the smooth operation of the Internet. The IETF addresses transport/session protocols for streaming media. Their work relevant to MPEG-4 includes audio-video elementary stream payloads, generic MPEG-4 payload formats, the Real Time Protocol (RTP) - a transport protocol for real-time applications, an RTP profile for audio and video conferences with minimal control, and a Real Time Streaming Protocol (RTSP).

3rd Generation Partnership Project (3GPP and 3GPP2)

The 3rd Generation Partnership Project (3GPP)[viii] defines standards for 3rd generation mobile networks and services starting from GSM-based systems. 3GPP2 does the same for CDMA-based systems. Both pay much attention to mobile multimedia.

In their Wireless terminal specification, 3GPP and 3GPP2 use MPEG-4 simple visual profile for video, MPEG-4 file format for multimedia messaging and RTP, RTSP for streaming protocols and control.

Internet Streaming Media Alliance (ISMA)

The Internet Streaming Media Alliance creates a set of vertical specifications for Internet Streaming. ISMA has chosen specific MPEG-4 Audio and Visual Profiles and Levels (see "What is MPEG-4"), and augmented these with IETF transport specifications to create cross-vendor interoperability for video on the Internet. The first ISMA-compliant implementations have started to emerge. (www.isma.tv).

ISMA recommends streaming protocols for internet, including MPEG-4 simple visual and advanced simple profiles for video, MPEG-4 High Quality Audio Profile for audio, MPEG-4 file format for file storage, and RTP, RTSP for streaming protocols and control.

MPEG-4 TECHNICAL OVERVIEW

This section presents an overview of the technical advantages of MPEG-4 and some of the latest information on extensions and the MPEG-4 upgrade path. For a much more detailed technical overview of MPEG-4 the official "MPEG-4 Overview" is recommended. Please visit http://www.m4if.org/resources for access to this document.

MPEG Advances

The best way to understand MPEG-4's new paradigm is by comparing it to MPEG-2.

In the MPEG-2 world, content is created from various resources such as moving video, graphics, text. After it is "composited" into a plane of pixels, these are encoded as if they all were moving video pixels. At the consumer side, decoding is a straightforward operation.

MPEG-2 is a static presentation engine: if one broadcaster is retransmitting another broadcaster's coverage of an event, the latter's logo cannot be removed, also for example, viewers may occasionally see the word "live" on the screen when a broadcaster is showing third party live footage from earlier in the day. You can add graphic and textual elements to the final presentation, but you cannot delete them.

The MPEG-4 paradigm turns this upside down. It is dynamic, where MPEG-2 is static. Different objects can be encoded and transmitted separately to the decoder in their own elementary streams. The composition only takes places *after decoding* instead of *before encoding.* This actually applies for visual objects and audio alike, although the concept is a little easier to explain for visual elements. In order to be able to do the composition, MPEG-4 includes a special scene description language, called BiFS, for *Binary Format for Scenes.*

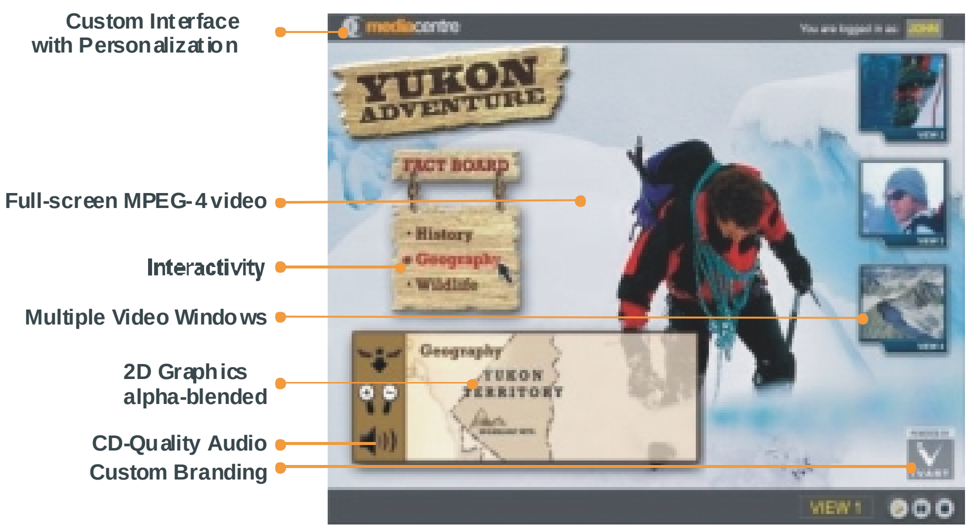

Figure 4: An MPEG-4 scene using multiple objects and elementary streams (Source: iVAST[ix])

The BiFS language not only describes where and when the objects appear in the scene, it can also describe *behavior* (make an object spin or make two videos do a cross-fade) and even *conditional behavior* – objects do things in response to an event, usually user input. This brings the interactivity of MPEG-4. All the objects can be encoded with their own optimal coding scheme – video is coded as video, text as text, graphics as graphics – instead of treating all the pixels as moving video, which they often really aren't.

As all the coders in MPEG-4 are optimized for the appropriate data types, MPEG-4 includes efficient coders for audio, speech, video and even synthetic content such as animated faces and bodies.

What are profiles and levels?

MPEG-4 consists of a large number of tools, not all of which are useful in any given application. In order to allow different market segments to select subsets of tools, MPEG-4 contains profiles, which are simply groups of tools. For example, the MPEG-4 Advanced Simple visual profile contains ¼ pel motion compensation, B-frames, and global motion vectors, but it does not contain shape coded video.

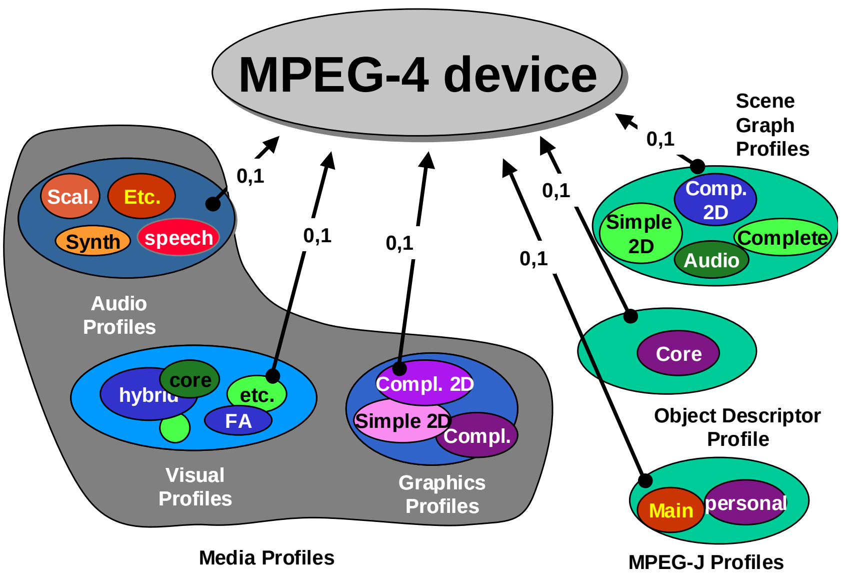

Figure 5: Profiles and Levels

Profiles allow users to choose from a variety of toolsets supporting just the functionality they need. Profiles exist at a number of levels, which provide a way to limit computational complexity, e.g. by specifying the bitrate, the maximum number of objects in the scene, audio decoding "complexity units," etc.

The concept of MPEG-2 Video Profiles has been extended to include the Visual, Audio and Systems parts of the standard, so that all the tools can be appropriately "subsetted" for a given application domain.

What are the parts of the MPEG-4 standard?

MPEG-4 consists of closely interrelated but distinct individual Parts, that can be individually implemented (e.g., MPEG-4 Audio can stand alone) or combined with other parts.

The basis is formed by Systems (part 1), Visual (part 2) and Audio (part 3). DMIF (Delivery Multimedia Integration Framework, part 6) defines an interface between application and network/storage. Conformance (part 4) defines how to test an MPEG-4 implementation, and part 5 gives a significant body of Reference Software, that can be used to start implementing the standard, and that serves as an example of how to do things.

Part 7 of MPEG-4 defines an optimized video encoder (in addition to the Reference Software, which is a correct, but not necessarily optimal implementation of the standard)

More recent parts added into MPEG-4 are:

- Part 8: Transport is in principle not defined in the standard, but part 8 defines how to map MPEG-4 streams onto IP transport.
- Part 9: Reference Hardware Description", Phase 1 Hardware Accelerators, Phase 2 Optimized Reference Software integration through Virtual Socket
- Part 10: Advanced Video Coding (as discussed below)
- Part 11: Scene description (to be split off from part 1)
- Part 12: ISO Media File Format.
- Part 13 : IPMP Extensions.
- Part 14 : MP4 File Format (based on part 12).
- Part 15 : AVC File Format (also based on part 12).
- Part 16 : AFX (Animation Framework eXtensions) and MuW (Multi-user Worlds).

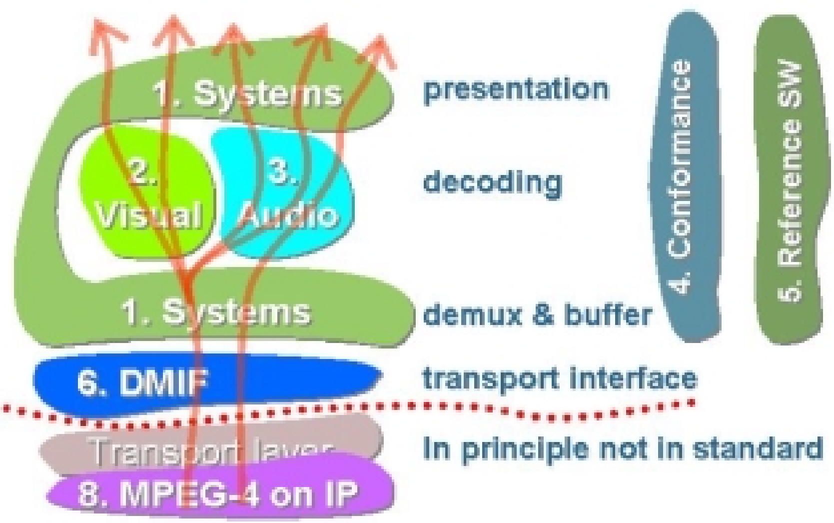

Figure 6 : The parts of MPEG-4. The arrows represent the flow of bits through the MPEG-4 system.

The Importance of Interoperability

Interoperability is the capability of products from different vendors to seamlessly work together. Interoperability is the goal of standards like MPEG-4.

Competition thrives and consumers benefit when multiple vendors' products interoperate. The MPEG committee's contribution to interoperability is to publish the specification and conformance points, such as profiles and levels. Other groups, notably the MPEG-4 Industry Forum and the Internet Streaming Media Alliance go farther. Both organizations sponsor interoperability programs and certification processes that allow vendors to test their products before they are marketed and consumers to know they are purchasing a product that will interoperate with other MPEG-4 products. Customers get choice and quality; vendors get secure customers and a larger market place.

Responsible Upgrades in MPEG-4

MPEG-4 is a dynamic standard. The first parts were published in early 1999, and work is ongoing. Changes to existing parts of the standard are always done in a backward-compatible way, as MPEG does not want to render already deployed systems non-conformant. This means that changes are usually done in the form of additions, sometimes as an amendment to an existing part of the standard, sometimes as a new part. There are basically two types of such additions: those that add functionality that wasn't there before and those that improve already existing functionality.

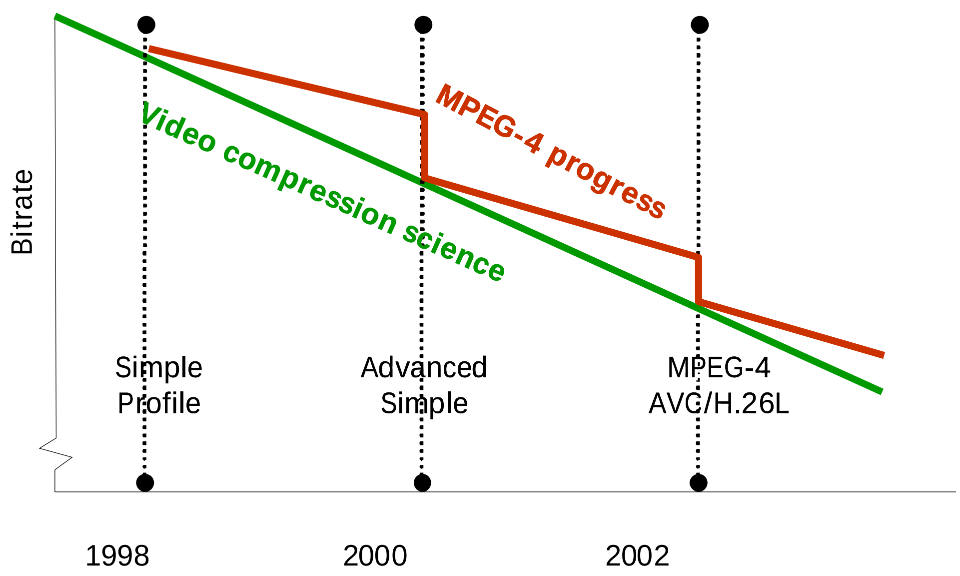

Figure 7: MPEG - A predictable, responsible upgrade strategy

Examples of the first kind are support for Multi-User Worlds (MUW) and the Animation Framework eXtension (AFX). A notable example of the latter is MPEG-4 Advanced Video Coding (AVC), or MPEG-4 part 10, also known as H.264, H.26L or JVT codec. MPEG-4 AVC is the codec that is developed in a collaboration with the Video Coding Experts Group (VCEG) of the ITU-T, the International Telecommunications Union[x]. The joint group that is responsible for this codec is called Joint Video Team (JVT)

MPEG-4 is a toolbox, as stated earlier. As the market moves, the requirements for the toolbox evolve as well, and the standard development work follows these requirements. Although not at the pace that some people believe, compression technology is still progressing, both for audio and video coding. In order for standards like MPEG-4 to be of most use to the market, interoperability and stability need to be combined with solid performance. MPEG standards, when they are issued, are always state-of-the-art, created through the collaboration of the world's best experts. Adversaries of open standards often describe the speed of standardization as "sluggish." This may be true for other standards, but not for MPEG. Aggressive schedules are always met, and the rigor and quality of so many experts collaborating far and away outweighs any theoretical speed disadvantage.

Unlike some of its proprietary competitors, MPEG exercises deliberate constraint in publishing new versions. Markets are not developed by releasing a new codec every 6 months; it takes time to build an interoperable eco-system that includes a rich variety of hardware and software tools. Case in point: MP3, or MPEG1/2 Layer III Audio. MP3 is not the most efficient audio compression codec anymore. But the advantages of using the interoperable format are so obvious that support for the format is still growing. Even today, MP3 is being integrated in car stereos, DVD players, CD players, etc. This is why MPEG employs a responsible upgrade strategy for its standards.

When applied to video coding, the principle of responsible upgrades means that there must be evidence for major increases in efficiency before undertaking the specification of a new video codec. Such evidence turned up last year, when MPEG did an evaluation of promising coding techniques, and subsequently started its work in the Joint Video Team, building on the work done in ITU-T's Video Coding Experts Group. The result, MPEG-4 Advanced Video Coding (to be also published as H.264), will fit right into the MPEG-4 architecture.

Extensions to MPEG-4 Advanced Audio Coding

MPEG-4 Advanced Audio Coding (AAC) is the most powerful audio codec licensed today. The near CD quality MPEG-1 Layer 2 audio codec (used in many digital video broadcasting apps) delivers high-quality stereo at 128kbit/s/channel while MPEG-4 AAC (Advanced Audio Coding) offers the same quality at 64kb/s/channel. MPEG-4 audio is capable of coding 5.1-channel surround sound very effectively and even allows transmission of wavefield audio data that will extend the listening sweet spot to the whole room in future sound systems.

The original MPEG-4 AAC has recently been extended with a technique called Spectral Bandwidth Replication, which gives spectacular bandwidth savings for apps like Internet Audio and digital broadcast. MPEG-4 AAC with SBR can deliver high quality stereo audio at a mere 48kbit/s. The SBR extension is both forward and backward compatible: an existing MPEG-4 AAC decoder can decode the extended signal (without the enhancement) and a decoder with SBR understand a signal that makes no use of the technique.

For digital audio broadcasting, MPEG-4 AAC is becoming the codec of choice. The satellite-based XM Radio uses AAC-SBR and Digital Radio Mondiale (DRM), that starts broadcasting in 2003, also uses AAC.

MPEG-4 Audio is inherently scalable. If, for example, a transmission uses an error-prone channel with limited bandwidth, an audio stream consisting of a small base layer and a larger extension layer provides a robust solution. Strong error protection on the base layer (adding only little overhead to the overall bitrate) makes sure there is always a signal, even with difficult reception. The extension layer (with little error protection) and base layer together give excellent quality in normal conditions. Any errors lead only to a subtle degradation of quality but never in a total interruption of the audio stream.

THE MPEG-4 MARKET

Business Opportunities

MPEG-4 is more than just another standard, rather it is the standard that creates business opportunities not only for established market participants, but also for newcomers to digital video and multimedia. MPEG-4 saves money and makes money.

Reducing Costs Through Multi-Vendor Support

MPEG-4 creates business opportunities in that it reduces costs to those deploying the media standard. Market participants in today's economic climate are much more receptive to introducing and deploying flexible, scalable and multi-purpose solutions and indeed much more focused on cost savings and reductions. This is where a unique key advantage of MPEG-4 comes into play. Something that none of the proprietary technology offerings will ever guarantee – a multi-vendor model that delivers competitive pricing guaranteeing a sensible economic solution.

Whilst being a "one-stop-shop" as a standard, there are many MPEG-4 vendors, specializing in the diverse opportunities and applications MPEG-4 is offering. In other words, the content/creator and distributor is in complete control as to which vendor for which element of the production and delivery chain is chosen.

Reducing Cost and Increasing Revenues Through MPEG-4 Object Orientation

As mentioned earlier, MPEG-4's object orientation gives rise to several advantages – many of which lead to reduced costs and new revenue opportunities. MPEG-4's advanced compression capabilities reduce the storage and bandwidth costs associated with digital media distribution. The result is greater capacity on a single CD or DVD, more digital television channels over the same bandwidth, and the ability to distribute rich media content over DSL, cable modems and 3G wireless networks.

MPEG-4's object orientation also enables the presentation of interactive, multimedia applications. With MPEG-4, service providers can sell customers an entirely new class of programming including networked gaming, interactive eLearning programs, and media-on-demand. Combined, MPEG-4 supports a compelling business argument in support of the deployment of the technology in environments that will not only save money but will also earn money.

MPEG's Success Proves Market Viability

Proponents of proprietary alternatives will mention interoperability in the multi-vendor world as problematic, but previous MPEG standards have already proven them wrong. The MPEG-4 community is working together, driven by market forces: vendors of different and complementary solutions are partnering to offer best of breed solutions. The formal interoperability work in e.g. M4IF is the tip of the vast market-based iceberg of work by companies working bilaterally to guarantee their products will interoperate.

MPEG-4 allows different vendors to specialize in their category of excellence. Hardware or software , real time and non real time encoders will all encode content for specific applications, but it can all be read by the same decoders.

Critics of MPEG-4 often cite the (lack of) licensing terms, particularly for MPEG-4 Visual. Licensing terms for the different MPEG-4 parts are available in one-stop-shop license pools, and support the market's adoption of MPEG-4 in various business models – from broadcast and streaming to mobile and packaged media distribution. The terms are designed to be fair and reasonable to both the licensees and licensors. Further, MPEG-4 licensors have all pledged to make licenses to their essential patents available to the market on reasonable and non-discriminatory terms.

Market Indicators

MPEG Market Penetration

In the past decade, MPEG technologies have fuelled the transition from analog to digital media delivery worldwide by providing the underlying foundation for today's mass-market digital distribution platforms. The unprecedented adoption of DVD technology over the last few years is evidence of the impact MPEG technology has had on home entertainment.

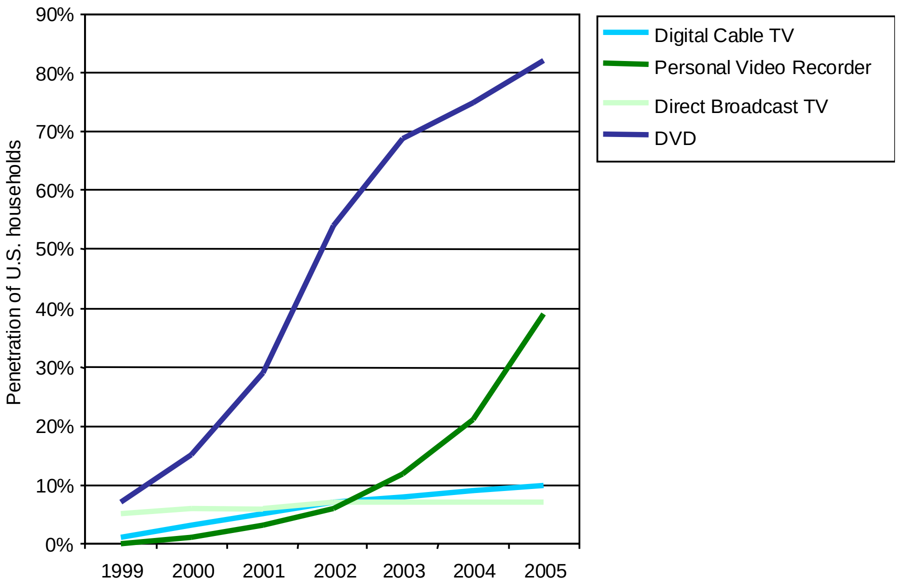

Figure 8: US Market Penetration – growth from 1999-2005 (The Yankee Group[xi])

Other trends include digital cable TV, direct broadcast satellite TV, personal video recording (PVR), high-definition television, and MP3 online digital audio – each contributing to new ways consumers experience media.

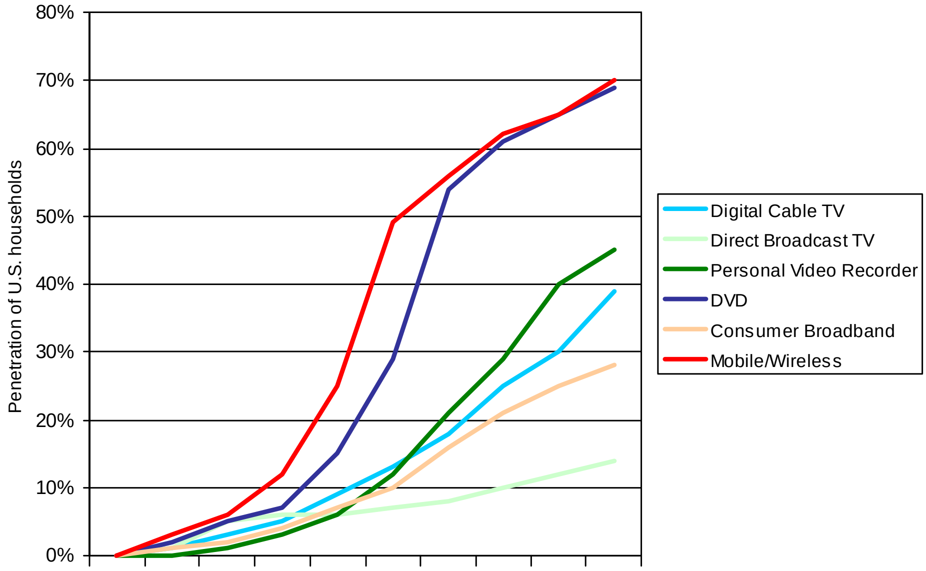

Figure 9: US Market Penetration – growth from year 0 (introduction) – 10 (The Yankee Group)

New Distribution Channels and Media Experiences

MPEG-4 enables the distribution of rich media across broadcast, broadband, wireless or wired networks – something never before technically or economically viable.

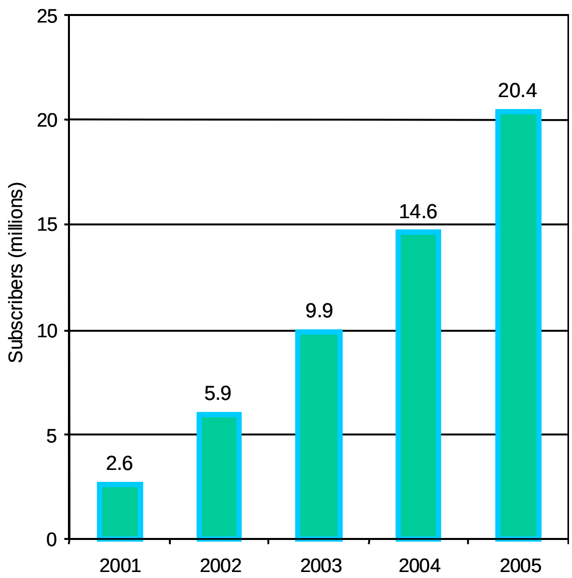

Figure 10: U.S. Video-On-Demand Subscriber Forecasts (The Yankee Group)

Taken together, MPEG-4 improves and accelerates digital media distribution – further contributing to the growth of consumer applications such as video-on-demand (VOD), mobile multimedia distribution, interactive packaged media and streaming media.

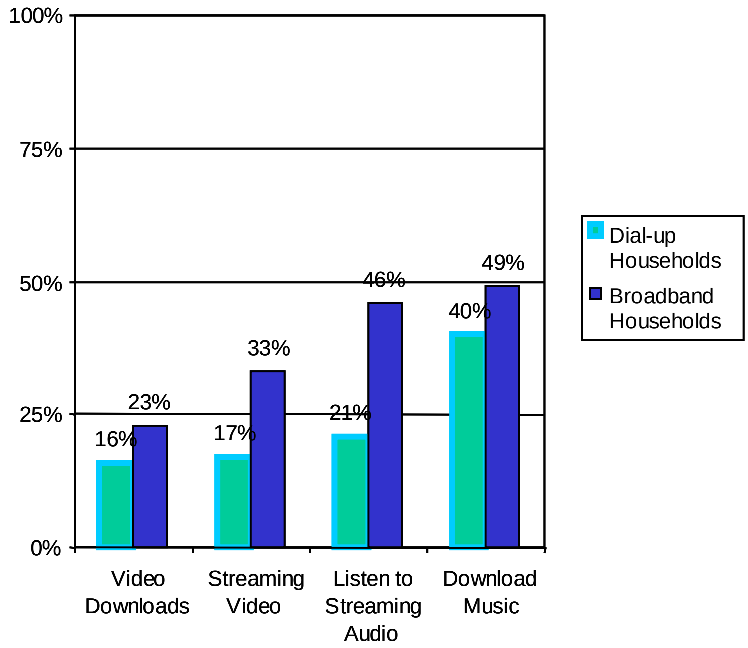

Figure 11: Broadband Media Access (The Yankee Group)

As broadband media access accelerates, consumer behavior shows signs of adopting new ways of accessing and using media. Streaming and downloading video are slowly but certainly becoming a part of the fabric of today's digital home. As with MP3, MPEG-4 provides a single media standard for accessing media globally – including feature length films and high quality (AAC) music albums. As a single standard, it ensures interoperability between media created by content owners and the playback capabilities on consumer devices – whether PCs, set-top boxes, DVD players, mobile terminals, or other devices.

CLARIFYING COMMON QUESTIONS

Who licenses MPEG-4 technology?

ISO/IEC MPEG is the group that makes MPEG standards. MPEG never deals with licensing. It asks of companies that propose technologies for the standard to commit to licensing their patents on Reasonable and Non-Discriminatory Terms and Conditions (also called "RAND").

The exact details of the licensing model are outside the scope of this paper, Please visit the M4IF Patents page for links to the latest information at this level, including links to the licensing terms themselves. (http://www.m4if.org/patents/)

There are several MPEG-4 "patent pools" or "joint licensing schemes," corresponding to the different parts of the MPEG-4 standard. MPEG LA (www.mpegla.com) will license MPEG-4 Visual and will also take care of a joint license for MPEG-4 Systems. Dolby has announced a program on behalf of several MPEG-4 AAC licensors (www.aac-audio.com), and a more comprehensive MPEG-4 Audio licensing program is expected to be announced soon.

These joint licensing schemes are not carried out on behalf of ISO, MPEG or M4IF, nor are they, or do they need to be, officially blessed by any such organization. There is no "authority" involved in licensing, it is a matter of private companies working together to offer convenience to the market.

What is the role of M4IF in licensing?

The MPEG-4 Industry Forum has written in its statutes that it shall not license patents or determine licensing fees. It does not share in the license royalties. M4IF has acted as a catalyst, promoting the use of patent pools. M4IF has among its members licensors, licensees and other entities which have an interest in fair and reasonable licensing. M4IF believes that the initial steps to get access to the necessary licenses have now been completed. Further tuning will likely be necessary as a better understanding of the market is gained.

Should we wait for MPEG-4 AVC/H.264?

In July 2001, after MPEG reviewed evidence from formal subjective tests, MPEG decided that scientific progress warranted the addition of a new Video codec to the MPEG-4 standard. The most promising development appeared to be "H.26L,"" a draft standard that had been in the works in ITU Study Group 16, developed by the Video Coding Experts Group (VCEG). MPEG accepted the invitation by ITU-T to start a joint effort, building on H.26L, which will result in a single, world-class video coding standard across ITU and ISO.

This standard will be ready by the end of 2002, and its audio and video quality will beat all other video and audio coding systems in existence, including several proprietary systems recently deployed. The new standard will be an extension of the MPEG-4 toolbox, because it fits in the MPEG-4 architecture, can make use of the MPEG-4 File Format, and can be combined with MPEG-4's other tools such as those for audio coding.

MPEG-4 Advanced Video Coding / H.264 makes use of the latest research in video coding. The coding methods rely on the fact that computational power and memory have become cheaper than 5 years ago, meaning that coding methods can be more complex than could previously be accommodated in hardware and software environments.

The new standard will not replace the current MPEG-4 (part 2) codecs. First, part 2 is being deployed right now, and part 10 is still under development. Second, part 2 is already part of some other standards and specifications. Third, in some environments, MPEG-4 Simple Visual will remain a great complexity/performance trade-off. In many environments, support for part 10 may be added to support part 2.

A common myth has it that when a standard is ready, quality gets frozen. This is not true, and MPEG-2 clearly shows how false that claim of some of MPEG-4's competitors is. MPEG only standardizes DEcoders and the bitstream syntax – which means that ENcoders are left to the engineering skill of the manufacturer. When MPEG-2 was ready in 1995, 6 Mbit/second was required to transmit a decent-quality broadcast signal. Today, this can be done in the 2 Mbit range – without any change to already deployed decoder boxes.

The same applies to MPEG-4 today. Implementations are rapidly improving in quality, because encoder manufacturers are competing to offer good quality. The same will also apply to MPEG-4 part 10. Competition in an interoperable eco-system is the best guarantee for quality.

Should companies wait for MPEG-4 AVC to start doing MPEG-4? Certainly not. MPEG-4 is a great proposition today, with its Systems framework and its highly efficient audio and visual coders.

What is the relation of MPEG-4 visual and the DivX codec?

DivX5 is an implementation of MPEG-4 Advanced Simple Visual Profile. DivXNetworks is also working on file format compliance.

Is Microsoft Windows Media an MPEG-4 codec?

Microsoft was one of the first companies to deploy an MPEG-4 Video codec in its Windows Media platform, in previous versions. It doesn't seem to be present in the latest version, Windows Media 9. Some developers will know Microsoft's contribution to the MPEG-4 Reference Software, one of the two implementations of the MPEG-4 Visual standard that developers can download from ISO's website (The other implementation is from the European project 'MoMuSys').

The future is all downloadable software codecs, why do we need a standard?

There are many environments in which downloading codecs is not possible. The future is video and multimedia on many different devices, with very many totally different uses. While the internet is growing exponentially, and streaming media and video on demand are poised to be large applications, the future is also about wireless connectivity on phone and PDA or slate type devices. Most often these are constrained by size to implement video decoding in hardware, and embedded systems.

Is MPEG-4 based on Quicktime?

The file format of MPEG-4 (MP4) is based on the QuickTime architecture. The rest of the MPEG-4 standard was developed independent of QuickTime. QuickTime started supporting MPEG-4 with its version 6, which includes Simple Visual profile and AAC.

MPEG-4 and other technical specifications

Multi-vendor support ensures market driven solutions. Standards like MPEG's have a potential for broad industry support. Proprietary solutions can only succeed if they are adopted by large market segments, which has not happened with existing technologies. The table below gives a comparison of MPEG-4 against most commonly used multimedia formats on the Internet today.

	MPEG-4	**Windows Media**	**Real**	**Flash**
Audio/Video Codec	Standards based; multi-vendor support.	Proprietary	Proprietary, but supports automatic download of MPEG-4 plug-in.	Proprietary + proprietary Real and QuickTime formats.
Interactivity	Highly interactive.	Limited	Yes, via SMIL.	Highly interactive.
Digital Rights Management	Interfaces to proprietary DRM. More interoperable DRM under development in MPEG-4 and MPEG-21	Microsoft DRM	Content access control	No
Real-time stream control	Yes	Yes	Yes	No
Synchronization	Audio, video and all other objects can be tightly synchronized with high accuracy	Tight synchronization between audio and video	Tight synchronization between audio and video	No synchronization between scene and streams
Broadcast capable	Yes, including interactive features	A/V only	Scene must be unicast	No
Object model support	Video/audio and rich 2D/3D mixed media, synthetic graphics. DRM on separate streams.	Audio/Video only	Video/audio and mixed media through SMIL based protocol. No streaming of mixed media.	Video/audio and mixed media through proprietary protocol.
Graphic Objects	Yes	No	No	Yes
Transport	Support exists for HTTP, UDP, RTP/RTSP, MPEG-2TS, mobile	HTTP, UDP, RTP/RTSP, mobile	HTTP, RTP/RTSP, mobile	HTTP
PC, Set Top Box, Wireless	Yes	Yes	Yes	No

MPEG-4 and Interactive TV

Alongside significantly less bandwidth for the same quality, the native support for Interactivity is a key difference between MPEG-4 and the MPEG-2 technology currently broadly deployed in digital television systems.

In every case, making Interactive TV work in an MPEG-2 based environment, means that operators need to adopt one or more proprietary solutions, or solutions based on technologies not native to MPEG, and add them productively to an MPEG-2 delivery environment. This has led to the emergence of several competing proprietary add-on technologies competing for the business of ITV operators.

Each operator has a unique composite solution of technologies, usually determined by their MPEG compression platform, their Conditional Access System and their Middleware platform, e.g. OpenTV. This has led to the emergence of several incompatible vertical solutions and markets. The problem with vertical markets, not only in the business sense but also in the technology sense, is that at the end of the day, end-users don't benefit from them, and service deployment is slowed-down. Several attempts to dissolve them into the horizontal market have taken place and are meeting great resistance from this sort of "economic gravity" which makes vertical markets inevitable without open standards.

One very promising technology is the DVB standard for interactive TV APIs, Multimedia Home Platform (MHP), described in more detail below. In the United States, MHP has its equivalents in the Java-based Digital Applications Software environment (DASE), an Advanced Television Systems Committee (ATSC) activity and in OCAP, the Open Cable Application Platform specified by the OpenCable consortium, which is based on MHP.

The mainstream of the Broadcast Industry likes Java, because unlike the host of other proprietary and flavored web-standards based approaches (e.g. MediaHighway, Liberate, OpenTV), it offers content creators and providers and service operators a chance to 'write once, run many times' of the same content, which is itself indispensable to creating a horizontal market.

This paper will not compare MPEG-4 to these technologies and operator-specific or platform-specific architectures. We will give a few examples of how the power of MPEG-4 can be easily added to complement the most important standards in this area. As an example of a platform based on procedural content, we examine ways in which MPEG-4 can add value to MHP. Then in the other track of Interactive TV, away from Java, there is significant interest in the W3C work on newer generations of its meta-languages based on HTML, e.g. XML and XHTML. In fact the MHP platform supports both.

MPEG-4 and MHP

An often-asked question about MPEG-4 is how it relates to the Multimedia Home Platform specification of DVB (Digital Video Broadcasting)[xii].

The first thing to understand is that there are two relevant groups of DVB specifications. The first, DVB 1.0, is the transport foundation of the DVB family of standards. This specification spells out how to implement DVB-compliant MPEG streams. There, MPEG-4 is seen as a logical evolution, and one which will be more efficient when DVB services are to be delivered over IP. DVB has announced it is considering the adoption of MPEG-4 AVC/H.264 and MPEG-4 AAC with SBR (see above) extensions by 2003.

The second DVB, also sometimes referred to as DVB 2.0, addresses the Multimedia Home Platform and a variety of next generation delivery applications, including Copy Protection and Copy Management and delivering DVB services over IP. The Multimedia Home Platform (MHP) defines a generic interface between interactive digital applications and the terminals on which those applications execute. The MHP specification specifies how to download applications and media content, typically delivered over a DVB compliant transport stream, and optionally in the presence of a return channel.

MPEG-4 is a natural companion to MHP applications, with low bitrate video and scene representation formats streamed or delivered over IP to set top boxes. The application, interaction and synchronization models of MPEG-4 allow more dynamic content to be added to MHP-type of applications.

Because MPEG-4 can be carried by MPEG-2 transports we can achieve a very fine grain synchronization between the broadcast program and the MPEG-4 multimedia content. Integrating MHP with MPEG-4 can enable object-based interactive digital television.

The combination of MHP and MPEG-4 provides the ability to develop very flexible and rich interactive applications for the interactive broadcast domain. The MPEG-4 features can be introduced smoothly and gradually, in a backwards-compatible manner.

MPEG-J and DVB-J

The MHP architecture is defined in terms of three layers: resources, system software and applications. Typical MHP resources are those elements that can be called upon by applications to perform certain functions, for example MPEG processing, I/O, CPU, memory and graphics handling. The system software presents a standardized abstract view of the resources of the platform to the applications, thus enabling "platform independence." An "application manager" is provided to manage the interaction between these elements.

Generic Application Program Interfaces (APIs) are specified by DVB-MHP, based around DVB-J, which includes the Java Virtual Machine (VM) as originally specified by Sun Microsystems. MHP applications can only access the resources of the platform via these specified APIs, a feature which guarantees the stability of the platform and its robustness against "rogue" applications. These APIs are specified by the DVB Technical Module (TAM) and are tested for conformance with the MHP specification through the use of agreed test applications.

MPEG-J(ava) offers a set of functionalities complementary to those offered by DVB-J. MPEG-J in MPEG-4 is a set of Java APIs that may be present on a MPEG4 terminal. MPEG-J applications (MPEGlets), which are sent as part of the presentation, use the MPEG-J API's to control the capabilities of the MPEG-J terminal. Java packages that must be in the terminal that supports MPEG-J include java.lang; java.io; java.util.

MHP's "DVB-J" consists of a generic Java platform that is similar to PersonalJava 1.2a. It supports "Xlets" (but not applets), and has a very limited subset of java.awt -- the widget set is removed. Additionally, it includes JMF 1.1, Java TV 1.0, and a number of DVB-specific APIs for accessing TV-specific data and controlling TV-specific functionality.

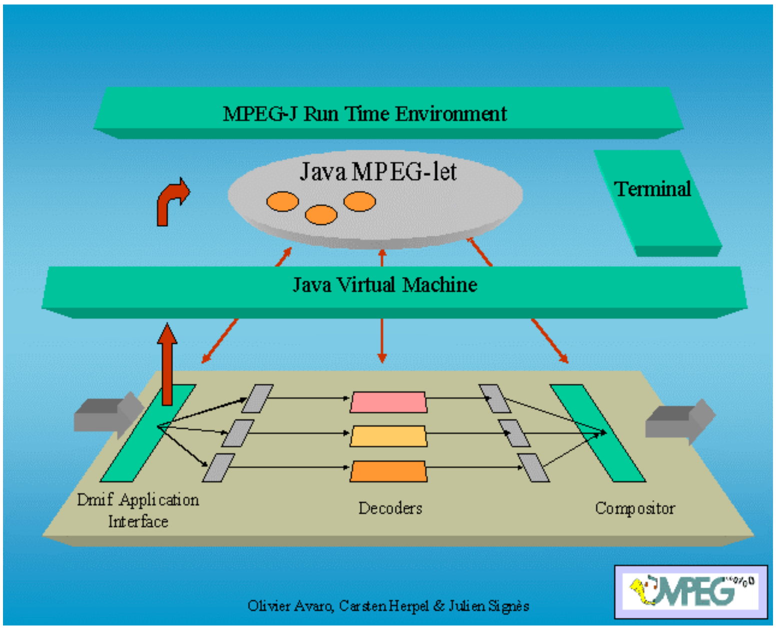

Figure 12:The MPEG-J reference architecture

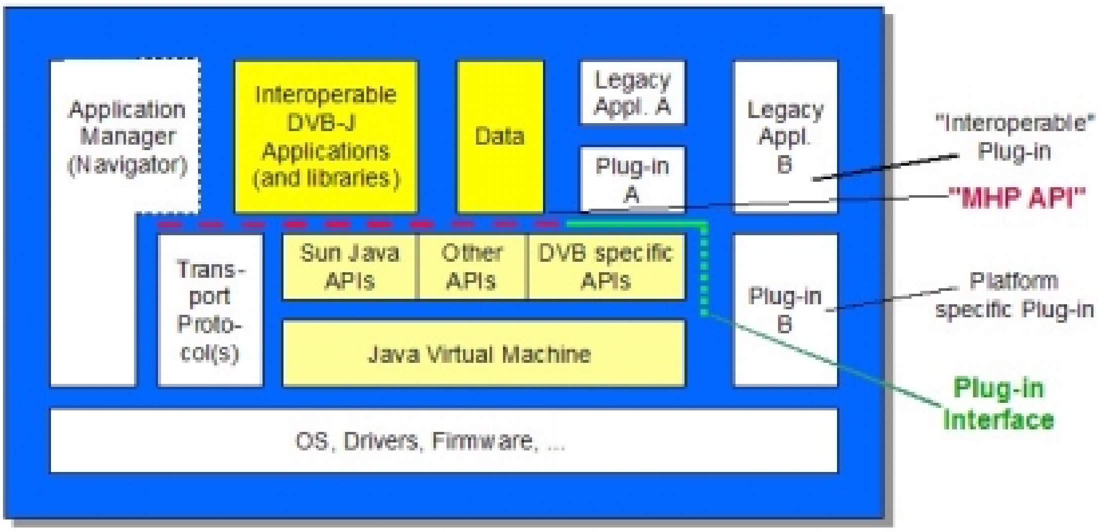

Figure 13: MHP reference architecture showing role of DVB-J

MPEG-4 and SMIL, SVG

SMIL is the Synchronized Multimedia Integration Language[xiii] of the W3C and SVG is W3C's Scalable Vector Graphics specification[xiv]. A comparison of MPEG-4 and SMIL+SVG capabilities follows below. MPEG-4 provides a rich multimedia experience, in which interactivity, streaming, and various mixed media, including graphics objects, are combined seamlessly. SMIL and SVG, as currently proposed for use by 3GPP, provide somewhat similar functionality, with notable differences, as SMIL is more declarative in nature and MPEG-4 is more procedural. The comparison mostly concerns MPEG-4 BIFS (the Binary Format for Scenes) and the Object Descriptor framework of MPEG-4, which takes care of the synchronization between the different objects .

Requirement	MPEG-4	SMIL+SVG
Spatial and temporal composition of text, graphics, images and streamed media (audio and visual streams)	Very simple to very complex composition. 2D and 3D profiles	Only 2D composition
Flexible synchronization models of different objects (co-start, co-end, ...)	Yes	Yes
Broadcast-grade synchronization of all objects on a rigid timeline (e.g. A and V)	Yes	No
Streaming scene description	Yes	No
Compression of scene description	Yes	No
Dynamic scenes (add/ remove objects, etc)	Yes	No
Streamed animation of scene components	Yes	No
Broadcast capable	Yes	No
DRM tightly coupled with scene (e.g. can protect streams independently)	Yes	No

MPEG-4's Textual Format: XMT

Originally, MPEG-4 only contained a binary scene description language. Later, it became clear that it would be helpful to add a textual representation as well, in the form of XMT, the eXtensible MPEG-4 Textual format. XMT is an XML-based language, like SMIL. MPEG has been careful to build XMT as compatible with SMIL as possible, to aid interoperability in media distribution. Another goal was to build compatibility with the X3D specifications for interactive 3D content (also known as VRML or Virtual Reality Modeling Language). MPEG-4's scene description model is based on the textual VRML language, to which MPEG added real-time behavior, 2D support and a binary representation for efficiency. Lastly, there is a relation between XMT and MPEG-4, which is based on XML schema language; further detail would be beyond the scope of this document.

In the course of defining the textual format, MPEG-4 was also extended with the flexible timing models that SMIL uses. The so-called "flextime" support was added to the broadcast-type of time stamp-based, rigid MPEG-2 type of synchronization.

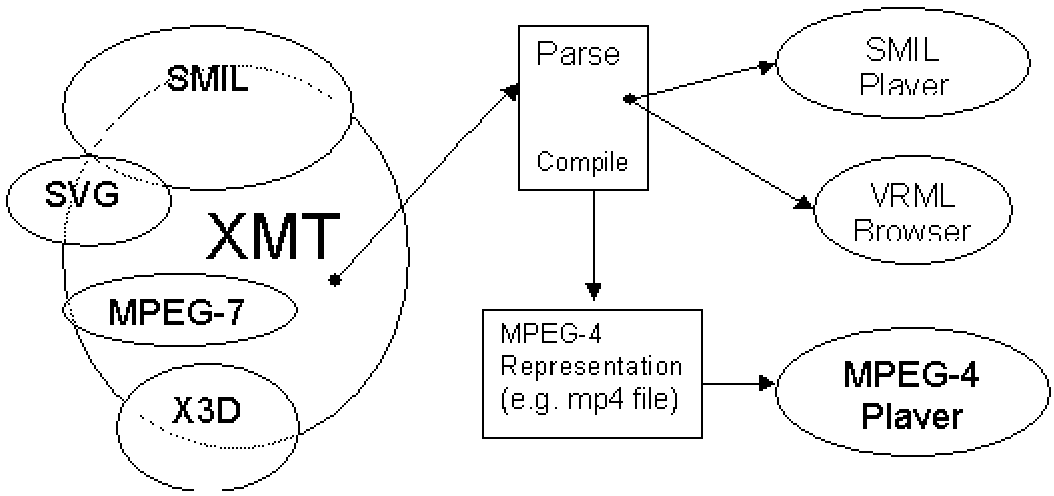

Figure 14: Relation of MPEG-4 XMT to other specifications.

The textual format and the binary format are largely dual representations of the same information. In most situations, one would want to deliver scene description information in binary form, as that is much more efficient. For exchanging scenes between authors or storing scenes inside a single organization in a way that is understood by multiple tools, a textual format is a useful tool. It is easy to go from text to binary representation, and while the other way is just as easy in theory, it is harder to do so in a ways that is meaningful for an author, much like a decompiled program can be hard to read.

THE MPEG-4 INDUSTRY FORUM

The MPEG 4 Industry Forum represents more than 100 companies from diverse industries evenly distributed across North America, Europe and Asia, addressing MPEG-4 adoption issues that go beyond the charter of ISO/IEC MPEG.

M4IF is vital to the success of the MPEG-4 standard, since the work done by MPEG is necessary but not sufficient. In its endeavors to get MPEG-4 widely adopted, M4IF picks up where MPEG stops. There are more than just technical issues to resolve; for example, M4IF does much marketing work and the recent public debate over the licensing terms for MPEG-4 visual profile was greatly assisted by the Forum's industry-wide public discussion on the pros and cons of the announced schemes. In addition, M4IF has an advanced program of cross-vendor product interoperability testing.

The M4IF provides the industry with a means to work in a pre-competitive, post-standardization environment to help building out the MPEG-4 ecosystem, Other activities of the forum include product certification, several working groups, access to MPEG committee members, and an annual conference (WEMP4).

Join the forum

Membership in M4IF will put you in touch with your future clients, partners, suppliers and competitors. It will put your company's name on the list of the top 100 companies at the leading edge of MPEG-4. For the price of a one-page advertisement in a trade magazine daily, you join the forum and might meet your next customer there.

Help Drive Success

The M4IF has a unique and broad spectrum of members, coming from all industry segments, all individuals who are focused on MPEG-4 and have decisive roles in their companies, the time is now to communicate your value proposition to other members.

To join M4IF, to find out what activities will benefit your company, visit: http://www.m4if.org.

ACKNOWLEDGEMENTS

This document was produced by Martin Jacklin from contributions by Tim Schaaff of Apple Computer, Sebastian Moeritz of dicas digital image coding[xv], Yuval Fisher of Envivio[xvi], U. Georg Ohler of Fraunhofer IIS, Jennifer Toton of iVAST[xvii], Rob Koenen of InterTrust[xviii] (M4IF President), Jan van der Meer of Philips[xix], Olivier Avaro of France Telecom[xx], Professor Klaus Diepold of the Technical University of Munich[xxi], and other members of the M4IF Board of Directors and Marketing Workgroup.

For further reading please refer to the M4IF resources page at http://www.m4if.org/resources.

Comments are welcome at papers@m4if.org

---

[i] http://www.isma.tv

[ii] http://www.m4if.org

[iii] http://www.iso.org

[iv] http://www.iec.org

[v] http://www.harmonicinc.com

[vi] http://www.tandbergtv.com

[vii] http://www.ietf.org

[viii] http://www.3gpp.org

[ix] http://www.ivast.com

[x] http://www.itu.int

[xi] http://www.theyankeegroup.com

[xii] http://www.dvb.org and http://www.mhp.org

[xiii] http://www.w3.org/AudioVideo/

[xiv] http://www.w3.org/TR/SVG/

[xv] http://www.dicas.de

[xvi] http://www.envivio.com

[xvii] http://www.ivast.com

[xviii] http://www.intertrust.com

[xix] http://www.philips.com

[xx] http://www.francetelecom.fr

[xxi] http://www.ldv.ei.tum.de