Understanding MPEG-4:
Technologies, Advantages, and Markets

An MPEGIF White Paper

MPEG-4: The Unified Framework of High-Performance, Open International Standards Enabling New Services and Greater Quality for Video, Audio, and Multimedia

Building on the enormous success of the MPEG-1 and MPEG-2 standards, MPEG-4 offers new audio and video codecs with twice the performance of MPEG-2, enabling new services ranging from HDTV satellite broadcasting to mobile video and games on handheld devices.

In addition to excellent audio and video codecs, MPEG-4 offers a comprehensive unifying framework of tools to combine them with 2D and 3D graphics, animation, text, and other objects in interactive rich media experiences.

In this paper, you will learn about:

- What MPEG-4 Actually Is
- Why Manufacturers and System Operators Have Chosen MPEG-4
- Applications and Markets for MPEG-4
- How MPEG-4 Relates to Other Standards
- What's in the MPEG-4 Standard and How It Evolves
- MPEGIF and the MP4 Logo Program
- Answers To Common Questions on MPEG-4:
  - Licensing
  - How MPEG-4 Relates to Other Codecs
  - Interactive TV and MHP

mp4 The Media Standard

The MPEG Industry Forum
http://www.mpegif.org
+1 510-744-4025 phone
+1 510-608-5917 fax

mpeg Industry Forum

Understanding MPEG-4:
Technologies, Advantages, and Markets

An MPEGIF White Paper

MPEG-4: The Unified Framework of High-Performance, Open International Standards Enabling New Services and Greater Quality for Video, Audio, and Multimedia

Document Number mp-in-40182

Copies of this document may be freely downloaded from MPEGIF web site at http://www.mpegif.org

For further reading please refer to the MPEGIF resources page at http://www.mpegif.org/resources.php.

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@mpegif.org

Editor: Robert Bleidt, Fraunhofer IIS

Contributors: Glenn Bulycz, Apple
Farzin Aghdasi, Pelco
Mikael Bourges-Sevenier, Mindego

This document is based on the original 2002 white paper produced by Martin Jacklin from contributions by Tim Schaaff of Apple, Sebastian Moeritz of dicas digital image coding, Yuval Fisher of Envivio, U. Georg Ohler of Fraunhofer IIS, David Price and Fadi Malak of Harmonic, Jennifer Toton of iVAST, Rob Koenen of InterTrust, Jan van der Meer of Philips, Olivier Avaro of France Telecom, Professor Klaus Diepold of the Technical University of Munich, and other members of the MPEGIF Board of Directors and Marketing Workgroup.

 MPEGIF thanks Apple Computer, Envivio, Inc., Harmonic, NTT DoCoMo, Panasonic, Pelco, Samsung, Streamcrest Associates, Sony, Tandberg TV, and others for the use of the photographs and illustrations they have provided.

Table of Contents

What Is MPEG-4?

A Unified Framework for Advanced Multimedia

MPEG-4 is a family of open international standards that provide tools for the delivery of multimedia. These tools include both excellent codecs for compressing conventional audio and video, and those that form a framework for rich multimedia – combinations of audio, video, graphics, and interactive features.

Excellent Conventional Codecs

MPEG-4's conventional audio and video codecs provide the highest quality and compression efficiency available today, and have been adopted as the foundation of many new media products and services.

MPEG-4's latest video codec is AVC, the Advanced Video Codec. Also identically standardized as ITU H.264, the AVC codec represents the latest developments in video coding, offering a typical compression rate half that of MPEG-2 for similar perceived quality. This dramatic improvement has led to AVC becoming the new standard for video transmission, being employed in most new video products and services where quality and compression efficiency are paramount. New HDTV satellite broadcasting and DSL video services will use AVC, as will the Sony PlayStation Portable and Apple QuickTime 7 player. AVC will also be used in video broadcasting to mobile handsets using the DVB-H, DMB, and MediaFlo systems, and specified in the HD-DVD and BluRay high-definition optical disc standards.

AVC is intended to be practical when implemented on the latest generation of high-performance hardware and processors. For applications where hardware cost or power considerations make implementing AVC difficult, MPEG-4 offers the Simple and Advanced Simple Profile codecs. These codecs offer good performance while using less complex encoder and decoder architectures. They are commonly used for 3G wireless videophony, digital still camera or convergence devices, and for security or intranet video applications.

These codecs are usually coupled with AAC, MPEG-4's family of general-purpose audio codecs. The core AAC codec offers excellent quality at stereo bitrates above 128 Kb/s. Compatible extensions to AAC, the HE-AAC and HE-AAC v2 codecs, improve its quality at lower bitrates, while maintaining compatibility with existing AAC decoders.

AAC is being used not only in television broadcasting, where it is the codec for Japan's ISDB digital TV system, but also in music players and distribution services, such as Apple's iPod and iTunes.

The low-bitrate quality improvements of the HE-AAC codec has enabled new digital music broadcasting services, such as XM satellite radio and the music download services of mobile carriers KDDI and Orange. The latest version, HE-AAC v2, incorporating MPEG-4's Parametric Stereo tool, has just been standardized by 3GPP for future music streaming services.

Codec	Features	Typical Applications and Users
AVC – Advanced Video Codec (MPEG-4 Part 10)	Highest-Performance video codec for demanding applications	HDTV Broadcasting – DirecTV, bSkyb, Premiere Mobile Multimedia – DMB, DVB-H, MediaFlo systems Internet Video – Apple QuickTime Gaming – Sony PSP UMD Disc
SP - Simple Profile, ASP - Advanced Simple Profile (MPEG-4 Part 2)	High-Performance video codec with scalability and error resilience features	3G Wireless Videophony – DoCoMo, Hutchinson-Whampoa Intranet Video – Envivio, vBrick Internet Video – Apple QuickTime Digital Cameras – Panasonic, Samsung, Sanyo
AAC – Advanced Audio Codec	High-Performance audio codec for excellent quality at moderate bitrates	Portable Music – Apple iPod, iTunes DTV Broadcasting – ISDB (Japan)
HE-AAC – High-Efficiency AAC	High-Performance audio codec for superior quality at bitrates below 48 Kb/s	Satellite Radio – XM Radio Mobile Handset Download – KDDI, Orange
HE-AAC v2 (HE-AAC + Parametric Stereo)	Highest-Performance audio codec for excellent quality at bitrates below 48 Kb/s	Mobile Handset Streaming - 3GPP

Table 1. Popular MPEG-4 Audio and Video Codecs

Framework for Rich Interactive Media

MPEG-4's rich media tools include codecs for combining audio and video with text, still images, animation, and both 2D and 3D vector graphics into interactive and personalized media experiences. MPEG-4 includes both a scripting language for simple interaction and a Java variant, MPEG-J for more elaborate programming.

While video and audio are coded with MPEG-4's excellent conventional codecs, 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. Additionally, 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 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.

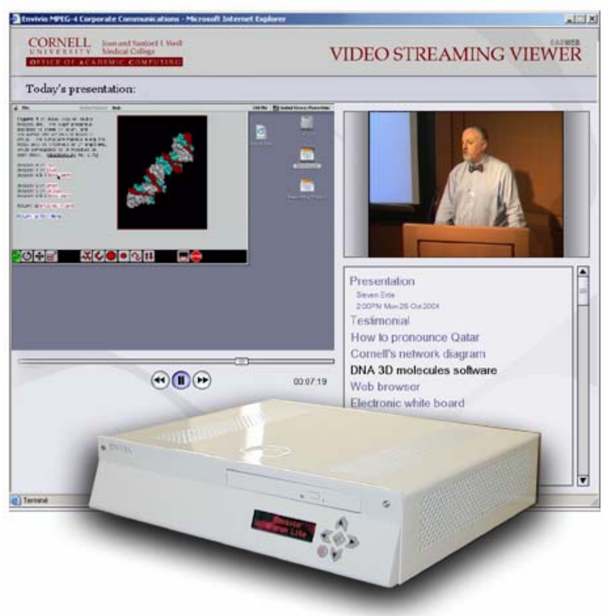

Figure 1. MPEG-4's rich media tools allow the tight integration of graphics, stored video, and a live presentation in this educational application. Courtesy Envivio, Inc.

Because MPEG-4's rich media tools may be 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 their moves, or to personalize an enhanced video program. MPEG-4 also opens up new revenue opportunities through the integration of back-office systems including transaction and e-commerce systems.

With MPEG-4, advanced interactive programming can be authored to seamlessly integrate audio and video with 2D and 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 also 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 previously impossible.

Truly a "framework of the future", MPEG-4's rich media tools enable the creation of content that is difficult or impossible to achieve with the proprietary tools available today. These tools are already in use for distributing complex corporate and educational multimedia presentations that require tight synchronization of video and graphics content, with entertainment applications expected as the creative community moves towards interactive content.

Why Have Manufacturers And Operators Chosen MPEG-4?

Excellent Performance

A codec's compression efficiency, or the bitrate required for a given perceived quality level, is one of its most important features. MPEG-4's latest audio and video codecs, AAC and AVC, provide outstanding compression efficiency, equaling or exceeding the performance of any codecs now available.

MPEG is constantly evaluating new algorithms and technology to see if they offer a practical and significant improvement over existing MPEG standards. Successful techniques are then incorporated as compatible extensions or improvements to existing standards, or if necessary are issued as complete new standards. An example of the latter is AVC, which is sufficiently different from MPEG-4's Part 2 video that it required development of Part 10 of MPEG-4.

Where possible, however, MPEG's philosophy is to extend rather than to replace. An example is the continual improvement in audio coding efficiency enabled by successive versions of the AAC codec. AAC was a significant improvement over MPEG-1's Layer 3 codec (MP3), but MPEG has extended AAC through the techniques in the HE-AAC and Parametric Stereo codecs, while preserving compatibility with existing AAC decoders. The upcoming MPEG Surround codec will improve coding efficiency even further, offering a four to six times improvement in bitrate over MP3 for multi-channel content.

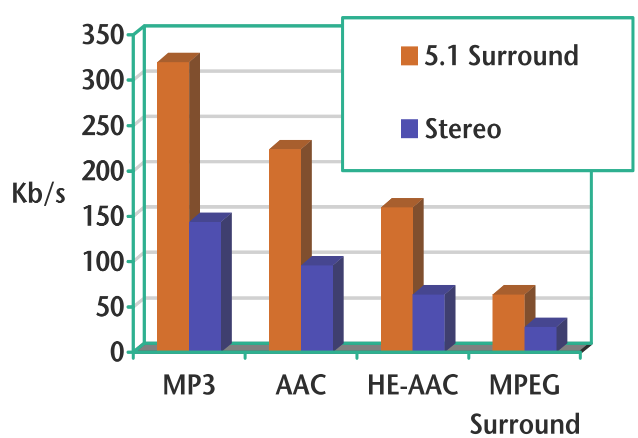

Figure 2. The evolution of MPEG Audio codecs has produced a 5 to1 reduction in bitrate for similar perceived quality.

Open, Collaborative Development to Select the Best Technologies

Key to this evolution in performance is MPEG-4's open and collaborative development process, where new techniques are proposed and evaluated not on business considerations, but strictly on their technical merit in a careful, thorough, and objective set of experiments and tests.

This development, like all international standards, is carried out in a public process by delegations from many countries, representing their top scientists and engineers. In MPEG's case, this development is done by the Motion Picture Experts Group, a working group of the ISO and IEC standards organizations, and typically involves hundreds of companies from tens of countries.

Once developed and proposed, new standards go through careful review and approval to insure that they are completely and unambiguously specified, so they can be implemented solely by following the standard. Additionally, test results, frameworks, and the source code to reference software implementations are made available so anyone can see how the codec was implemented, study why design choices were made, and understand how the codec operates.

Competitive but Compatible Implementations

With MPEG-4, manufacturers and technology developers collaborate to create the standard, but they compete to implement it. In MPEG-4, the bitstream format and decoder operations are carefully specified to allow complete interoperability, but manufacturers are not restricted in how they design encoders and decoders internally.

Decoder manufacturers, particularly those building decoders in hardware chips or devices, can take comfort in the knowledge that the bitstream format will not change. Encoder manufacturers can use their proprietary techniques and experience to continually improve coding efficiency and quality, knowing that their output bitstreams will be rendered by all decoders.

These same principles apply to MPEG-2, where history has shown them to be very worthwhile. Early MPEG-2 encoders required a 6 Mb/s bitrate for acceptable picture quality, while today's encoders can operate at 2 to 2.5 Mb/s, saving 60-65% of the original bandwidth costs, without any changes to the installed base of MPEG-2 decoders. A similar experience curve is expected for MPEG-4.

For users and operators, this competition offers several benefits. Besides potential bandwidth savings as encoders improve, they can choose between multiple suppliers for each component in their signal chain – encoders, servers, transmission equipment, and decoders – with different prices, performance levels, and business terms. Once chosen, they are not locked-in to a specific supplier, but can supplement or replace their components with MPEG-4 equipment from another supplier secure in the knowledge that they will operate together.

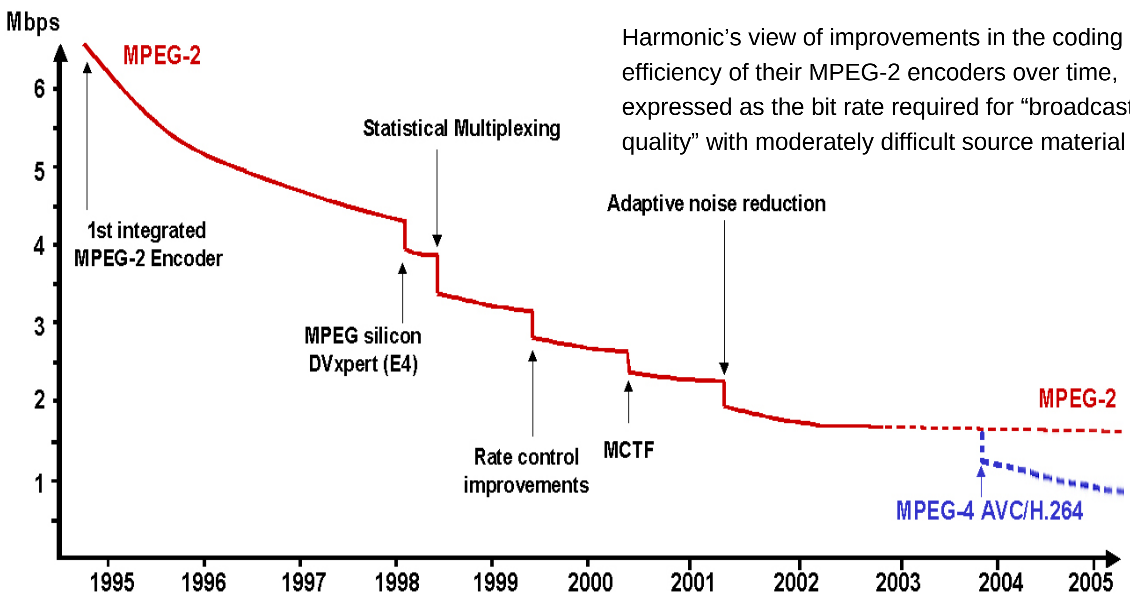

Figure 3. Improvements in coding efficiency are facilitated by MPEG-based competition. Courtesy Harmonic.

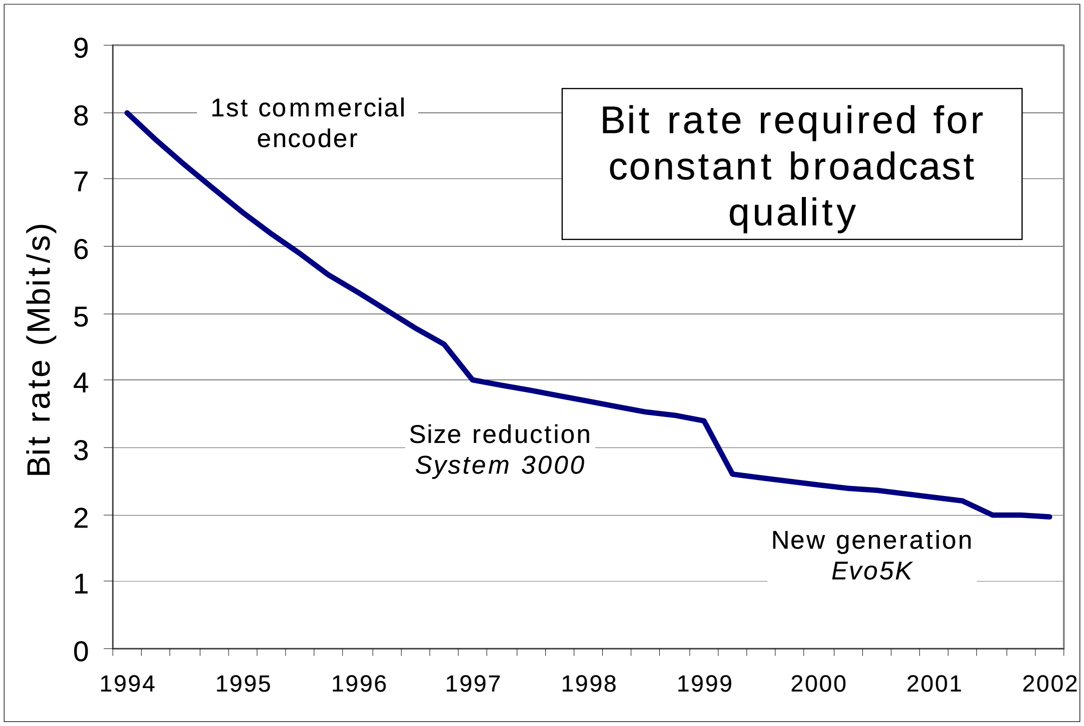

Figure 4. Improvements in coding efficiency are facilitated by MPEG-based competition. Courtesy Tandberg TV.

Lack of Strategic Control by a Supplier

In addition to the benefits of intense day-to-day competition among manufacturers to best implement the standard, MPEG-4 also benefits from the lack of strategic control by a supplier, as is common with proprietary codecs.

The specifications and tools needed to implement MPEG-4 are available to anyone, and there are no restrictions on how or where it is used, or limits on what hardware or software can be employed. MPEG-4 is not tied to any particular application program or service offering, and it may be used with any DRM system, operating system, software platform, or hardware chip, without any bundling, limitations, or conditions being placed on the user.

The business failure of an MPEG-4 supplier does not put the standard in jeopardy, and users may shift new purchases of MPEG-4 equipment to other suppliers without losing their investment in equipment or software already purchased. Conversely, the success of a MPEG-4 supplier does not mean research and development efforts on the core functionality of the codec will be shifted to product extensions into other areas, as is common when a proprietary standard matures.

A user's risk is also reduced through the stability of the standard. There are no time limits on the use of the MPEG-4 standard, and non-discriminatory patent licenses for MPEG-4 have guaranteed terms and renewals, along with strict limits on changes in royalties.

For MPEG-4 manufacturers, there are no restrictions, favoritism, or delays in accessing specifications as they are developed, and no requirements for certification or approval from another party before it is produced. (MPEGIF does operate voluntary interoperability and conformance programs, but these are not limits on what a manufacturer can do.)

Public, Known Development Roadmap

Since MPEG-4 development occurs in a public, international forum within established procedures, its development process is known and can be monitored by manufacturers and operators to aid their planning and development. Indeed, the breadth of the MPEG-4 standard offers both excellent performance today and a long-term framework for the future, without the switching costs or technology locks incurred with proprietary formats.

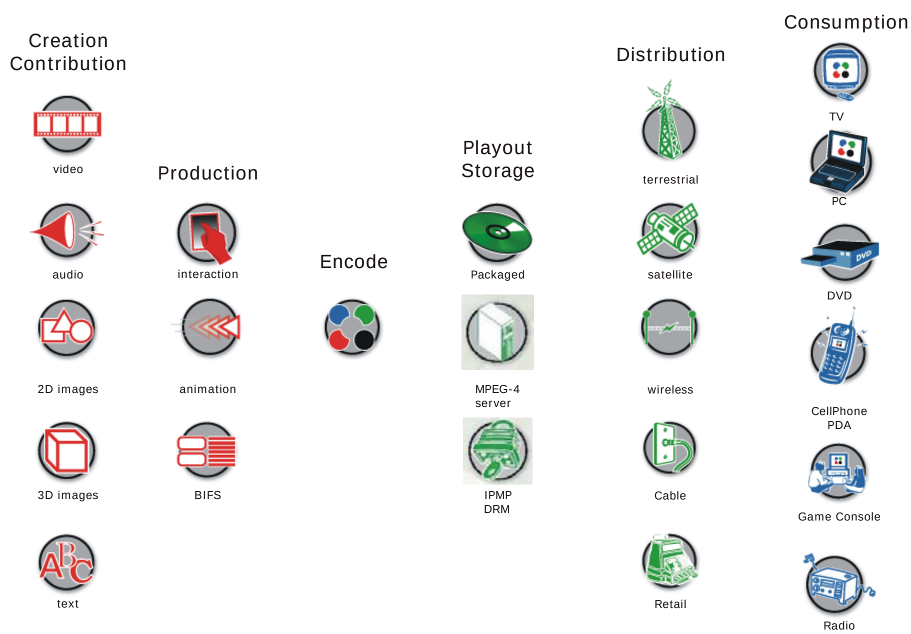

Figure 5. The MPEG-4 ecosystem liberates multimedia for delivery across any network to any user of any device

Encode Once, Play Anywhere

MPEG-4 is truly a multi-platform standard. It operates as easily on embedded hardware as it does in software, and it is designed to be independent of any transport medium, working as well over wireless networks as on legacy MPEG-2 transport streams. This means content providers can avoid duplicate production for different delivery media by encoding it once in MPEG-4, instead of in several delivery formats.

Flexible Integration with Transport Networks

MPEG-4 is designed to be transport independent, operating over both IP-based and traditional MPEG-2 transport stream networks. Thus, it can be easily transmitted over the Internet to PCs, or over MPEG-2 cable or satellite networks to set-top boxes. MPEG, MPEGIF, IETF, DVB, ISMA, and other organizations have developed standards for the carriage of MPEG-4 content

Further, MPEG-4 has been designed from the beginning to operate over lossy networks such as mobile 3G or WiFi, and incorporates scalability, error concealment, and error recovery techniques in its codecs so that program quality is preserved when changes in bandwidth or signal dropouts occur.

Popular transport protocols for MPEG-4 include:

- The MPEG-4 File Format. This is the standardized way to store MPEG-4 content as computer files, and includes MPEG-4's hinting mechanism to pre-packetize content to increase video server efficiency.
- The IETF RTP Protocols. This is a family of real-time streaming protocols standardized by Internet RFCs that allow MPEG-4 to be streamed over IP networks.
- MPEG-4 over MPEG-2 Transport Streams. This allows MPEG-4 content to be combined with MPEG-2 programming, if desired, and carried over existing cable TV or satellite networks and equipment.
- Audio-only Protocols. MPEG-4 also includes several protocols such as LATM/LOAS, ADIF, and ADTS that are popular for carrying audio-only programming.

Established Terms and Venues for Patent Licensing

Audio and video coding is a field where patents have been asserted and licensed for decades, and many organizations have substantial IP portfolios that apply to media codecs. The use of most MPEG codecs has involved the payment of royalties through patent pools to patent holders. With MPEG-4's patent pool organizations, a manufacturer or operator can be assured that he has a license to essential patents from the major IP holders with fixed, renewable prices and conditions.

Though patent pools and royalties have been attacked as a disadvantage of MPEG-4, it is actually an advantage compared to the patent uncertainty surrounding proprietary or "royalty free" media codecs, where patents from other organizations could be later asserted.

Though there is always a chance that an unknown inventor may someday assert a patent against an MPEG-4 manufacturer or user, it is more likely the inventor will join an established pool where the costs of licensing and royalty collection can be shared with many other licensors. His royalties would be distributed as a share of the pool's, and a user's royalties would be unchanged.

MPEG-4 Markets and Applications

Television Broadcasting

MPEG-2 is the current standard for digital television production and distribution, but doesn't offer enough compression for transmitting the hundreds of channels cable and satellite TV consumers will expect in high-definition. As more programming moves to HDTV, cable and particularly satellite distribution systems will begin using more efficient compression codecs such as AVC. Already, major satellite TV operators such as DirecTV, bSkyb, and Premiere have announced new high-definition services using AVC.

IP-based Television Distribution

With the rise in availability of DSL and other broadband IP networks, IP-based television distribution is now an alternative to cable, satellite, or over-the-air broadcasting. Because IP TV is inherently two-way, improved interactive and video-on-demand services are easily provided. But the limited bandwidth of many of these networks requires high-efficiency coding such as AVC to allow for multiple sets to be served over a single connection. This is also a natural application for AVC.

Thousands of AVC encoders have already been deployed in the field by both established and new encoder suppliers and a number of system operators are undertaking extensive trials with a view to launching services soon, most notably UK-based Video Networks Ltd., which has recently launched the first full revenue-bearing service based on AVC.

Figure 6. Sony's PlayStation Portable uses MPEG-4 AAC and AVC. Courtesy Sony.

Portable Gaming

Gaming devices are typically closed systems, where manufacturers are free to use whatever proprietary technologies meet their needs. Notably, Sony's new portable gaming platform, the PlayStation Portable, uses MPEG-4 Simple Profile to view video stored on memory cards, and AVC for video playback from its internal UMD optical disc drive, along with AAC audio encoding.

Figure 7. 3G Mobile Handset with MPEG-4 Two-Way Video Capability. Courtesy NTT DoCoMo

Mobile Communication and Entertainment

MPEG-4 Simple Profile video has been a part of international 3G mobile standards since videophone service was introduced in 2001. Today, mobile operators such as NTT DoCoMo and Hutchinson-Whampoa have deployed millions of MPEG-4 handsets to enable two-way video calls or watch video programming over 3G networks.

Additionally, MPEG-4 HE-AAC audio coding has been employed in several music download services, such from operators KDDI and Orange, and HE-AAC v2 with the Parametric Stereo tool has been selected by 3GPP as a new standard for music streaming.

The next phase of mobile entertainment is broadcasting to handsets, as opposed to point-to-point connections, to reduce the bandwidth and expense required. MPEG-4's AVC is a part of all three of the major systems being deployed for this service: DVB-H, DMB, and MediaFlo.

Internet Streaming

MPEG-4 was designed from the start for streaming, and it is now featured in systems from several manufacturers. Apple supports MPEG-4 Simple Profile video and AAC audio in its QuickTime platform and now includes support for AVC in QuickTime 7. Real Networks supports decoding of MPEG-4 content, and the popular DivX codec is also MPEG-4 compliant. A number of MPEG-4 vendors offer plug-ins for Microsoft's Windows Media Player that enable users to also watch MPEG-4 content in this player.

Figure 8. iPod portable music player using MPEG-4 AAC audio coding. Courtesy Apple.

Apple supports MPEG-4 natively in their QuickTime Streaming Server and Darwin Streaming Servers, available for free download or as source code, and in its new QuickTime 7 encoder and player. At the time of writing, QuickTime 6 has been downloaded one billion times since its introduction in summer 2002. Real Networks also supports MPEG-4 on its Helix servers.

Packaged Media

As industry consortia prepare standards for the next generation of optical discs for video distribution, the performance of AVC has led to its adoption in both of the leading disc formats. AVC is a codec specified for both the HD-DVD and Blu-Ray standards.

Video Conferencing

Video conference terminals are a natural application for MPEG-4's AVC codec, and recently market leaders Polycom, Tandberg, and Sony have all introduced AVC support in their products. AVC will enable increased video quality over the same connection compared to the earlier H.261 and H.263 standards, or a half-bandwidth connection may be used for similar quality.

Home Networking and PVRs

Though traditional standard-definition cable and satellite TV is likely to remain MPEG-2 based, given the huge installed base of MPEG-2 receivers, the consumer's home is moving from islands of digital video, as with current set-top boxes and personal video recorders, to interconnected home networks. In the future, consumers will be able to seamlessly switch from watching programming on their traditional TV to viewing it on a phone or PDA device, or downloading it to their car. Given the bandwidth constraints of home wireless networks and devices, content will likely be transcoded to new codecs such as MPEG-4 AVC when sent around the home. Standards consortia such as DVB and DLNA have adopted AVC for this purpose, as well as for future IP network delivery.

Digital Still Cameras and Convergence Devices

Digital still cameras now include "movie modes" for capture of short video sequences, and with the new affordability of high-capacity flash memory, it is possible to build camera-like Mobile Content Convergence Devices that include the functions of a still camera, camcorder, and music player in one device. Given its compression efficiency and multi-platform support, plus its freedom from platform bundling requirements, MPEG-4 is an ideal fit for these devices, and manufacturers Panasonic, Sanyo, Samsung, Audiovox, and Archos, among others, have all introduced portable convergence devices using MPEG-4.

Figure 9. Digital Convergence Device using MPEG-4 for flash-memory based recording. Courtesy Panasonic.

Satellite Radio

MPEG-4's HE-AAC audio codec, as well as its AAC and BSAC audio codecs, have been employed in several systems for satellite radio and multimedia broadcasting. XM Radio employs HE-AAC coding, while Digital Radio Mondiale employs HE-AAC as well as the MPEG-4 speech codecs CELP and HVXC. Korea's DMB system will employ AAC and BSAC audio coding and AVC video.

Security

Traditional video surveillance systems have employed time-lapse video recorders and more recently small Digital Video Recorders using some variant of motion JPEG or wavelet compression. These devices often must limit the video resolution and frame rate to provide a reasonable recording time, and they require proprietary video players or browser plug-ins for users

Figure 10. Combination CDMA2000 mobile handset and receiver for Satellite DMB service using MPEG-4 AVC and AAC. Courtesy Samsung.

to view stored content. Additionally, users desire better resolution and frame rates to examine security alarms or events.

Thus, several security video firms have begun offering improved video systems with recording of MPEG-4 video at full D1 resolution and frame rates of 25 or 30 frames per second. MPEG-4 coding dramatically reduces the storage costs: A typical installation with 3000 cameras might use 800 TB of storage to support 15 days of recording. MPEG-4's interoperability also allows users to combine equipment from different manufacturers in their systems and still be able to export the video in a universally readable format.

Figure 11. MPEG-4 coding enables this casino to store weeks of video with better quality for forensic analysis. Courtesy Pelco.

MPEG-4 and Related Standards

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

The Emmy Award-winning MPEG committee has built the foundations of digital content delivery with its highly successful standards since 1991 with the release of the MPEG-1 standard. Today, close to a billion devices support one of MPEG's standards.

MPEG-1

MPEG-1 was originally intended to enable video to be compressed for storage on optical disks of the era, and led to the Video CD format, hugely popular in Asia for video content distribution. MPEG-1 was also used in early video-on-demand systems and interactive media and has become almost universally supported by PC media players. MPEG-1 audio offers Layer 2 coding, used in the DAB digital radio standard, and Layer 3 coding, the MP3 format.

MPEG-2

MPEG-2 is the audiovisual standard most widely used for entertainment video applications. MPEG-2 enables digital television and DVDs, with hundreds of millions of 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 video decoding and encoding are more CPU-intensive than for MPEG-1.

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-4

The subject of this paper, MPEG-4 is the successor standard to MPEG-2, extending its application to IP-based and lossy networks, to rich multimedia presentations, and to objects and interactivity. MPEG-4 also has improvements in the core audio and video codecs over MPEG-2, notably the AVC, AAC, and HE-AAC codecs which have enabled so many new products and services.

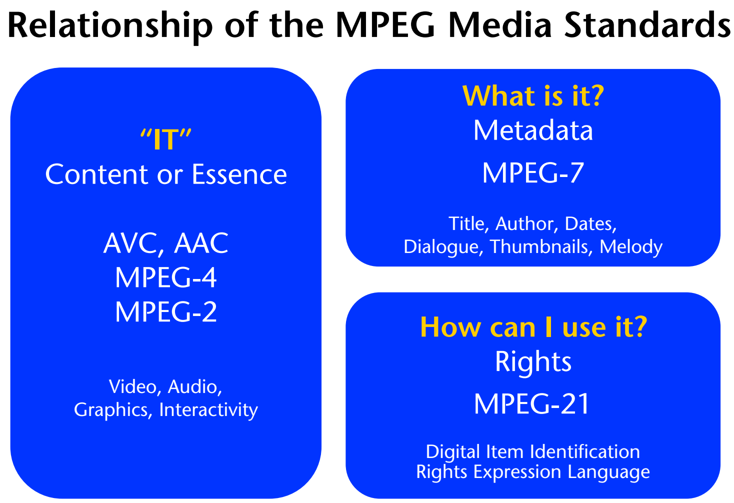

Figure 12. The Metadata and Rights Standards MPEG-7 and MPEG-21 are designed to complement coding standards MPEG-1, -2, and -4. Courtesy Streamcrest Associates.

MPEG-7 and MPEG-21

The related standards 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 stored locally, on-line, and in broadcasts.

For example, an overheard song can be captured by a mobile handset and converted locally to an MPEG-7 description. This description can then be transmitted to a search engine which offers a download of the song. 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 with fully interoperable coding there are still additional steps to insure that all the features of different networks work together. MPEG-21 is a framework which allows interoperability and portability of content.

Related Standards

Internet Engineering Task Force (IETF)

The Internet Engineering Task Force (IETF) 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 the Real Time Streaming Protocol (RTSP).

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

The 3rd Generation Partnership Project (3GPP) defines standards for 3rd generation mobile networks and services evolving from GSM-based systems. 3GPP2 does the same for CDMA-based systems. Both give 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. 3GPP has also recently adopted HE-AAC v2 for music delivery.

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 Are Profiles and Levels?"), and augmented these with IETF transport specifications to create cross-vendor interoperability for video on the Internet. ISMA uses the MPEG-4 file format for file storage, and the IETF protocols RTP and RTSP for streaming protocols and control.

There are two ISMA specifications: ISMA 1.0 is mature, conformance testing is in place, and ISMA-compliant implementations are available. It uses MPEG-4 Simple Visual and Advanced Simple profiles for video, and MPEG-4 High Quality Audio Profile for audio.

ISMA 2.0 updates the codec suite to use Advanced Video Coding (MPEG-4 Part 10 AVC or H.264) Baseline, Main, and High video profiles, and High-Efficiency AAC (HE-AAC) audio. Implementations are emerging and conformance testing is available.

The DVB Project

DVB's main transmission standards, DVB-S for Satellite, DVB-C for Cable and DVB-T for terrestrial, are used worldwide for television transmission and are the basis for many alternative standards. While DVB has been historically based on

MPEG-2 transport streams and video coding, DVB has recently adopted MPEG-4 AVC and HE-AAC codecs for future systems. MPEG-4's rich media framework is also closely related to DVB's Multimedia Home Platform (MHP) standard for digital interactive television middleware.

The MPEGIF Interoperability and Qualification Programs

MPEGIF operates two programs to aid members in establishing interoperability of their MPEG-4 implementations.

The **Interoperability Program** has been active since 2000 and currently conducts three test rounds per year. It is intended as an informal, confidential program where members can test their initial designs and implementations, and resolve interoperability problems.

The program organizes and executes interoperability tests around MPEG technologies, such as audio codecs, video codecs, file format, transport, and other systems elements. The activities of the program are worked out by consensus to the benefit of all members. For example, new features and test points are regularly added based on the interest of the companies who participate in the tests.

The main work of the group is done through conference calls and a private reflector. Typically, calls are held every two weeks during an active test round. Calls may be held less frequently or not at all during the periods between active test rounds.

The **Logo Qualification Program** was established in 2004 and offers members a procedure by which companies may qualify encoder & decoder products and earn the right to label their products with a logo provided by MPEGIF. MPEGIF also operates a website with a searchable directory of Qualified Products. The program is designed to meet the following goals:

- Promote “MPEG-4 Brand Awareness”
- Provide marketing value for MPEGIF members who participate in the program.

- Encourage participation in the MPEGIF Interoperability Working Group, and in the Forum in general.
- Provide an environment for companies to discover and resolve interoperability problems.
- Encourage growth of the industry by building confidence that products will interoperate.

The Logo Qualification Program is tightly linked to the MPEGIF Interoperability Testing program, and in fact participating companies are required to complete at least one round of testing in the Interoperability WG before submitting their product for qualification. An additional verification step is also required:

- Decoder companies must verify that the product correctly decodes all streams in the relevant Qualification Stream Set. This set is a combination of published conformance streams and streams posted to and tested within the Interoperability WG.
- Encoder companies must obtain Qualification Endorsements from three other companies with qualified decoder products. The endorsement is a statement from the decoder company that a relevant stream set generated by the encoder company was decodable without problems.

Note that this is essentially a Self-Qualification process. In order to complete the qualification procedure, a company will sign an agreement stating that the testing criteria have been met. The program stops short of full "Certification," which normally would require an independent third party to administer certification tests and make judgments on whether the product passes or fails. In this way MPEGIF is providing a framework and encouraging a base level of testing, however the details of testing and verification are the responsibility of the individual company, not that of MPEGIF.

The program will initially focus on MPEG-4 Audio and Video encoders/decoders, but may eventually be expanded to include other MPEG standards. Support for Profile/Level combinations is designed to be flexible in that new combinations may be added at any time based on interest from the membership, and also based on the evolution of testing in the Interoperability WG. Initially, the program includes AVC Main, Baseline, and Extended visual and High Efficiency AAC audio profiles.

The Qualified Products Directory may be searched, and additional information on the program found at http://logo.mpegif.org.

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.mpegif.org/resources.php for access to this document.

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 fundamentals of MPEG-4 are described by the parts on Systems (part 1), Visual (part 2) and Audio (part 3), along with the new part 10 on Advanced Video Coding. DMIF (Delivery Multimedia Integration Framework, part 6) defines an interface between an application and the network or storage. Conformance (part 4) defines how to test an MPEG-4 implementation, and part 5 offers a significant body of example Reference Software, that can be used to start implementing the standard.

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)

Other parts in MPEG-4 include:

- 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 Sockets
- Part 11: Scene Description (BIFS) and Application Engine (MPEG-J)
- Part 12: ISO Base 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 eXtension)
- Part 17: Streaming Text Format
- Part 18: Font Compression and Streaming
- Part 19: Synthesized Texture Streaming
- Part 20: Lightweight Application Scene Representation (Laser)
- Part 21: MPEG-J Graphical Framework eXtension (GFX)
- Part 22: Open Font Format

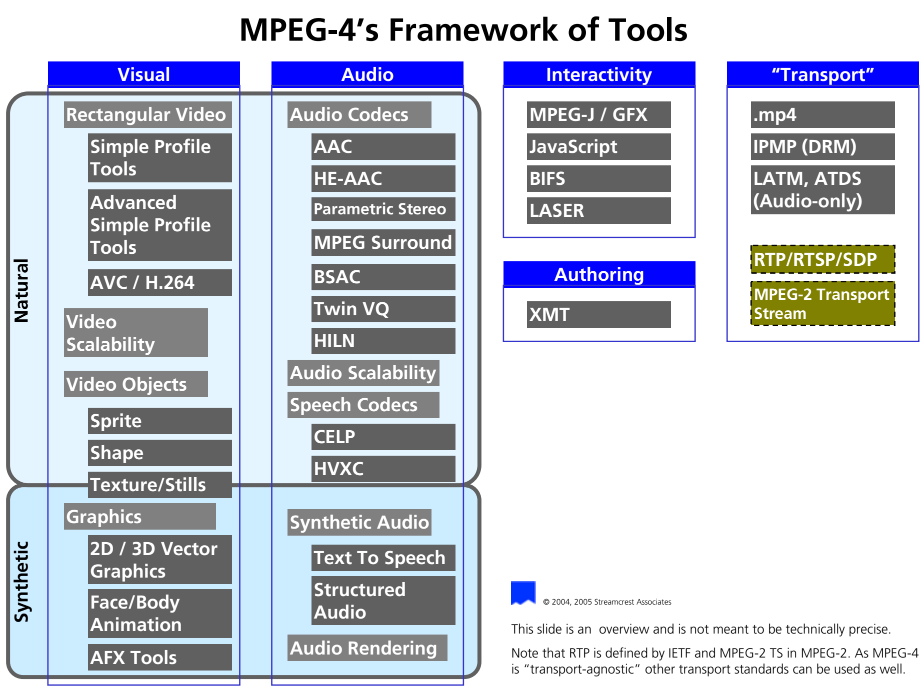

Figure 13. This is merely one way to classify MPEG-4's toolset. Courtesy Streamcrest Associates.

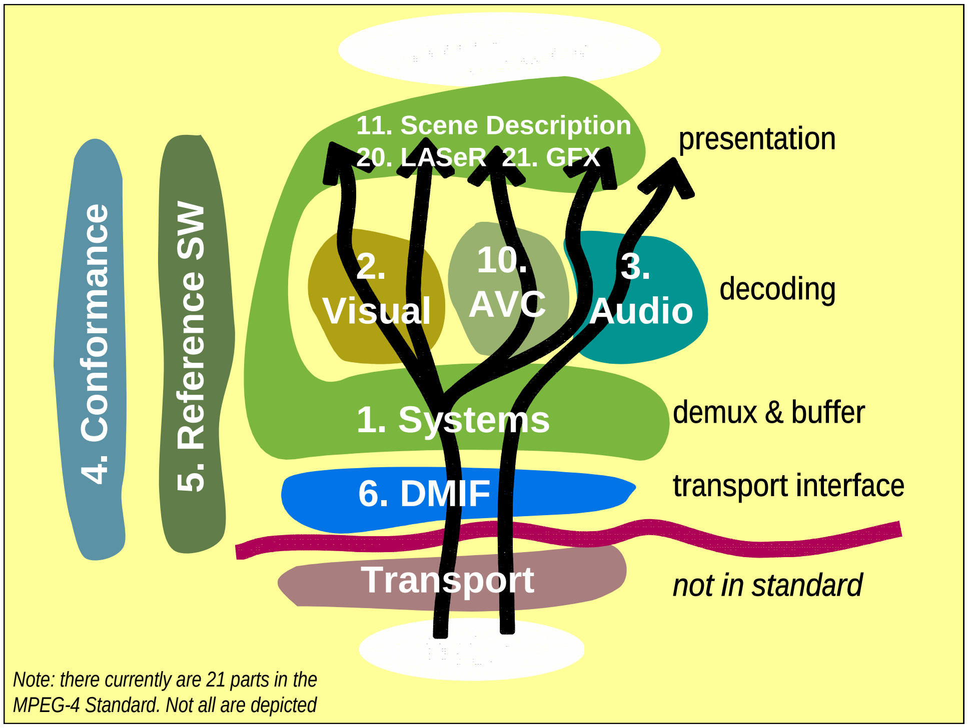

Figure 14. The parts of MPEG-4. The arrows represent the flow of bits through the MPEG-4 system.

What are Profiles and Levels?

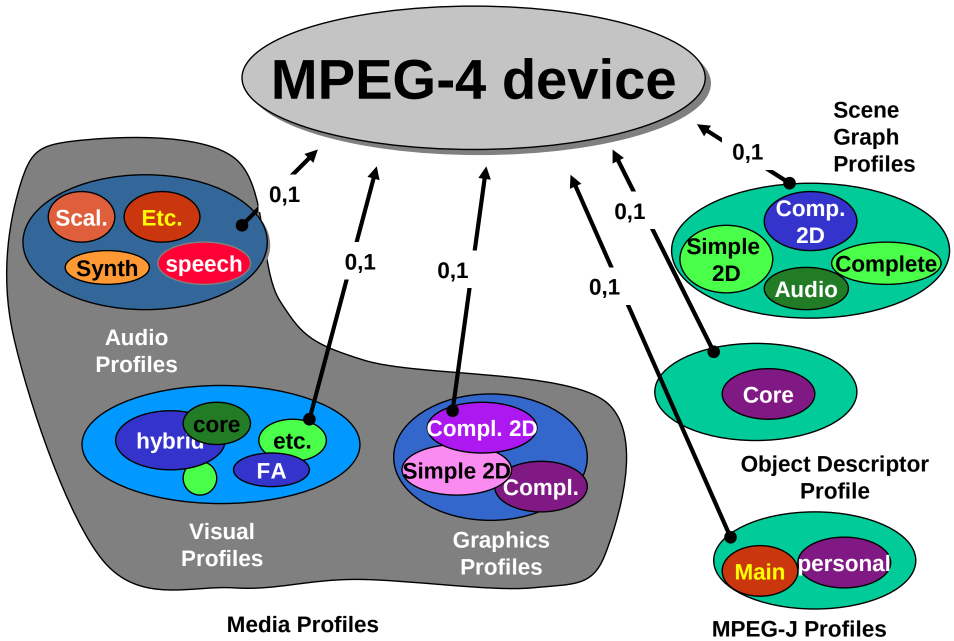

Figure 15. Profiles and Levels

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.

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.

Common Video Profiles	
Profile	Features
Simple Visual	Similar to MPEG-2 coding
Advanced Simple Visual	Adds support for B-frames, Global Motion Compensation, Interlace (at levels 4-5)
Core Visual	Adds support for binary shapes (video objects) and B-frames
AVC Baseline	Low Delay, Lower Processor Load
AVC Main	Supports Interlaced video, B-Frames, CABAC encoding
AVC Extended	Includes Error Resilience Tools, B-Frames
AVC High	Supports High-quality, High-resolution formats for Digital Cinema

Common Audio Profiles	
Profile	Features
High-Quality Audio	Includes the most popular AAC object type, AAC Low Complexity, and the CELP speech coder
Low Delay Audio	A variant of the AAC codec with ~20mS delay, suitable for high-quality conferences or conversations
High Efficiency AAC Audio	Adds Spectral Band Replication tool to improve coding efficiency at low bitrates
HE-AAC v2 Audio	Adds Parametric Stereo Tool to further improve coding efficiency at low bitrates

Table 2. Common MPEG-4 Profiles

MPEG-4's Rich Multimedia Framework

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

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

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 doing things in response to an event, usually user input. This makes the interactivity of MPEG-4 rich multimedia possible. 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. For applications that need more complex logic in response to an event, ECMAScript can be used or the Java language via MPEG-J.

Recently, MPEG has redefined the scene description based on W3C's SVG instead of VRML (Virtual Reality Modeling Language) that BIFS was based on. This lightweight scene representation - called Laser – is geared at 2D applications on limited resources devices such as mobile handsets.

BIFS and Laser are descriptions of a scene to compose audio-visual objects. As all scene descriptions, they only contain a limited number of features and sometimes features specific to some applications. To allow content creators more freedom in the composition and logic of their applications, a programming language is necessary. The Java-based Graphical eXtension Framework (GFX) provides a programmatic way to compose and to render audio-visual objects. GFX was designed around mobile entertainment applications such as 3D games enhanced with video.

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.

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 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 recognize 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.

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 was not present before and those that improve existing functionality.

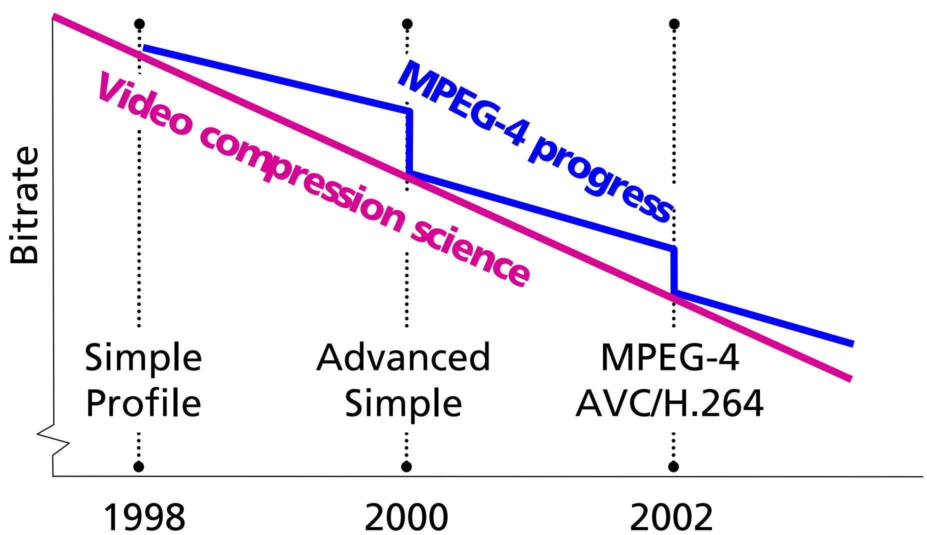

Figure 16. MPEG-4 - A predictable, responsible upgrade strategy

Examples of the first kind are support for fonts, synthesized textures and the Animation Framework eXtension (AFX). Examples of the latter are MPEG-4 Advanced Video Coding (AVC), Lightweight Scene Representation (Laser), and MPEG-J Graphical Framework eXtension (GFX).

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.

An example of this upgrade strategy is MPEG-4 Advanced Audio Coding (AAC). While AAC offers excellent quality at bitrates of 64 Kb/s per channel and higher, MPEG has extended it with a technique called Spectral Bandwidth Replication, which gives spectacular bandwidth savings for applications like Internet audio and digital broadcasting. MPEG-4 AAC with SBR, known as High Efficiency AAC, or HE-AAC, can deliver high quality stereo audio at a mere 48 Kb/s. The SBR extension is both forward and backward compatible: an existing MPEG-4 AAC decoder can decode a signal without using the technique, while a decoder with SBR uses the extended signal to enhance the upper octave of the signal.

For digital audio broadcasting, MPEG-4 AAC is becoming the codec of choice. Satellite-based XM Radio uses HE-AAC, as does terrestrial Digital Radio Mondiale.

MPEG-4 Audio is also 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.

Clarifying Common Questions

Who licenses MPEG-4 technology?

The Motion Picture Experts Group of the international standards bodies ISO and IEC develops and publishes MPEG standards. MPEG requires that companies proposing technologies for MPEG standards to commit to licensing their patents on Reasonable and Non-Discriminatory Terms and Conditions (also called "RAND"). Other than the costs of publication, there are no fees for using MPEG standards themselves, and neither MPEG nor MPEGIF is involved in patent licensing.

MPEG	MPEGIF	Patent Pools	Manufacturers
Develops MPEG Standards	Promotes MPEG standards Operates Interoperability and conformance programs	License Patents Essential to MPEG standards	Sell or license their implementations (systems, chips, or software) of MPEG standards

Table 3. Roles of Organizations in the IP Licensing Process

Patent pool organizations provide a single convenient point for licensing patents that are essential to implementing one of the MPEG standards. They operate independently of MPEG and MPEGIF, and each sets its own terms and royalties for patent licensing. Currently, MPEGIF is aware of two firms that operate pools for MPEG-4 patents: MPEG-LA and Via Licensing.

MPEG-4 Patent Pools		
Section of standard	Topic	Patent Pools
Part 1	Systems	MPEG-LA
Part 2	Video	MPEG-LA
Part 3	Audio	Via Licensing
Part 10	Advanced Video	MPEG-LA, Via Licensing

Table 4. Current MPEG-4 Patent Pool Consortia

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

These joint licensing schemes are not carried out on behalf of ISO, MPEG or MPEGIF, 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 MPEGIF in licensing?

The MPEG 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. MPEGIF has acted as a catalyst, promoting the use of patent pools. MPEGIF has among its members licensors, licensees, and other entities which have an interest in fair and reasonable licensing. MPEGIF continues to monitor terms and conditions of licensing, and to offer forums where parties can exchange information and opinions on licensing

With the release of Part 10 AVC, is Part 2 video coding obsolete?

AVC is currently MPEG-4's highest performance video codec, offering the lowest bitrate for a given quality among any MPEG or proprietary codecs, including MPEG-4 Part2.

MPEG-4 AVC 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.

In applications where coding efficiency is paramount, such as HDTV satellite broadcasting, the performance needed means Part 2 video, as well as MPEG-2 and other codecs, are not suitable. Part 2 will likely continue to be used in applications where power consumption is critical or bandwidth is less so, as in digital still cameras or webcams. Part 2 codecs will also remain important for object-coded rich multimedia systems where AVC is not yet supported.

Additionally, one must remember that Part 2 video is a part of several deployed standards, such as mobile videophony. The installed base of 3G mobile video handsets means that Part 2 video encoders will continue to be improved and future handsets will support Part 2 for interoperability. Standards continue to flourish for many years after their introduction – with MPEG-2 we are likely just now reaching the peak of annual production of MPEG-2 encoders and decoders, a decade after its release.

What is the relationship between MPEG-4 Visual and the DivX codec?

DivX5 is an implementation of MPEG-4 Advanced Simple Visual Profile. DivX Networks 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 previous versions of its Windows Media platform. Explicit support for MPEG-4 was removed from Windows Media several years ago. It is unknown to what extent the current version of Windows Media uses MPEG-4 concepts internally. Some developers will know of Microsoft's contribution to the MPEG-4 Reference Software, one of the two implementations of the part 2 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 mobile phone, PDA, or camera devices. In many of these devices, available memory and power supply requires a hardware decoder. Also, it is important to realize that standards such as MPEG-4 are not just about stable decoders that can be implemented in hardware. They also allow interoperability among all implementations, whether in software or hardware.

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. QuickTime 7 added support for MPEG4 Part 10 (AVC/H.264) on April 28, 2005.

I read a benchmark of MPEG-4 where it did poorly, how can you claim it is higher performance?

One of the virtues of MPEG-4 is that it is an open standard that can be implemented by anyone. While there are conformance and interoperability programs run by MPEGIF and other organizations to insure that products work together and correctly implement the standard, they are voluntary, unlike the tight controls imposed by proprietary codec licensing. So it is possible to test an immature or poor implementation of MPEG-4 against the latest proprietary codec, or even a mature MPEG-2 codec, and have MPEG-4 fair poorly.

One of the advantages of MPEG-4 is that encoders from different manufacturers will be of different quality, yet all will be compliant to the standard and decoded by any decoder. This variety means a user or system manufacturer can pick an implementation that has the cost and performance needed in his application, and he can switch encoders without making changes to his installed base of decoders or re-encoding existing content.

Benchmark tests are also complex to perform correctly because of the biases, fatigue, and other psychological effects of the observers. A viewing jury in a test conducted according to ITU R-500 sees a controlled set of clips that they rank without knowing which codec is being used, and the test clips are carefully chosen so there is a mix of motion, detail, and other features so the test represents all types of scenes that would be encoded. In less formal benchmarks, particularly those carried out by journalists or enthusiasts, these effects are often not considered correctly.

How does MPEG-4 Compare to Other Internet Media Formats?

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.

Table 5. Comparing MPEG-4 to Other Internet Media Formats

	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	Content access control
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	Yes

How Will MPEG-4 Be Used in Interactive TV?

Along with significantly less bandwidth for the same quality, the native support for interactivity is a key difference between MPEG-4 and the MPEG-2 technology broadly deployed in current 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 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' using 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.

What is the Difference between 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).

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. DVB is traditionally MPEG-2 based, but MPEG-4 is seen as a logical evolution, and one which will be more efficient when DVB services are to be delivered over IP. To this goal, DVB has included MPEG-4 Main Profile AVC and HE-AAC in its latest revision.

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.

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, and 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.

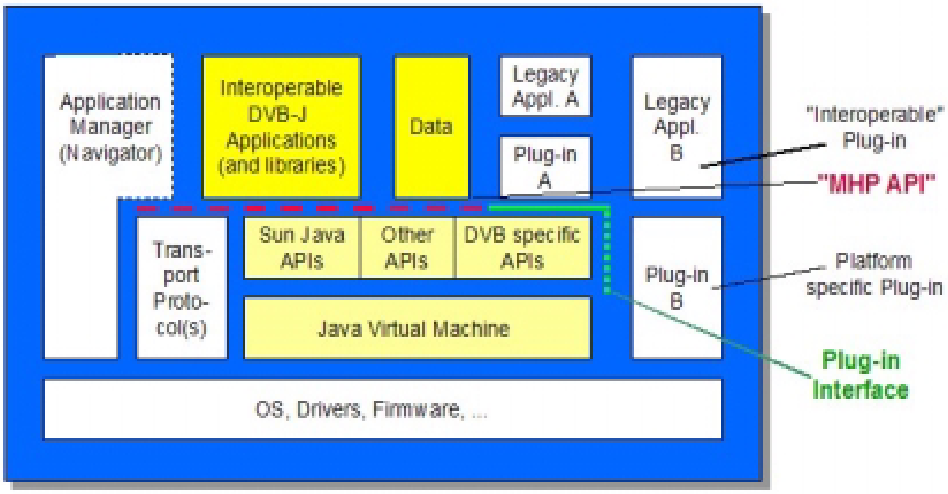

Figure 17. MHP reference architecture showing role of DVB-J

How Does MPEG-4 Compare to SMIL and SVG?

SMIL is the Synchronized Multimedia Integration Language of the W3C and SVG is W3C's Scalable Vector Graphics specification. 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.

Table 6. Comparing MPEG-4 to SMIL and SVG

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 (which is an XML extension of the older Virtual Reality Modeling Language (VRML)). MPEG-4's scene description model is based on the textual VRML language, to which MPEG added streaming behavior, 2D support and a binary representation for efficiency.

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.

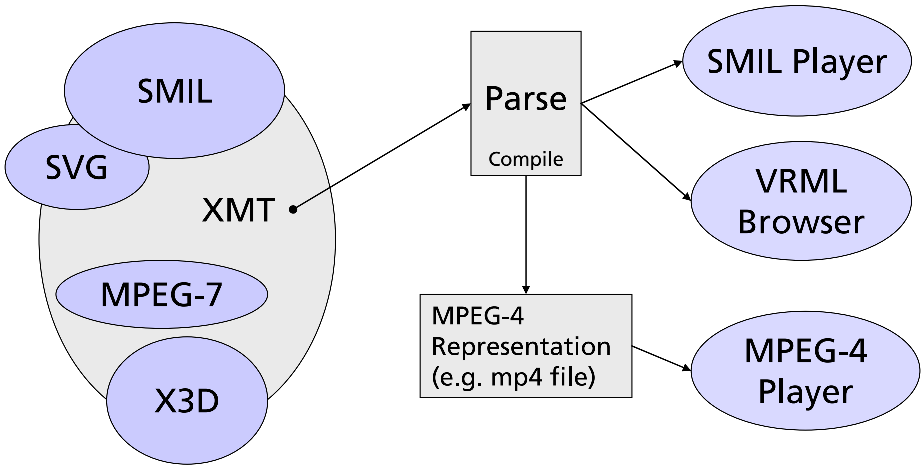

Figure 18. Relation of MPEG-4 XMT to other specifications.

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 Industry Forum

The MPEG Industry Forum represents more than 80 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.

MPEGIF 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 promote wide adoption of MPEG-4, MPEGIF picks up where MPEG stops.

The following is a list of MPEGIF's current activities:

- Promoting the emerging MPEG standards (MPEG-4, MPEG-7 and MPEG-21), and serving as a single point of information on technology, products and services for these standards;
- Carrying out interoperability tests, which lead to an ecosystem of interoperable products. Over 30 companies have tested their products in MPEGIF's MPEG-4 interop program;
- Developing and establishing an MPEG-4 Certification program, which comes with the right to carry MPEGIF's MP4 logo;
- Organization of and participation in many trade show events - MPEGIF has show floor presence together with some of its members, at shows such as NAB and IBC. In many other shows we organize panels and presentations.
- Organization of MPEG exhibitions and tutorials. MPEGIF has organized several Workshops and Exhibitions on MPEG-4, such as in Geneva (2000) and San Jose (June 2001 and June 2002);
- Establishing a forum for discussions that led to the formation of independent patent pools for licensing MPEG-4 patents on fair terms.

MPEGIF's website has a wealth of information on the MPEG-4 standard and is starting to collect information on MPEG-7 and MPEG-21. It has many links to external resources, and is updated daily with latest relevant news and press releases. Through the website, anyone can sign up for MPEGIF's News, Discussion, and Technology mailing lists.

Join the forum

Membership in MPEGIF 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 companies at the leading edge of MPEG-4. For the price of a one-page advertisement in a trade magazine, you may join the forum and might meet your next customer there.

Help Drive Success

MPEGIF 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 MPEGIF, or to find out what activities will benefit your company, visit: http://www.mpegif.org.

mpeg
Industry Forum