WHAT IS MPEG-4

Version 1.0 (11-September-2005)

By: Muhammad Imran (Islamabad, Pakistan)
i-m-r-a-n-1@gmail.com (without dashes i.e. imran1). to avoid spam search engines

Dedicated to "**Nazia Hassan"** & "**Zoheb Hassan"** Pop Legends of Pakistan

Preface

I caught site of this format in early 2004 and started learning about it and in few months I have quite in-depth knowledge about it. In mid of 2004 I thought why not I publish my learning for the betterment of people because when this format will hit the streets there should be some consolidated information about it. So I started collecting the information again and this was not simple as I was simply crazy to hunt every single piece of information I had learned on internet and recalling those pages and searching them back was very irritating because when u know u had seen a thing and it's a crucial piece of content but u just cant find that webpage again. So, for some pages it took several weeks to dig them again and put it here. Second thing was to collect all the material and compose in a proper flow not just dumping every thing in a word file.

When all this was done and the final shape of this document started appearing in Late 2004 my desperation started to overflow to find a code for the MP4 encoding. There were only three softwares and each of them following some type MP4 encoding scheme ("Profile" discussed later) and I was just stuck to encode in H264 (MP4-Base10) and no less than that. These softwares were Discret's Cleaner-XL, MainConcept and another tool Expert-H264 by PixelTools. Well Later two were incapable of producing files customizable for low bandwidths and were only to reduce file on some fix predefined MP4 profile Or Non Downloadable i.e. PixelTools. Cleaner-XL was that time king of the hill allowing good control on various level of MP4 but still not capable of H.264. Here I must say that the delay MainConcept caused for me was the most hurting because they were the most enthusiasts about it and the had a beta coder ready which was a limited version having many unbearable limitations so I waited and waited on false hopes that soon they would come out of beta and release some final products and this was longest wait of my life that they atleast didn't come out of beta for next 8 months and left to pursue them, along that my project also went in cold storage and I never got time gain to resume work on it.

The only Players were Quicktime and RealPlayer again both were incapable of playing H264, infect no player was available on internet who could recognize H264 so it would be true to say that H264 was a secret weapon which was known to Labs and being perfected by them. Infect its true because until 2004 H264 was under development and perfection.

So for more than year this document you are about to read remained unpublished because I had to make H264 video of Nazia-Zoheb and dedicate this research to them. But I just couldn't do this because of these hurdles. Nazia-Zoheb introduced pop in south asia because they were not inspired by stupid filthy indian culture but more by British and American that's why it made them true pop singer's which led them to set a yet unbroken record of highest sales in Asia, "Dum Dum Dee Dee" 10+ times consecutively at top in Japanese pop charts and several other achievements which need a separate whitepaper to describe.

So it was not until two months back that I purchased SonyEricsson-P900 smart phone which records movie in MP4. So this sparked an interest again in me to resume my project.

So here I am on 11-September-2005, Today I looked briefly on the internet and I first went MainConcept site and I was shunned to see that there site was saying "MainConcept has finally released the new version of their latest development in video codec technology: the H.264 Encoder v1.0 for Microsoft® Windows®!" http://www.mainconcept.com/products.shtml Hell man u started work on it in late 2003 and 2 years to reach from "H.264 Encoder v1.0 Beta" to "H.264 Encoder v1.0 Full" so I hit them at right time after 2 years… height of coincidence… or what.

Currently the best encoders available to my knowledge are…

1. Sorenson Squeeze 4
2. MainConcept
3. Discret Cleaner XL 6
4. Adobe Premiere Elements 1 (no MP4 encoder or plugin available by default)
5. Pixel Tools (current status unknown)

Media Players are…

1. Quicktime 6.5.1 and 7
2. RealPlayer (RealOne)

3. Mediaplayer 9 and 10 (with envivio or other third party plugin)

Webites for resources and info for MP4….
1. www.doom9.net, www.doom9.com, www.doom9.org
2. http://forum.doom9.org (Doom9's Forum > MPEG-4 and Co. > New A/V Formats - Codecs > MPEG-4 Information (including AVC/H.264))
3. www.apple.com (http://www.apple.com/macosx/tiger/h264.html)
4. http://www.pixeltools.com/h264_paper.html
5. http://www.mainconcept.com/h264_encoder.shtml

Now I think time has passed to publish Nazia-Zoheb MP4 H264 video on website because the format has become common to people but H264 is still not. But people will take it still an MP4 and another problem is that even if they will go on internet to see it, no player is still capable to play it neither Mediaplayer10, Quicktime6.5.1, RealPlayer10 or others so what would they see. I believe Quicktime-7 just released few day s back on internet would be capable of playing it and I believe this because Apple claims that MP4 in developed on the foundations of their Quicktime format or we can say MP4 is a derivation of quicktime. But Microsoft has dismissed this claim in a document which can be found on their site, I think its present in FAQs of their WMV-9 format. Now back to publishing Nazia-Zoheb video, so people have to download Quicktime7 and ….. I think I should just publish this document and u should should imagine that if there would have been any first H264 video from Pakistan or again south-asia it is "Pyaar ka Jaadu" (inspired by Michael Jackson's video for "Billie Jean") by Nazia (Late) and Zoheb and you would have watched it in Q2 of 2004.

So ladies and gentlemen do ahead and read this research dedicated to “Nazia Hassan (Late)” and “Zoheb Hassan” made in Pakistan’s. POP pioneers and legend in south asia.

If you have any corrections, comments, questions or even if u need my downloaded/saved resources (webpages) or you need pictures of nazia-zoheb from childhood to now or their web-resources, fanclubs etc please free to write me at i-m-r-a-n-1@gmail.com (without dashes i.e. imran1). I have put dashes to avoid spam search engines. Also please make the email subject meaningful so I won’t mistaken it as spam. Do not write at mi1400 yahoo its my junkyard.

This Literature is constructed as follows

WHAT IS MPEG-4?

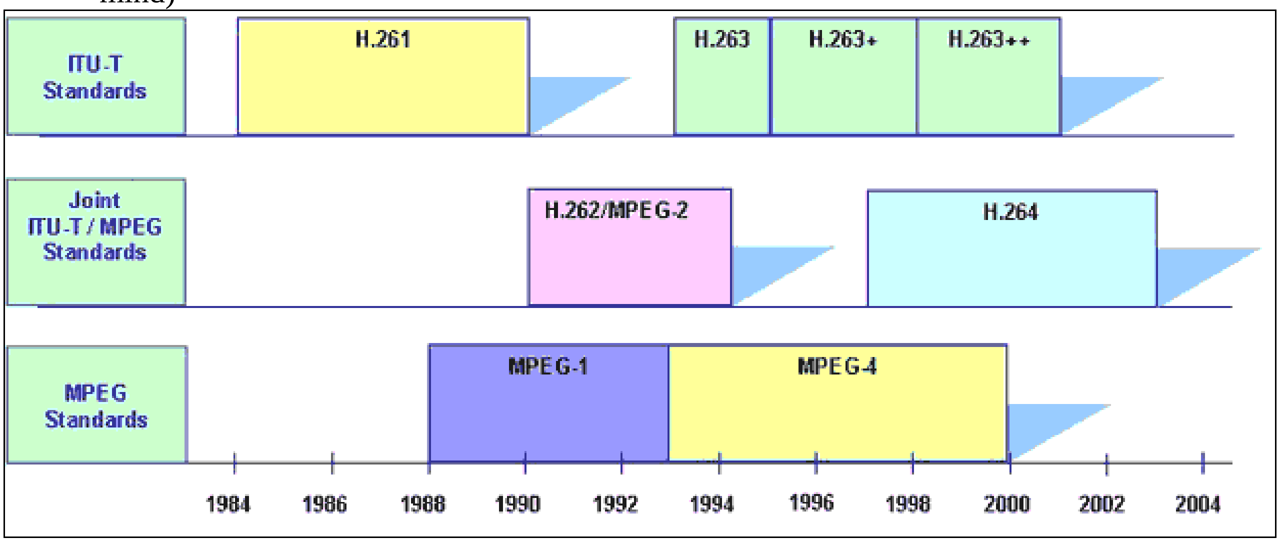
Advanced compression technology for video creation and playback called H.264/AVC (Advanced Video Coding), also known as MPEG-4 Part 10. This ultra-efficient, fully scalable video technology produces higher quality video at lower data rates for everything from 3G to HD (High-Definition).

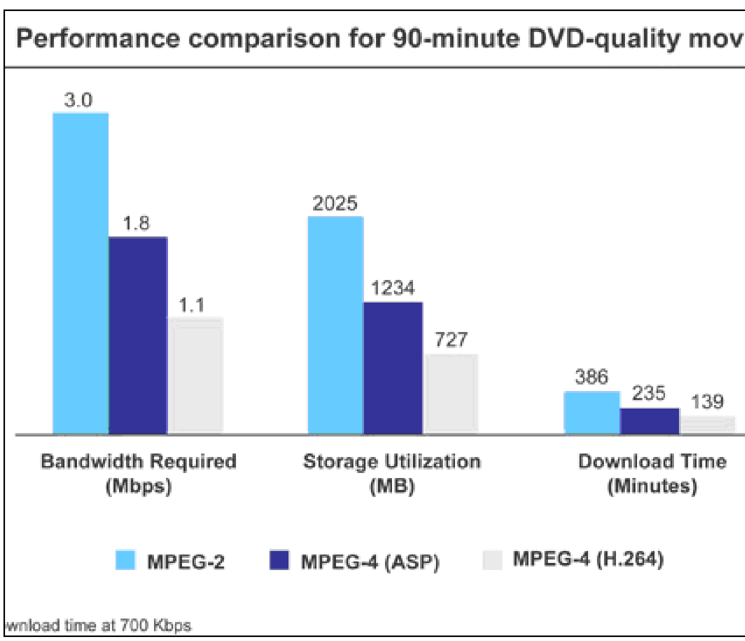
Stunning, More and Less

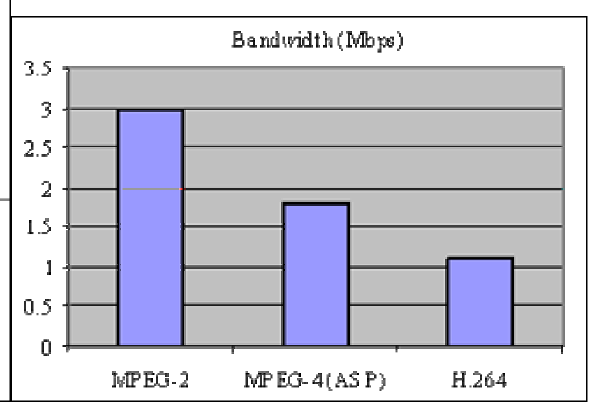
H.264 is the most advanced video coding standard available today. It uses many new coding techniques not available in MPEG2, MPEG4 and even in H.263. This new standard is taking the industry and later the world by storm. It is jointly developed by the ISO MPEG and ITU standardization bodies. Not only is H.264/AVC very efficient, providing extremely high quality in smaller files, but H.264/AVC is also scalable, producing video for everything from 3G for mobile phones to High Definition TVs (HDTV).

H.264/AVC can create

- Great-looking 3G mobile content at 50-160 Kbps
- Excellent Standard Definition (SD) video at 800-1500 Kbps
- Beautiful HD (High-Definition) video (1280x720, 24p) at 5-7 Mbps
- Full HD (High-Definition) video (1920x1080, 24p) at 7-9 Mbps. (HD can be understood using HD-TV in mind)







So at today's SD DVD data rates, H.264/AVC can deliver full HD. In fact, H.264/AVC was ratified by the DVD Forum for inclusion in the next-generation HD DVD format. Full HD H.264/AVC video plays back on today's desktop computers. H.264/AVC produces even better results than the H.263 codec. A few years ago, the

International Organization for Standardization selected the QuickTime file format as the basis for MPEG-4. Get ready for QuickTime with H.264/AVC to change the digital video universe.

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. 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.

Mobile streaming and download There is significant expectation for mobile multimedia services associated with the new 2.5G and 3G mobile service networks. The problem is that for much of the coverage area the peak bandwidth in these networks is generally around 144Kbps with individual users sustaining connections of about 40Kbps. Delivering quality video over this type of connection is problematic and may not meet the consumer expectation. HE AAC is well suited to solve this problem by being able to provide consumer-grade download and streaming audio services within today's bandwidth. 48Kbps HE AAC provides CD-quality stereo programming while at 32Kbps

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.

Internet audio has become wildly popular in recent years, specifically in the MP3 format. But what most listeners don't realize is that MP3's compression technology is more than a decade old. In those ten years, many advances in perceptual audio coding and compression have been achieved. In early 1998 the Video Coding Experts Group (VCEG – ITU-T SG16 Q.6) issued a call for proposals on a project called H.26L, with the target to double the coding efficiency (which means halving the bit rate necessary for a given level of fidelity) in comparison to any other existing video coding standards for a broad variety of applications. The first draft design for that new standard was adopted in October of 1999. In December of 2001, VCEG and the Moving Picture Experts Group (MPEG – ISO/IEC JTC 1/SC 29/WG 11) formed a Joint Video Team (JVT), with the charter to finalize the draft new video coding standard for formal approval submission as H.264/AVC [1] in March 2003.

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.

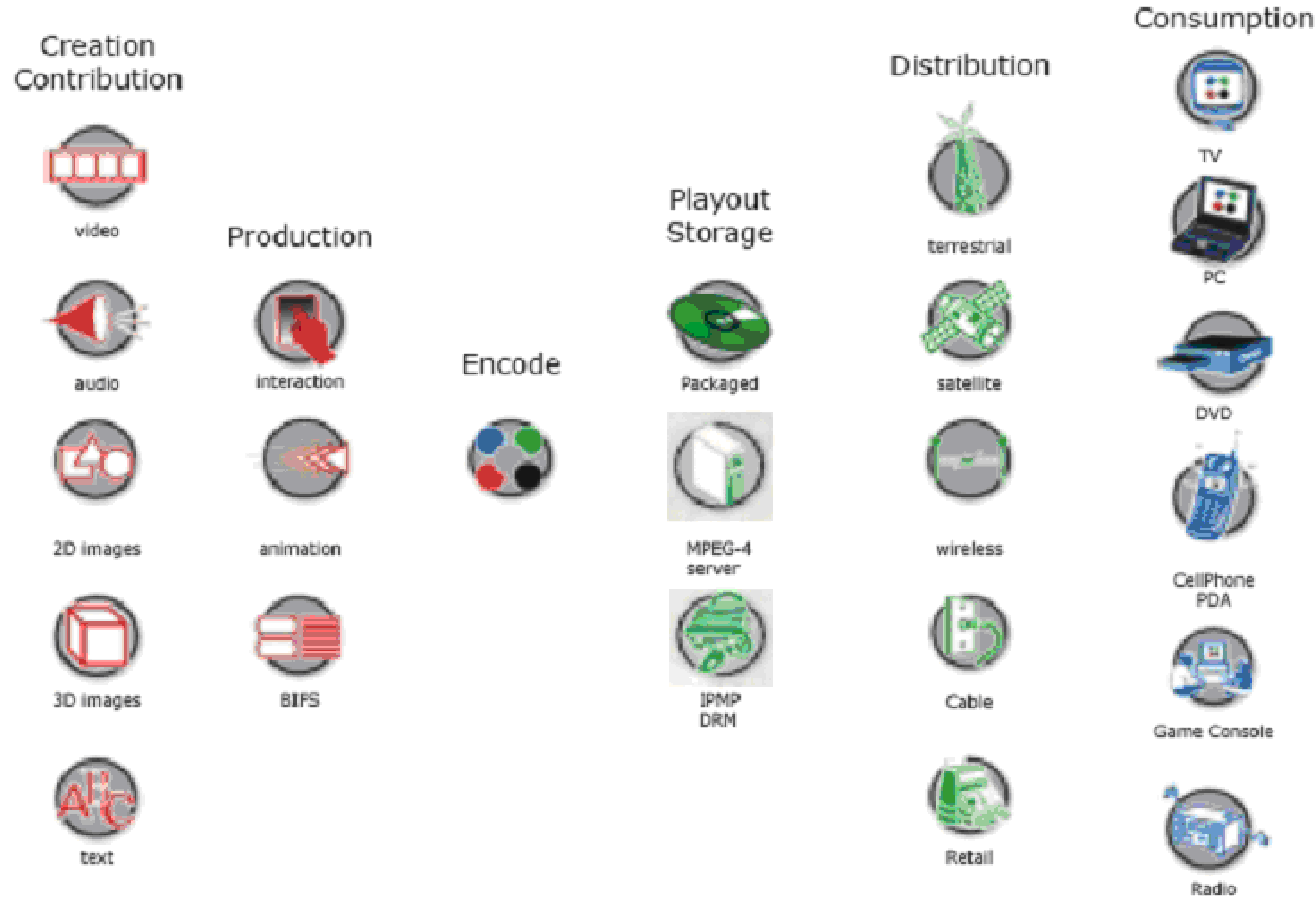


Figure 1: The MPEG-4 ecosystem liberates multimedia for delivery across any network to any user of any device

MPEG-4 is an international standard (ISO/IEC 14496) developed by MPEG (Moving Picture Experts Group). The MPEG-4 specification was finalized in October 1998 and became an International Standard in early 1999. MPEG-4 Version 2 became an International Standard early in 2000. Several extensions have been added since, including MPEG-4 AVC.

MPEG-4 differs from MPEG-2 in an important way. MPEG-2 codes a video stream as video even if it contains graphics and text overlays. MPEG-4 separates the graphics and text overlays into separate streams which are re-combined by the decoder.

For more information on MPEG-4 see the MPEG-4 Industry Forum website.

MPEG-4 Parts

The MPEG-4 standard is divided into a number of parts:

- Part 1: Systems
- Part 2: Visual
- Part 3: Audio
- Part 6: DMIF (Delivery Multimedia Integration Framework), which defines an interface between the application and network/storage.
- Part 4: Conformance, which defines how to test an MPEG-4 implementation,
- Part 5: Reference Software, including an example of how to do things.
- Part 7: Optimized video encoder description

More recent parts added to MPEG-4 are:

- Part 8: How to map MPEG-4 streams onto IP transport.
- Part 9: Reference Hardware Description
- **Part 10: Advanced Video Coding (AVC)**
- Part 11: Scene description
- 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).

Applications of MPEG-4 include digital television; interactive graphics applications (synthetic content) and interactive multimedia eg via the Internet. It is also likely to be used as one of the video codecs for the **"next generation blue laser optical disc"** http://www.disctronics.co.uk/technology/hddvd/index.htm. MPEG-4 provides features that are not available with MPEG-2

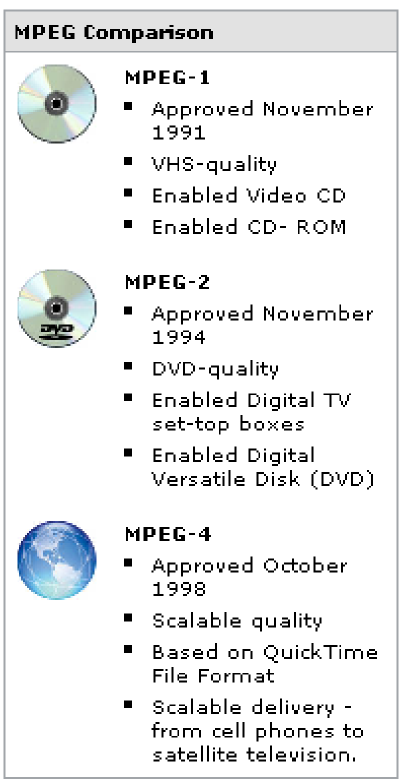
VIDEO

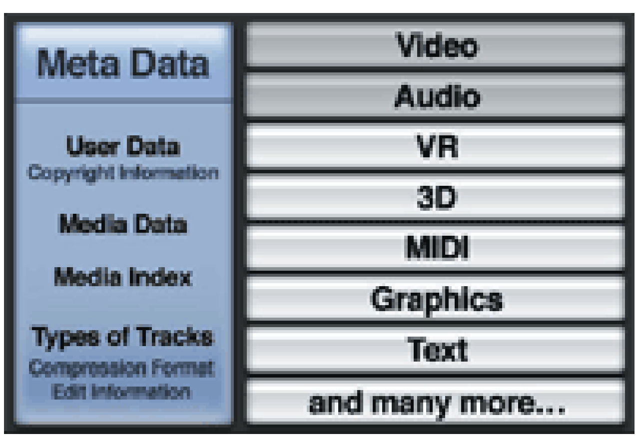
MPEG-4 Industry Forum mp4 **AVC**

Advanced Video Coding (AVC)

H.264 Codec

H.264 or "MPEG-4 AVC (Advanced Video Coding)" or H.26L

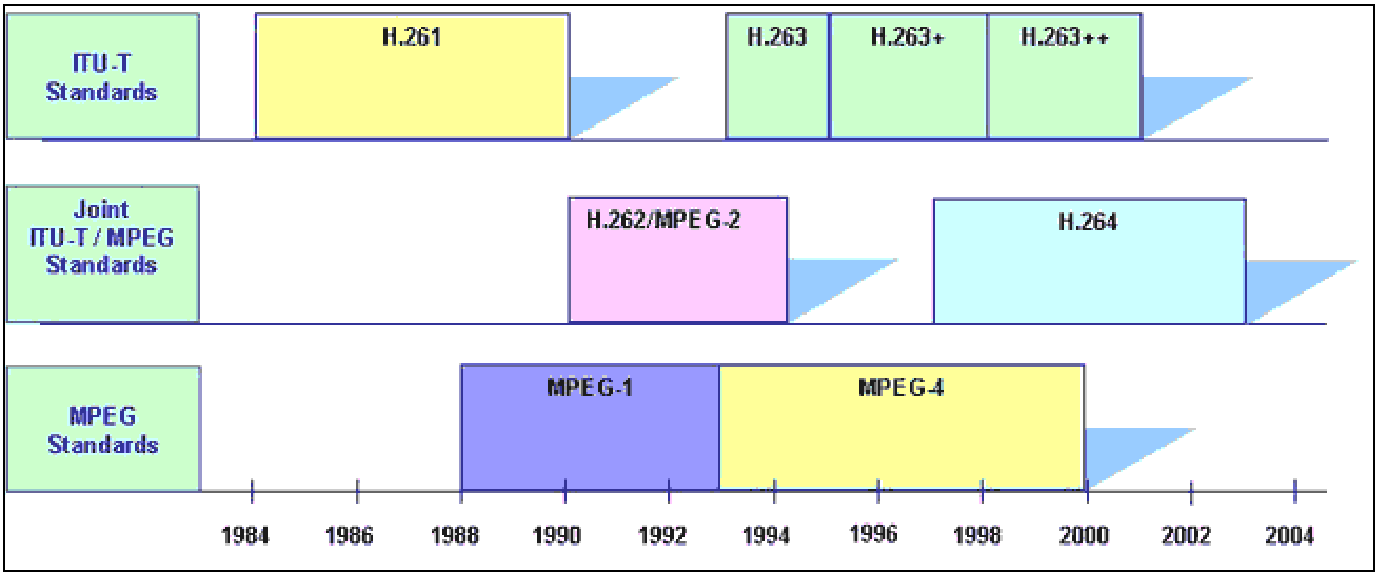




A key component in the Video Compression systems from DSP Research is the very efficient and compact H.264 codec which runs on the DM642. This codec has been developed over a period of two years by DSP Research's parent company W&W Communications. This codec implementation is first to market and is also available for licensing for your own hardware. Please see http://www.wwcoms.com for more details.

The H.264 / MPEG4 AVC Standard

H.264 is the most advanced video coding standard available today. It uses many new coding techniques not available in MPEG2, MPEG4 and H.263. This new standard is taking the industry by storm and was jointly developed by the ISO MPEG and ITU standardization bodies. This chart shows the evolution of video coding standards.



For those of you who want to learn more about video compression and the H.264 standard, the following site is an excellent place to start:

H.264/MPEG-4 AVC Introduction

For you who are already familiar with video compression standards and want to know more about the H.264 standard, the following publication (19 pages) is an excellent source:

H.264/MPEG-4 AVC Overview (PDF) (2.986KB) (search from internet)

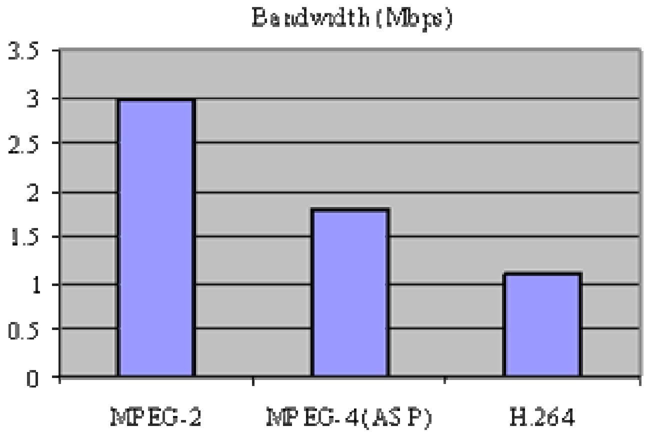
A draft of the whole H.264 Standard (264 pages) from the 'ITU-T 7th meeting: Pattaya, Thailand, 7-14 March 2003', is also downloadable below:

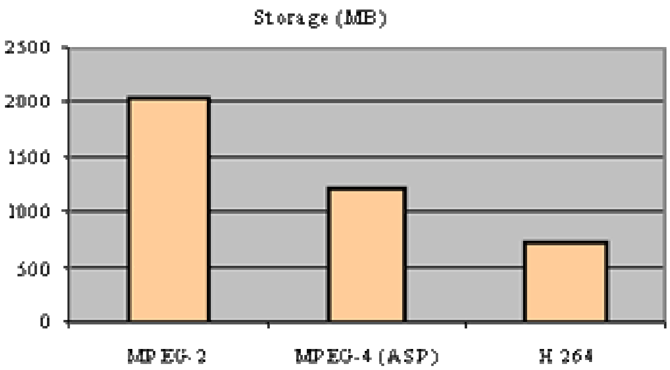
The whole H.264 Standard (draft from 7-14 March 2003) (PDF) (1,816KB) (search from internet)

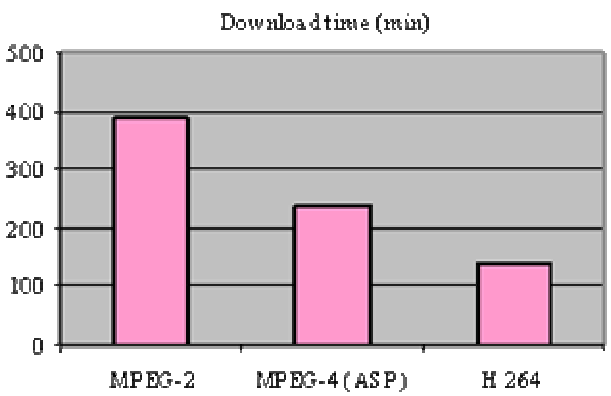
H.264 Performance

The new H.264 video compression standard is the first to be ratified by both the ISO/IEC and ITU-T. It raises the bar considerably compared to MPEG4-ASP and yields better picture quality while significantly lowering the bit rate. This allows lower overall system cost, reduced infrastructure requirements and enables many new video applications. H.264 is considered the greatest achievement in video compression in the past 10 years and is already finding its way into many mainstream video applications.

Below follow some practical comparisons between different coding technologies for 90 min of DVD quality video:


Bandwidth (Mbps)
3.5
3
2.5
2
1.5
1
0.5
0
MPEG-2
MPEG-4(ASP)
H.264


Storage (MB)
2500
2000
1500
1000
500
0
MPEG-2
MPEG-4 (ASP)
H 264


Download time (min)
500
400
300
200
100
0
MPEG-2
MPEG-4 (ASP)
H 264

New Apple iChat with Mac OS X Tiger
Note: MSN Messenger 6.2 uses JPEG2000 codec which is inferior.

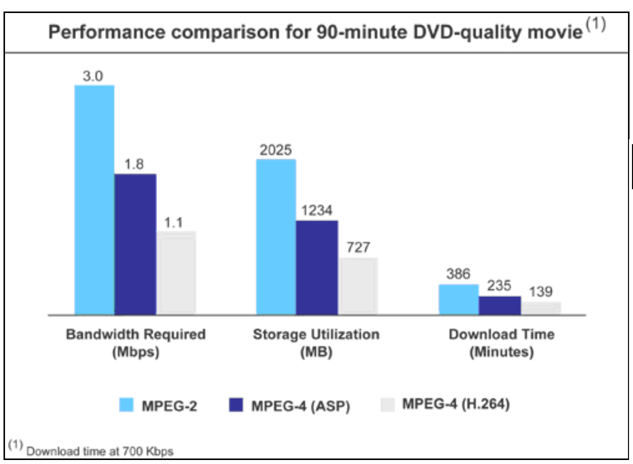
H.264 Coding Efficiency

H.264/AVC offers a significant improvement in coding efficiency compared to other compression standards such as MPEG-2. Independent lab tests show a 50 to 70 percent improvement in bit rate reduction, making it much more cost effective to deliver broadcast-quality video over cable, satellite, and telecom networks.

As with MPEG-2, H.264/AVC is based on block transforms and motion compensated predictive coding. H.264/AVC leverages today's processing power to provide improved coding techniques, including multiple reference frames and several block sizes for motion compensation, intra-frame prediction, a new 4x4 integer transform, a 1/4 pixel precision motion compensation, an in-the-loop de-blocking filter, and improved entropy coding.

- Increased coding efficiency provides the same picture quality at half the bit rate.
- Error resilience tools (included) maintain video quality in error-prone transmissions.

A performance comparison for a 90-minute DVD-quality movie is shown in the following chart.

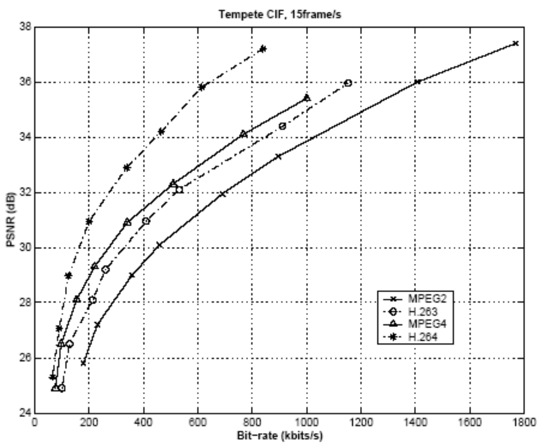


Rate Saving Compare to Other Standards

Table 1: Average rate savings compare to other standards

Codec	MPEG4	H.263	MPEG2
H.264	39%	49%	64%
MPEG4	-	17%	43%
H.263	-	-	31%

Performance Comparison

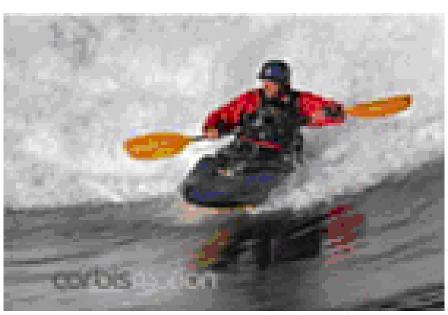


The following tables summarize the savings in offline and live encoding using H.264, MPEG-4, and MPEG-2 encoders.

H.264 Encoding Feature "Evolution"

The following table compares H.264 with both MPEG-2 and MPEG-4 feature sets.

		Standards		
		MPEG-2	MPEG-4 ASP	MPEG-4 H.264
Features	I, P, B-frames	✓	✓	✓
	Interlace	✓	✓	✓
	Coding	Huffman	Huffman	Huffman or Arithmetic
	Block size	fixed 16x16	fixed 16x16	variable down to 4x4
	¼ pixel		✓	✓
	GMC		✓	
	Loop Filter (aka deblocking fliter)			✓
	Slice-based motion prediction			✓
	Multiple reference frames			✓
	MB AFF (improved interlaced management)			✓
	RDO (Rate Distortion Optimisation)			✓
	WP (Weighted Prediction)			✓
	Switching pictures (for fast change channel)			✓



Original DV Source
File size: 208,039 kb
Bitrate: 28,208.68

DVD-Quality MPEG-2
File size: 54,526 kb
Bitrate: 7,393.36

H.264, Main Profile
File size: 6,174 kb
Bitrate: 837.15

Recently the MPEG-4 Visual (MPEG-4 part 2) standard [5] has also begun to emerge in use in some application domains of the prior coding standards. It has provided video shape coding capability, and has similarly worked toward broadening the range of environments for digital video use. In early 1998 the Video Coding Experts Group (VCEG – ITU-T SG16 Q.6) issued a call for proposals on a project called H.26L, with the target to double the coding efficiency (which means halving the bit rate necessary for a given level of fidelity) in comparison to any other existing video coding standards for a broad variety of applications. The first draft design for that new standard was adopted in October of 1999. In December of 2001, VCEG and the Moving Picture Experts Group (MPEG – ISO/IEC JTC 1/SC 29/WG 11) formed a Joint Video Team (JVT), with the charter to finalize the draft new video coding standard for formal approval submission as H.264/AVC [1] in March 2003.

Performance of the deblocking filter for highly compressed pictures. Left: without deblocking filter, right: with deblocking filter.

Figure 1A shows a picture encoded in MPEG2 at 2400 kbps (kilobits per second) – in essence an infinite number of bits for such a small format. MPEG2 encoding at this bitrate is essentially lossless, so the resulting quality closely matches that of the original source data.

In **Figure 1B**, the bitrate has been reduced to 400 kbps, and you can begin to see some fuzziness in the frame as the quantization has been made coarser to drop the bit rate; when you view the corresponding video clip, you will also see some smearing of fast panning motions.

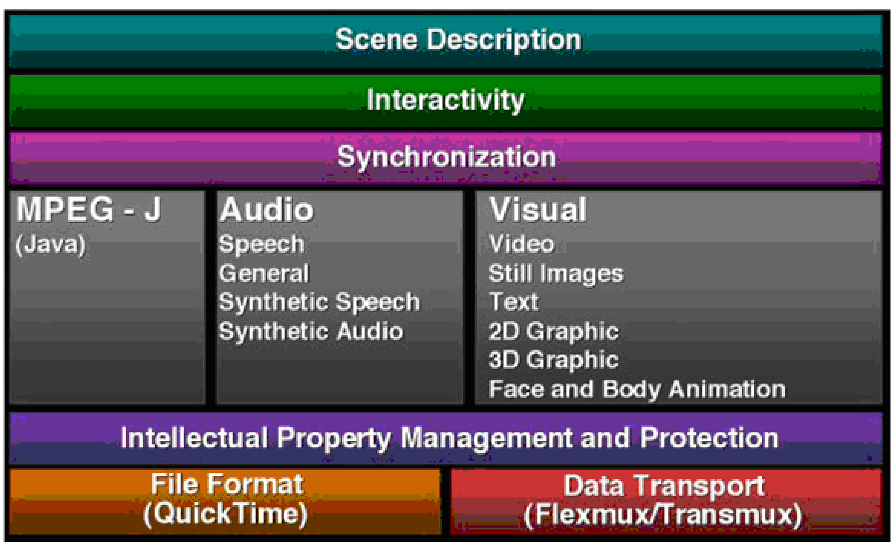
When the MPEG2 bitrate is further dropped to 100 kpbs in **Figure 1C**, things begin to fall apart. You begin to see blockiness at the macro block level as some macro blocks can only be resolved as uniform (DC) values, and any fast motion is distorted.

The value of H.264 is most obvious at low bit rates. In figure 2, you can see the difference of encoding via MPEG-2 and H.264 at 100 kpbs. The tremendous improvement in quality produced by H.264 is self-explanatory.

Figure 2A. MPEG-2

Figure 2B. H.264

Another way to look at this is to compare the bit rates needed by MPEG-2 and H.264, for similar quality images. In our judgment, figure 1b (MPEG-2 at 400 kpbs) and Figure 2b (H.264 at 100 kbps) show very similar quality. While comparing quality is very subjective, this is consistent with PixelTools evaluations of many tests – generally showing a 3-4-fold decrease in bit rate from MPEG-2 for the same level of quality.



The components of MPEG-4

Sequence	Coding Gain
IRENE	53%
MOBILE & CALENDAR	58%
PARIS	42%
STUDENTS	45%
TEMPETE	41%
Average Gain	49.2%

Table 3. Coding gain of H.264 over H.263 Baseline Profile for various CIF sequences.

Sequence	Coding Gain
AKIYO	51%
CAR PHONE	40%
CLAIRE	48%
COAST GUARD	38%
CONTAINER	41%
FOREMAN	42%
IRENE	50%
MISS AMERICA	44%
MOBILE & CALENDAR	48%
NEWS	56%
PARIS	52%
TEMPETE	37%
AVERAGE GAIN	45.6%

Table 4. Coding gain of H.264 over H.263 Baseline Profile for various QCIF sequences.

Sequence	Coding Gain
IRENE	29%
MOBILE & CALENDAR	31%
PARIS	22%
STUDENTS	24%
TEMPETE	20%
Average Gain	25.4%

Table 5. Coding gain of H.264 over H.263 Conversational High Compression Profile for various CIF sequences.

Sequence	Coding Gain
AKIYO	25%
CAR PHONE	20%
CLAIRE	34%
COAST GUARD	19%
CONTAINER	22%
FOREMAN	20%
IRENE	26%
MISS AMERICA	25%
MOBILE & CALENDAR	22%
NEWS	28%
PARIS	22%
TEMPETE	17%
AVERAGE GAIN	23.3%

Table 6. Coding gain of H.264 over H.263 Conversational High Compression Profile for various QCIF sequences.

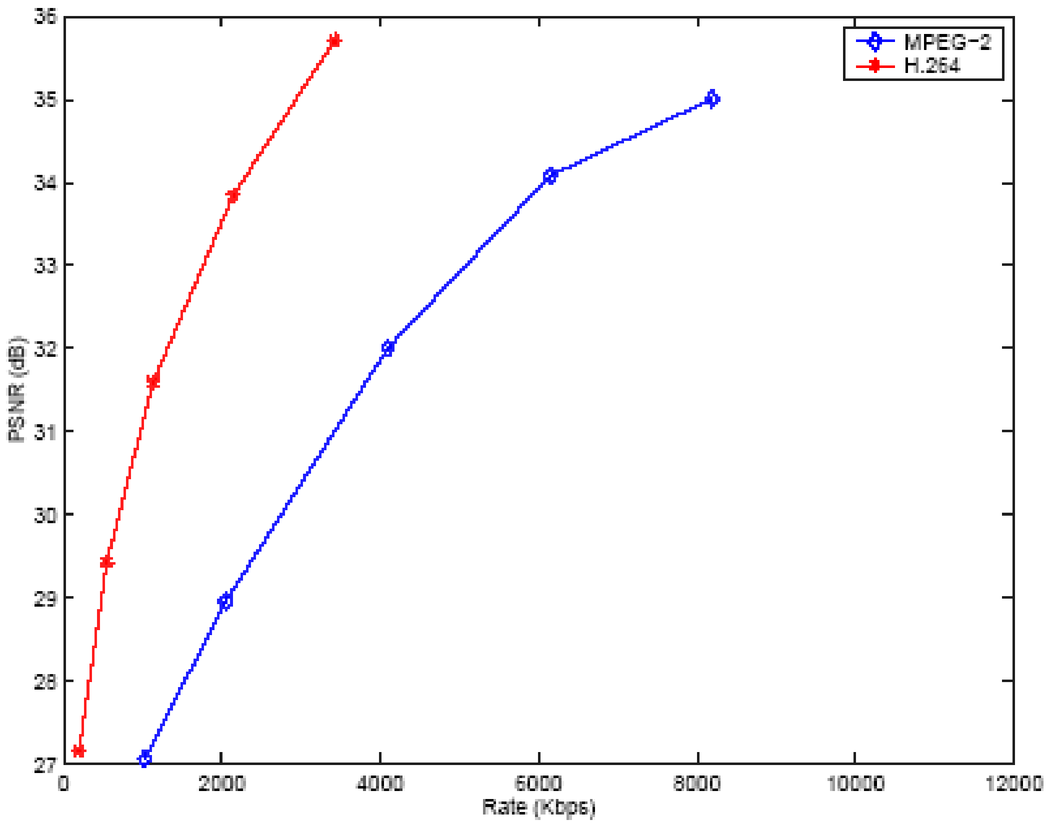
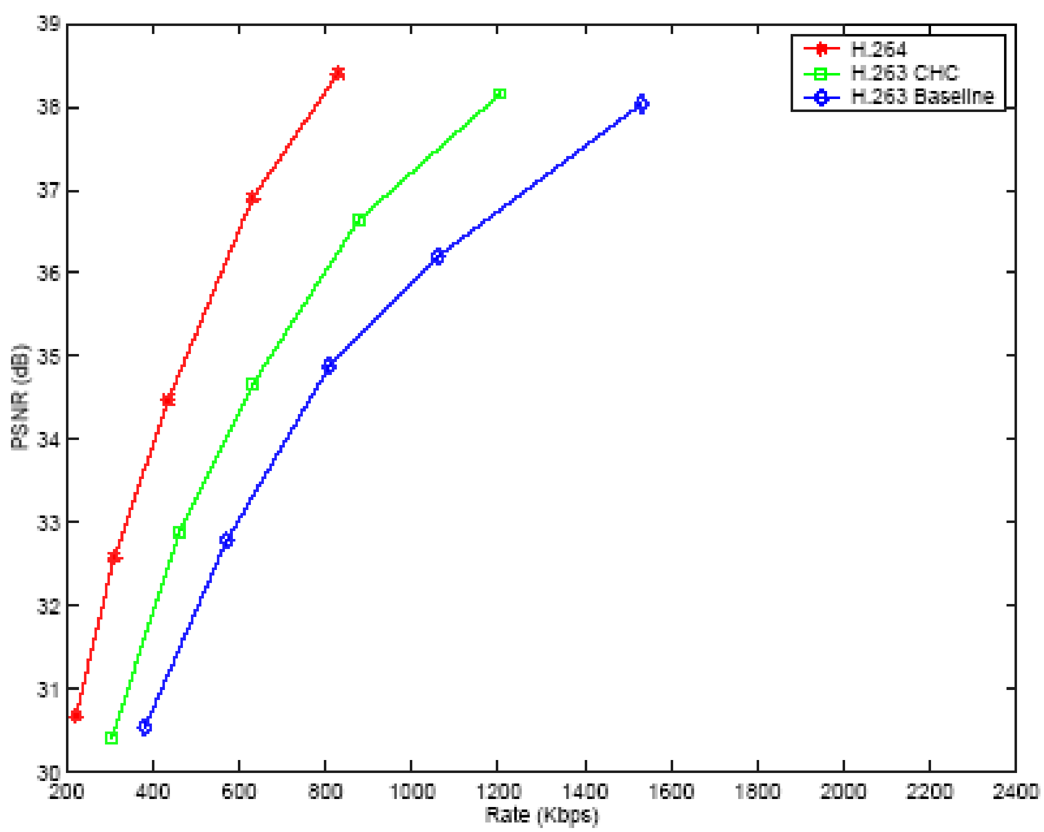


Fig. 2. Rate-PSNR curves of FLOWER GARDEN sequence encoded using H.264 and MPEG-2 encoders.



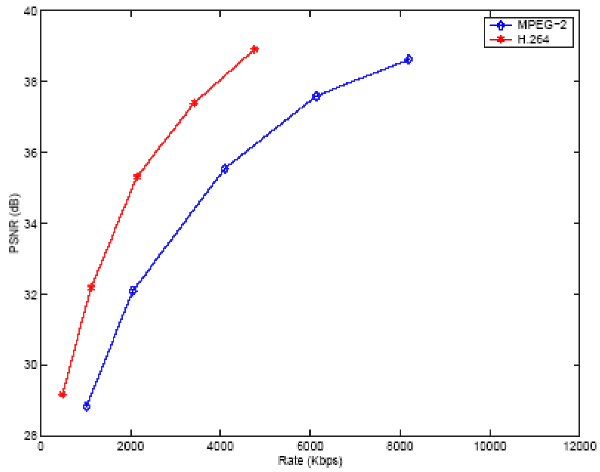


Fig. 1. Rate-PSNR curves of CAROUSEL sequence encoded using H.264 and MPEG-2 encoders.

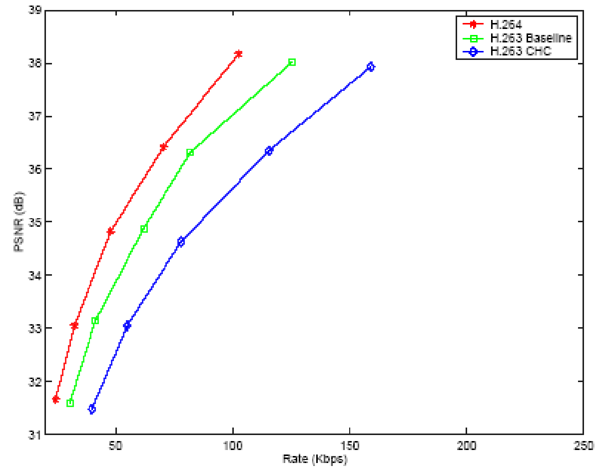


Fig. 4. Rate-PSNR curves of QCIF FOREMAN sequence encoded using H.264, H.263 Baseline and H.263 CHC encoders.

H.264 or "MPEG-4 AVC (Advanced Video Coding)"

The recently added MPEG-4 part 10, also known as MPEG-4 AVC (Advanced Video Coding) or H.264, offers significantly greater compression than its predecessors. It can provide DVD-quality video at under 40% of the bit rate of MPEG-2 and is considered promising for full-motion video over wireless, satellite, and ADSL Internet connections. It is also one of the video codecs that have been provisionally chosen for blue laser HD DVD.
H.264 makes use of spatial, temporal, and psycho-visual redundancies to improve video coding efficiencies. The additional features contained in H.264 include the following:

- Variable block-size motion compensation with small block sizes down to 4 x 4 pixels.
- Quarter-sample-accurate motion compensation for smoother motion.
- Motion vectors over picture boundaries as opposed to within the previously-decoded reference picture.
- Multiple reference picture motion compensation, compared with only one previous picture for earlier codecs.
- Decoupling of referencing order from display order for more flexibility and removal of the extra delay associated with bi-predictive coding.
- Decoupling of picture representation methods from picture referencing capability.
- Weighted prediction allowing the motion-compensated prediction signal to be weighted and offset by amounts specified by the encoder. This can dramatically improve coding efficiency for scenes containing fades, and can be used flexibly for other purposes as well.
- Improved "skipped" and "direct" motion inference for better coding of video containing global motion.
- Directional spatial prediction for intra coding, for extrapolating the edges of the previously-decoded parts of the current picture and applying it in regions of pictures that are intra coded. This improves the quality of the prediction signal, and also allows prediction from neighbouring areas that were not coded using intra coding.
- In-the-loop deblocking filtering to reduce artefacts.

Video compression using H.264/MPEG-4 AVC generates video at bit rates that are typically two to four times lower than previous MPEG-2 and MPEG-4 compression standards, at equivalent video quality levels.

Comparison of formats

The following table compares MPEG-4 H.264 with MPEG-4 ASP and MPEG-2. H.264 offers a number of new features for improved quality and lower bit rates, which include those listed in the table.

	MPEG-2	**MPEG-4 ASP**	**MPEG-4 H.264**
I, P, B frames	Yes	Yes	Yes
Coding	Huffmann	Huffmann	Huffman or arithmetic
Block size	16 x 16	16 x 16	down to 4 x 4
Quarter pixel resolution		Yes	Yes

Deblocking filter			Yes
Sliced-based motion prediction			Yes
Multiple reference frames			Yes
Weighted Prediction			Yes
Switching pictures			Yes
Bit rate comparison	100%	60%	40%

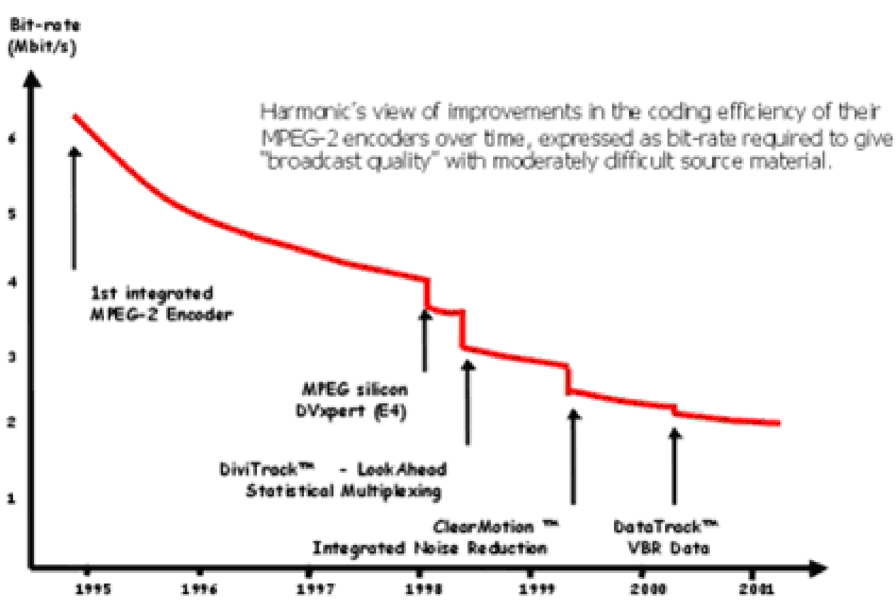
Interactive scenes, objects, elementary streams

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 backoffice systems including transaction/ e-commerce systems.

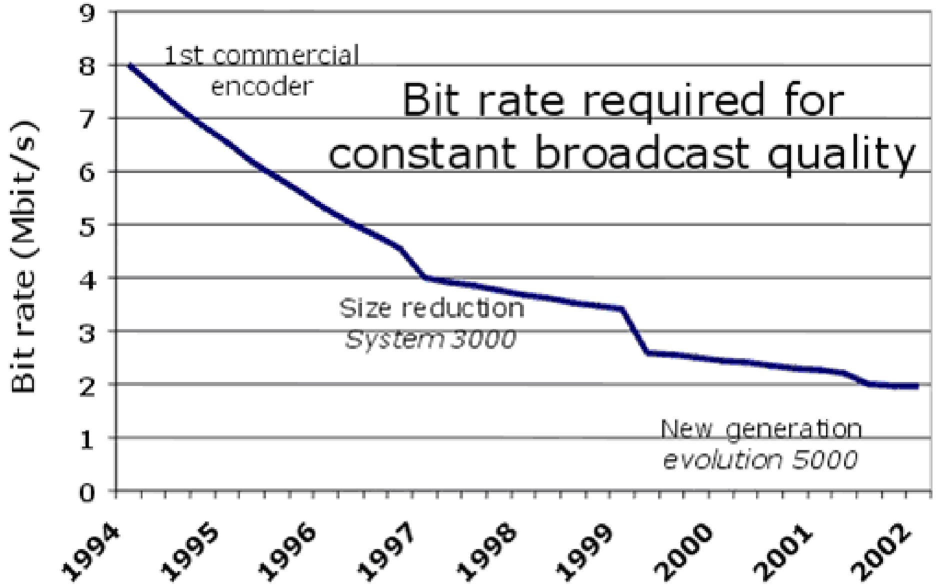
With MPEG-4, advanced interactive programming can be authored seamlessly integrating audio/video with 2D, 3D objects, animation and interactivity. For example, a viewer can navigate a sporting event's course from a 3D map, select information about aspects of the program, listen to commentary within a 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.

Saves money, makes money

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.



Reproduced with the kind permission of Harmonic Inc.



Reproduced with the kind permission of TANDBERG Television

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.

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 Advances

1. 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.

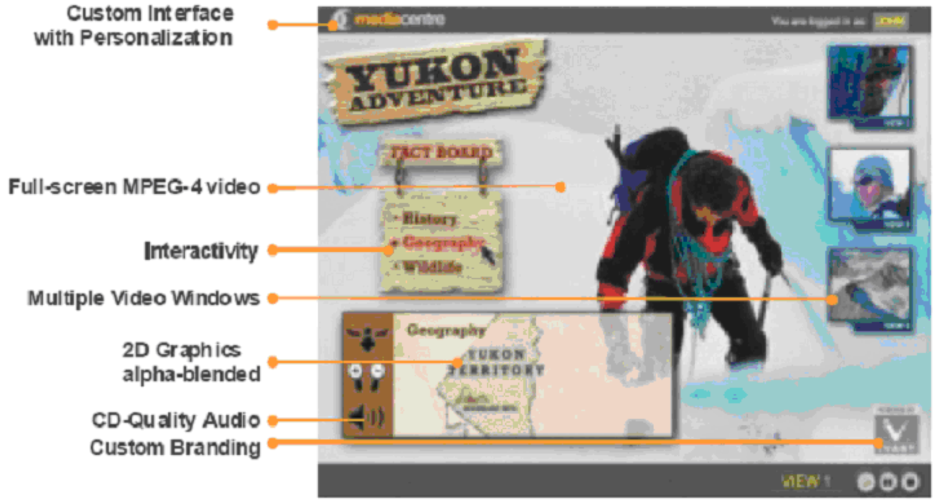
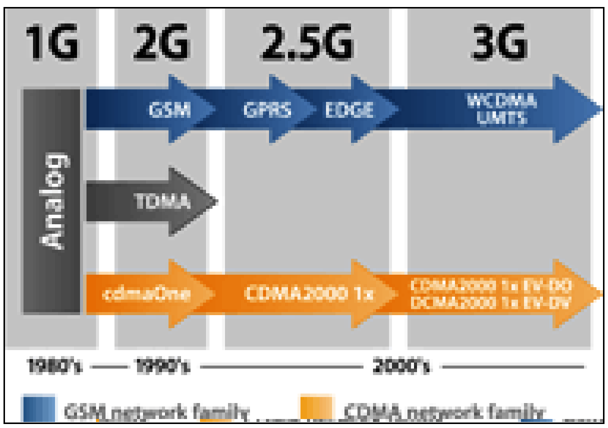
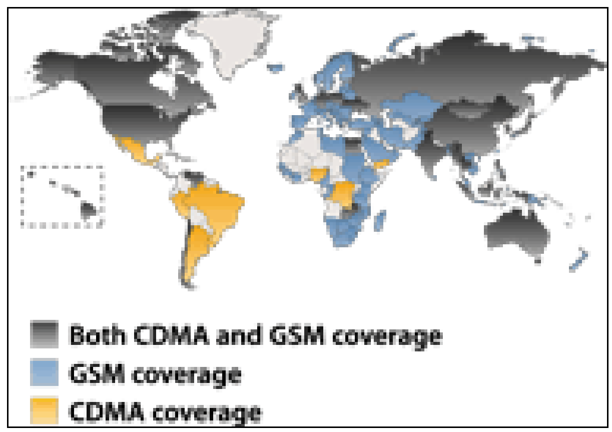
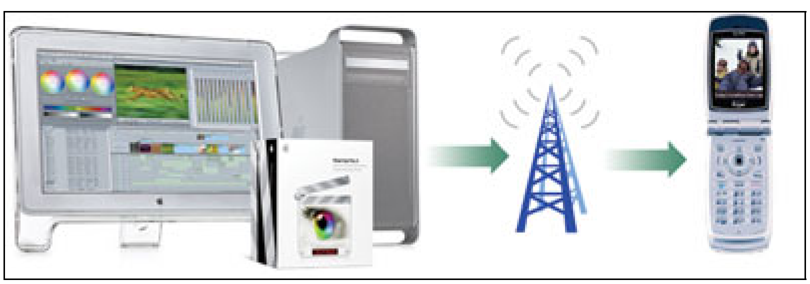


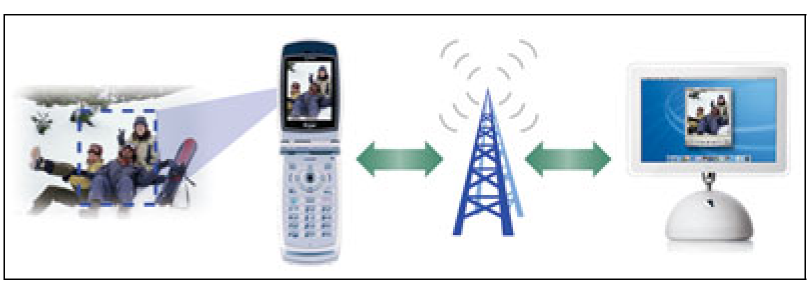
Figure 4: An MPEG-4 scene using multiple objects and elementary streams (Source: iVAST[ix])

2. 3GPP and 3GPP2 are based on the MPEG-4 standard









In order to ensure interoperability between products from all players in the mobile telecommunications industry, the 3GPP and 3GPP2 specifications call out specific types of media that can exist in .3gp and .3g2 files.

3GPP 3GPP2 Network Type GSM CDMA2000 Video MPEG-4, H.263 MPEG-4, H.263 Audio AAC, AMR AAC, AMR, QCELP Text 3G Text 3G Text File Format Basis QuickTime QuickTime File Extension .3gp .3g2

	3GPP	3GPP2
Network Type	GSM	CDMA2000
Video	MPEG-4, H.263	MPEG-4, H.263
Audio	AAC, AMR	AAC, AMR, QCELP
Text	3G Text	3G Text
File Format Basis	QuickTime	QuickTime
File Extension	.3gp	.3g2

Simple Profie (SP) Advanced Simple Profile (ASP)

The first version of MPEG-4 was MPEG-4 SP (Simple Protocol) which was intended for low bandwidth applications such as dial up Internet access.

MPEG-4 SP uses both intra-frame and intra-field compression for both progressive and interlace display and employs both DCT and DPCM (differential pulse code modulation) compression to ensure optimum compression for every type of video.

MPEG-ASP (Advanced Simple Protocol) includes a number of additional features including

- Motion compensation with 1/4 picture element accuracy
- Bi-directional prediction
- Total picture area motion compensation

FLEXIBLE AND SCALABLE MPEG-4 video coding is characterised by superior flexibility and scalability. Its usefulness spans from low and intermediate to high bit rates, delivering a competitive advantage to other video coding standards and proprietary video coding technologies. To support the diversity of potential application spaces, MPEG-4 offers a variety of so-called “profiles”: well-defined subsets from the toolbox, useful for specific applications. In natural video coding, the most commonly used profiles are Simple Visual Profile (SP) and Advanced Simple Visual Profile (ASP).

MPEG-4 VISUAL CODING PROFILES The MPEG-4 market has converged on two important visual profiles for natural video:

• **MPEG-4 Visual – Simple Profile (SP)** is designed primarily for low processing power coding, low latency and use in less-than-ideal transmission circumstances. Ideal for real-time desktop software encoding, mobile and wireless devices, video telephony and video–conferencing.

• **MPEG-4 Visual - Advanced Simple Profile (ASP)** offers the best MPEG-4 coding performance and can be deployed into more demanding environments where video decoding at intermediate and higher bit rates is required. ASP is also ideal for broadcasters wishing to provide wide range of programming qualities suited to a variety of platforms, e.g. Broadcast, Internet and Mobile Phones. Both profiles are already implemented in a wide range of software players, hardware, silicon, embedded systems and consumer devices: mobile phones, handheld computers, PDAs, personal video jukeboxes, and consumer digital still- and video cameras.

BENEFITS OF PROFILES

• The scalable choice between MPEG-4 profiles supports the diversity of digital video applications, avoiding costly production of a number of different multimedia formats.
• Service providers can use broadband and narrowband platforms for the distribution of their content, such as wireless networks, digital television and the Internet.
• Profiles allow users to choose tool-sets supporting just the functionality they need.

COMPARISON OF MPEG-4 VSP AND ASP

Simple Profile	*Advanced Simple Profile*
The basis of all MPEG-4 video profiles; to some extent compatible with H.263	Based on Simple Profile, adds advanced coding tools to SP
Suited for low bit rates, from 10 kbit/s upwards, and for low latency applications.	Supports a wide range of bit rates from narrowband (56 kbit/s) to broadband (300-750 kbit/s), and broadcast SD to HD (1-8+ Mbit/s).
Industrial quality control, low complexity desktop video, mobile video, use close circuit video surveillance, teleconferencing or video telephony, more.	Broadcast, Unicast and Multicast applications, advanced Internet Streaming, media asset management, browsing, VOD, education, security, more
Adopted by 3GPP for wireless video streaming and ISMA (Level 0) for narrowband internet streaming.	Adopted by ISMA (Level 1) for broadband internet streaming. Used in many consumer devices such including some DVD players and PDAs.
Includes error resilience tools	Includes interlace Support
Real-time software encoding easy to achieve	Better coding efficiency at higher quality levels and bitrates.
Can be Implemented in small or resource-constrained devices.	Encoder and decoder is more complex; encoding and decoding in software possible on modern personal computers
Coding of Intra (I) and Predicted (P)Frames	Adds support for coding of B-Frames (Bidirectionally interpolated)
½ Pixel accurate motion compensation	¼ Pixel accurate motion compensation
Block-based (16x16, 8x8 blocks) motion compensation	Adds global motion compensation (GMC) with up to 6 parameter affine model

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.

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.

OR

Profiles and levels are interoperability/Conformance points. In order to make sure that MPEG-4 products work with other MPEG-4 products from other vendors, conformance points are developed. These points specify items such as

- Tools that can be used
- Bitrates
- Image Sizes
- Number of objects

With out these Profiles and Levels, there would be no way to know that one product could work with another. For more information on MPEG-4 visual profiles and levels, see http://www.m4if.org/resources/profiles/index.html.

What are the different parts of the MPEG-4 Standard?

Currently MPEG-4 is broken down into 8 Separate parts, some of which are still under development and are not available for purchase from ISO.

- **ISO/IEC 14496-1 (Systems)**
  Contains tools such as BiFS, Object Descriptors, FlexMux, MP4 File Format, etc.
- **ISO/IEC 14496-2 (Visual)**
  Includes natural and synthetic coding as well as Facial and Body Animation.
- **ISO/IEC 14496-3 (Audio)**
  Including Speech coding, General Audio Coding, Structured Audio, Text to Speech interface, Parameteric Audio.
- **ISO/IEC 14496-4 (Conformance)**
  Specifies tests to be performed to verify whether bitstreams and decoders meet the requirements of parts 1, 2, 3, and 6.
- **ISO/IEC 14496-5 (Reference Software)**
  Unoptimized software implementation of the MPEG-4 specification.
- **ISO/IEC 14496-6 (Delivery Multimedia Integration Framework**
  Provides a means for transparent access and delivery of content irrespective of delivery technologies.
- **ISO/IEC 14496-7 (Optimised software for MPEG-4 tools)** -- Under development
- **ISO/IEC 14496-8 (4 on IP framework)** -- Under developement
  A framework for transmitting MPEG-4 over IP neworks

Where can I purchase the standard?

The MPEG-4 Standard (ISO/IEC 14496) can be purchased online at the ISO website.

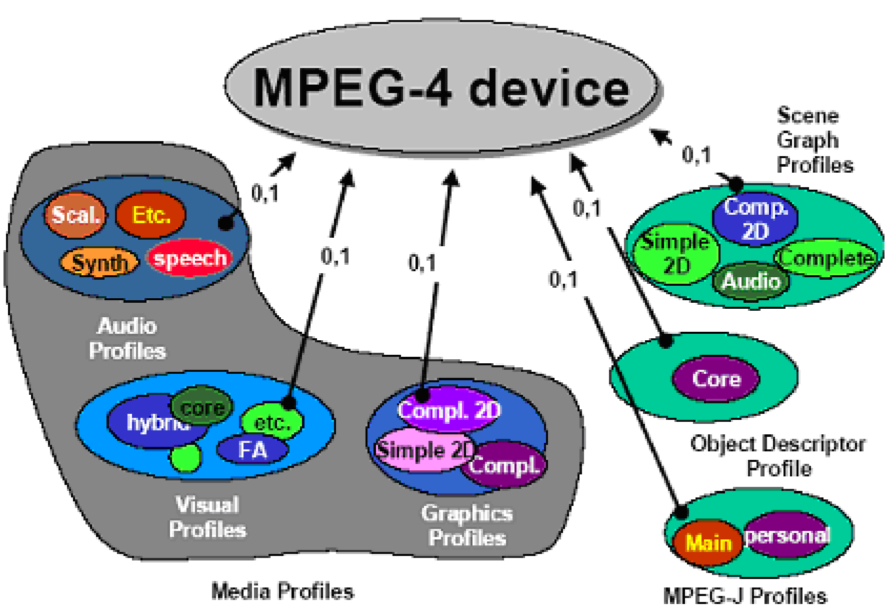


Figure 5: Profiles and Levels

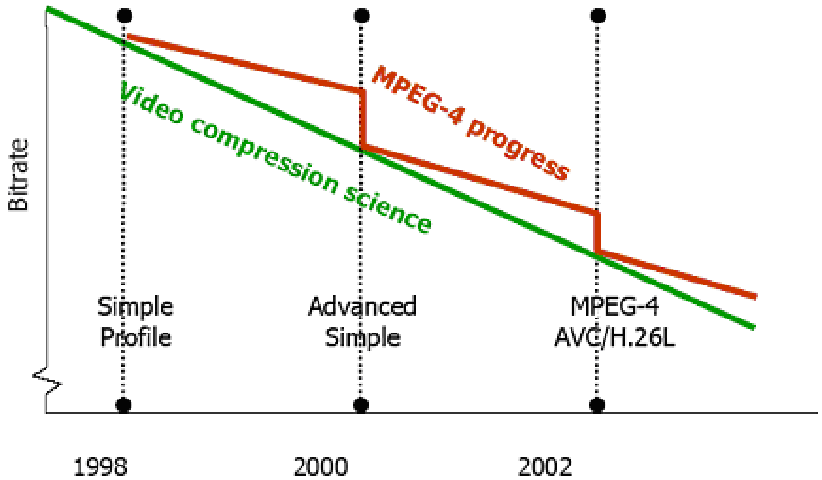


Figure 7: MPEG - A predictable, responsible upgrade strategy

AUDIO

MP4 Advanced Audio Coding

Because of its exceptional performance and quality, Advanced Audio Coding (AAC) is at the core of the MPEG-4, 3GPP, and 3GPP2 specifications and is the new audio codec of choice for Internet, wireless, and digital broadcast arenas. AAC provides audio encoding that compresses much more efficiently than older formats such as MP3, yet delivers quality rivaling that of uncompressed CD audio.

AAC was developed by the MPEG group that includes Dolby, Fraunhofer (FhG), AT&T, Sony, and Nokia—companies that have also been involved in the development of audio codecs such as MP3 and AC3 (also known as Dolby Digital).

MPEG-2 AAC is the audio format utilized in the Japanese Digital Broadcast system, known as ISDB (Integrated Services Digital Broadcasting). MPEG-2 AAC is also the basis of the audio coding technology used by XM Radio, one of two satellite radio services currently operating in the United States.

Move Over MP3
Internet audio has become wildly popular in recent years, specifically in the MP3 format. But what most listeners don't realize is that MP3's compression technology is more than a decade old. In those ten years, many advances in perceptual audio coding and compression have been achieved. AAC takes full advantage of these advances, resulting in higher quality output at lower data rates, allowing even modem users to hear a difference.

When compared side-by-side, AAC proves itself worthy of replacing MP3 as the new Internet audio standard. Take a look at these AAC advantages over MP3:

- Improved compression provides higher-quality results with smaller file sizes
- Support for multichannel audio, providing up to 48 full frequency channels
- Higher resolution audio, yielding sampling rates up to 96 kHz
- Improved decoding efficiency, requiring less processing power for decode

The Data Speaks for Itself
In numerous comparison tests, AAC comes out on top. Check out these impressive results:

- AAC compressed audio at 128 kbps (stereo) has been judged by expert listeners to be "indistinguishable" from the original uncompressed audio source.*
- AAC compressed audio at 96 kbps generally exceeded the quality of MP3 compressed audio at 128 kbps. AAC at 128 kbps provides significantly superior performance than does MP3 at 128 kbps.*
- AAC was the only Internet audio codec evaluated in the range "Excellent" at 64 kbps for all of the audio items tested in EBU listening tests.*

* Information provided by Dolby Labs.

Advanced Audio Coding (AAC) is a wideband audio coding algorithm that exploits two primary coding strategies to dramatically reduce the amount of data needed to convey high-quality digital audio. First, signal components that are "perceptually irrelevant" and can be discarded without a perceived loss of audio quality are removed. Next, redundancies in the coded audio signal are eliminated. Efficient audio compression is achieved by a variety of perceptual audio coding and data compression tools, which are combined in the MPEG-4 AAC specification.
The MPEG-4 AAC standard incorporates MPEG-2 AAC, forming the basis of the MPEG-4 audio compression technology for data rates above 32 kbps per channel. Additional tools increase the effectiveness of AAC at lower bit rates, and add scalability or error resilience characteristics. These additional tools extend AAC into its MPEG-4 incarnation (ISO/IEC 14496-3, Subpart 4).
The MPEG-4 AAC patent license grants rights for multiple **MPEG-4 AAC Object Types**, including

- AAC LC (Low Complexity)
- AAC LTP (Long-Term Prediction)
- AAC Scalable, and ER AAC LD (Low Delay).
- High Efficiency AAC

What is low Complexity AAC?

Low Complexity AAC (Advanced Audio Coding) is similar to standard AAC except that prediction is not used and there is a lower order Temporal Noise Shaping (TNS) filter.

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.

MPEG-4 High Efficiency AAC (HE AAC) MPEG-4 High Efficiency AAC is the combination of MPEG AAC and the SBR Bandwidth Extension amendment, which was finalized during the March 2003 MPEG meeting. The amendment is based on Coding Technologies' SBR (Spectral Band Replication) technology which Coding Technologies and its customers deploy the technology under the aacPlus™1 brand name. In June 2001. Since then, it has gained market momentum by delivering CD-quality stereo at 48Kbps and excellent quality stereo at 32Kbps. In late 2001, its value was recognized by MPEG and it is now on track to become a core profile for MPEG-4 audio called “High-Efficiency AAC.”

MPEG-4 High Efficiency AAC (HE AAC) is not a replacement for AAC, but rather a superset which extends the reach of high-quality MPEG-4 Audio to much lower bit rates. High Efficiency AAC decoders will decode both plain AAC and the enhanced AAC plus SBR. The result is a backward compatible extension of the standard which nearly doubles the efficiency of MPEG-4 Audio. SBR is a unique bandwidth extension technique that enables audio codecs to deliver the same listening experience at around half the bit rate. As a result, High Efficiency AAC delivers CD-quality stereo at 48Kbps and 5.1 channel surround sound at 128Kbps. This level of efficiency is ideal for Internet content delivery and fundamentally enables new applications in the markets of mobile and digital broadcasting.

In addition, MPEG has also recognizes a “low-power” decoder variant for HE AAC. Developed jointly by Panasonic, NEC and Coding Technologies, this low-power decoding method requires 40% less processing power and decodes HE AAC bitstreams with only a slightly reduced audio quality. The availability of both low-power and high-quality decoders for HE AAC enables the standard to run on the widest possible range of processors in mobile and portable device applications.

Several independent tests have been conducted over the past 2 years to demonstrate the value of MPEG-4 HE AAC compared to ‘normal’ AAC, other audio coding standards and proprietary codecs. These tests show that MPEG-4 HE AAC offers a significant benefit over these proprietary codecs and over AAC without extensions, and places it clearly as the most efficient audio codec “in existence”.

NOTE: aacPlus is a trademark of Coding Technologies AB.

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.

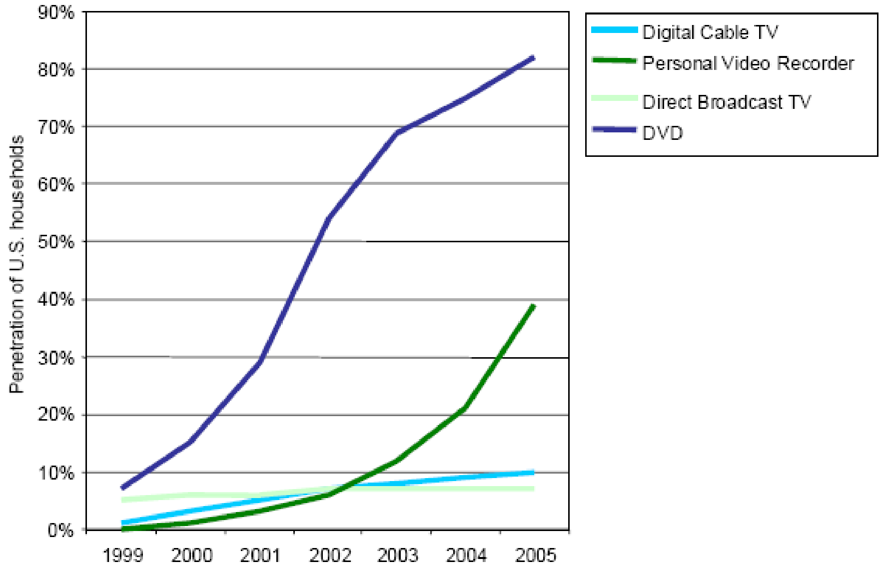


Figure 8: US Market Penetration – growth from 1999-2005 (The Yankee Group[xi])

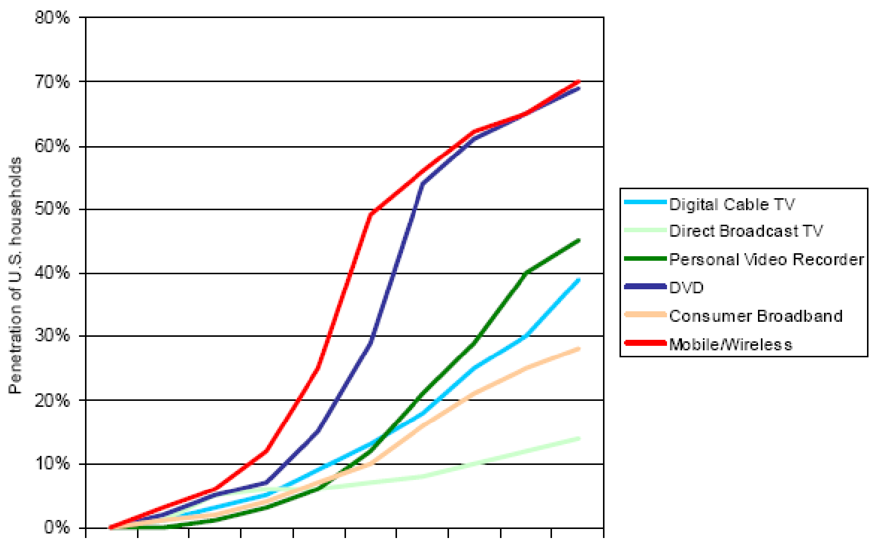


Figure 9: US Market Penetration – growth from year 0 (introduction) – 10 (The Yankee Group)

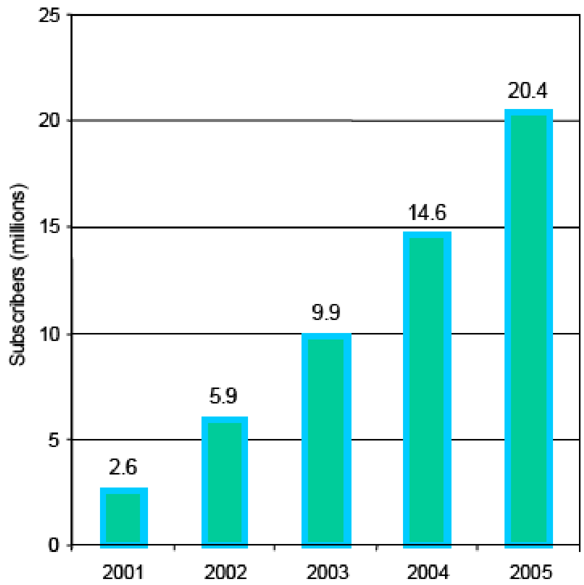


Figure 10: U.S. Video-On-Demand Subscriber Forecasts (The Yankee Group)

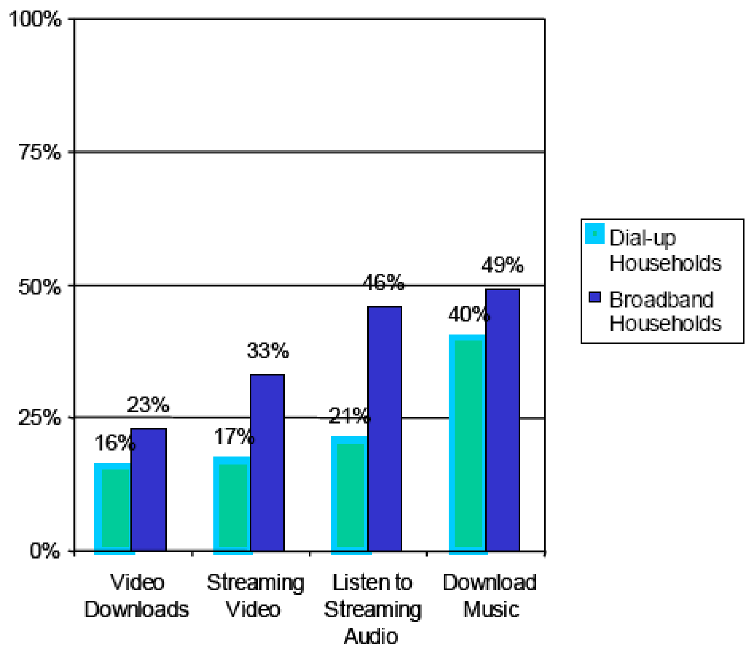


Figure 11: Broadband Media Access (The Yankee Group)

AUDIO CAPABILITIES OF MP4

MOST POWERFUL AUDIO AVAILABLE TODAY MPEG-4 Advanced Audio Coding (AAC) is the most powerful audio codec available today. The near CD quality MPEG-1 Layer 2 audio codec (used widely in digital video applications) 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 and 7.1-channel surround sound very effectively and will allow transmission of wave field audio data that extends the listening sweet spot to the whole room in future sound systems. **HIGH EFFICIENCY AAC** The original MPEG-4 AAC has recently been extended with a technique called Spectral Bandwidth Replication, which gives spectacular bandwidth savings for applications like Internet Audio and digital broadcast. MPEG-4 AAC with SBR also known as “High Efficiency AAC” can deliver high quality stereo audio at a mere 48kbit/s.

FEATURES OF MPEG-4 AAC Advanced Audio Coding offers the following features: • **High Quality Compression**: At a data rate of 96 kbps for a 44.1kHz/16bit stereo signal, AAC provides CD audio quality. Furthermore, AAC fulfils the requirements for studio sound quality, as defined by the European Broadcasting Union (www.ebu.ch), at a data rate of 64 kbps per channel. High Efficiency AAC offers stereo, near-CD quality at a mere 48 kbps • **Multichannel Support**: In addition to mono and stereo, AAC supports various surround sound configurations (e.g. 5.1 or 7.1 channels), up to 48 audio channels. • **Low Computational Complexity**: Most AAC encoder implementations are real-time capable. Low computational demands make AAC the ideal codec for high performance audio archiving as well as for mobile applications. • **Wide Application Range**: AAC supports a large set of audio sample rates, ranging from 8 kHz up to 96 kHz. Thus it ideal for high quality audio in many applications with limited channel or memory capacities.

Is that what h.264 is, mp4 format??No. h.264 is an extension of MPEG4 video. There are many profiles in MPEG4. The most well known at this time are Simple and Advanced Simple Profile. As opposed to SP MPEG4 or ASP MPEG4 h.264 is AVC MPEG4. It is however important to note where the differences between SP and ASP MPEG4 are small AVC MPEG4 is almost a different beast entirely

This means SP-MPEG4, ASP-MPEG4 these two are somewhat similar
While AVC-MPEG4 called as H.264 also MPEG4-Base-10 also H.26L is far much superior than above two.

SP: Simple Profile
ASP: Advanced Simple Profile
AVC: Advanced Video Coding

MPEG Listening Test MPEG conducted a listening test in the course technology selection for the MPEG-4 Audio Extension 1. For the MPEG test, experienced listeners used the MUSHRA blind test method to relatively rank the items compared to a known unencoded reference. As is typical for MPEG testing, items known to be difficult to properly encode (e.g. harpsichord, glockenspiel, pitchpipe, male German speech, castanets) were used. Even with this high bar, HE AAC displayed good absolute performance and provided a clear improvement to AAC.

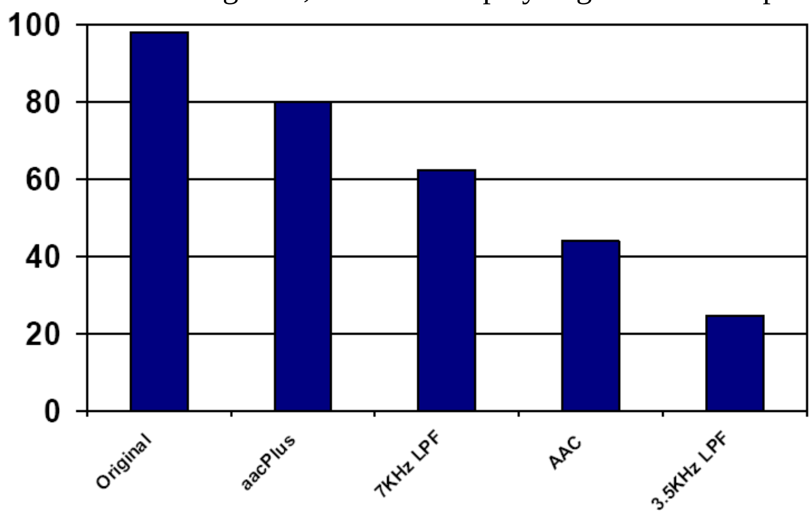


Figure: MPEG test result at 24Kbps mono
Note: LPF = Low-pass Filter

Digital Radio Mondiale Listening Test Digital Radio Mondiale is an international industry consortium creating the new global standard for digital broadcasting in the AM bands below 30 MHz. Their testing for technology selection targeted 24Kbps mono and also used the MUSHRA method with average broadcasting content. HE AAC again demonstrated superior performance, even when compared to AAC at 32Kbps.

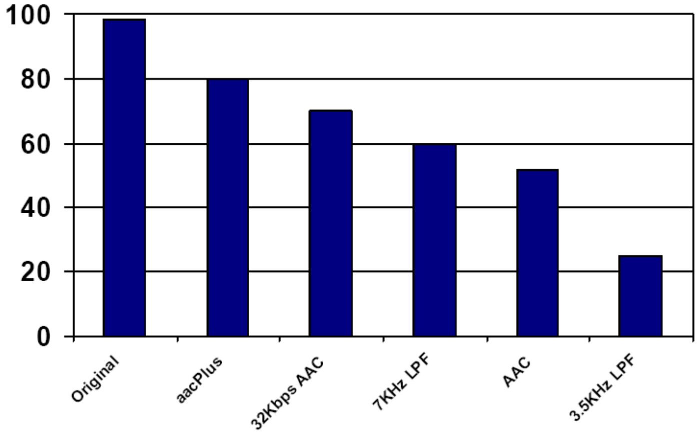


Figure: Digital Radio Mondiale listening test results for 24Kbps mono
Note: "AAC Wideband" is AAC at 32Kbps mono

European Broadcasting Union Internet Audio Listening Test In 2002, the European Broadcasting Union (EBU) conducted its second round of Internet audio tests. This suite of tests compared a variety of codec at several bitrates with Coding Technologies' aacPlus included in the mix at 48Kbps. These results show aacPlus (now standardized as HE AAC) as the clear winner, significantly outperforming proprietary competitors and improving over other standards. The EBU report also went on to credit the SBR technology in particular (used in both the aacPlus and the mp3PRO submissions) as being the only fundamental enhancement to audio compression as compared to the same suite of tests two years earlier.

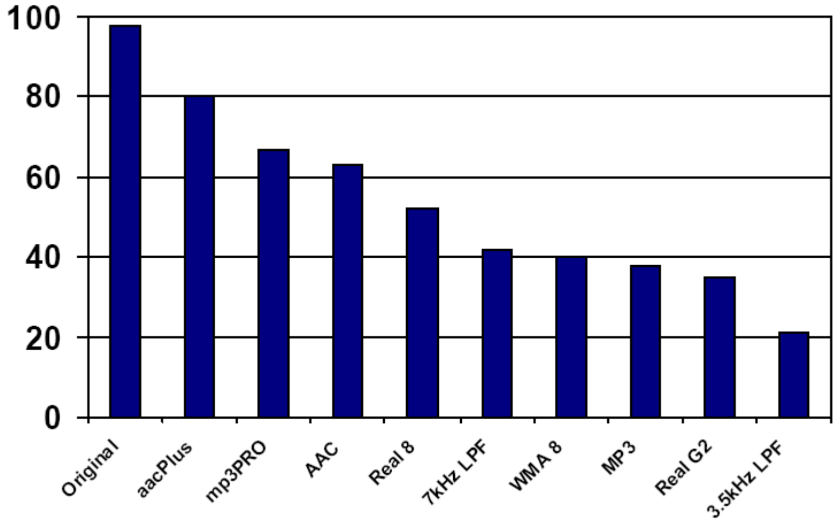


Figure: EBU listening test results for 48Kbps stereo

Markets The unique capability of MPEG-4 HE AAC to achieve high-quality at very low bit rates not only enhances existing markets but also enables new markets for digital audio. Where bandwidth is constrained, the value of "High-efficiency AAC" is magnified. It not only enhances audio-only services, but also video services like digital TV. By coupling HE AAC with MPEG-4 Video, more bits can be allocated to the video signal without degrading the quality of the audio signal. This is true for mono, stereo, and multichannel applications. Combined with the new MPEG/ITU Advanced Video Coding standard (included in MPEG-4 as part 10), even more significant gains in quality are possible.

Mobile streaming and download There is significant expectation for mobile multimedia services associated with the new 2.5G and 3G mobile service networks. The problem is that for much of the coverage area the peak bandwidth in these networks is generally around 144Kbps with individual users sustaining connections of about 40Kbps. Delivering quality video over this type of connection is problematic and may not meet the consumer expectation. HE AAC is well suited to solve this problem by being able to provide consumer-grade download and streaming audio services within today's bandwidth. 48Kbps HE AAC provides CD-quality stereo programming while at 32Kbps, it provides excellent quality stereo programming. These bit rates combined with the growing market for subscription audio services show a strong business opportunity starting in 2003 on into 2004 and beyond.

Digital broadcast via satellite and cable aacPlus gained its first commercial success with XM Satellite Radio. Over a given transponder bandwidth, aacPlus allowed XM Radio to offer more channels at a higher quality than the competing Sirius Radio system which uses the proprietary PAC codec from Lucent. aacPlus was also selected by the Digital Radio Mondiale consortium as part of the standard for Shortwave and AM digital radio. This heritage has brought credibility for High-efficiency AAC in the open standards world of digital satellite and cable broadcasting. As operators look to enhance their services with more channels or with high-definition, the efficiency of HE AAC gives them more options to either consolidate audio bandwidth to make room for more video or to layer more audio services like multi-lingual and 5.1 surround. Since Spectral Band Replication (SBR) is being added to the MPEG-2 standard as well, operators have the flexibility to use HE AAC within the MPEG-2 or MPEG-4 context as desired.

SBR Technology SBR or Spectral Band Replication was developed by Coding Technologies as a generic method to significantly enhance the efficiency of perceptual audio codecs like MPEG Layer-3 (mp3) and MPEG AAC. SBR does not replace the core codec, but rather operates in conjunction with it to create a more efficient superset that can cut the required bit rate in half. MPEG-4 Audio uses SBR in conjunction with AAC to create the “High-efficiency AAC” profile which Coding Technologies has given the name “aacPlus”.

Present in both the encoding and decoding process, SBR leverages the correlation between the low and high frequencies in an audio signal to describe the high-end of the signal using only a very small amount of data. This SBR data describing the high-frequencies is coupled with the low-frequency compressed data from the AAC codec. Once combined, the complete HE AAC bitstream contains enough data to recreate the original signal.

For example, to create 48Kbps stereo HE AAC, the encoder generates two signals: an MPEG AAC signal at about 42Kbps and a SBR signal at about 6Kbps. The SBR signal is then placed into the MPEG AAC auxiliary fields as defined in MPEG-4 and sent out as a complete 48Kbps MPEG-4 HE AAC bitstream.

Since the SBR data is placed within the AAC auxiliary fields, the enhanced signal will be accepted by both an existing AAC and a new HE AAC decoder. If sent to an AAC decoder, only the low-frequency audio signal will be recognized and decoded. If sent to an HE AACdecoder, the SBR and the AAC will be decoded to recreate the full frequency signal. This technique makes the new Profile forward compatible with AAC. Also, since the HE AAC decoder contains a full-fledged AAC decoder, it is able to decode both the “Plain AAC” and “Highefficiency AAC” MPEG-4 Audio profiles. This combination makes HE AAC backward compatible with AAC.

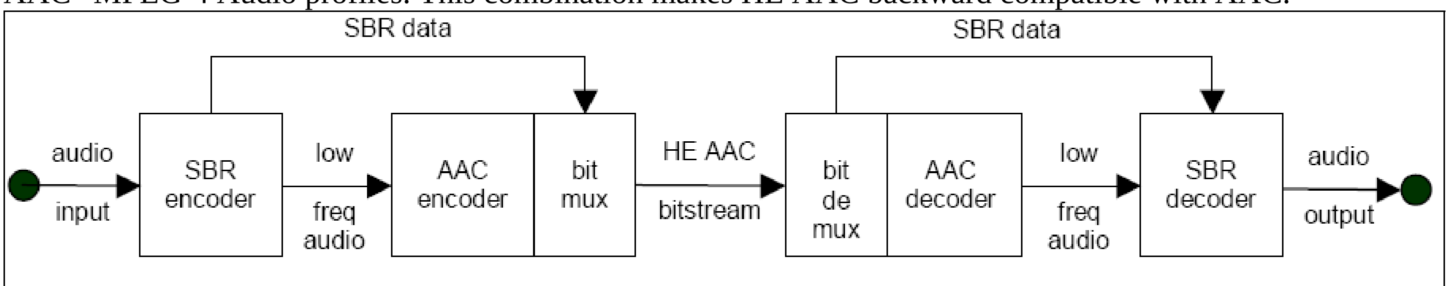
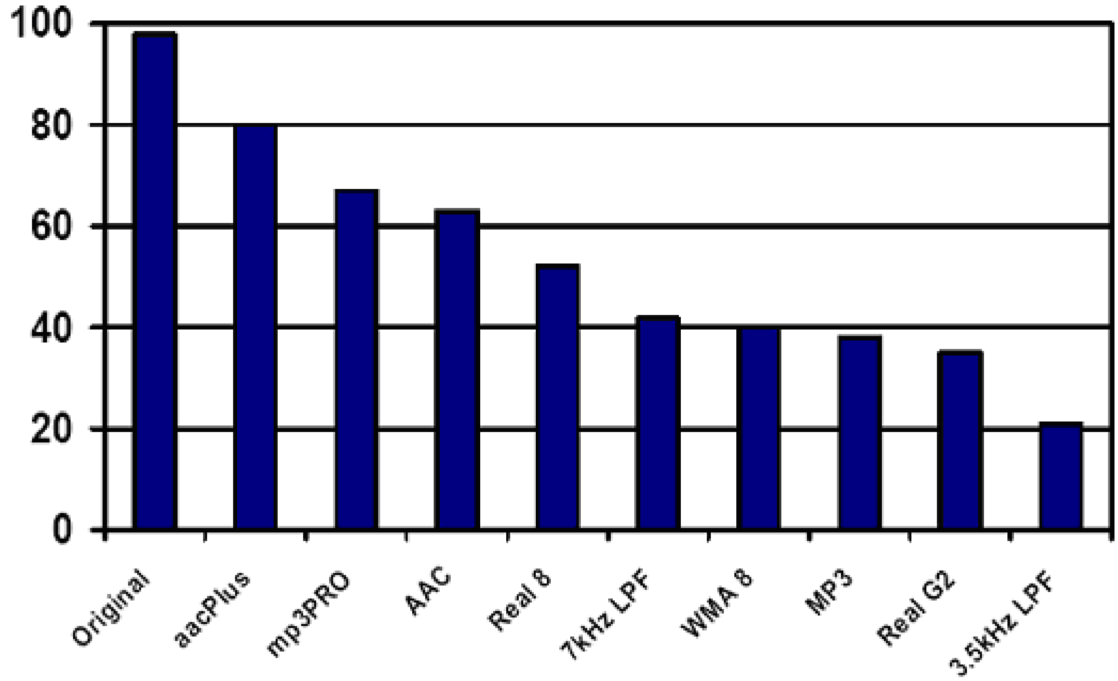


Figure: MPEG-4 aacPlus codec structure showing the integration of SBR and AAC

How does MPEG-4 video compare to Windows Media Video 8 codec or other state-of-the-art codecs? It is difficult to measure the compression performance of a standard, as the quality experienced by a user depends on the quality of each encoder and decoder implementation and what parts of the standard they use. Current consensus indicates that although MPEG-4 has the capability of providing a savings in bit rate relative to older standards such as MPEG-2, other design efforts such as Windows Media Video 8 codec and the ITU-T H.26L project have demonstrated significantly better solutions given advances in codec engineering and associated efficiency. As a result, MPEG is working with the ITU-T to develop a new codec for use in MPEG-4. This new codec will be incompatible with that currently defined in MPEG-4.



OBJECT ORIENTED INTERACTIVITY
MPEG-4 was built from the ground up to allow for object-oriented interactivity. The MPEG-4 standard embraces the concept of object-based coding and incorporates not just audio and video, but also 2D and 3D objects, animation, text, interactivity, data and more. Because MPEG-4 is object-based, it is possible to construct multimedia scenes which revolutionize the possibilities of interactive media. With MPEG-4, interactive programming seamlessly integrates 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

POWERFUL INTERACTIVE MEDIA AUTHORING To compose these various media elements, an MPEG-4 presentation consists of audio and/or visual objects and scene description information, called BiFS (binary format for scenes). The precise description of the object layout and interaction is critical given that the scene composition and rendering in MPEG-4 are done on the presentation level. MPEG-4's scene description tools also bring interactivity to the content, as each object can be enabled with unique functionality. Thus, with object-based interactivity, all media elements – whether audio, video, 2D or 3D graphics, text, animations, etc. – can be easily authored and managed. 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 is impossible with other technologies and standards.

SERVER-CLIENT INTERACTION The power of MPEG-4 object-based interactivity is that the interactivity can be local (watching a live stream or playing back from a stored file such as on a CD-ROM or DVD), or remote – interacting with a media server. Client-side interactivity includes object-to-object interaction (e.g. one object can trigger another) or user interaction (e.g. a user clicks on an object and it zooms to the foreground).

BENEFITS OF OBJECT ORIENTED INTERACTIVITY

• **Complete Standards-based Implementation**: To facilitate interoperability across various MPEG-4 servers and terminals, the presentation layer, communication and messaging protocols are defined by the standard
• **Flexibility and Creative Control**: The experience of the media consumer and the user interface can be completely controlled by the content creator. The user only needs to deal with a single, consistent user interface.
• **Create Once, Distribute Everywhere**: Unique objects within an MPEG-4 scene can be swapped in or out for ease in re-purposing content • **Delivery over Any Network**: The separation of the media objects' description and presentation enables the ability to deliver interactive MPEG-4 data across broadcast, broadband or wireless networks
• **Delivery to Any Device**: MPEG-4 set of technologies enable content flow seamlessly across media delivery networks to these end devices whether Given the proliferation of presentation devices cell phones, PDAs, PCs, or set-top boxes
• **Integrated Security and Copyright Protection**: MPEG-4 addresses the concerns of content providers by integrating interfaces to digital rights management systems deeply into the system.
• **Powerful Digital Asset Management**: Treating meta-data as an additional associated object within the scene allows for a complex set of archival capabilities.

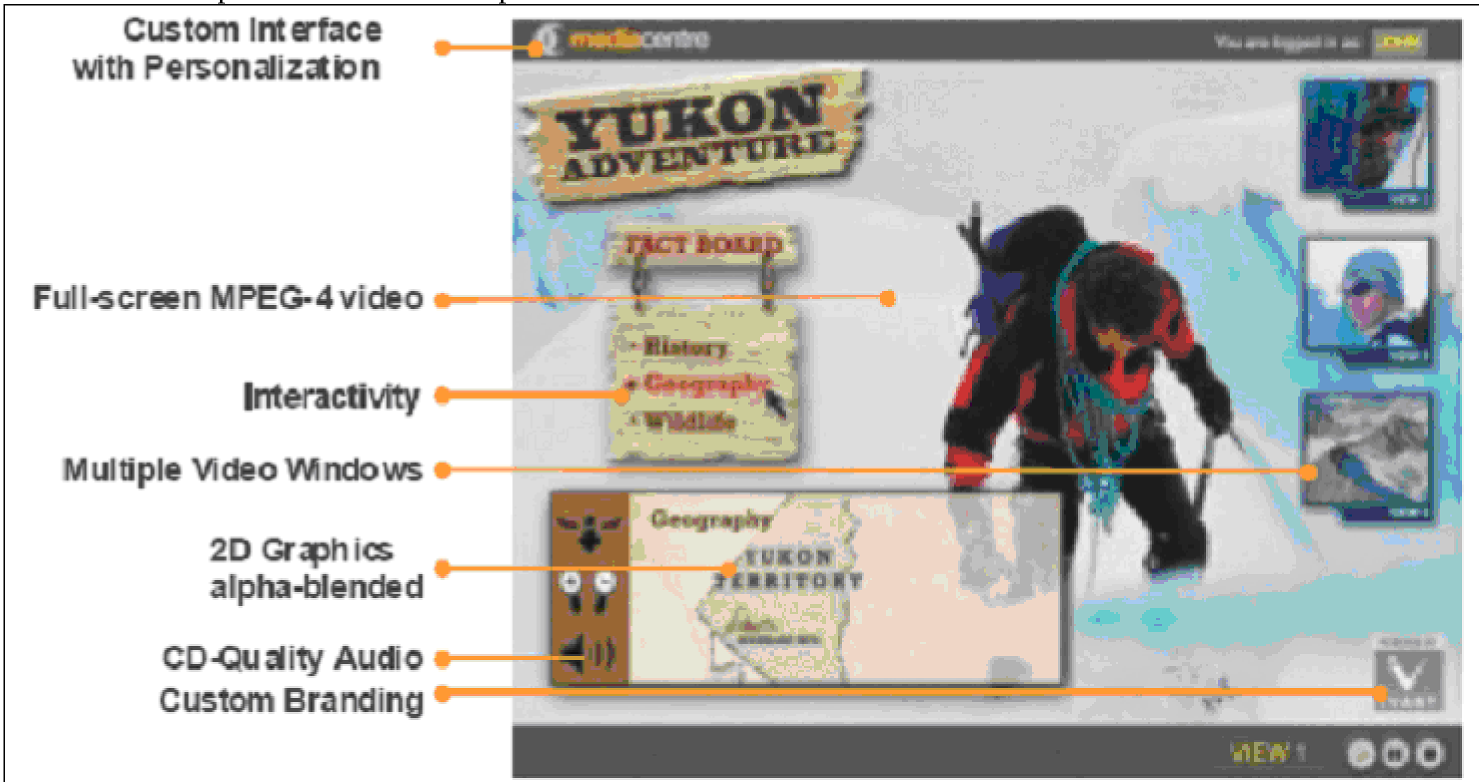


Figure 4: An MPEG-4 scene using multiple objects and elementary streams (Source: iVAST[ix])

BENEFITS OF AN OPEN STANDARD Open standards are part of your everyday life. When you switch on your TV or pick up the phone, when you load A4 or letter-sized paper in a photocopier, you are benefiting from open standards. A multi-vendor model automatically creates a larger market and delivers what users want: choice. In addition, using an open standard like MPEG-4 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.

What is 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.

An immeasurable amount of multimedia information is available today—in digital archives, on the Web, in broadcast data streams, and in personal and professional databases—and this amount continues to grow. Yet, the value of that information depends on how easily we can manage, find, retrieve, access, and filter it. The transition between two millennia abounds with new ways to produce, offer, filter, search, and manage digitized multimedia information. Broadband is being offered with increasing audio and video quality, using ever-improving access speeds on both fixed and mobile networks. As a result, users are confronted with numerous content sources. Wading through these sources, and finding what you need and what you like in the vast content sea, is becoming a daunting task. On the other end of the usage spectrum, the task at hand is becoming equally challenging. For content providers who have many (digital and analog) sources, the effort to find the right piece of content increases with the amount available. The question of identifying and managing content isn't just restricted to database retrieval applications such as digital libraries but extends to areas like broadcast channel selection, multimedia editing, surveillance, and home-entertainment devices. MPEG-7—developed by the Moving Picture Experts Group (MPEG)—Addresses this content management challenge.

Example An image sensor can produce visual data in the form of objects with associated physical measures and time information. A processing device can locally or remotely process these to verify if certain programmed conditions are met. A video recording device could receive descriptions of the multimedia information associated with a program that would enable it to record, for example, only news with the exclusion of sports. We could describe products from a company so that a machine could respond to unstructured queries from customers.

What is 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.

MPEG-21 context

The appetite for content consumption and easier access to information continues to increase rapidly. Access devices are also evolving so that we have a wide choice of terminal and network capabilities. The result is that people, both in their personal and professional lives, are increasingly creators as well as consumers of digital media. They thus demand solutions that deliver accessible and advanced multimedia creation and consumption on many platforms. These new content providers and the traditional media sources share a core set of concerns: management of content, repurposing content based on user preferences and device capabilities, protection of rights, protection from unauthorized access/ modification, protection of privacy of providers and consumers, and so on. The MPEG-21 Multimedia Framework initiative aims to meet these challenges by enabling the use of multimedia resources across a wide range of networks and devices.

MPEG-21 is the newest of a series of standards being produced by the Moving Picture Experts Group (see http://mpeg.telecomitalialab.com), more formally known as ISO/IEC JTC 1/SC 29/WG11. MPEG has a long history of producing multimedia standards: MPEG-1 (1993), MPEG-2 (1996), MPEG-4 (1999), and MPEG-7 (2001).

MPEG-21 objectives

A key assumption of MPEG-21 is that every human is potentially an element of a network involving billions of content providers, value adders, packagers, service providers, consumers, and resellers. Thus, besides client-server-based applications, peer-to-peer networking and the resulting flexibility of user roles has been an underlying part of MPEG-21 thinking since the early days of the standardization process.

FAQ

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. *[NOTE: source of this text is a document from 19/11/2002]*

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.

Should I wait for MPEG-7 and skip MPEG-4 for my streaming needs?

MPEG-7 does not compress multimedia data for later playback. Instead, it is a standard for describing multimedia content. For more information on MPEG-7, see Overview of the MPEG-7 Standard and MPEG-7 Alliance.

I tried viewing an MPEG-4 video created by one company on another company's viewer and it crashed? If MPEG-4 is a standard why did it not work?

This can be due to a host of reasons. An MPEG-4 decoder should **never** crash. A decoder must be resilient to all error conditions. In this case, the problem could be in the encoder, in the decoder, or the bitstream could simply be corrupted.

M4IF is carrying out interoperability tests to make sure the products of different vendors do indeed work together seamlessly.

Can I do interlaced video with the simple visual profile i.e. MP4-SP?

No. Interlaced Video is not supported in the Simple Profile. It is supported in Advanced Simple from Level 4 and up.

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.

I have heard that the MPEG-4 file format is the same as quicktime, is that true?

No. Although Apple's QuickTime file format was adopted as the basis for the MPEG-4 file format, it has gone though many changes in order to support all the functionality of MPEG-4. Apple has however, played a key role in the development of MP4, the MPEG-4 file format.

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.

Question: Did Apple invent MPEG-4? Is MPEG-4 the same as QuickTime?

Microsoft's Answer: Many companies have contributed to the development of MPEG-4. Microsoft, for example, has made more than 100 formal contributions to the MPEG-4 standardization process and has patents that are relevant to MPEG-4 video implementations.

MPEG-4 is not QuickTime, nor was it developed by Apple. The MPEG-4 standards body did originally select a file format or "container" based on QuickTime, but that proposal was later modified and the MPEG-4 file container is not compatible with QuickTime. The MPEG-4 file container is a very small part of the overall MPEG-4 standard.

Which visual profiles support arbitrarily shaped objects?

The following visual profiles support arbitrary shaped objects

- Core
- Main
- N-bit
- Advanced Coding Efficiency
- Core Scalable

The following visual profiles **do not** support arbitrary shaped objects

- Simple

- Simple Scalable
- Advanced Real Time Simple
- Advanced Scalable Texture
- Advanced Simple
- Fine Granularity Scalable

Where can I find a list of supported tools for a profile?

The visual specification (ISO/IEC 14496-2) contains a list of all tools that can be used for a particular profile. An abridged version, listing the most popular profiles is also available at the M4IF website.

I have heard of a Level 0 for Simple Profile, what is it?

Level 0 of the simple profile was designed to target to the wireless industry where screen sizes are small, and processing power is low. Level 0 has the same functionality as the other Simple profile levels but with the following additional restrictions.

a. The maximum frame rate shall be 15 frames per second;
b. The maximum f_code shall be 1;
c. The intra_dc_vlc_threshold shall be 0;
d. The maximum horizontal luminance pixel resolution shall be 176 pels/line;
e. The maximum vertical luminance pixel resolution shall be 144 pels/VOP;
f. If AC prediction is used, the following restriction applies : QP value shall not be changed within a VOP (or within a video packet if video packets are used in a VOP). If AC prediction is not used, there are no restrictions to changing QP value.

Level 0 is available in the Streaming Video Profile Amendment.

Does MPEG-4 Video support H.263?

Yes. H.263 Baseline (No Annexes) is incorporated as part of MPEG-4. It is known as 'short header' inside the specification.

I have heard that MPEG-4 visual is wavelet based. Is that true?

Yes and No. The visual specification still uses the Discrete Cosine Transform (DCT) for its video compression, however, the still texture part uses the Discrete Wavelet transform.

I have heard rumours that MPEG is adopting H.26L to replace MPEG-4, is that true?

No. MPEG has started talking with the video compression group inside ITU for joint compression development. Output from this joint venture would result in a new part to the MPEG-4 standard, to be finalized in early 2003. It will become part of the MPEG-4 framework.

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

CONCLUSION

SP-MPEG4, ASP-MPEG4 are two different but somewhat similar coding technologies having performance/ capabilities inferior to mighty H.264.

AVC-MPEG4, H.264, MPEG4-Base-10, H.26L are one and same thing and is entirely a different beast from above two and far much superior to above two. Also this format should NOT be pronounced/called only “MP4”, Because this could be a disgraceful act as above mentioned two profiles (SP & ASP versions of MP4) are “simple” ones.

So what I have learnt from this literature is that from all MP4 profiles in Audio and Video the “AAC-SBR” or “AAC with SBR” (Spectral Bandwidth Replication) encoding should be used for Audio for the target file and AVC should be used for Video for target file for the ultimate compactness.

Now ultimate compactness is achieved but this file will not be able to play on most software players even on Quicktime-6.5.1 I cant say anything about yet to release Quicktime-7 whether it would be able to play it. Realplayer10 (RealOne) cant play it..