Windows Media 9 Series

— a platform to deliver compressed audio and video for Internet and broadcast applications

Jordi Ribas-Corbera
Microsoft Windows Digital Media Division

Microsoft® Windows Media® 9 Series is a set of technologies that enables rich digital media experiences across all types of networks and devices. These technologies include an encoder to author the multimedia content, a server to distribute the content, a Digital Rights Management (DRM) system to let content owners set usage policies, and a variety of players to decode and render the content on personal computers and other consumer electronic devices. These components are built on top of a programmable and extensible platform that enables partners to build tailored applications and services.

This article provides a high-level overview of the technologies in Windows Media 9 Series, with a particular focus on the different audio and video codecs available. Applications and services for broadcast (e.g., IP datacasting via DVB) are also discussed.

Introduction

Windows Media 9 Series is the latest generation of digital media technologies developed by Microsoft [1]. Although the origins of Windows Media focused on streaming compressed audio and video over the Internet to personal computers, the vision moving forward is to enable effective delivery of digital media through any network to any device.

A wide range of applications

Fig. 1 illustrates a variety of examples of how Windows Media technology is being used today. In addition to Internet-based applications (subscription services, video-on-demand, Web broadcast, and so on), content compressed with Windows Media codecs is being consumed by a wide range of wired and wireless consumer electronic devices (such as mobile phones,

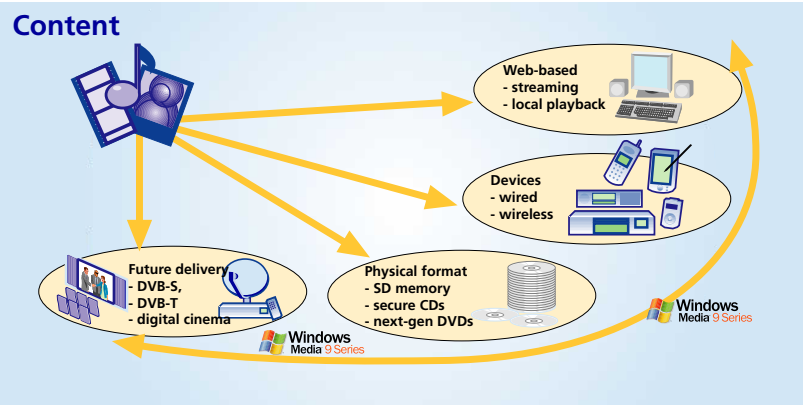

Figure 1
Examples of current Windows Media technology applications

EX1053
Roku V. Media Chain
U.S. Patent No. 10,515,191

DVD players, portable music players, and car stereos) [2]. Windows Media content can also be delivered to consumers in physical formats—for instance, using the Secure Digital (SD) memory card [3], or on CD or DVD using the emerging HighMAT™ format [4].

In the terrestrial and satellite broadcast space, a recent project at the International Broadcasting Convention (IBC) demonstrated how to deliver Windows Media 9 Series content via DBV-T and DVB-S [5]. As another example, Windows Media technology is also used to compress movies in high definition and multi-channel audio, and these films are being projected in public digital theatres in the USA.

End-to-end delivery

All of the applications mentioned so far require a set of building blocks or components that permit the deployment of complete end-to-end solutions. The fundamental components of Windows Media 9 Series are illustrated in *Fig. 2* and can be classified in three steps: authoring, distribution and playback.

Authoring

Authoring is the process of creating and encoding digital media. The basic encoding software provided by Microsoft is called Windows Media Encoder 9 Series. It is a flexible encoder that can compress audio and video sources for live or on-demand streaming by using the Windows Media codecs.

At the same time, there are alternative encoding solutions provided by third parties that are built on top of the Windows Media porting kits (e.g., hardware encoders from companies such as Optibase, Tandberg Television, Texas Instruments) or the Windows Media software development kits (SDKs) (e.g., software encoders from companies like Accom, Adobe, Avid, Discreet and Sonic Foundry).

Distribution

The distribution over the Internet of content, compressed with Windows Media codecs, is generally done by a Windows Media Services server. Windows Media Services version 4.1 is an optional component in Windows 2000 Server, and Windows Media Services 9 Series is expected to be an optional component in Windows Server 2003.

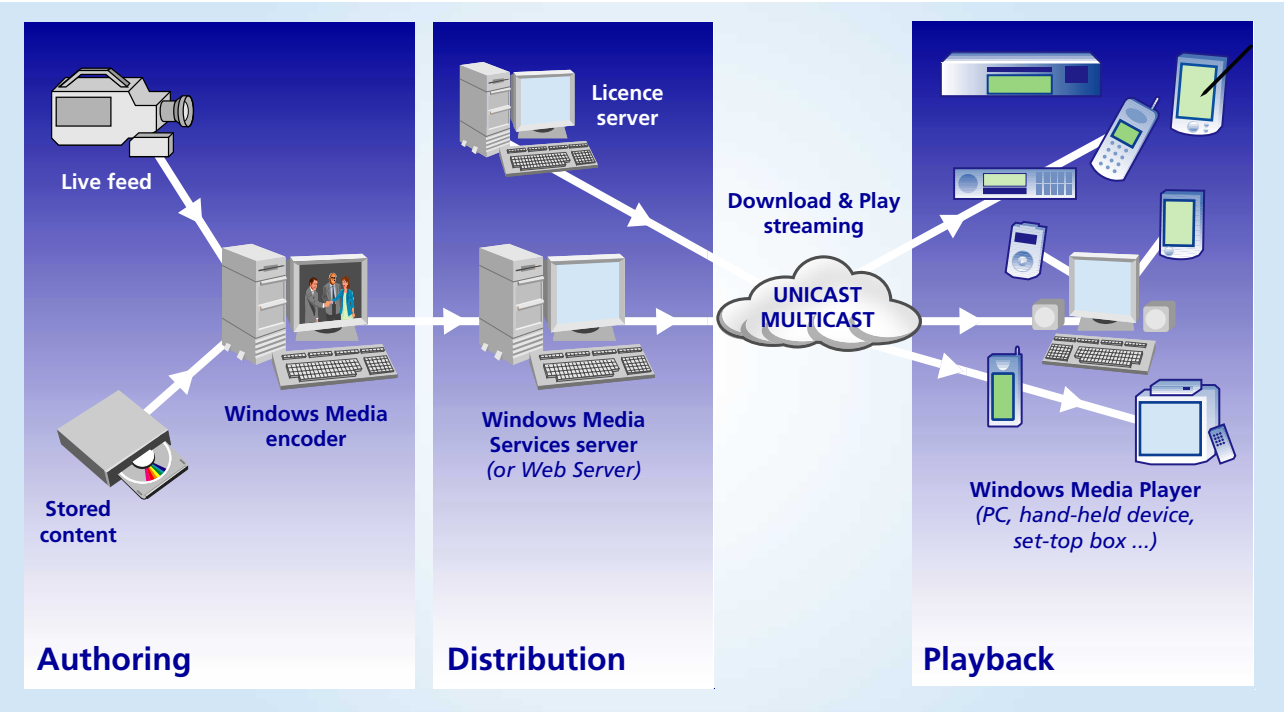

Figure 2
End-to-end delivery of Windows Media content: authoring, distribution and playback. The DRM system protects content, based on policies set by the owner.

The new server supports more features for advertising and corporate scenarios and is twice as scalable – it doubles the number of customers that can receive a media clip at the same time.

A server can either stream the clip (transmit it with as little delay as possible) or download it (transmit and store it) into a user's playback device. The transmission of the clip can be performed live (for news, sports, concerts or similar events) or on demand (for music videos, movies-on-demand and so on). When streaming a media clip, the server adapts its throughput and re-transmits lost packets intelligently, using feedback from the network quality metrics. For on-demand streaming, the latest server takes advantage of the additional bandwidth available (above the average bit-rate of the clip) to reduce the start-up delay. In addition, such a server reduces the likelihood of losing the connection (which manifests as playback "glitches" and re-buffering to the viewer) by sending more data to the playback device, so that when there is network congestion the device can continue playing.

A robust and scalable server is essential for Internet delivery, but obviously having a solid network connection is also critical. The latter is addressed by content delivery networks (CDNs) such as Akamai, Digital Fountain and SMC. The combination of robust servers and networks result in TV-like experiences that are remarkably superior to those of Internet streaming in the recent past.

As mentioned earlier, one can also deliver Windows Media content via physical formats such as CD or DVD, or through other networks such as DVB. We will discuss transmission of Windows Media over DVB in more detail later in this article.

Playback

The final step of the end-to-end delivery process is playback, which consists of decoding and rendering the compressed data in the user's device. On a personal computer, Windows Media Player and a variety of other players built by third parties (such as MusicMatch Jukebox or RealOne Player) can decode and play Windows Media streams and files. As illustrated in *Fig. 2*, a wide variety of consumer electronic devices can also play Windows Media content [2]. Like in the encoder case, third parties can build such hardware playback devices on any platform by using Windows Media porting kits.

Digital Rights Management

A critical component of the end-to-end delivery is Digital Rights Management (DRM), which we represent by the licence server in *Fig. 2*. DRM works across the three steps of the media delivery system, so we will discuss it separately in this subsection.

The DRM technology used by Windows Media lets owners encrypt their media products and services and specify the usage rules and policies. For example, the owner may decide that the user can play the digital media until a certain date, or can play it a given number of times. Or the owner can let the user copy the digital media to a certain number (and type) of devices.

In the typical Internet scenario, the content owner encrypts the (compressed) digital media stream using DRM. When a viewer selects the stream, the playback device connects to the licence server, which offers a licence for the content. The viewer then decides whether to accept the terms and price of the licence and, if so, the licence is downloaded into the viewer's device. Then the viewer is able to decrypt and play the content according to the terms of the licence.

Designing a complete DRM service is a challenging project. The system needs to be secure (with the capability of upgrading quickly), flexible (it must accommodate the desires of the content owner and device manufacturers), and user-friendly. Windows Media DRM fulfils the requirements of many scenarios and is one of the leading solutions in the market.

Platform Components

Windows Media provides a development platform in addition to a set of components for authoring, distributing and playing digital media. Windows Media Encoder, Windows Media Services and Windows Media

Player meet the requirements of a fair number of applications, but they are essentially examples that showcase the capabilities of the platform. These components are built on top of SDKs that can also be used by third parties to develop their own encoders, servers and players, tailored to specific applications. By providing a platform with state-of-the-art compression and delivery technology, Windows Media enables third parties to create their own leading-edge customized solutions [6].

Figure 3
The Windows Media 9 Series SDKs

Fig. 3 shows the Windows Media 9 Series SDKs. Software such as Windows Media Player and the MusicMatch Jukebox are built using the Windows Media Player SDK. A large number of companies have downloaded and licensed one or more of these SDKs to build their own digital media systems.

In addition, companies can build Windows Media-based software and hardware solutions by using the porting kits described later in this article.

Audio and video codecs

Windows Media audio and video codecs are key components for the authoring and playback of digital media. The table below lists the audio and video codecs that ship in Windows Media 9 Series. Each codec uses different technology and bit-stream syntax and, consequently, is not compatible with the others; e.g., the bit-streams of Windows Media Audio 9 Professional cannot be decoded by a Windows Media Audio 9 decoder, and vice-versa. Older codecs (such as Windows Media Video 8, ISO MPEG-4 Video and others) are also supported for backward compatibility, but they are not listed here as the focus of this article is on the new 9 Series codecs.

Windows Media Audio 9 Codecs	Windows Media Video 9 Codecs
Windows Media Audio 9 Windows Media Audio 9 Professional Windows Media Audio 9 Lossless Windows Media Audio 9 Voice	Windows Media Video 9 Windows Media Video 9 Screen Windows Media Video 9 Image

Windows Media Audio (WMA) 9

The Windows Media Audio 9 codec is the most popular audio codec in Windows Media and is commonly known as WMA. The decoder (or bitstream syntax) was frozen more than four years ago, and only the encoder has been improved since then. WMA 9 is the third iteration of backward-compatible improvements. Maintaining backward compatibility has been critical to support the consumer electronics manufacturers who choose to build devices that play WMA.

The new WMA encoder enhances the one-pass constant-bit-rate (CBR) encoding mode (the only mode supported by WMA in previous versions), using improved rate control and masking algorithms. It adds new two-pass and variable-bit-rate (VBR) modes that further enhance quality over the one-pass mode.

For any of the codecs, one-pass CBR encoding is required for live encoding and transmission, while two-pass CBR is appropriate when encoding off-line for on-demand streaming. VBR modes are recommended when the compressed clips will be downloaded into the user's device (for download-and-play applications). Even

though VBR-encoded clips can also be streamed (with the new server), the bit-rate fluctuations within the clips are usually high and their transmission requires long buffering delays. There is also a peak-constrained VBR encoding mode to create bitstreams meant to play back on devices with a constrained reading speed. The WMA 9 codec supports all of these encoding modes.

WMA 9 also supports a large list of encoding settings for mono and stereo audio, with bit-rates ranging from 5 kbit/s to 320 kbit/s and sampling rates ranging from 8 kHz to 48 kHz. At the typical CD sampling rate (44.1 kHz), we found that most users select rates between 48 kbit/s to 128 kbit/s to achieve CD-like quality, depending on their sensitivity to compression artefacts and bandwidth availability. A small percentage of critical listeners may require higher quality to be satisfied, which is why higher encoding rates are also provided.

There has been a lot of discussion about which audio codec technology produces the best quality. There are quite a few good audio codecs available, and many opinions. We find that experts still differ on what is the best compressed sound. Some prefer codecs that preserve more bandwidth, and hence produce a richer sound, while tolerating some distortion at high frequency. Others prefer codecs that generate a more muffled sound with minimum high-frequency distortion.

The audio quality of the latest WMA 9 has not been evaluated by independent studies yet. Prior versions have been extensively evaluated and have been ranked at the top of some studies. For example, WMA 8 was recently chosen over MP3 and RealAudio 8 in a study by the "golden ears" of *Sound & Vision* [7]. Other studies have reached different conclusions for a variety of reasons [1], including differences in content and testing conditions, or sometimes simply because the listeners had other subjective preferences. As a result, audio experts in organizations usually base their decisions on their own subjective tests.

A new objective for Windows Media 9 Series has been to develop compression technology that goes beyond CD quality. The first major step in that direction is the introduction of the Windows Media Audio 9 Professional codec.

Windows Media Audio (WMA) 9 Professional

The WMA 9 Professional codec is the first Windows Media codec for audio that supports high-resolution (up to 24 bits per audio sample, and sampling rates of up to 96 kHz), and multiple channels (up to eight discrete channels) for typical 5.1 or 7.1 speaker configurations in high-end home systems or commercial digital theatres.

An important application for this codec is the encoding of multi-channel music and movie sound tracks at Internet broadband rates, where there is currently no suitable, widely deployed codec in the industry. For example, the minimum bit-rate of Dolby's AC-3 codec is 384 kbit/s (for 5.1 channels), which provides very little bit-rate for video on DSL/cable connections. WMA 9 Professional can encode 5.1 channels at as low as 128 kbit/s, and 192 kbit/s appears to be the "sweet spot" for the technology. As a result, there is enough bandwidth left for encoding video for broadband delivery.

As with the WMA 9 codec, there are also five encoding modes for WMA 9 Professional: one-pass CBR, two-pass CBR, one-pass VBR, two-pass VBR, and peak-constrained VBR. In addition, WMA Professional 9 allows near-lossless compression at its highest-quality VBR setting.

Windows Media Audio (WMA) 9 Voice

Another new codec in Windows Media 9 Series is WMA 9 Voice. This codec compresses mono-only audio at very low bit-rates, which is useful for transmitting digital media through low-rate modem or ISDN connections. The supported bit-rates range from 4 to 20 kbit/s and the sampling rates range from 8 to 22 kHz. At this point in time, WMA 9 Voice only supports one-pass CBR encoding.

1. One of the most comprehensive audio codec studies (which included WMA 8) was recently performed by several members of the EBU [8]. Unfortunately, WMA 8 was compared to other codecs at different settings – in some cases the bandwidth or sampling rate (in Hz) for WMA 8 was set to more than twice that of others. We hope to work more closely with these EBU members in the future to avoid these drawbacks. We believe that future tests will likely include the latest WMA 9 codec.

When compressing audio at very low bit-rates, typical transform-based codecs will generally produce better quality on music, while CELP-based (or harmonic) codecs will produce better quality on speech. WMA 9 Voice is a unique hybrid codec that uses an automatic classifier to detect voice and music, and applies the appropriate coding mode for each segment. When the content contains voice and music, the mode selected depends on the type of audio that dominates. The encoder also provides a manual mode so that the user can select the desired mode for any given segment.

The voice encoding mode is based on a new advanced algorithm. The music mode uses essentially WMA transform technology. Consequently, this codec will provide state-of-the-art quality for both types of audio content, while previous codecs will generally only do well with one type.

Windows Media Audio (WMA) 9 Lossless

The final audio codec in Windows Media 9 Series is a mathematically lossless codec. This codec is competitive with other state-of-the-art lossless audio codecs such as Monkey Audio, and can compress a wide variety of audio sources, from CD resolution and sampling rates up to 24-bit, 96 kHz, 7.1 channel audio.

WMA 9 Lossless is integrated into Windows Media Player 9 Series (in the CD copy function) and can achieve compression ratios of about 2:1 for stereo content. Multi-channel, high-resolution audio clips can often be compressed losslessly with higher ratios.

Selecting appropriate audio codecs

For typical stereo broadcast and Internet broadband applications, the standard WMA 9 codec is recommended. If the audio or film contains a high-resolution or multi-channel track, then one should consider WMA 9 Professional.

The other audio codecs are useful for somewhat more limited scenarios. WMA 9 Voice targets very low rate audio applications (such as modem or ISDN delivery), and WMA 9 Lossless is useful for audio archival.

Windows Media Video (WMV) 9

The WMV 9 codec is the most popular video codec in Windows Media 9 Series and is based on technology that can achieve state-of-the-art compressed video quality from very low bit-rates (such as 160x120 at 10 kbit/s for modem applications) to very high bit-rates (1920x1080 at 6 to 20 Mbit/s for high-definition video). The codec supports all five CBR and VBR encoding modes.

WMV 9 achieves 15 to 50 percent compression improvements over version 8, and the improvements tend to be greater at higher bit-rates. For instance, *Fig. 4* shows a peak signal-to-noise ratio (PSNR) quality plot versus bit-rate for WMV 9, WMV 8 and Microsoft's ISO MPEG-4 video codec (simple profile). The source consisted of 13 typical MPEG clips (i.e., "Stefan", "Akiyo", "Coastguard", "News", "Mobile & Calendar", etc).

Abbreviations

ANSI	American National Standards Institute
ASF	(Microsoft) Advanced Streaming Format
CBR	Constant Bit-Rate
CDN	Content Delivery Network
CPU	Central Processing Unit
DSL	Digital Subscriber Line
DRM	Digital Rights Management
DVB	Digital Video Broadcasting
DVB-S	DVB - Satellite
DVB-T	DVB - Terrestrial
IBC	International Broadcasting Convention
ISDN	Integrated Services Digital Network
ISMA	Internet Streaming Media Alliance
ISO	International Organization for Standardization
ITU	International Telecommunication Union
ITU-T	ITU - Telecommunication Standardization Sector
MMX	(Intel) Pentium-based chip technology
PSNR	Peak Signal-to-Noise Ratio
QCIF	Quarter Common Intermediate Format
SD	Secure Digital
SDK	Software Development Kit
VBR	Variable Bit-Rate
WMA	(Microsoft) Windows Media Audio
WMV	(Microsoft) Windows Media Video

We set a fixed quantization step size for all codecs and used the same mode selection strategy, as it is usually done in MPEG and ITU standard tests. Even though PSNR is by no means an exact measure of video quality, the plot conveys that the visual compression gains also translate to PSNR gains.

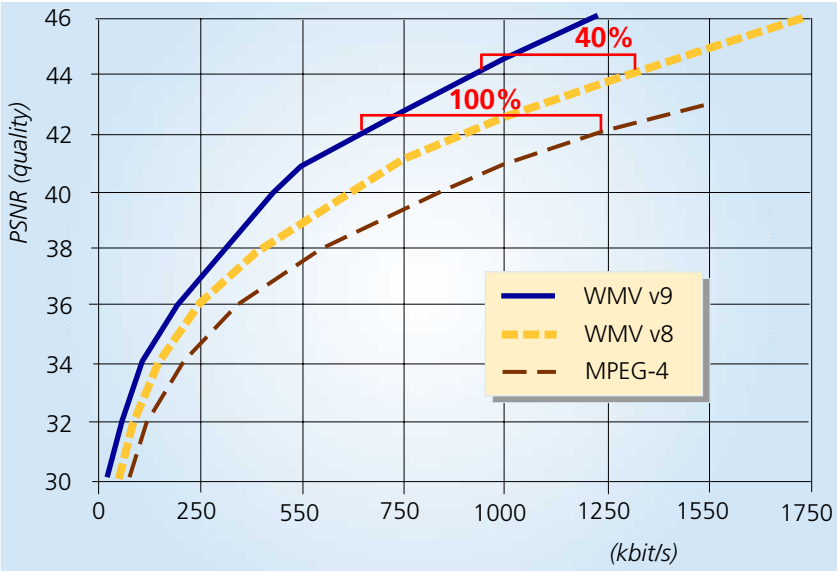

Figure 4
PSNR versus bit-rate for WMV 9, WMV 8 and Microsoft's ISO MPEG-4 video codec, on MPEG clips

With the compression efficiency of WMV 9, one can achieve broadcast-quality BT.601 video at about 2 Mbit/s, and high-quality, high-definition video (e.g., 720p) at high-end broadcast or DVD rates (e.g., 4 to 6 Mbit/s). All broadcast formats are supported, including the high-definition 720p and 1080i variants. The codec includes native interlace compression tools. In addition to 4:2:0, it also supports the 4:1:1 sampling structure to maintain the odd and even field chroma separately in interlaced video (4:2:0 mixes the chroma values between both fields).

Since there are different applications that require different levels of complexity for WMV, we have defined several profiles and levels for interoperability. For example, the "simple profile and low level" supports up to QCIF resolution, 96 kbit/s, and 15 frames/sec (fps), and targets low-end hand-set devices. The "main profile and main level" targets the standard-definition broadcast scenario (the functional equivalent of MPEG-2's MP@ML). The "main profile and high level" is appropriate for high-definition applications (equivalent to MPEG-2's MP@HL). WMV 9 bitstreams targeted to higher-end applications (e.g., standard-definition and high-definition broadcast) are referred to as WMV 9 Professional bitstreams.

The video quality of the latest WMV 9 codec has not yet been evaluated by independent studies. Nevertheless, prior versions of the codec have been shown to provide competitive compression efficiency by such studies. For example, WMV 8 was selected as the video codec that produced the best quality in the latest video codec review of *DV* magazine [9]. In addition, the preliminary conclusions of the recent EBU tests [10] also gave an edge to WMV 8. Like in the audio case, there are plenty of opinions and studies on which codec produces the best video quality, so once again it is recommended that experts do their own tests and reach their own conclusions.

WMV 9 and video compression standards

A typical question that arises is whether the compression efficiency provided by WMV 9 is superior to that of prior standards like MPEG-2 and MPEG-4, or even of the emerging H.264. This question is difficult to answer because standards only define the bitstream syntax and decoder semantics. Therefore, different implementations can produce different quality results. The same could be said about WMV 9, since we anticipate that future (backward-compatible) versions of the encoder built by hardware vendors will likely enhance the compression efficiency over our current version.

Nevertheless, in order to provide some indication of quality comparisons, we conducted internal tests using the well-known Minerva C250 MPEG-2 hardware video codec and the recently-released QuickTime 6 MPEG-4 codec (with the more advanced ISMA profile 1 compatibility level). In those tests, using the same encoding settings, WMV 9 achieved similar quality to MPEG-2 and MPEG-4 with only 1/3 and 1/2 of the bit-rate, respectively [2]. Although there may be better implementations of MPEG-2 and MPEG-4, these large compres-

2. To obtain a DVD data disk with the uncompressed source, the bitstream results, and instructions for duplicating these tests independently, please send an e-mail to stevsk@microsoft.com.

sion gains suggest that WMV 9 provides a significant quality advantage (or bandwidth savings) over codecs that are compliant with these standards. Indeed, recent independent studies have also concluded that the compression efficiency of an early (pre-beta) version of WMV 9 [11] and even of WMV 8 [10] is superior to that obtained by solutions based on MPEG-2 or MPEG-4.

H.264 [12] is a joint ITU-ISO video compression standard that is expected to be completed by May 2003. Since the interoperability process usually continues for months after the standard is completed, it will take some time before competitive standard-compliant products appear on the market. Therefore, it is still too early to reach any solid conclusions on the quality differences between H.264 and WMV 9. Nevertheless, since the rate-distortion-optimized reference codec – developed by the Joint Video Team of ITU-ISO [13] – is believed to provide very high quality, some companies have already made their own initial tests. For example, a fairly comprehensive study, performed several months ago, concluded that the video quality achieved by H.264 and WMV 9 is similar [11], although both codecs have been undergoing enhancements since then.

A concern with H.264 is the high computational complexity required for encoding and decoding. For example, some initial studies show that in order to achieve the compression benefits, the encoding computational complexity needs to increase by an order of magnitude over MPEG-4 simple profile [14]. In addition, they suggest that the decoder complexity of H.264 (main profile) will be about three times higher than MPEG-4 (simple profile) [14].

On the other hand, the decoding complexity of the main profile of WMV 9 is relatively close to that of our (MMX-optimized) MPEG-4 simple profile codec. To be more concrete, decoding with WMV 9 is only about 1.4 times slower, which can be verified easily by using both codecs in the Windows Media Player (or using other MPEG-4 simple profile decoders). Even though one should not draw parallels on such complexity analyses, this information suggests that H.264 decoding complexity could be twice that of the WMV 9 codec, or at least that there is a significant decoding computational benefit of WMV 9 over H.264.

Video smoothing

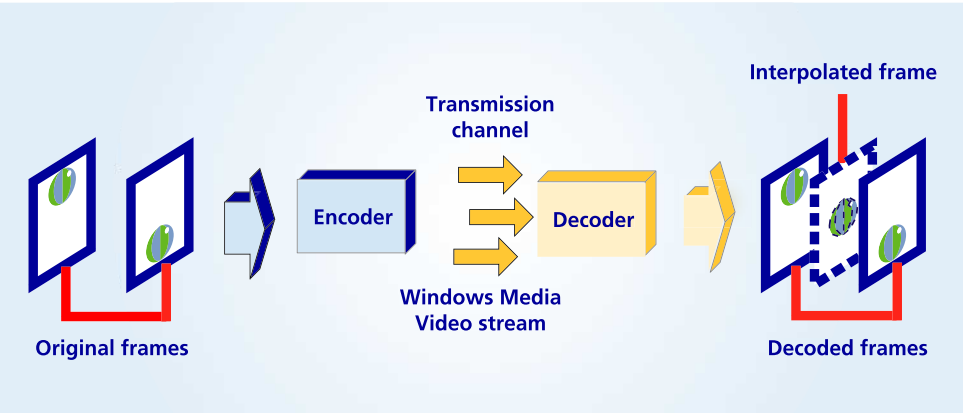

Figure 5
Video smoothing interpolates missing frames at the decoder side using optical flow and warping techniques

A new feature of WMV is the ability to interpolate missing frames after decoding. This feature is called "video smoothing" in Windows Media 9 Series, and it is commonly referred to as *frame interpolation* in the literature [15]. The video smoothing algorithm uses an advanced optical flow estimation technique (on a per-pixel basis), along with warping, to synthesize new frames. *Fig. 5* provides an illustration of the process. The feature is CPU-intensive and is only triggered when there are enough CPU cycles available. For example, end users typically need a CPU operating at 733 MHz or higher to interpolate a video clip at 320x240 pixels resolution from 10 to 30 fps.

This feature is particularly useful at very low bit-rates, where the full-frame rate is difficult to achieve during encoding, and where the low-rate compression artefacts will mask occasional interpolation errors. Video smoothing will remove the jerkiness associated with such low-rate video, and hence it will improve the video quality. Alternatively, a content provider can encode a video clip at a lower frame rate (e.g., 12.5 fps) and a lower bit-rate on purpose, and then let video smoothing up-sample at the decoder side (e.g., up to 25 fps).

Windows Media Video (WMV) 9 Screen

The WMV 9 Screen codec is the next version of a highly efficient engine for palletized computer-screen video compression. This codec is used for capturing applications from a computer desktop for the purposes of creating a demonstration. The entire desktop can be coded and transmitted at rates as low as 28 kbit/s, although when there are natural images embedded in the desktop application, the required bit-rate is usually around 100 kbit/s.

This new version of the codec improves both picture quality and CPU usage when there is motion and natural images, with respect to the earlier version. It supports one-pass VBR and one-pass CBR encoding. A future version will incorporate motion compensation to handle embedded videos on the desktop.

Windows Media Video (WMV) 9 Image

The last new codec in Windows Media 9 Series is WMV 9 Image. This codec enables a user to combine a set of still images to create a video clip using fading, panning, zooming and other effects. One can understand this compression technique as a video codec where the I-frames in the bit stream are followed by a set of motion and transition instructions for each frame (rather than P-frame data).

Selecting appropriate video codecs

In the vast majority of cases, including broadcast applications, the WMV 9 codec is the correct choice.

The WMV 9 Screen and WMV 9 Image codecs are used in more limited (albeit interesting) scenarios, such as when the user wants to compress the video of the computer screen or create a video out of a set of still images.

Windows Media and consumer-electronic devices

Windows Media made inroads in the consumer electronic space when the syntax for WMA was frozen more than four years ago. With a fixed bit stream, a large number of companies started implementing WMA in their products. There are currently more than 170 types of devices (DVD players, CD players, personal digital assistants such as PocketPC, portable music players, car stereos, and so on) that support WMA [2].

Video codec technology has improved significantly during the last few years. Even though we envision new algorithms that can provide incremental benefits, their computational cost appears to be very high. Therefore, we believe that these advances will not be commercially viable for quite some time, perhaps five years or longer. Consequently, the WMV bit stream syntax will be frozen for the first time in version 9, and future versions of WMV (e.g., versions 10, 11, etc.) will be backward compatible.

The process for freezing the WMV 9 bit stream was done carefully with much feedback from chip manufacturers. The objective was to achieve the highest quality with the minimum computation requirement. Algorithms that provided a very minor quality improvement, but required much computation, were either re-designed or occasionally eliminated. As a result, the quality-to-computation balance of WMV 9 is quite compelling. Even in software, one can decode and render some 1080p content in real-time using WMV 9 on a high-end computer (e.g., 2.8 GHz processor) without any hardware acceleration. Nevertheless, we are actively working with ATI and NVIDIA to incorporate WMV decoding in their graphic cards through Microsoft DirectX® Video Acceleration, which should guarantee robust playback for all content at all bit-rates in even lower-end machines.

A common misconception for customers not familiar with Windows Media-based devices is that Windows Media technologies can only be implemented on computers running Microsoft Windows® operating systems. In reality, Windows Media is a generic digital media format that is operating system (OS) and platform independent. The different Windows Media components are broadly licensed and can be ported using Windows Media Porting Kits (which provide specifications and ANSI C source code) to non-OS hardware devices, graphic cards, Linux, Mac, etc. Further, the specification and licence for the Windows Media file container, called Advanced System Format (ASF), is open and publicly available [16], and organizations such as the

Secure Digital (SD) consortium have adopted Windows Media components (in this case, ASF and WMA) in their specifications [3].

Consequently, many partners and licensees compete among themselves to provide better and cheaper chips and OEM solutions that support Windows Media on a variety of platforms. For example, there have been more than 60 licensee companies that have ported WMA to a variety of chips and devices during the last few years. More recently, more than 40 companies have licensed WMV porting kits, and some prototype chips and set-top boxes utilizing the WMV 9 codec have already been demonstrated. Companies that have licensed both WMA and WMV, and have publicly announced it, include ARM, ATI, Cirrus Logic, Equator, ESS Technology, LSI Logic, MEI, Tandberg Television, Texas Instruments, ST Microelectronics and Zoran. Some of these solutions were showcased at the 2002 International Broadcasting Convention, including a chip from Equator playing high-definition 720p WMV. Equator's chip will be used in set-top boxes built by Pace Micro Technology and other manufacturers.

Broadcast applications

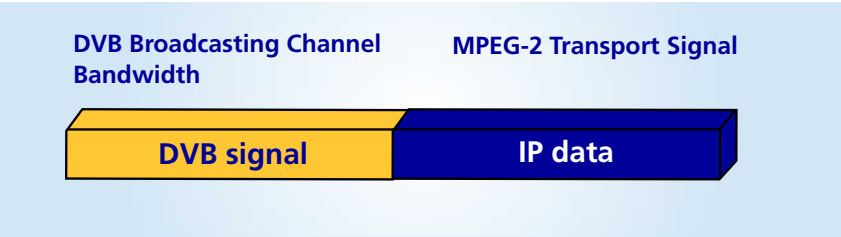

Figure 6
Broadcasting clips compressed with Windows Media 9 Series via DVB

Some standards, like DVB, provide for new technologies to be integrated and to create additional services and revenue. In the private data section of the MPEG-2 transport stream, one can easily encapsulate IP data (for example, audio and video compressed with Windows Media codecs), as shown in *Fig. 6*. This data can then be transmitted via DVB along with the standard DVB signal.

Windows Media is starting to be used for such broadcast-like applications over DVB. In this section we discuss a couple of examples of recent efforts.

Fig. 7 illustrates the first example. NTL Broadcast and Tandberg Television developed a DVB-based system for broadcasting sports (Eurosport) and news (ITN) media, compressed with Windows Media 9 Series. This system was demonstrated live for several days at the 2002 International Broadcasting Convention (IBC) [5].

The British Eurosport signals originated at NTL's Digital Media Centre near London, and the media was encoded live using a prototype Windows Media 9 Series encoder (which supports WMA 9 and WMV 9) built by Tandberg Television.

ITN news stories were fed via update-detection software from Gee Broadcast to a Tandberg Television Format transcoder that also compressed the clips into Windows Media. The encoded files were transmitted to the Crawley Court satellite teleport, south-west of London, into the NTL Broadcast's "store and forward" system.

Both types of Windows Media streams were IP-encapsulated by a source media router (SMR) from SkyStream and fed into a DVB-standard multiplexer. They were then transmitted via DVB satellite

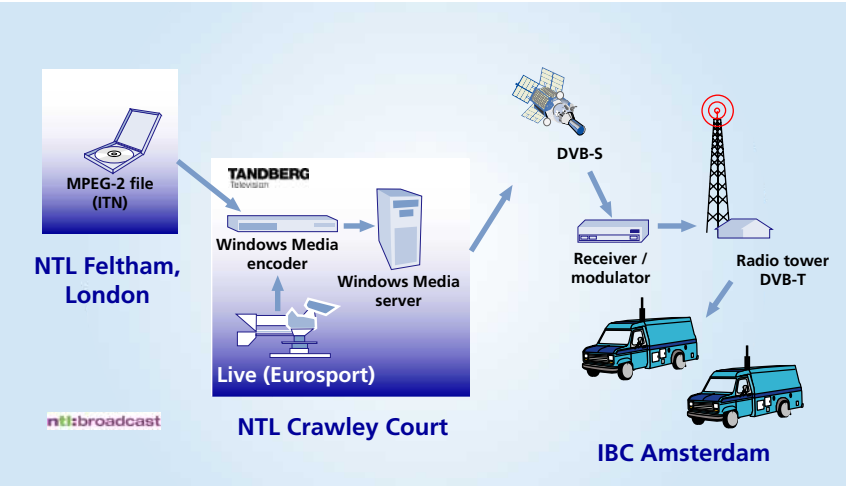

Figure 7
Broadcasting Windows Media over DVB at IBC

(DVB-S) to Amsterdam and from there via DVB terrestrial (DVB-T) to be received in moving vehicles at IBC. The vehicles transported attendees from the convention site to the city, and were equipped with on-board computers using Windows Media decoders connected to a monitor. The end-to-end system was integrated by NTL Broadcast.

This project demonstrated the ability to use the standard DVB infrastructure for mobile television transmission using Windows Media 9 Series.

Another interesting example is the on-demand movie system being deployed by LuxSat International in several countries. Movies are encoded using Windows Media 9 Series and delivered through IP datacasting over DVB-S to a user's hard drive. The user will have as many as 100 movies to choose from, and the available movies will be refreshed daily using a FIFO approach.

Summary

The vision of Windows Media 9 Series is to deliver compressed digital media content to any device over any network. Solutions and services are provided by ecosystems of partners that either use the basic components in Windows Media 9 Series (Windows Media Encoder for authoring, Windows Media Services for distribution, Windows Media DRM technology for usage rights assignment, and Windows Media Player for playback), or build their own components on top of the Windows Media hardware porting kits or software SDKs.

Windows Media 9 Series provides a variety of state-of-the-art audio and video codecs for different applications. The WMA 9 and WMV 9 codecs provide good quality for broadcast applications at relatively low bandwidth. For instance, one can achieve standard-definition broadcast quality at about 2 Mbit/s, and high-definition broadcast quality at typical DVD or high-end broadcast rates. The new WMA 9 Professional codec can encode multichannel or high-resolution audio tracks with bit-rates as low as 128 kbit/s.

Windows Media 9 Series is being adopted by consumer electronics manufacturers, where open and broad licensing programs have resulted in a large number of partners that have ported Windows Media technology to many devices on multiple platforms.

Finally, business opportunities are also emerging in the broadcast industry, based on new services that can be provided using Windows Media 9 Series in conjunction with DVB standards.

Bibliography

[1] Official Windows Media website:
http://www.microsoft.com/windows/windowsmedia/default.asp

[2] Windows Media website for Consumer Electronics:
http://www.microsoft.com/windows/windowsmedia/conselec.asp

[3] Official website of the SD association:
http://www.sdcard.org/

[4] Website for HighMAT CD and DVD format:
http://www.microsoft.com/windows/windowsmedia/Consumelectronics/highmat.asp

[5] "Video on the move", the IBC Daily, Saturday 14 Sept 2002, page 1
(electronic version under IBC Daily News at http://www.ibc.org).

[6] Windows Media partners website:
http://www.microsoft.com/windows/windowsmedia/partner.asp

[7] David Ranada: **Facing the codec challenge**
Sound & Vision, pp. 98-100, July 2002.

[8] BPN 049: **The EBU Subjective Listening Tests on Low Bit-rate Audio Codecs**
EBU, Sept. 2002.

Jordi Ribas-Corbera received an *Enginyer Tècnic de Telecomunicacions* degree from *Escola d'Enginyeria La Salle* (Barcelona) in 1990, an M.Sc. in Electrical Engineering from the University of California (Irvine) in 1992, and a Ph.D. in Electrical Engineering (Systems) from the University of Michigan (Ann Arbor) in 1996.

During the summer of 1994, he worked in the Advanced Video Processing Laboratory at the NTT Human Interface Labs, Yokosuka, Japan. From 1996 to 2000, he was with the digital video department at Sharp Labs of America, Camas, Washington State, where he did research on data compression and optimized the coding performance of several Sharp products. He joined Microsoft Corp. in February 2000, where he is currently the group program manager for the Codec and DSP team in the Windows Digital Media division. His team has developed multiple compression and signal processing products for Windows Media, such as the Windows Media Audio (WMA) and Windows Media Video (WMV) codecs that are part of the Windows Media Player and other consumer electronic devices.

Dr Ribas has participated in the development of compression standards such as ISO/MPEG-4, ITU-T/H.263+, and ITU-T/H.264. His research interests are in image processing and information theory, especially data compression. He is the author of numerous contributions to standards and peer-reviewed technical papers in academic conferences and journals, and has been invited to speak at a number of conferences and seminars, such as at Cable Labs, the EBU, NAB and the SMPTE.

Jordi Ribas received the Young Investigator Award in the 1997 IS&T/SPIE International Conference on Visual Communications and Image Processing, and the Sharp Labs President's Award in 1999. He has been awarded five patents and has others pending.

[9] Ben Waggoner: **Web video codecs compared**
DV Magazine, Nov. 2001.

[10] Geir Ove Rapp: **Main EBU B/VIM results and conclusions**
EBU specialized meeting on Audio/Video Coding Technologies, 5-6 Sept. 2002.

[11] J. Bennet and A. Bock: **Comparison of MPEG and Windows Media Video and Audio Encoding**
White Paper, Tandberg Television, Sept. 10, 2002.

[12] R. Schäfer, T. Wiegand and H. Schwarz : **The emerging H.264/AVC standard**
EBU Technical Review No 293, January 2003.

[13] Document JVT-D15: **Joint Model number 3**
Joint Video Team of ISO/IEC, MPEG and ITU-T VCEG, Klagenfurt, Austria, 22-26 July, 2002
(available via anonymous ftp from **ftp://ftp.imtc-files.org/jvt-experts/**).

[14] Document JVT-D153 (M. Ravassi, M. Mattavelli and C. Clerc): **JVT/H.26L decoder complexity analysis**
Joint Video Team of ISO/IEC MPEG & ITU-T VCEG, Klagenfurt, Austria, 22-26 July, 2002
(available via anonymous ftp from **ftp://ftp.imtc-files.org/jvt-experts/**).

[15] J. Ribas-Corbera and J. Sklansky: **Interframe interpolation of cinematic sequences**
Journal of Visual Communication and image representation, Vol. 4, No. 4, Dec 1993, pp. 392-406.

[16] Specification and licence for Windows Media file format (ASF):
http://www.microsoft.com/windows/windowsmedia/WM7/format/asfspec11300e.asp

Acknowledgements

The author would like to thank his colleagues in the Microsoft Windows Digital Media Division for their contributions and feedback to this paper. In particular, Jon Billings, Jim Beveridge, Terrence Dorsey, Tricia Gill, Pat Griffis, Paul Johnson, Rich Lappenbusch, Dr Bruce Lin, Dr Ming-Chieh Lee, Amir Majidimehr, Steve Sklepowich, and Dr Wei-ge Chen. In addition, the author is grateful to the multiple colleagues in partner companies that contributed to this article.