A Brief History of Video Coding

Marco Jacobs and Jonah Probell
ARC International
{marco.jacobs, jonah.probell}@ARC.com

Almost everyone in the industrialized world uses video compression many times each day. Whether you're watching Oprah on TV, funny clips on YouTube, Disney on DVD or homemade movies of your family, all video storage and delivery mechanisms rely heavily on compression technology.

As early as 1929, Ray Davis Kell described a form of video compression and was granted a patent for it. He wrote, "It has been customary in the past to transmit successive complete images of the transmitted picture. […] In accordance with this invention, this difficulty is avoided by transmitting only the difference between successive images of the object." Although it would be many years before this technique would actually be used in practice, it is still a cornerstone of many video compression standards today.

We'll take a trip down memory lane in this article, describing video compression standards over the years, after a brief introduction to television signals in general. We'll describe the challenges in designing video subsystems that support these video standards and conclude with a description of the ARC Video 401 Subsystem, highlighting how it addresses the challenging video market demands.

CRTs and Analog Broadcast

The first video systems evolved from oscilloscopes. They were essentially oscilloscopes with a second dimension, able to display multiple lines on the screen. The systems were analog and used cathode ray tubes (CRTs). Engineers soon realized that this CRT-based display device could present a signal modulated on a transmitted radio frequency carrier. Thus television was born.

Color television was first developed by RCA in the late 1940s. The television sets in most consumers' living rooms remained black and white for some time to come, however, so backward compatibility was important. Color information was cleverly modulated on the same signal, which allowed black-and-white TV sets to display an intelligible black-and-white representation from a video signal that contained color information.

In 1935 the British government defined high-definition television (HDTV) as having at least 240 lines. Today the de facto definition of HDTV is at least 720 lines.

Interlacing

A notable early form of compression was interlacing, where fields of even lines and odd lines are transmitted alternately. Interlacing is useful in two ways: it can either halve the required bandwidth or double the vertical resolution.

The phosphor glow in early CRTs didn't persist long enough to display clear images at lower frame rates; by sending every other line, the screen could be (partially) refreshed within the available time. The method relies on the tendency of the human visual system to integrate parts into a coherent whole.

Interlacing became a de facto standard and is still used today, though mainly for legacy reasons.

Recording

Video was first recorded on open-reel magnetic tape by the BBC in 1955. Consumer-friendly cassettes were standardized in the 1970s and adopted in home entertainment in the 1980s. This gave rise to the first standards war that was settled by consumers (VHS vs. Beta).

In the 1970s, broadcasters sued VCR makers for enabling consumers to make unauthorized recordings of broadcast content. This went to the Supreme Court of the United States, where it was decided that the ability to record content is "fair use" and is legal.

Digital Video Processing

Digital video processing has been developed to handle each of the techniques of broadcast, interlacing, and recording as well as Mr. Kell's technique of predictive compression. So many standards for digital video compression are in wide use today that television and digital video device makers find it necessary and preferable to use programmable digital video processors such as the ARC Video subsystem within their designs. We will give an overview of several notable compression standards and techniques below.

H.120

In 1984, the International Telecommunication Union (ITU) standardized recommendation H.120, the first digital video coding technology standard. H.120 used differential pulse code modulation (DPCM), scalar quantization and variable-length coding techniques to transmit NTSC or PAL video over dedicated point-to-point data communications lines. The application target then and for all ITU H.xxx standards since is video conferencing. H.120 is essentially deprecated today.

H.261

Practical digital video compression started with the ITU H.261 standard in 1990. The target was to transmit video over ISDN lines, with multiples of 64 kbit/s data rates and CIF (352x288-pixel) or QCIF (176x144-pixel) resolution. The standard was a pioneering effort and used a hybrid video coding scheme that is still the basis for many video coding standards today.

Hybrid video coding combines two methods:

1. Motion from frame to frame is estimated and compensated for by predicting data from previously coded frames.
2. The residual difference after prediction is encoded by decorrelating the data in the spatial domain through transformation to the 2D frequency domain. The transformed data are quantized, a stage in which information is lost, after which the data are encoded with a lossless compression scheme such as Huffman coding or an arithmetic coder.

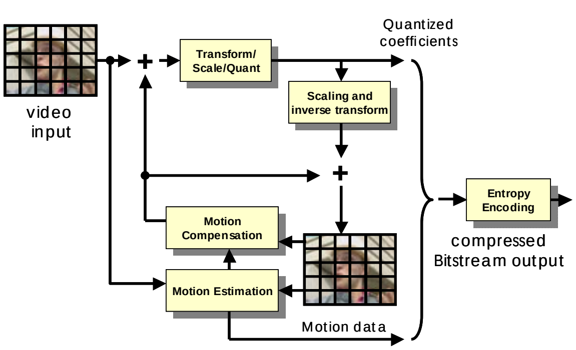

Figure 1 - Hybrid Video Encoder

H.261 uses 4:2:0 data sampling, in which there are twice as many luminance samples as chrominance samples. The human eye is more sensitive to light intensities than to color. H.261 has 16x16-pixel macro blocks with motion compensation, an 8x8-pixel discrete cosine transform (DCT), scalar quantization, zigzag scanning and Huffman-based variable-length entropy coding. The standard operates at 64–2048 kbit/s. H.261 is still in use but has largely been overtaken by H.263.

The organizational structure and processes used by the H.261 international committee to develop the standard have remained the basic operating process for subsequent standardization efforts.

Motion JPEG

JPEG is a widely used method for compression of still images (photographs), standardized in 1992. JPEG stands for Joint Photographic Experts Group. The standard is widely used in digital still cameras, mobile phones and the World Wide Web. Motion JPEG encodes video data as a sequence of independently coded JPEG images. Since there's no motion compensation, the temporal redundancy intrinsic to video is not exploited, resulting in a lower compression ratio than can be achieved with inter-frame-predicted, motion-compensated coding. Motion JPEG is often used in digital still cameras to capture video sequences. Digital cinemas and video editing systems also frequently use Motion JPEG. Since the video frames are encoded individually, Motion JPEG video streams can be edited without decompression and recompression, vastly speeding up the editing process and making functions such as frame-by-frame reverse play much faster.

Though JPEG is well standardized, Motion JPEG is not standardized by any industry-wide organization. There is no document that describes the exact format to which Motion JPEG-encoded bitstreams should adhere.

MPEG-1

The MPEG-1 standard (1992) was designed to achieve acceptable video quality at 1.5Mbit/s and 352x288/240-pixel resolution. MPEG-1 is playable in almost all PCs, VCD players and DVD players. Unfortunately, MPEG-1 supports only progressive-scan (non-interlaced) pictures, while the widely used NTSC and PAL video formats are interlaced. This missing key feature helped spur the development of MPEG-2, which does support interlaced signals.

MPEG-2 / H.262

The MPEG-2 / H.262 standard was jointly developed by the ISO and ITU standardization organizations and approved in 1993. The standard supports standard definition (SD, 720x576/480-pixel) resolutions as well as high-definition (HD) video signals with a pixel resolution of 1920x1080. The standard's support for HD stopped development of the MPEG-3 standard. MPEG-2 / H.262 is the most widely used video compression standard today. DVD players, personal video recorders, camcorders, distribution networks, set-top boxes receiving DVB-T/S/C signals, and ATSC, the American digital television standard, all use MPEG-2 / H.262.

The MPEG-2 / H.262 standard was designed with quality vs. complexity trade-offs meant to deliver studio-quality digital video at rates ranging from 3–10Mbit/s for SD (D1 pixel rate) video. An MPEG-2 decoder is backward compatible with the MPEG-1 standard.

Digital Video (DV)

The DV video coding standard developed by the IEC and standardized in 1994 targets primarily camcorders that store the video bitstream data on tape. DV does not use motion compensation, but instead encodes the individual frames at a fixed bitrate of 25Mbit/s. Combined with the audio and error detection and correction data, this results in a very high 36Mbit/s data rate. DV efficiency is better than Motion JPEG and is similar to I-frame-only MPEG-2 / H.262. The I-frame-only approach was chosen for easy video data editing. The standard also provides means to efficiently fast-forward or fast-reverse playback from tape.

H.263

The H.263 standard (1995) developed by the ITU was a big step forward and is today's dominant video conferencing and cell phone codec. The primary design target was video conferencing at low bit rates for mobile wireless communications. Especially for progressive video, H.263 quality is superior to all prior standards at all bit rates. At very low bit rates, video quality is better by a factor of two compared to MPEG-2 / H.262. H.263 has different versions and there are many annexes. H.263 has a somewhat tangled relationship with MPEG-4, since the two standards were partially developed at the same time.

H.263 is used in H.324, H.323 and H.320 video conferencing standards as well as in World Wide Web-distributed video content that plays with Macromedia's Flash 7 player (such as YouTube or Google Video). The 3GPP cell phone standardization project includes the H.263 video codec for transmission of video to and from mobile phones.

RealVideo

RealNetworks was one of the first successful commercial companies to sell tools for streaming digital audio and video over the Internet. Version 1 of the RealVideo codec was introduced in 1997. This first version was based on the H.263 codec, but versions 8 and higher use a proprietary video codec. Version 10 was introduced in 2004 and has been in use since.

MPEG-4 SP/ASP

MPEG-4 standardization began in 1995 and has continually been enhanced with new profiles including many novel coding concepts such as interactive graphics, object and shape coding, wavelet-based still image coding, face modeling, scalable coding and 3D graphics. Very few of these techniques have found their way into commercial products. Later standardization efforts have focused more narrowly on compression of regular video sequences.

The MPEG-4 video standard has been designed to allow a wide range of compression quality vs. bit rate trade-offs. The MPEG-4 Simple Profile is very similar to H.263. The Advanced Simple Profile (ASP) adds support for SD video, interlaced tools and additional tools to further increase compression efficiency, such as quarter-pel motion estimation and global motion compensation.

DivX / XviD

The popular proprietary DivX coding standard and file format are based on the MPEG-4 ASP profile. They are unrelated to the poorly conceived and commercially unsuccessful DivX video disk rental service. DivX encoders and decoders have been released into the public domain and are developed and supported by DivX and the Xvid project.

On2 VPx

On2 Technologies is a publicly traded company that designs proprietary video codec technology, which it licenses to its customers. The On2 codec designs are known as VP3 through VP7. In 2004, Macromedia (now Adobe) selected On2's VP6 for use in the Flash 8 video codec. Adobe's Flash

Video is frequently used as the vehicle for presenting video in web pages. The popular YouTube and Google Video websites use this technology. Ebay's Skype has licensed the On2 video codec for video conferencing over IP, and XM has licensed On2's coding standard for satellite broadcasts. On2 claims its technology carries no third-party patent licensing fees.

On2 has released an open-source implementation of its VP3 coding standard.

WMV9 / VC-1

Microsoft developed the Windows Media Video version 9 video codec. The codec was initially proprietary, but was later standardized by the SMPTE organization. SMPTE announced the formal release of the SMPTE 421M video codec standard in 2006. VC-1 is the standard's informal name. WMV9 / VC-1 is a hybrid video codec that uses conventional motion compensation, the DCT transform and Huffman coding, similar to H.263 and MPEG-4. However, WMV9 / VC-1 uses 4x4-pixel blocks, which are smaller than the 8x8-pixel blocks found in previous standards. VC-1 is characterized as an alternative to the latest ITU-T and MPEG video codec standard known as H.264/MPEG-4 AVC; however, studies have shown VC-1 to compress less well than H.264. The codec is computationally less demanding than H.264 and has been adopted by both HD DVD and the Blu-ray Disc Association as a mandatory video standard for players and optional codec for video disc sellers. The codec is also frequently used on the Internet and in the Xbox 360 gaming console. Today's high-end PCs can decode high-definition WMV9 / VC-1 bitstreams in real time in software.

A large number of patent holders have made royalties claims against WMV9 / VC-1. Though many proprietary coding standards (such as RealVideo, DivX and On2 Technologies' standards) are designed to avoid the use of patented technologies from third parties, WMV9 / VC-1 does not provide this cost-saving benefit.

H.264 / MPEG-4 part 10 / AVC

Where the original MPEG-4 standardization committees put a lot of effort into novel media coding techniques, H.264 took a "back to basics" approach. The goal was to compress video at twice the rate of previous video standards while retaining the same picture quality. The standard was originally known as H.26L or JVT, for the Joint Video Team, in which the ISO and ITU organizations worked together to complete the standardization. H.264 is the ITU name for the standard; MPEG-4 part 10, Advanced Video Coding (AVC) is the ISO name. Due to its improved compression quality, H.264 is quickly becoming the leading standard; it has been adopted in many video coding applications such as the iPod and the Playstation Portable, as well as in TV broadcasting standards such as DVB-H and DMB. Portable applications primarily use the Baseline Profile up to SD resolutions, while high-end video coding applications such as set-top boxes, Blu-ray and HD-DVD use the Main or High Profile at HD resolutions. The Baseline Profile does not support interlaced content; the higher profiles do.

AVS

The government of the People's Republic of China initiated development of the Audio Video Standard, or AVS. The compression codec for digital audio and video was standardized in 2005. One of the key factors driving development was to remove the need for Chinese consumer electronics manufacturers to pay royalties to companies that hold patents on internationally standardized compression technology. AVS coding efficiency is comparable to H.264, while its computational complexity is lower. The AVS standard will likely be used in some set-top box applications for IPTV and mobile broadcast TV within mainland China. Major semiconductor companies have announced support for the standard, but today it's unclear whether AVS will succeed in becoming a widely adopted video coding standard.

A Brief Future of Video Coding

Current standardization efforts center around extending today's standards rather than developing completely new video coding methods. There's an effort to add coding of multiple views to H.264 for 3D video. Scalability, which allows streams to be encoded once but transmitted and decoded at different resolutions and frame rates, is also an area of research and standardization. Both MPEG-2 and MPEG-4 have scalable profiles, but these have not been widely adopted by the industry. It remains to be seen whether an H.265 standard will follow the H.261 to H.264 series of increasingly complex advanced standards. The major goal would likely be a further 50% savings in bandwidth. One way to achieve this is to focus on a perceptual quality metric, replacing the quantitative peak-signal-to-noise ratio metric that has been used in the past.

In contrast to Moore's law for integrated circuits, compressed video bandwidth is not likely to halve at a fixed rate.

Which Codec?

Many video coding standards have been developed and are in use today. Which video codecs to support in a particular device depends on the exact application, but tomorrow's connected and open devices must all be able to decode many standards and simply play any bitstream. The following issues arise during the design of the video subsystem:

- Coding standard: The subsystem must be able to encode or decode the applicable video coding standards. There are many standards, and within each standard there are often variations (different versions and profiles).
- File format: video coding bitstreams are encapsulated in different file formats, which the video subsystem must be able to read and decode.
- Image resolution: video bitstreams contain compressed images with a particular display resolution. Standard definition TV signals for instance have 720 pixels per line, while portable media player displays often have 320 pixels per line.
- Frame-rate: the number of frames per second stored in the bitstream differs per application. Low-bandwidth wireless video conferencing for instance often uses 15 frames per second, while good quality video requires 30 frames per second.
- Bitrate: MPEG-2 based DVDs use up to 10Mbit/s while H.264 standard definition material is often stored at 1.5Mbit/s. Higher bitrates typically require more processing power to decode.
- Latency: some applications, like video conferencing, require very low latency transmission. A video recording application can have a much longer latency before the compressed images are stored in the memory system.
- Error resilience and robustness: corrupt bitstreams should still be decoded as well as possible, and under no circumstances should the video subsystem lock up.
- Pre/post processing: before or after the coding of the video, high quality algorithms for image filtering, scaling and frame rate conversion can greatly improve the quality of the images presented on the display.
- Synchronization: audio and video (and subtitles) must be timed correctly in order to prevent any distracting misalignment.
- Processing power: different video standards and processing techniques require different amounts of computational resources. The video subsystem should be capable of supporting the most demanding use case.

When selecting a video processing system for your device, it's ideal to choose a video subsystem that can handle many video coding standards and processing techniques, and is software programmable and firmware upgradeable for maximum flexibility to address the issues described above.

The ARC® Video 401 Subsystem

At ARC International we designed the ARC Video 401 Subsystem to support a wide range of codecs (Motion JPEG, MPEG-2, MPEG-4 SP/ASP, DivX/Xvid, H.264 / MPEG-4 AVC, VC-1 / WMV9) at up to standard definition resolutions. The subsystem consumes very little energy, and has a very small die size for low-cost implementations. The system is fully software programmable and easy to integrate into a system-on-a-chip (SoC). It is based on a member of the ARC 700 configurable processor family, extended with an entropy decoding accelerator, a SIMD engine and a tightly coupled DMA controller optimized for video data transfers.

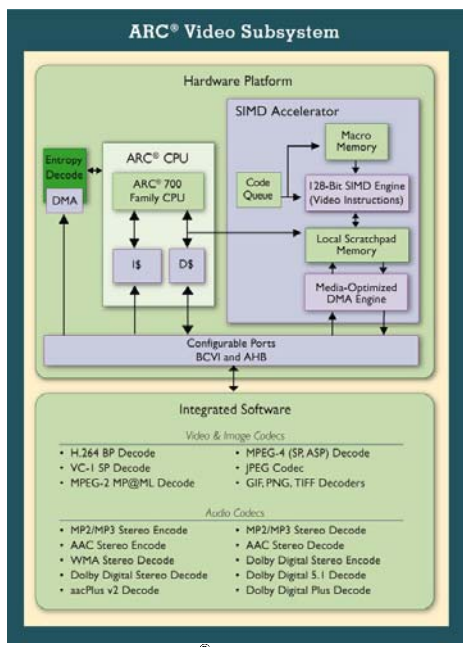

Figure 2 – ARC® Video 401 Subsystem

The member of the ARC 700 family configurable CPU has seven pipeline stages, allowing high clock rates. It can perform host-processing functions, such as running an operating system, act as the controller in the application, and decode audio. The

ARC Video 401 Subsystem extends the ARC 700 family core with a single instruction multiple data (SIMD) engine that operates on 128-bit vectors of data. A single instruction performs up to 16 operations on parallel data. The SIMD engine can effectively operate autonomously, in parallel with the ARC 700 family core. Especially recent video coding standards have grown more complex, requiring a large amount of fine-granularity decision-making (e.g. H.264 operates on 4x4 blocks) and include many control operations. The ARC Video 401 Subsystem's architecture efficiently supports these video-coding applications that are both control and data intensive. The generic instruction set architecture of the SIMD engine supports a wide variety of applications. Specific instructions have been added to accelerate the required deblocking algorithms in H.264 and VC-1. Without these instructions, these kernels would consume a disproportionate number of compute cycles. The SIMD unit's high-performance processing pipeline is complemented with a DMA system that moves data to and from system memory autonomously. A hardwired entropy decode accelerator supports the Huffman-based MPEG-1/2/4 and VC-1 video coding standards, as well as Context Adaptive VLC (CAVLC), Exp-Golomb, and other bit-level decoding required for H.264 Baseline profile.

The ARC Video 401 Subsystem retains the must-have traits of very low-power consumption and small silicon die-size. Due to the ARC Video 401 Subsystem's software programmable architecture, the video coding and processing applications can continuously be further optimized and extended, guaranteeing a future proof design and protecting the customer's investments.

About the Authors

Marco Jacobs *is ARC's Director of Multimedia Solutions for Europe and Asia. Prior to joining ARC, Marco held management positions at Silicon Hive and BOPS and was a software architect at Philips. He can be reached at* marco.jacobs@arc.com.

Jonah Probell *is Director of Multimedia Solutions for ARC International. His interests are in microprocessor design and multimedia system architecture. He can be reached at* jonah.probell@arc.com.

About ARC International plc

ARC International is the world leader in configurable media subsystems and CPU/DSP processors. Used by over 140 companies worldwide, ARC's configurable solutions enable the creation of highly differentiated system-on-chips (SoCs) that ship in hundreds of millions of devices annually. ARC's patented subsystems and cores are smaller, consume less power, and are less expensive to manufacture than competing products.

ARC International maintains a worldwide presence with corporate and research and development offices in Silicon Valley and St. Albans, UK. For more information visit www.ARC.com. ARC International is listed on the London Stock Exchange as ARC International plc (LSE: ARK).