Whitepaper:

The H.264 Advanced Video Coding (AVC) Standard

What It Means to Web Camera Performance

Introduction

A new generation of webcams is hitting the market that makes video conferencing a more lifelike experience for users, thanks to adoption of the breakthrough H.264 standard. This white paper explains some of the key benefits of H.264 encoding and why cameras with this technology should be on the shopping list of every business.

The Need for Compression

Today, Internet connection rates average in the range of a few megabits per second. While VGA video requires 147 megabits per second (Mbps) of data, full high definition (HD) 1080p video requires almost one gigabit per second of data, as illustrated in Table 1.

Table 1. Display Resolution Format Comparison

Format	Horizontal Pixels	Vertical Lines	Pixels	Megabits per second (Mbps)
QVGA	320	240	76,800	37
VGA	640	480	307,200	147
720p	1280	720	921,600	442
1080p	1920	1080	2,073,600	995

Video Compression Techniques

Digital video streams, especially at high definition (HD) resolution, represent huge amounts of data. In order to achieve real-time HD resolution over typical Internet connection bandwidths, video compression is required. *The amount of compression required to transmit 1080p video over a three megabits per second link is 332:1!* Video compression techniques use mathematical algorithms to reduce the amount of data needed to transmit or store video.

Lossless Compression

Lossless compression changes how data is stored without resulting in any loss of information. Zip files are losslessly compressed so that when they are unzipped, the original files are recovered. Since a single bit error in a PowerPoint file or an application can render it unusable, lossless compression is great for maintaining data integrity. However, in both of these examples the compression ratios are only perhaps 2:1 or 5:1.

Lossy Compression

The MP3 file format for music is a great success story for the use of lossy compression. While MP3 files are not exact copies of music stored on a CD, they have been good enough to have basically rendered audio CD obsolete.

In the case of video, lossy compression is valuable, since there is much more information in images than the eyes or brain can absorb. By understanding how we process visual information, lossy compression algorithms can reduce the amount of data that is transmitted in a way that has negligible impact to the viewer.

Redundancy Encoding

In most video images, there are some elements that are relatively uniform, such as a blank whiteboard or a wall. Video compression algorithms take advantage of *spatial redundancy* by grouping pixels into blocks. When blocks are relatively uniform, they can be represented by far fewer bits than if each pixel were represented individually.

Also, there are usually periods when there is very little or no change from one frame to another – such as the background behind a speaker in a video call. Video compression algorithms take advantage of *temporal redundancy* by only transmitting the differences between consecutive frames.

Role of Industry Standards

Video Compression

Many competing compression standards have been used over the years – think Betamax versus VHS, and HD DVD versus Blu-ray – creating confusion and frustration in the market. Adherence to standards is great for interoperability and easy setup.

In the case of video, industry standards are clearly valuable in telecommunications, broadcasting, and content distribution applications, since they make it possible for video compressed by one device or system to be decompressed on another device or system. An overview of historical ITU Telecommunication Standardization Sector (ITU-T) standards as well as their specifications and applications is provided in Table 2.

Table 2. Compression Standards, Specs & Applications

ITU-T Recommendation	Year Ratified	Target Resolutions	Target Bit Rates	Target Applications
H.261	1988	352Q288 (CIF) 176Q144 (QCIF)	40 kbps - 2 Mbps	ISDN videophones
H.262	1995	720Q480 720Q576 1280Q720 1920Q1080	1 - 25 Mbps	SD/HD Broadcast, DVD, HDV
H.263	1996	128Q96 176Q144 352Q288 704Q576 1408Q1152	20 kbps – 4 Mbps	Videoconferencing MMS Streaming Internet Video
H.264	2003	128Q96 up to 4,096Q2,304	64 kbps up to 25 Mbps	Videoconferencing Broadcast Blu-ray Disc DV & Mobile phone cameras

The goal behind the H.264 standard was to provide high quality video at considerably lower bit rates than previous standards. At the same time, the design needed to be not too complex or expensive to implement. A secondary goal was to make the standard flexible enough to be implemented across a variety of applications, networks and systems.

Video standards have hit a sweet-spot based on the compression/decompression technologies available currently. Because technology is roughly doubling in performance every 18 months, every few years the industry is able to make a quantum leap in performance, as has been done with H.264.

The H.264 vastly improves compression performance over standards such as MPEG-2 and MPEG-4 Visual. Figure 1 shows a comparison of a video frame using MPEG-2, MPEG-4 Visual, and H.264 standards, with compression at the same bit rate.

Figure 1. Video Frame Comparison of Several Compression Standards

MPEG-2

MPEG-4 Visual

H.264

Images provided courtesy of Iain Richardson/Vcodex Ltd.

USB

The Universal Serial Bus (USB) standard allows all kinds of peripherals to be easily attached to PCs and other electronic devices, and operate in a plug-and-play manner. The standard defines different classes of devices – such as webcams, printers, keyboards, and speakers – and the connectors for those classes. The USB Video Class (UVC) and the USB Audio Class (UAC) govern devices capable of streaming audio and video, like webcams.

The USB Implementers Forum (USB-IF) has standardized how H.264 encoding cameras should be supported. For the existing installed base of PCs running Windows operating systems such as Windows Vista and Windows 7, Mac OS-X, and Linux, there is a set of H.264 extensions to current standard, UVC 1.1. Cameras and applications that support these extensions can work together without a driver. The first cameras and applications that support these requirements are about to hit the market.

A new version of the UVC, version 1.5 is approaching ratification. UVC 1.5 will make support for encoding cameras even more integrated in the PC operating system and will make it easier for application developers to take advantage of the unique capabilities of an H.264 encoding camera. Microsoft has announced it will support UVC v1.5 natively in Windows 8.

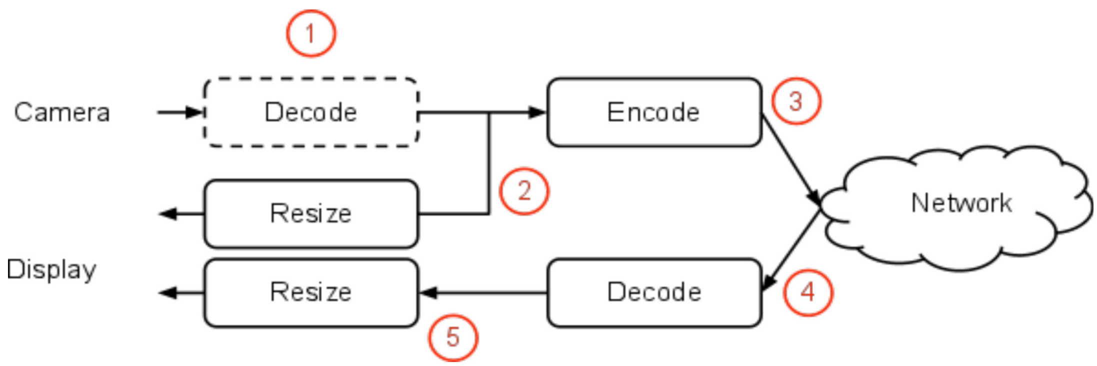
Compression/Decompression Encoding

The compression/decompression (CODEC) encoding process can be handled by software on the CPU, or through the camera directly. This section compares both methods in terms of process efficiency, CPU utilization, power consumption, and video image quality.

Efficiency

Figure 2 shows the process flow for a two-way video calling application using a non-encoding webcam; a description of the process follows.

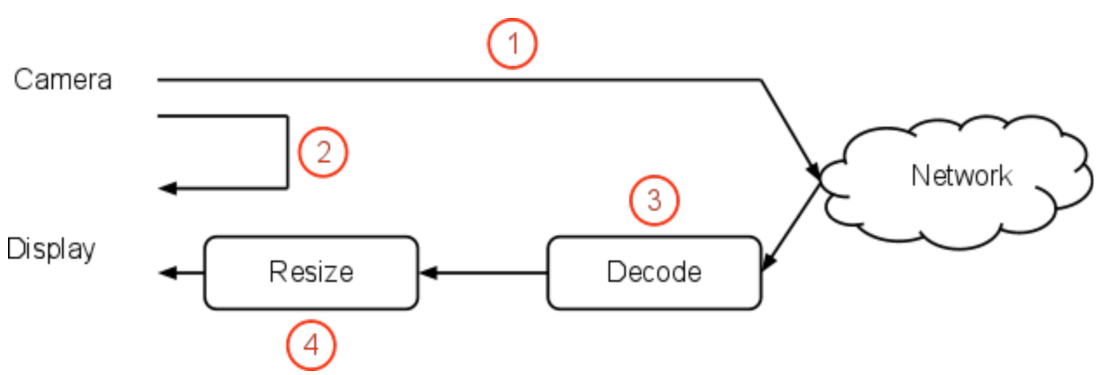
Figure 2. Two-way Video Calling – Non-Encoding Webcam



1. Video is captured from camera in an uncompressed or MJPEG format. At higher resolutions like 720p, MJPEG is used and it must be decoded or uncompressed. This is the decode block shown in dashed lines. For efficiency, the capture resolution is matched to resolution of video you want to send into the network e.g. VGA or 720p.
2. To show a preview of the video, the video must be resized to the preview window size.
3. For transmission, the video must then be encoded into the transmission format which could be H.264, another standard like H.263, or a proprietary codec like VP7 or VP8.
4. The received video must be decoded into a compressed format.
5. The uncompressed video is resized for display.

Figure 3 shows the process flow for a two-way video calling application using an encoding webcam; a description of the process follows.

Figure 3. Two-way Video Calling – Encoding Webcam



1. H.264 encoded video is captured at the correct resolution and bit rate for transmission from the encoding camera.
2. An uncompressed video is captured at the correct resolution for display from the encoding camera.
3. At the far end, received video must be decoded or uncompressed.
4. Finally, uncompressed video must be resized for display

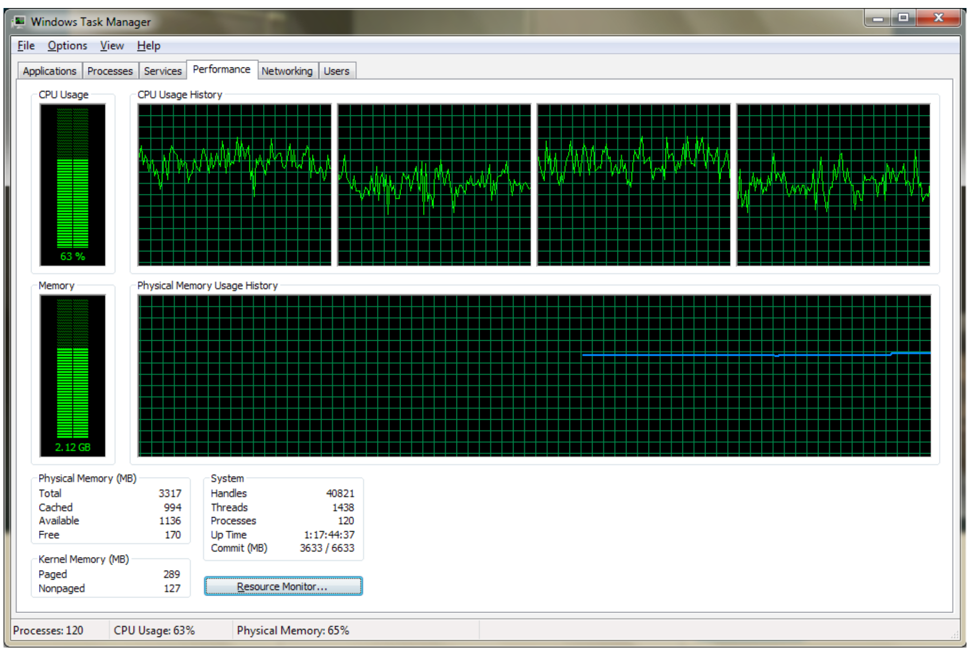
As can be seen, use of an H.264 encoding webcam is more efficient.

CPU Utilization

We conducted some limited in-house testing to compare the performance of non-coding and encoding cameras. *Note: In other testing, CPU usage may vary depending upon camera, notebook or computer, video conferencing client and video transmission format.*[i]

Figure 4 shows a screenshot of the performance screen of Windows Task Manager using a Dell Latitude E6410 notebook with Core i5 processor and a typical non-encoding camera; the notebook is running Skype, transmitting 720p video, and receiving 480p (VGA).

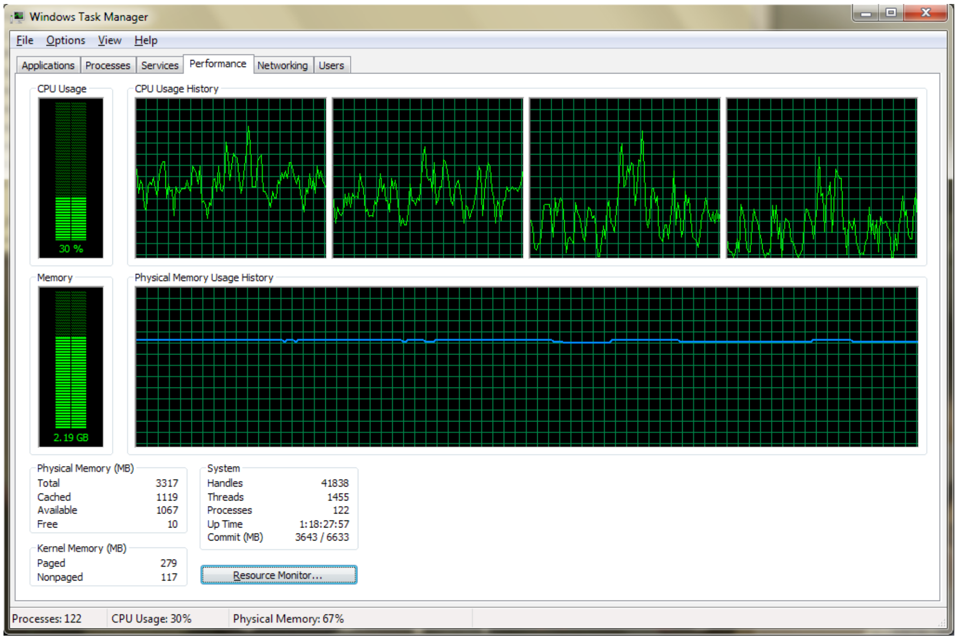
Figure 4. CPU Usage – Non-Encoding Webcam with Software Compression



Since the CPU is being used to encode and decode video, *over 60% of the CPU (average, over four cores) is being used for the video call,* as shown in Figure 4.

Figure 5 shows the same notebook using an encoding webcam, the Logitech C920.

Figure 5. CPU Usage – Encoding Webcam



As shown in Figure 5, CPU utilization is roughly 30% – *a savings of 50% over the traditional webcam scenario* – resulting in much more CPU being available for other applications, such as PowerPoint or a shared white-board application.

Because of the asymmetrical nature of video compression – encoding is much more CPU intensive than decoding – the contrast would be more stark if 1080p video was used in this comparison. In fact, the Dell Core i5 CPU might not be able to send 1080p when relying on software encoding, but runs well when using an encoding camera capable of 1080p.

Power Utilization

The Dell Core i5 processor in the notebook used for this testing has a maximum power consumption of 35 watts. With the non-encoding webcam, the encode task uses about 60% of the CPU or over 20 watts, as shown in Figure 4.

H.264 encoding cameras consume less power by offloading compute intensive operations into a small dedicated hardware block contained in the camera. The H.264 encoder in the C920 camera consumes about one watt. As shown in Figure 5, the encoding webcam only uses about 30% of the CPU or about 10 watts.

Lower power usage typically translates to cost savings for your business and longer battery life in notebooks for on-the-go video calling.

Conclusion

H.264 encoding cameras can help deliver a better user experience each and every time they are used. But delivering a great video camera is more than adding a H.264 encoder to a webcam. The Logitech C920 HD Pro Webcam delivers the highest video quality at every step of the way:

- Wide field of view (78 degrees) covers group calls without having to reposition camera
- True widescreen video without cropping and zooming for all widescreen and HD resolutions
- High-quality sensor (1/3" 3 MP HD) and Logitech RightLight™ 2 technology improves visual quality in low light and backlit situations
- Autofocus and a Carl Zeiss® quality lens allow detailed documents or visuals to be shared close up during calls
- Omni-directional dual digital microphones designed to capture sound in larger spaces like offices and small conference rooms
- Multiple mounting options: LCD screen, notebook, tabletop or wall using attached clip; tripod using embedded thread
- Plug-and-play USB connectivity
- Works with most popular video calling applications

For More Information

For more information about Logitech, its webcams and other business products, please visit:
www.logitech.com/business.

To order Logitech's business products, contact your reseller or call 1-866-283-4521.

[i] Note: Testing was conducted in-house at Logitech; this testing was limited, and merely designed to illustrate differences between a non-encoding camera and an encoding camera. In other testing, CPU usage may vary, depending upon camera, notebook or computer, video conferencing client and video transmission format.

Images shown in Figure 1 are provided courtesy of Iain Richardson/Vcodex Ltd.