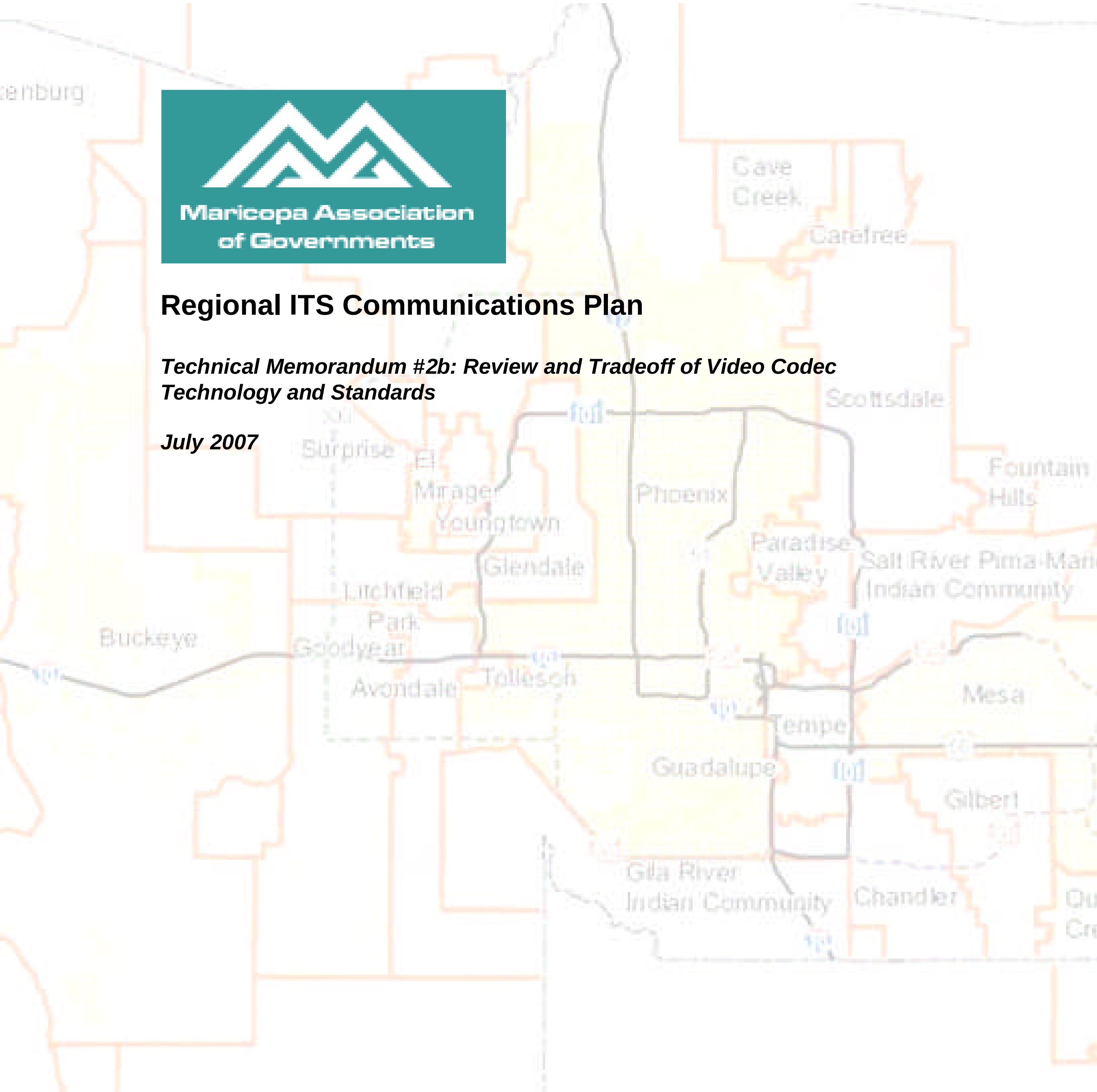


Maricopa Association of Governments

Regional ITS Communications Plan

Technical Memorandum #2b: Review and Tradeoff of Video Codec Technology and Standards

July 2007

Table of Contents

1.0 Review Video Compression Standards and Recommend Appropriate Standard Supporting Interoperability

1.1 The "Basics" of Video CODEC Technology

Video codec technology provides two main functions in the encoder. First, the analog video at the input of the device is converted from analog to digital. Second, the digital signal is compressed using a multiple mathematical processes. The codec receiver is responsible for providing the inverse operation, and its output is a recreation of the analog video.

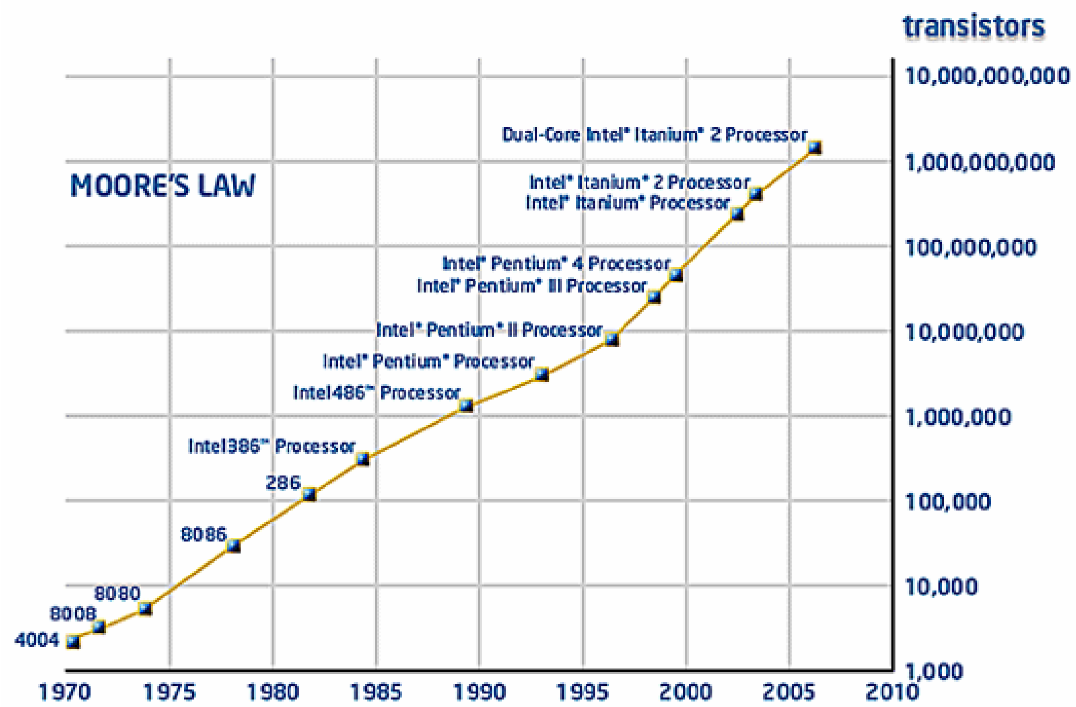
CODEC technology (otherwise known as compression/decompression) enables a reduction in digital electronic file size by identifying similar values within the file, and representing those values by a code rather than by each discrete value within the file. A number of "contiguous" and "like" values can then be represented by a reduced number of symbols, which later can be expanded (decoded) to match content representation of the original file. A good example of a codec is the PKZIP or WINZIP utility found on most computers. If a file needs to be transmitted across a network or stored on electronic media, it is preferred that the file is first compressed and then transmitted or stored to disk. The advantages offered by using compression include reduced bandwidth during file transmission and reduced storage space.

Compressing video signals requires more dynamic and special case algorithms for sending the video "file". Different techniques for encoding the video signal exist. In order to understand video compression, it is important to examine the electronic structure of the video signal. Using the NTSC video signal as a reference, we have 30 frames per second. Each frame can be thought of as a still picture which can then be compressed.

Different approaches for encoding each of the pictures (frames) will yield different results in terms bandwidth (file size) required and the ability to faithfully re-construct the original signal without losing object definition or detail. Video codecs are classified as being lossy or non-lossy, depending on the performance at the mathematical algorithm embodied in hardware/software. Video codecs use complex mathematical transforms which operate against the data set (file or video pixel information) to achieve a reduction in file size. The most common transforms utilized for video applications include the Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT) and more recently Simple Integer Transform algorithms. As a rule, more efficient compression algorithms are greater in mathematical complexity, which correlates to increased processor utilization for encoding/decoding operations.

Video codecs are currently supported in hardware, software and firmware. Video surveillance operations typically require hardware or software in support of "real" time command and control. Most internet video applications do not require real time decode nor camera control and are typically software based. Regardless of whether a codec algorithm is hardware or software based, the encoding and decoding of video signals is processor (CPU) intensive rather than RAM intensive. Moore's Law describes the prediction that the processing power of computer chip will double approximately every 18 months. Current technological trends associated with multi-core processors are creating computer platforms which are becoming capable of concurrently decoding multiple video streams, processing operating system and application software tasks. Current offerings include single, dual and quad-core processors, with eight core processors in development. Intel's roadmap for multi-core processing is currently focused on competing with Sun's micro-architecture, and plans to scale up to 32 cores. Figure 1.1-1 indicates the application of Moore's law with respect to Intel's processor line up.

Figure 1.1-1 Application of Moore's Law with Respect to Intel's Processor Line-up. (Ref. Intel)



The main implication of enhanced processing capability is that general purpose processors will become powerful enough to support multi-channel video decode capability in software without additional dedicated, special purpose hardware assistance. Testing performed on a single core, Pentium 4 processor supported by 1GB RAM, produced CPU loading of 97% when four (4) simultaneous MPEG2 transport streams were decoded in software. This left the workstation essentially useless for other tasks associated with other application software as may be required for ITS applications. More efficient processing means lower software introduced latency and better overall workstation performance. The following investigation into video codec standards and recommendations are based primarily on hardware/firmware implementations that are currently available as Common-Off-The-Shelf (COTS) products.

Regardless of the codec algorithm supported, the basis for video compression falls back to the scanning system incorporated for each video frame.

1.2 Video CODECs

Image, Video and Audio Compression standards have been specified and released by two main groups since 1985:

- ◊ ISO - International Standards Organization: JPEG, MPEG.
- ◊ ITU - International Telecommunications Union: H.261 - 264.

A third organization, the Society of Motion Pictures and Television Experts (SMPTE), is recently driving efforts for collaborative work with the private sector to produce CODEC specifications.

A compression artifact is the result of an aggressive data compression scheme applied to an image, audio, or video that discards some data which is determined by an algorithm to be of lesser importance to the overall content but which is nonetheless discernible and objectionable to the user. Artifacts in time-dependent data such as audio or video are often a result of the latent error in lossy data compression.

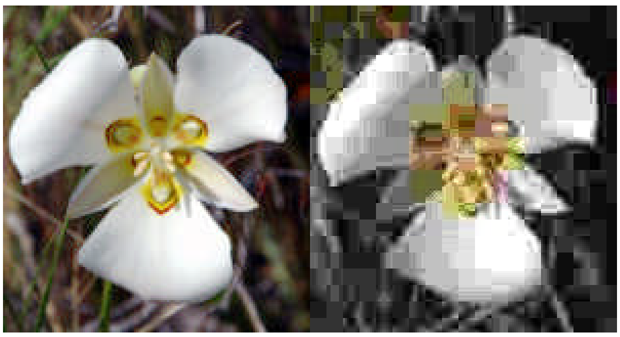
Technically speaking, a compression artifact is a particular class of data error that is usually the consequence of quantization in lossy data compression. Where transform coding is used, they typically assume the form of one of the basis functions of the coder's transform space.

When using the Discrete Cosine Transform for block-based coding, as in JPEG-compressed images, several types of artifacts can appear, including contouring and posterizing in otherwise smooth gradients, staircase noise along curving edges, "mosquito noise" around edges, and/or checkerboarding in "busy" regions (sometimes called quilting or blockiness). Figure 1.2-1 illustrates MPEG compression artifacts (mosquito noise) on calendar letters when the train is moving within the video frame.

Figure 1.2-1 example of mosquito noise around calendar lettering and sheep's wool due to movement of train and ball in foreground. (ref. NIST, Mosquito noise in MPEG-compressed video: test patterns and metrics)

To stop the build-up of compression artifacts, most compression systems occasionally send an entire compressed frame without prediction or differencing, usually at the start of a shot and at regular intervals thereafter. In MPEG picture coding, these are known as "I-frames", with the 'I' standing for "intraframe compression." Figure 1.2-1 illustrates compression artifacts, specifically blocking and color loss.

Figure 1.2-1 illustrates the effects of compression artifacts, specifically blocking and color loss.

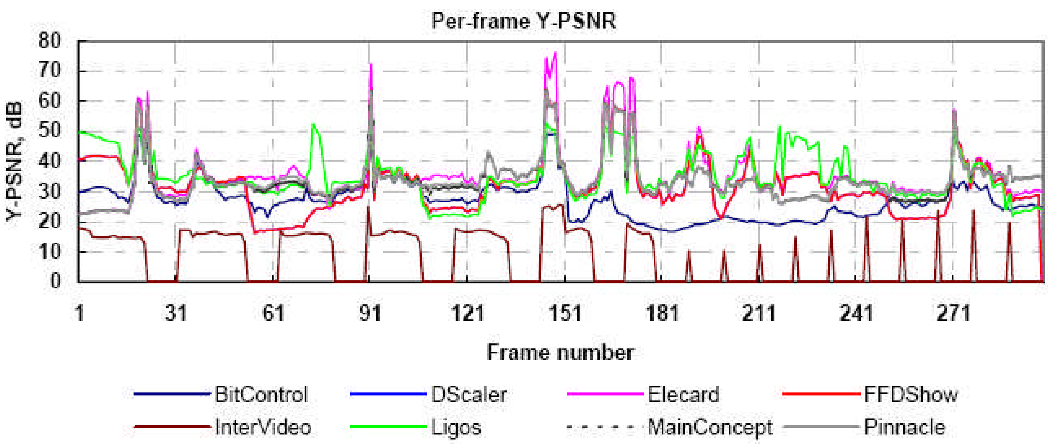


Peak Signal to Noise Ratio (PSNR) – measurement of amount of noise between the original source video and a decompressed copy of the original. Video quality is a subjective measure and cannot easily be quantified. One method of approximating a measure of quality is the signal to noise ratio between encoded/decoded video and the original raw source.

Studies performed by the National Institute of Standards and Technology (NIST) have proven a direct correlation between subjective video quality and measured PSNR levels recorded. The overall effect of PSNR is a cumulative system measurement. The measurement ultimately represents the noise introduced by the codecs compression algorithm, operating noise floor of codec electronics (Signal to Noise & Distortion (SINAD)), and associated link budget losses associated with end-to-end transmission of the video signal. Figure 1.2-2 Illustrates increasing values of PSNR between the encoded/decoded and original picture.

Figure 1.2-2 illustrates PSNR levels from left to right: original picture, PSNR = 33.52 dB, PSNR = 26.56 dB and PSNR = 20.56 dB.

Figure 1.2-3 illustrates PSNR measurements of a sample video segment decoded in software. (Ref. MPEG forum)



Picture 14. Per-frame Y-PSNR, bit inversion probability 10^{-5}, the "Foreman" sequence

Table 1.2-1 Timeframe Associated with CODEC Development & Release (ref. CMP Video Design Line)

<table>
<tr><th>Year</th><th>1985</th><th>1986</th><th>1987</th><th>1988</th><th>1989</th><th>1990</th><th>1991</th><th>1992</th><th>1993</th><th>1994</th><th>1995</th><th>1996</th><th>1997</th><th>1998</th><th>1999</th><th>2000</th><th>2001</th><th>2002</th><th>2003</th><th>2004</th></tr>
<tr><td>ISO/IEC 10918</td><td></td><td colspan="9">JPEG</td><td colspan="10"></td></tr>
<tr><td>ISO/IEC 15444-3:2002</td><td colspan="10"></td><td colspan="8">MJPEG, JPEG2000</td><td colspan="2"></td></tr>
<tr><td>ITU-T standards</td><td colspan="6">H.261</td><td colspan="3"></td><td colspan="7">H.263</td><td colspan="4"></td></tr>
<tr><td>ITU-T/ISO joint standards</td><td colspan="6"></td><td colspan="4">H.262 / MPEG2</td><td colspan="3"></td><td colspan="6">H.264 / MPEG4 Part 10 AVC</td><td></td></tr>
<tr><td>ISO standards</td><td colspan="4"></td><td colspan="5">MPEG1</td><td colspan="6">MPEG4</td><td colspan="5"></td></tr>
<tr><td>Microsoft /SMPTE</td><td colspan="16"></td><td colspan="4">WM-9 / VC-1</td></tr>
<tr><td>Apple/ITU_T</td><td colspan="16"></td><td colspan="4">QuickTime / H.264</td></tr>
</table>

All of these CODEC standards can be used to encode captured raw digital video data into a compressed bitstream and decode a compressed bitstream into raw digital video data for display. Each CODEC standard allows for some degree of trade-off between the three key parameters, *bit rate*, *resolution* (including frame rate), and *quality*. Bit rate is measured in bits/second. Resolution is measured in pixels/line x lines/frame at frames/second. Of additional importance is the ability for video compression algorithms to compensate or predict motion within subsequent fields of video. Algorithms that provide this functionality provide significant reduction in transmitted bandwidth, but require more processing power to handle the mathematics involved in the reconstruction of the video signal.

Providing motion compensated video from a video encoder-decoder system requires the support of additional "frame" types. The following descriptions provide a summary of frame function as related to compressed digital video transmission.

Frame Type	Functions Associated with Frame
Intra-Frame (I)	A frame having no reference for prediction
Inter or Predictive (P) Frame	A frame based on a previous frame
PB Frame and Improved PB Frame	A frame representing two frames and based on a previous frame
B Frame	A frame based on two reference frames, one previous and one afterwards
EI Frame	A frame having a temporally simultaneous frame which has either the same or smaller frame size
EP Frame	A frame having two reference frames, one previous and one simultaneous
D Frame	Contain only DC-frequency data, and intended for use in fast visual searches

Table 1.2-2 provides ten (10) specific items that must be provided to ensure the greatest possibility for interoperability between video codecs. Specifically, all of the parameters are required to match in order to provide hardware interoperability.

Table 1.2-2 Critical Operating Features Supporting CODEC Interoperability

Feature	Example
Standard Supported	JPEG, MJPEG, JPEG200, MPEG, H.261, H.263++, MPEG2, MPEG4 ASP, H.264
Transform Type Utilized	DCT, DWT, Integer Transform
Video Input Standard	EIA-170, NTSC, PAL, SECAM, ATSC
Level Supported	MPEG2 Low Level 352*288 @ 30 fps $\leq$ 4 Mbps, or H.264 Levels 1 - 5.1
Profiles Supported	Simple, Main, Advanced Simple, etc.
Stream Type	Elementary, Program, Transport
SNMP	Interoperability w/360 Surveillance's Chameleon for locating CODECs
IP Addressing	Static preferred for Chameleon, IPv4, IPv6
Physical Layer (Interface)	100Base-TX, 100Base-FL, SONET, EIA530
Network Protocols	TCP, UDP, IP, IGMP, SAP, SNMP, RTP, HTTP, NTP, SMTP, RSVP, etc.

1.2.1 JPEG

The Joint Photographic Experts Group (JPEG - ISO/IEC 10918) was designed for still continuous-tone grayscale and color images (still frame compression). JPEG images may be of any resolution and color space, with both lossy and lossless algorithms available. JPEG offers image compression ratios that range from 2:1 to 100:1, which is directly related to image quality as the compression ratio increases. Typically, the threshold of visible difference between the source and reconstructed images is somewhere between a 10:1 and 20:1 compression ratios. JPEG is widely accepted as a standard and consequentially supported by a large number of imaging software and web browser platforms.

Besides the compression level, the image itself also has an impact on the resulting compression ratio. For example, a white wall may produce a relatively small image file (and a higher compression ratio), while the same compression level applied on a very complex and patterned scene will produce a larger file size, with a lower compression ratio. JPEG uses the Discrete Cosine Transform (DCT) method for compression. Granularity of the codec is limited to an 8x8 block size.

1.2.2 MJPEG/JPEG2000

Motion JPEG offers video as a sequence of JPEG images. MJPEG is perhaps the most commonly used compression method used in consumer based, network video systems. A network camera captures individual images (frames) and compresses them into JPEG format. Motion JPEG uses the core DCT compression standard and uses a digital "container" for delivery of consecutive compressed frames. A container format is a computer file format that can contain various types of data, compressed by means of standardized audio/video codecs. The container file is used to identify and interleave the different data types. Simpler container formats can contain different types of audio codecs, while more advanced container formats can support multiple audio and video streams, subtitles, chapter-information, and meta-data (tags) - along with the synchronization information needed to play back the various streams together. Container formats include the .avi and .mov file extensions.

The network camera can capture and compress, for example, 30 such individual images per second and then make them available as a continuous flow of images over a network to a workstation, where it is decompressed. At a frame rate of approximately ≥16 fps, the viewer starts to “perceive” full motion video. Full motion NTSC video occurs at 30 frames per second. As each individual image is a complete JPEG compressed image, they all have the same guaranteed quality, determined by the compression level chosen for the network camera or video server.

Unfortunately, MJPEG does not have any way to take advantage of frame-to-frame redundancies to improve compression, as does MPEG (or other motion compensated algorithms). MJPEG sends all picture information from all frames as output. This results in a compromise between network bandwidth versus image resolution and frame rate. Use of MJPEG codecs leads to transmission of the video frames at less than 30 fps (required for full motion video) or at a substantially lower resolution. The end result being that the decoded video provided is of lower resolution than competing motion compensated algorithms, or that full motion is not achieved.

Typical resolutions and frame rates associated with MJPEG are: 160x120 or 320x240, operating at 10, 12 or 15 frames/second. MJPEG operating at SIF resolution (352*480) at 30 fps yields a video stream at approximately 4 Mbps. MJPEG uses the DCT method for sequential frame compression.

A more recent change is the support for the newer JPEG-2000 standard. This has moved away from the DCT compression used in "older” JPEG and MJPEG algorithms towards "wavelet" compression of image data. JPEG 2000 is a new image coding system that uses state-of-the-art compression techniques based on wavelet technology. JPEG 2000 refers to all parts of the ISO/IEC standard which is currently published. ISO/IEC 15444-3:2002 references the use of Motion JPEG 2000. The JPEG2000 and MJPEG2000 series of video codecs employ relatively modern algorithms, but do not have the necessary backing of COTS equipment to support deployment in ITS applications. JPEG2000 uses the Discrete Wavelet Transform (DWT) method for video compression.

1.2.3 H.261

ITU-T H.261 was the first video compression and decompression standard developed for videoconferencing. H.261 supports motion prediction unlike MJPEG, which allows codec operation at lower bandwidths. H.261 is intended for low bandwidth applications of p x 64 kbps, where p=1 to 30. H.264 allows only I & P frame types. H.264 codecs provides a self-contained video streams, and multiplex control data and audio.

Parameter	CIF	QCIF
Active Resolution	352*288	176*144
Frame Rate Refresh	29.97 Hz	
YCbCr Sampling Structure	4:2:0	

The H.261 series of video codecs are also considered to be ageing algorithms, and do not have the necessary backing of COTS equipment to support deployment in ITS applications. H.261 uses the DCT method for video compression and provides integer (pixel) motion compensation accuracy.

1.2.4 H.263

ITU-T H.263 improves upon H.261 by providing improved video quality at lower bit rates. The H.263 compression technique targets a fixed bit rate video transmission. The downside of having a fixed bit rate is that when an object moves, the quality of the image decreases. H.263 was originally designed for video conferencing applications and not for surveillance where details are more crucial than fixed bit rate.

The encoder provides a self-contained video stream, which is multiplexed with control data and audio. H.263 represents multiple standards, identified as H.263+ and H.263++. The baseline specification of H.263 includes support for the following frame types:

- ◊ Intra or I Frame – a frame having no reference frame for prediction
- ◊ Inter or P Frame – a frame based on a previous frame
- ◊ PB Frame and Improved PB Frame – a frame representing two frames and based on a previous frame
- ◊ B Frame – a frame based on two reference frames, one previous and one afterwards
- ◊ EI Frame – a frame having a temporally simultaneous frame which has either the same or smaller frame size
- ◊ EP Frame – a frame having two reference frames, one previous and one simultaneous

Table 1.2.4-1 H.263 Layer Definitions

Parameter	16CIF	4CIF	CIT	QCIF	SQCIF
Active Resolution	1408*1152	704*576	352*288	176*144	128*96
Frame Rate Refresh	29.97 Hz				
YCbCr Sampling Structure	4:2:0				

The H.263 series of video codecs employ relatively modern algorithms, but do not have the necessary backing of COTS equipment to support deployment in ITS applications. H.263 uses the DCT method for video compression and offers integer (1 pel) and half-pixel motion compensation accuracy.

1.2.5 MPEG1

MPEG1 was released in 1993 and intended for storing digital video onto CDs. Therefore, most MPEG1 encoders and decoders are designed for a target bit-rate of about 1.5 Mbps at CIF resolution. For MPEG1, the focus is on keeping the bit-rate relatively constant at the expense of a varying image quality, typically comparable to VHS video quality. The frame rate in MPEG1 is locked at 25 (PAL)/30 (NTSC) fps.

MPEG1's basic principle is to compare two compressed images to be transmitted over the network. The first compressed image is used as a reference frame, and only parts of the following images that differ from the reference image are sent. The network viewing station then reconstructs all images based on the reference image and the "difference data". MPEG1 provides motion prediction in its algorithm.

MPEG1 Operating Parameters	
Horizontal Resolution	≤ 768 samples
Vertical Resolution	≤ 576 Scan Lines
Picture Area	≤ 396 Macroblocks
Picture Rate	≤ 30 Frames per Second
Bit Rate	≤ 1.856 Mbps

Real-time Transport Protocol (RTP) per RFC2250 supports encapsulation and delivery of MPEG1 compressed video over IP networks.

The MPEG1 series of video codecs employ relatively modern algorithms, but do not have the necessary backing of COTS equipment to support deployment in ITS applications. MPEG1 uses the DCT method for video compression and offers integer (pixel) and half-pixel motion compensation accuracy.

1.2.6 MPEG2

MPEG-2 was approved in 1994 as a standard and was designed for high quality digital video (DVD), digital high-definition TV (HDTV), interactive storage media

(ISM), digital broadcast video (DBV), and cable TV (CATV). The MPEG-2 project focused on extending the MPEG-1 compression technique to cover larger pictures and higher quality at the expense of a lower compression ratio and higher bit-rate. The frame rate is locked at 25 (PAL)/30 (NTSC) fps just as in MPEG-1.

MPEG2 supports four levels, which specify resolution, frame rate, coded bit rate, etc. for a given profile. MPEG2 supports six different profiles, which specify the coding syntax (algorithm) to be used. Most chip sets support the Main Profile. Table 1.2.6-1 provides details of the MPEG2 Profiles and Levels.

Table 1.2.6-1 MPEG2 Profiles Indicating Maximum Bit Rate (ref. Video Demystified)

Level	**Profile**					
	Non-Scalable				**Scalable**	
	Simple	Main	Multiview	4:2:2	SNR / Spatial	High
High	n/a	80	130 (both layers) 80 (base layer)	300	n/a	100 (all layers) 80 (middle & base layers) 25 (base layer)
High 1440	n/a	60	100 (both layers) 60 (base layer)	n/a	60 (all layers) 40 (middle & base layers) 15 (base layer)	80 (all layers) 60 (middle & base layers) 20 (base layer)
Main	15	15	25 (both layers) 5 (base layer)	50	15 (both layers) 10 (base layer)	20 (all layers) 15 (middle & base layers) 4 (base layer)
Low	n/a	4	8 (both layers) 4 (base layer)	n/a	4 (both layers) 3 (base layer)	n/a

Table 1.2.6-2 MPEG2 Levels Indicating Maximum Bit Rate (ref. Video Demystified)

MPEG2 Level	**Resolution Supported**	**Maximum Frame Rate**	**Bandwidth Requirement**
Low Level	352*288	60 fps	3 - 8 Mbps
Main Level	720*576	60 fps	4 – 50 Mbps
High 1440 Level*	1440*1080	30 fps	15 – 100 Mbps
High Level*	1920*1080	30 fps	25 – 300 Mbps

Note: MPEG2 High 1440 and High levels provide High Definition (HD) content, where Low and Main levels provide Standard Definition (SD).

MPEG2 supports three different types of streams: elementary, program and transport. The MPEG2 Transport stream is tailored for communicating or storing one or more programs of MPEG2 compressed data and also other data in relatively error-prone environments. The MPEG2 Program stream is tailored for relatively error-free environments and applications such as writing to a CDROM or DVD. While both Program and Transport stream have been deployed in ITS applications, it is recommended that Transport stream be used for distribution over long distances. The MPEG2 Transport stream additionally supports multiple video and data signals (with different synchronization) multiplexed into the same stream.

Real-time Transport Protocol (RTP) per RFC2250 supports encapsulation and delivery of MPEG2 compressed video over IP networks. Additional IP network support includes the following protocols: RSVP, RTSP, RTP & RCP.

The MPEG2 series of video codecs employ modern algorithms, and have the necessary backing of COTS equipment to support deployment in ITS applications. MPEG2 continues to be the algorithm of choice for video distribution to public media/news agencies. MPEG2 also supports compression of High Definition video signals. MPEG2 uses the DCT method for video compression and offers half-pixel motion compensation accuracy.

1.2.7 MPEG4 (Part 2/ASP)

MPEG-4 is a major development from MPEG-2, primarily from advanced feature sets and ability to provide higher quality at a reduced bandwidth. MPEG4 Part 2 has 11 different profiles defined within the specification with a maximum data rate ≤ 38.4 Mbps. Table 1.2.7-1 identifies the all of the MPEG4 Part 2 profiles. The majority of MPEG4 Part 2 devices capable of supporting ITS deployments operate using the MPEG4 ASP specification.

Table 1.2.7-1 MPEG4 Part 2 Profiles Indicating Resolution and Maximum Bit Rate (ref. Video Demystified)

MPEG4 Part 2	Typical Resolution	Maximum Bit Rate
Main	BT.709 BT.601 CIF	38.4 Mbps 15 Mbps 2 Mbps
Core	CIF QCIF	2 Mbps 384 kbps
Advanced Core	CIF QCIF	2 Mbps 384 kbps
N-Bit	CIF	2 Mbps
Simple	CIF CIF QCIF	384 kbps 128 kbps 64 kbps
Advanced Simple (ASP)	BT.601 352*576 CIF CIF CIF QCIF QCIF	8 Mbps 3 Mbps 1.5 Mbps 768 kbps 384 kbps 128 kbps 128 kbps

Advanced Real Time Simple	CIF CIF CIF QCIF	2 Mbps 384 kbps 128 kbps 64 kbps
Core Scalable	BT.601 CIF CIF	4 Mbps 1.5 Mbps 768 kbps
Simple Scalable	CIF CIF QCIF	256 kbps 128 kbps 128 kbps
Advanced Coding Efficiency	BT.709 BT.601 CIF CIF	38.4 Mbps 15 Mbps 2 Mbps 384 kbps
Fine Granularity Scalable	BT.601 352*576 CIF CIF QCIF QCIF	8 Mbps 3 Mbps 768 kbps 384 kbps 128 kbps 128 kbps

The MPEG4 series of video codecs employ modern algorithms, and have the necessary backing of COTS equipment to support deployment in ITS applications. MPEG4 bolsters diversity in support from many 3rd party software providers, including Microsoft Media Player, Apple QuickTime and others. 360 Surveillance's Chameleon currently supports the MPEG4 ASP specification. MPEG4 Part 2 uses both the DCT and DWT methods for video compression. MPEG4 Part 2 offers ¼ pel motion compensation accuracy.

1.2.8 H.264 or MPEG-4 (Part 10/AVC)

The two groups behind H.263 and MPEG-4 joined together to form the next generation video compression standard: called H.264 or MPEG-4 Part 10. The intent is to achieve very high data compression. This standard would be capable of providing good video quality at bit rates that are substantially lower than what previous standards would need, and to do so without so much of an increase in complexity as to make the design impractical or expensive to implement.

The H.264 standard includes the following six sets of capabilities, which are referred to as profiles, targeting specific classes of applications:

- **Baseline Profile (BP)**: Primarily for lower-cost applications demanding less computing resources, this profile is used widely in videoconferencing and mobile applications.
- **Main Profile (MP)**: Originally intended as the mainstream consumer profile for broadcast and storage applications, the importance of this profile faded when the High profile was developed for those applications.
- **Extended Profile (XP)**: Intended as the streaming video profile, this profile has relatively high compression capability and some extra tricks for robustness to data losses and server stream switching.

- **High Profile (HiP)**: The primary profile for broadcast and disc storage applications, particularly for high-definition television applications (this is the profile adopted into HD DVD and Blu-ray Disc, for example).
- **High 10 Profile (Hi10P)**: Going beyond today's mainstream consumer product capabilities, this profile builds on top of the High Profile — adding support for up to 10 bits per sample of decoded picture precision.
- **High 4:2:2 Profile (Hi422P)**: Primarily targeting professional applications that use interlaced video, this profile builds on top of the High 10 Profile — adding support for the 4:2:2 chroma sampling format while using up to 10 bits per sample of decoded picture precision.
- **High 4:4:4 Profile (Hi444P) [deprecated]**: This profile builds on top of the High 4:2:2 Profile — supporting up to 4:4:4 chroma sampling, up to 12 bits per sample, and additionally supporting efficient lossless region coding and an integer residual color transform for coding RGB video while avoiding color-space transformation error. **Note:** The High 4:4:4 Profile is being removed from the standard in favor of developing a new improved 4:4:4 profile.

Table 1.2.8-1 H.264 Profile Descriptions (ref. Video Demystified)

Profile Options	Baseline	Extended	Main	High	High 10	High 4:2:2	High 4:4:4
I and P Slices	Yes	Yes	Yes	Yes	Yes	Yes	Yes
B Slices	No	Yes	Yes	Yes	Yes	Yes	Yes
SI and SP Slices	No	Yes	No	No	No	No	No
Multiple Reference Frames	Yes	Yes	Yes	Yes	Yes	Yes	Yes
In-Loop Deblocking Filter	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CAVLC Entropy Coding	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CABAC Entropy Coding	No	No	Yes	Yes	Yes	Yes	Yes
Flexible Macroblock Ordering (FMO)	Yes	Yes	No	No	No	No	No
Arbitrary Slice Ordering (ASO)	Yes	Yes	No	No	No	No	No
Redundant Slices (RS)	Yes	Yes	No	No	No	No	No
Data Partitioning	No	Yes	No	No	No	No	No
Interlaced Coding (PicAFF, MBAFF)	No	Yes	Yes	Yes	Yes	Yes	Yes
4:2:0 Chroma Format	Yes	Yes	Yes	Yes	Yes	Yes	Yes
4:2:2 Chroma Format	No	No	No	No	No	Yes	Yes
4:4:4 Chroma Format	No	No	No	No	No	No	Yes
8 Bit Sample Depth	Yes	Yes	Yes	Yes	Yes	Yes	Yes
9 and 10 Bit Sample Depth	No	No	No	No	Yes	Yes	Yes
11 and 12 Bit Sample Depth	No	No	No	No	No	No	Yes
8x8 vs. 4x4 Transform Adaptivity	No	No	No	Yes	Yes	Yes	Yes
Quantization Scaling Matrices	No	No	No	Yes	Yes	Yes	Yes

Separate Cb and Cr QP control	No	No	No	Yes	Yes	Yes	Yes
Monochrome Video Format	No	No	No	Yes	Yes	Yes	Yes
Residual Color Transform	No	No	No	No	No	No	Yes
Predictive Lossless Coding	No	No	No	No	No	No	Yes

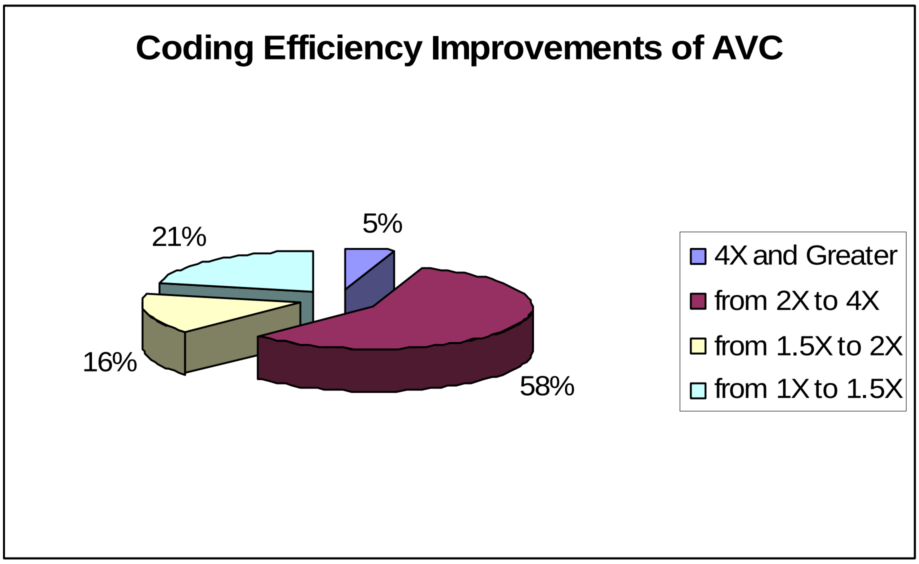
Table 1.2.8-2 H.264 Level Descriptions (ref. Video Demystified)

Level number	**Max macro-blocks per second**	**Max frame size (macro-blocks)**	**Max video bit rate (VCL) for Baseline, Extended and Main Profile**	**Max video bit rate (VCL) for High Profile**	**Max video bit rate (VCL) for High 10 Profile**	**Max video bit rate (VCL) for High 4:2:2 and High 4:4:4 Profile**	**Examples for high resolution / frame rate in this profile**
1	1485	99	64 kbit/s	80 kbit/s	192 kbit/s	256 kbit/s	128x96/30.9 176x144/15.0
1b	1485	99	128 kbit/s	160 kbit/s	384 kbit/s	512 kbit/s	128x96/30.9 176x144/15.0
1.1	3000	396	192 kbit/s	240 kbit/s	576 kbit/s	768 kbit/s	176x144/30.3 320x240/10.0
1.2	6000	396	384 kbit/s	480 kbit/s	1152 kbit/s	1536 kbit/s	176x144/60.6 320x240/20.0 352x288/15.2
1.3	11880	396	768 kbit/s	960 kbit/s	2304 kbit/s	3072 kbit/s	352x288/30.0
2	11880	396	2 Mbit/s	2.5 Mbit/s	6 Mbit/s	8 Mbit/s	352x288/30.0
2.1	19800	792	4 Mbit/s	5 Mbit/s	12 Mbit/s	16 Mbit/s	352x480/30.0 352x576/25.0
2.2	20250	1620	4 Mbit/s	5 Mbit/s	12 Mbit/s	16 Mbit/s	720x480/15.0 352x576/25.6
3	40500	1620	10 Mbit/s	12.5 Mbit/s	30 Mbit/s	40 Mbit/s	720x480/30.0 720x576/25.0
3.1	108000	3600	14 Mbit/s	17.5 Mbit/s	42 Mbit/s	56 Mbit/s	1280x720/30.0 720x576/66.7
3.2	216000	5120	20 Mbit/s	25 Mbit/s	60 Mbit/s	80 Mbit/s	1280x720/60.0
4	245760	8192	20 Mbit/s	25 Mbit/s	60 Mbit/s	80 Mbit/s	1920x1088/30.1 2048x1024/30.0
4.1	245760	8192	50 Mbit/s	62.5 Mbit/s	150 Mbit/s	200 Mbit/s	1920x1088/30.1 2048x1024/30.0
4.2	522240	8704	50 Mbit/s	62.5 Mbit/s	150 Mbit/s	200 Mbit/s	1920x1088/64.0 2048x1088/60.0
5	589824	22080	135 Mbit/s	168.75 Mbit/s	405 Mbit/s	540 Mbit/s	1920x1088/72.3 2560x1920/30.7
5.1	983040	36864	240 Mbit/s	300 Mbit/s	720 Mbit/s	960 Mbit/s	1920x1088/120.5 4096x2048/30.0

H.264 is relatively new to the marketplace. Chipsets supporting both Standard Definition and High Definition encoding and decoding are currently available. As indicated in Table 1.2.8-2, the H.264 algorithm will support HDTV. H.264 is currently the most modern compression algorithm to date, but may not have the necessary backing of COTS equipment to support deployment in ITS applications at this time. H.264 has currently been implemented in temperature hardened devices by Delta and Teleste. Additionally, Sony and SmartVue provide (non-hardened) H.264 compatible IP camera systems. H.264 software implementations include Apple QuickTime, and is currently being implemented in products from vendors such as 360 Surveillance's Chameleon. The ability to fully deploy IP compatible, H.264 based video codecs for ITS applications should take place by Q4'07. H.264 uses the simple Integer Transform method for video compression and offers ¼ pel motion compensation accuracy.

Figure 1.2.8-3 illustrates test results from NIST related to the improvement of H.264 when compared to both MPEG2 and MPEG4 (Part 2).

Figure 1.2.8-3 Coding Efficiency Gains in AVC/H.264 relative to MPEG-2 and MPEG-4. The percentage is based on a count of conditions of test clip and comparison codec (both MPEG-2 and MPEG-4.). The percent is the fraction of the statistically conclusive test conditions for which the indicated improvement was measured. (ref. NIST, *Subjective testing methodology in MPEG video verification)*



1.2.9 VC-1 (SMPTE/Windows Media Player WM-9)

SMPTE recently adopted the VC-1 video codec as a project supporting internet video. Essentially, the VC-1 codec utilizes the same technology as the H.264 specification with one major exception. The VC-1 codec does require the same video filtering during the recomposition of the video signal. Essentially, VC-1 is a lighter (mathematically less complex) version of H.264 at the sacrifice of image fidelity. This fact is currently being used to help market the codec, by saying that the VC-1 decoder operates with less latency than H.264. No comparative analysis between H.264 and VC-1 related to PSNR have been identified, however observation of VC-1 video streams @ approximately 500 kbps showed indications of motion artifacts. Furthermore, WM9 which supports VC-1 is the native streaming video API supported by Windows Server 2003 and newer products.

The VC-1 documents are SMPTE 421M-2006, "VC-1 Compressed Video Bitstream Format and Decoding Process" - the Standard itself, as well as two supporting Recommended Practices, SMPTE RP227-2006 "VC-1 Bitstream Transport Encodings" and SMPTE RP228-2006 "VC-1 Decoder and Bitstream Conformance".

VC-1 supports three profiles: Simple, Main, and Advanced. VC-1 content is transport-independent and container-independent, allowing delivery over MPEG-2 and real-time transfer protocol (RTP) systems as well as advanced systems format (ASF). Table 1.2.9-1 provides the details for each level and profile supported by the VC-1 standard.

Table 1.2.9-1 VC-1 Profiles and Levels (ref. Microsoft)

Profile	Level	Maximum Bit Rate	Representative Resolutions by Frame Rate (Format)
Simple	Low	96 kilobits per second (Kbps)	176 x 144 @ 15 Hz (QCIF)
	Medium	384 Kbps	240 x 176 @ 30 Hz 352 x 288 @ 15 Hz (CIF)
Main	Low	2 megabits per second (Mbps)	320 x 240 @ 24 Hz (QVGA)
	Medium	10 Mbps	720 x 480 @ 30 Hz (480p) 720 x 576 @ 25 Hz (576p)
	High	20 Mbps	1920 x 1080 @ 30 Hz (1080p)*
Advanced	L0	2 Mbps	352 x 288 @ 30 Hz (CIF)
	L1	10 Mbps	720 x 480 @ 30 Hz (NTSC-SD) 720 x 576 @ 25 Hz (PAL-SD)
	L2	20 Mbps	720 x 480 @ 60 Hz (480p) 1280 x 720 @ 30 Hz (720p)
	L3	45 Mbps	1920 x 1080 @ 24 Hz (1080p)* 1920 x 1080 @ 30 Hz (1080i) 1280 x 720 @ 60 Hz (720p)

	L4	135 Mbps	1920 x 1080 @ 60 Hz (1080p)* 2048 x 1536 @ 24 Hz

1.3 Market Analysis of COTS Video Codec Equipment

KCS Systems, Inc. performed a marketing survey of Common-Off-The-Shelf (COTS) video CODEC hardware. Specifically, research was performed to identify two major end items: a) CCTV cameras with integrated CODEC, and b) Stand-alone video CODEC models. Data sheets from each piece of equipment found were electronically gathered (.pdf, .doc. &.xls) for incorporation into master spreadsheets where equipment specifications could be viewed in a side-by-side manner. The worksheets were subsequently formatted to provide rough demographics associated with each product type. All types of IP equipment were compared for indoor vs. outdoor and codec type based on standard.

1.3.1 Integrated CCTV – CODEC Solutions

The sample size for integrated CCTV/CODEC solutions numbered 109 different products. IP based web cameras were intentionally not included in the survey, as the study focused on professional video equipment intended for use in ITS environments. IP cameras were further grouped by identification as indoor and outdoor, as well as form factor. Form factors included three groups: a) fixed cameras, b) dome cameras and c) PTZ cameras (barrel + PTZ). Figure 1.3.1-1 illustrates the distribution of Integrated CCTV/CODEC solutions available and indicates that fixed CCTVs make a majority of the sample.

Figure 1.3.1-1 Categorization of Integrated CCTV/CODEC Solutions in Sample

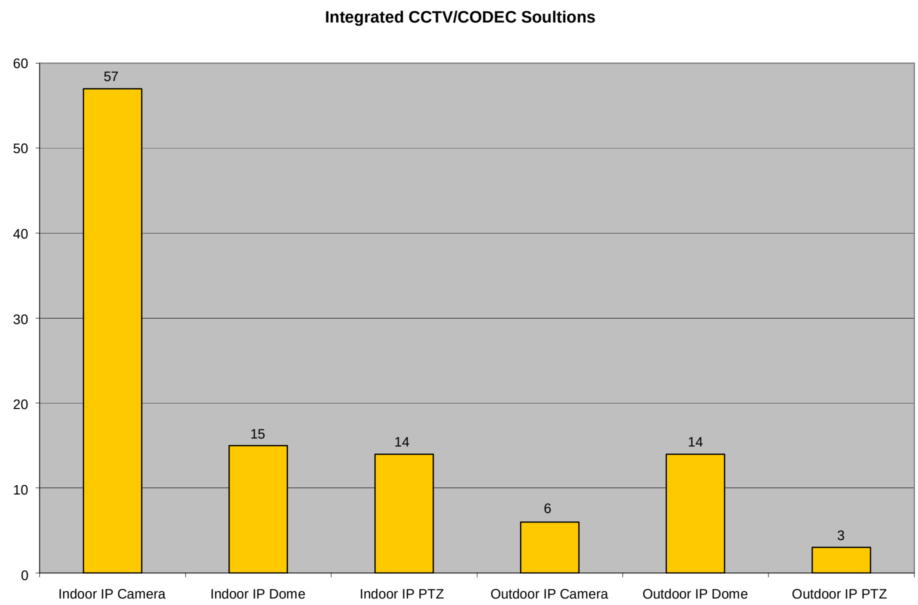
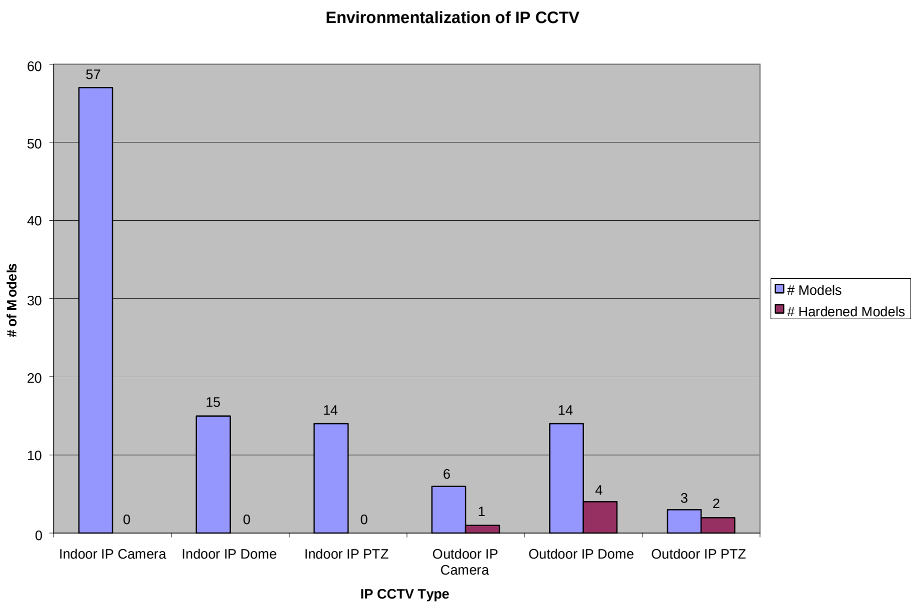


Figure 1.3.1-2 illustrates the actual number of temperature hardened IP CCTV solutions that come close to or meet the -32° to +74°C NEMA range. Results indicate that only seven (7) models from the sample should even be considered for deployment in ITS environments.

Figure 1.3.1-2 Temperature Hardened Versions of the IP CCTV

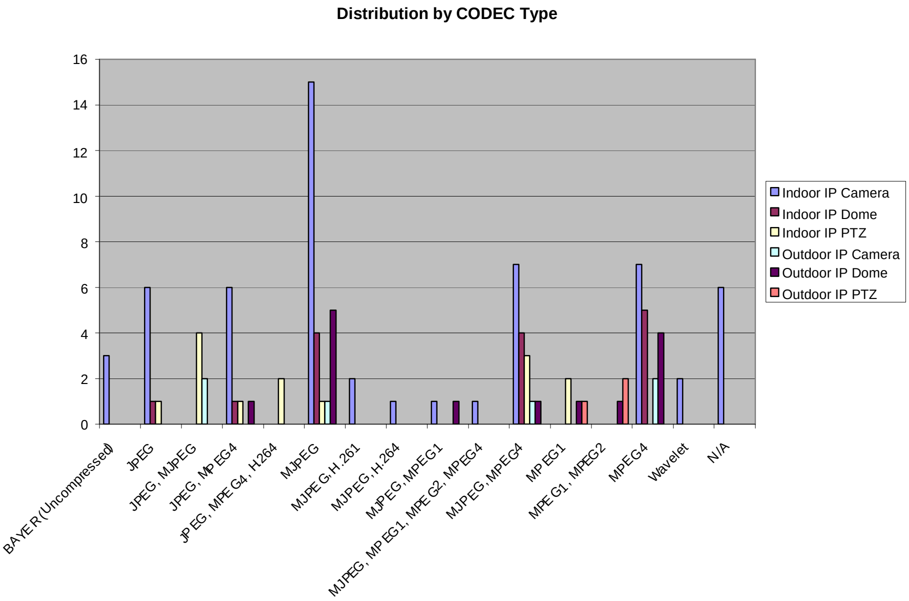


In the sample set, a number of IP CCTV devices provided support for multiple video codecs. Currently, some of the devices are capable of providing snap shot (JPEG) still images and sending over the network concurrent with the full motion video (MPEG4, etc.). Models supporting multiple codec specifications for example, versions 1 thru 4 of the MPEG specification, are typically procured with the selection of one (1) codec algorithm from the factory. Loading different codec algorithms requires flashing (re-programming) the firmware that is in the device. Switching between codec algorithms "on-the-fly" is not readily supported, and is cost prohibitive to implement, maintain and control.

Any device which included digital wrappers such as ".avi" or ".mov", as well as JPEG2000 in the codec description were included with the MJPEG category totals. Devices identified as operating with "wavelet" compression were left as a separate category, since a determination can't be made as to the proper standards family the codec would belong to (either JPEG2000 or MPEG4 Part2). Similarly, the Bayer format is considered as uncompressed digital video, and is left in its own category. The N/A category indicates that a determination of compression algorithm could not be identified from the vendor's specification sheet, and those values are also tabulated as a separate category. Figure 1.3.1-

3 illustrates the number of IP CCTV product indexed by form factor and codec algorithm used.

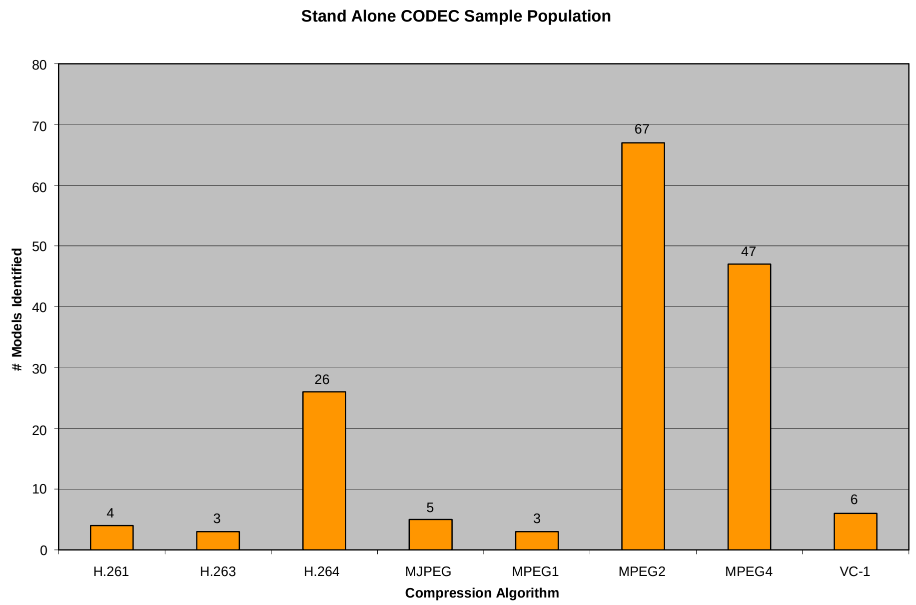
Figure 1.3.1-3 IP CCTV Models Indexed by Form Factor and Codec Algorithm



1.3.2 Stand-Alone Video CODECs

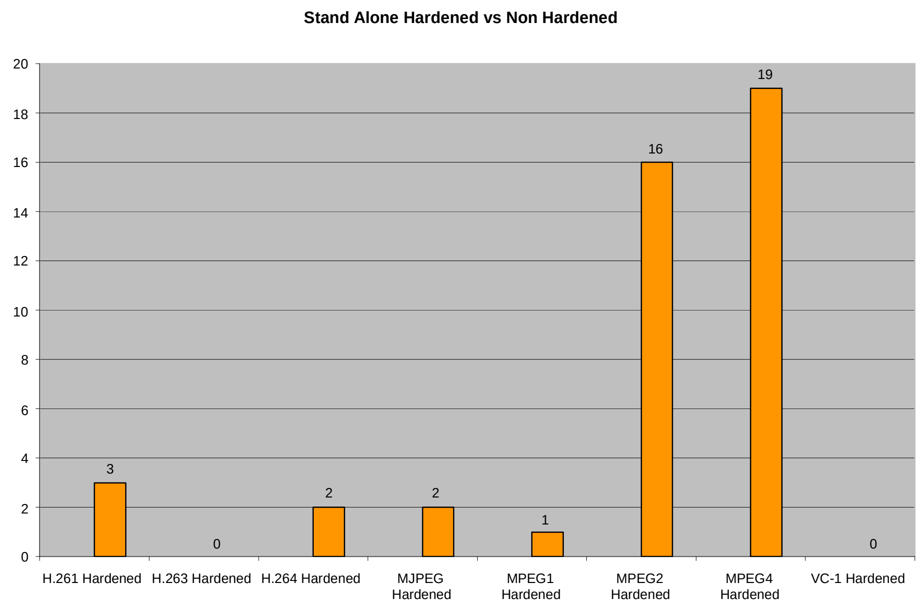
The sample population for stand-alone video codecs included 161 different models. Models identified for inclusion in the survey included units from the professional broadcasting, teleconferencing, military and ITS industry sectors. Stand-alone video codecs were grouped as indoor versus outdoor, and by compression algorithm supported. Results indicate that MPEG2 has the greatest overall number of models available, but places 2nd behind MPEG4 when the number of environmentalized models are summed. Both MPEG2 and MPEG4 hardware are proven within the ITS industry, and offer the greatest probability of interoperability based on vendor support for the standards. H.264, while being the newest codec standard, ranks 3rd in overall numbers for models available, and is supported by a number of industry power players such as Cisco, Tandberg and Thompson Grass-Valley. H.264 is currently being offered in an environmentalized version by Teleste. Teleste is one in a group of manufacturer's that compete for ITS market influence. It is not unlikely that other product manufacturers such as Cohu, Cornet, CoreTec, IFS, Optelecom and VBrick will begin to provide competing hardware. At the time of the market survey, no VC-1 hardware was available. Figure 1.3.2-1 illustrates the distribution of stand-alone codecs by algorithm.

Figure 1.3.2-1 Distribution of Stand Alone Codecs by Algorithm Supported



The numbers of temperature hardened stand alone codecs occur more often than with the integrated CCTV/CODEC solutions surveyed. Figure 1.3.2-2 illustrates the number of temperature-hardened codecs from the overall number of codecs in the same algorithm group.

Figure 1.3.2-2 Occurrence of Hardened Codecs within Algorithm Groups



1.4 Video CODEC Recommendation Supporting a Regional ITS Center-to Center Network

Selection of video processing and control equipment is crucial to the day-to-day operations associated with integrated Freeway and Arterial Management Systems. Consideration of current operations as well as expansion projects and interoperability projects, require that technology is chosen and deployed in such a manner that ensures the best long-term cost objectives. In order to meet this objective, technology needs to be assessed not only for its technical viability, but also on the institutional needs of the organization and on supporting interoperable features of the technology.

Currently, video compression technology is on the move. There are a number of CODEC implementations, each supporting different compression algorithms and related standards. Within a particular compression standard, multiple profile and level definitions exist which define the operating characteristic of that codec type. Implementation of profiles and levels from a particular vendor depends upon factors such as time-to-market, specification maturity and market acceptance. Ultimately, these factors may result in a partial implementation of the standard at the product level, resulting in limited modes of operation and restricting full interoperability.

In order to evaluate video codecs, a comparative matrix was established. Each video codec was weighed by eleven (11) different categories that are considered to be critical to the evaluation. Each category in the matrix is worth 10 points each. Table 1.4-1 provides the comparison between codecs currently under evaluation for the MAG RCN project.

- Bandwidth Utilization – effectiveness of compression algorithm and ability to support both narrowband and wideband operations
- HD Support – standard's ability to support High Definition (HD) resolutions
- Support $\leq$ 64 kbps – support for remote locations using dial-up technology such as Integrated Services Digital Network (ISDN)
- DCT=5, DWT=7, IT=9 – relative weighting based on comparison of transform algorithm supported. Basis rests in codec efficiency, PSNR of recovered signal, and market acceptance related to standard.
- Motion Compensated Algorithm – motion compensation as based on prediction accuracy.
- Product Availability – number of available COTS equipment from vendors included in study sample
- Temperature Hardened Products – number of available COTS equipment from vendors included in study sample
- Obsolescence – technology and equipment considered to be at end of life
- Chameleon Integration – ability to support 360 Surveillance's Chameleon, as an indication of product interoperability
- Prior ITS Deployments – market support for CODECS of a particular type

- Cost – associated cost of COTS equipment

Table 1.4-1 IP Codec Trade-off Matrix

Compression Standard	Bandwidth Utilization	HD Support	Support ≤ 64 kbps	DCT=5, DWT=7, IT=9	Motion Compensated Algorithm	Product Availability	Temperature Hardened Products	Obsolescence	Chameleon Integration	Prior ITS Deployments	Cost	Total Points
JPEG	3	4	2	5	0	5	4	7	7	0	5	42
MJPEG	3	4	2	5	0	7	2	5	7	7	4	46
JPEG2000	4	7	3	7	0	2	1	5	5	0	4	38
H.261	5	1	8	5	4	1	3	1	5	5	6	44
H.263++	6	5	8	5	5	1	0	3	3	1	5	42
MPEG1	4	3	8	5	5	1	1	3	10	3	5	48
MPEG2	5	9	1	5	8	7	8	8	10	7	7	75
MPEG4 (Part2)	6	7	8	7	8	8	9	7	10	7	8	85
H.264	9	10	9	9	10	5	2	9	8	0	6	77
VC-1	9	10	9	9	9	0	0	7	1	0	6	60

The results from the technology matrix clearly favor three technologies, H.264, MPEG2 and MPEG4 Part2. Each of these three codec standards are based on open standards protocols. A number of different manufacturers that are providing support for these protocols in products that are oriented towards deployment in ITS environments. This will guarantee competition in the availability of the devices over the long term.

What the technology matrix fails to express, is the technology trend associated with the adoption of the H.264 technology. This technology is clearly the best overall performing codec solution. H.264 is relatively new, and is gaining popular support amongst internet users. Due to H.264's current deployment in cable set top boxes which are providing high definition for the home, this technology is on the verge of popular consumer acceptance. It is for these reasons that H.264 should be provided as the codec of choice for new ITS implementations.

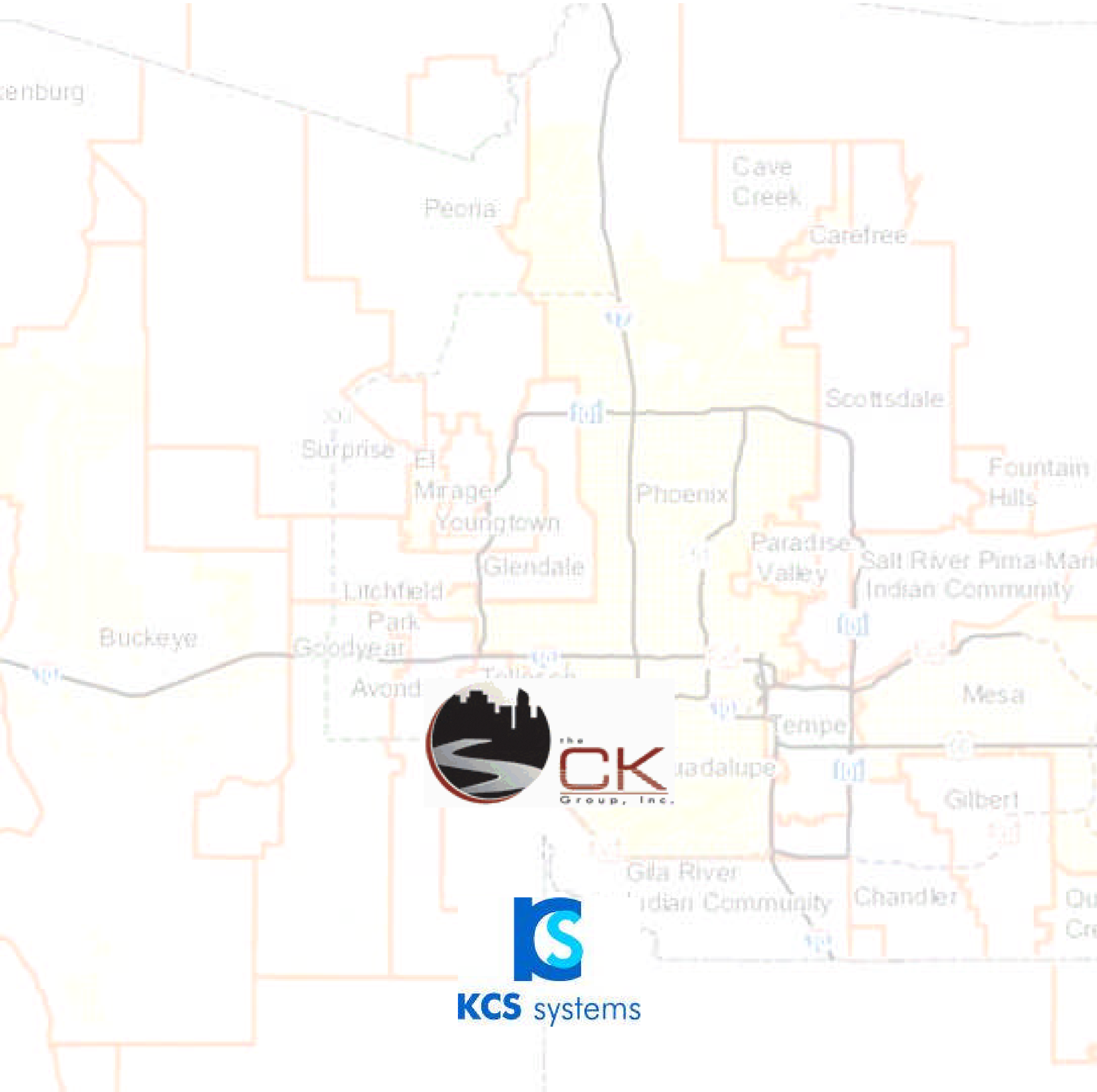
2.0 References

"Video Demystified, Fourth Edition, A Handbook for the Digital Engineer", 2005, Keith Jack, ISBN 0-7506-7822-4

Institute of Electrical and Electronics Engineers (IEEE), IEEE Transactions on Circuits and Systems for Video Technology, July 2003, *"Overview of the H.264/AVC Video Coding Standard"*, Thomas Wiegand, Gary Sullivan, Gisle Bjontegaard and Ajay Luthra

Stanford University, Information Systems Laboratory, *“Network Aware h.264/AVC Region of Interest Coding for a Multi-Camera Wireless Surveillance Network”*, Pierpaolo Baccichet, Xiaoqing Zhu and Bernd Girod

Indigo Vision Group plc, *“IP Video CCTV and H.264”*, Barry Keepence



the CK Group, Inc.

KCS systems