

# Understanding IP Video for

**D**ealers

**U**sers

**M**anagers

**I**nstallers

**E**ngineers

**S**alespeople

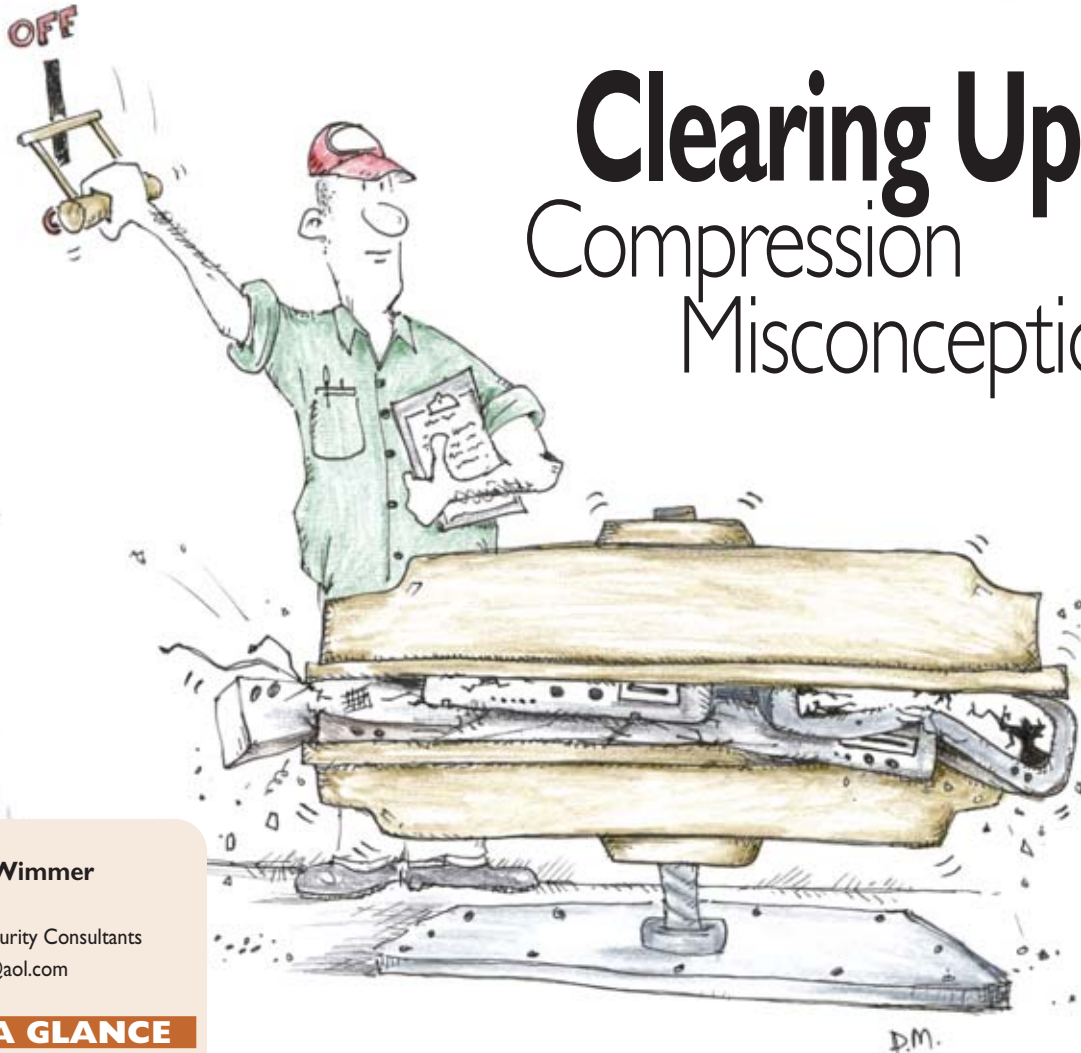
Presented by

**SecuritySales**  
& INTEGRATION

Part 3 of 4

Brought to You by





## Clearing Up Compression Misconception

### By Bob Wimmer

Principal  
Video Security Consultants  
cctvbob@aol.com

### AT A GLANCE

- Three forms of bandwidth compression are Common Intermediate Formats (CIF) sizing; standard video compression (.jpeg, .mpeg); and images per second (ips) transmitted
- Bit rate is how much physical space an image occupies in 1 second on a network; the higher the bit rate, the more space it requires
- The majority of IP cameras used in surveillance incorporate JPEG, MJPEG or MPEG-4 for their compression method
- Lossy compression, which is used in most IP cameras, takes advantage of two principles: irrelevant reduction and redundancy reduction

**As the size of broadband pipelines continue to increase, the amount of data seems to expand right along with them. Therefore, the bandwidth consumed by IP video system transmission continues to be paramount. Fortunately, there are numerous compression solutions available to help.**

**W**elcome to Part III of the latest in *Security Sales & Integration's* acclaimed "D.U.M.I.E.S." series: "Understanding IP Video for D.U.M.I.E.S." Brought to you by Pelco, this four-part series has been designed to educate readers about networked video — the next phase of surveillance technology following the quantum leap from analog to digital CCTV systems. "D.U.M.I.E.S." stands for dealers, users, managers, installers, engineers and salespeople.

In Part 1 of this year's series (see "Using Camera Specs to Solve IP Application Issues" in the *March issue*), we discussed the design of cameras, their specifications and how it relates to

IP cameras as well as their networks. We also discussed how to determine the quality of cameras and whether they are designed for indoor or outdoor applications.

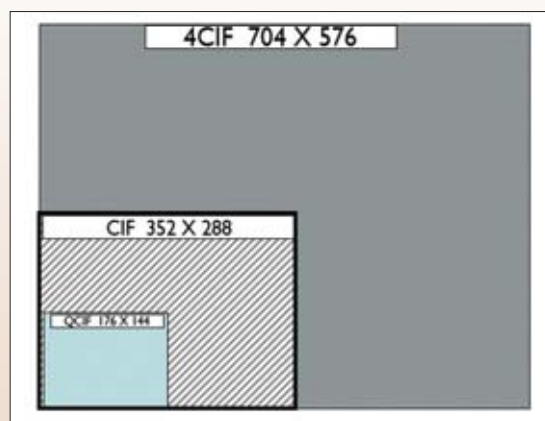
In Part II (see "Clear Eye for the IP Video Guy" in the May issue), we investigated ways to enhance IP camera image quality and basic system layouts. The material also addressed how image file sizes as well as bit rates affect both system bandwidth requirements and the amount of storage required for different applications.

This month's lesson takes it all a step further by explaining the need for video bandwidth compression and the different methods used in reducing system bandwidth for use on any IP network.

### CIF Is 1 of 3 Forms of Compression

There are three basic forms of bandwidth compression found in the industry: 1. Common Intermediate Formats (CIF) sizing; 2. Standard video compression methods (.jpeg, .mpeg); and 3. Images per second (ips) transmitted.

## Common Intermediate Formats



The pixel size of an image is normally referred to as the CIF size. CIF is a standard video format used in videoconferencing. CIF formats are defined by their resolution and standards both above (2CIF/4CIF) and below (QCIF/SQCIF) with CIF being the original established resolution reference. The original CIF is also known as Full CIF (FCIF).

CIF was discussed in Part II of this series; however, as a refresher or for those who missed that installment, a quick review can help you better understand the complete process incorporated into IP camera setups.

## Bit Rates at 30 ips

| Format                  | Resolution    | Mbps  |
|-------------------------|---------------|-------|
| SQCIF (Sub Quarter CIF) | 128 X 96      | 4.4   |
| QCIF (Quarter CIF)      | 176 X 144     | 9.1   |
| CIF (Full CIF, FCIF)    | 352 X 288     | 36.5  |
| 4CIF (4 X CIF)          | 704 X 576     | 146   |
| 16CIF (16 X CIF)        | 1,408 X 1,152 | 583.9 |

The pixel size of an image is normally referred to as the CIF size. CIF is a standard video format used in videoconferencing. CIF formats are defined by their resolution and standards, both above (2CIF/4CIF) and below (QCIF/SQCIF), with CIF being the original established resolution reference. The original CIF is also known as Full CIF (FCIF). The bit rates shown above are for *uncompressed* color frames.

### A Little Bit About Bit Rates

Bit rate is defined as how much physical space an image occupies in 1 second on a network. The higher the

transmission medium and has the usual physical meaning).

For example, a single camera incorporating a 4CIF (high quality, uncompressed color image @ 30 ips) using a DSL (1.5Mbps) would take in excess of 20 minutes to transmit over the network.

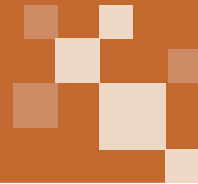
And, as we all are aware of, adding more cameras and/or increasing the movement within video scenes demands greater bandwidth requirements of the network. This bit rate number can and is governed by the compression methods used by the IP camera.

### Compression Means Concessions

In its basic form, compression is the art process of removing information that is deemed irrelevant to the viewer. In this case, the viewers are dealers, systems integrators or anyone else who relies on high quality recorded images. The amount and type of information that is removed varies from system to system and can be controlled by system setup parameters.

Why do we need compression? To help answer this question, let's evaluate the requirements needed to transmit a single minute of uncompressed composite video to a remote location.

Without compression the ability to transmit video over a network would be next to impossible. Referring back to that 1 minute of uncompressed video, incorporating a DSL modem with a 1.5Mbps transmission speed, it would take more than 20 minutes to transmit from point A to point B.



## Spatial Reduction



**Spatial reduction is one of three redundancy reduction methods. It is the reduction of the correlation between neighboring pixel values. Redundancy reduction is accomplished by removing duplication from the signal source, which is found either within a single image or between multiple images of a video stream.**

In today's surveillance community, which depends on the ability to see this information remotely, that is completely unacceptable.

There has always been a tradeoff between the quality of the video and file size. If you require a high quality image, then you must deal with a large file size. On the other hand, if you are willing to settle for a lesser-quality image, your file size will be reduced lower than the required bandwidth to transport the video packets.

With a few exceptions, the majority of IP cameras used in surveillance incorporate JPEG, MJPEG or MPEG-4 for their compression method. However, there has recently been an increase of multiple video stream cameras entering the marketplace.

A quick description of the basic components involved in compression will help set the stage for some of the theory and explanations discussed later in this article.

### IP Cameras Look to Win With Lossy

In Part II of this series, we discussed *lossy* and *lossless* compression. For the most part, all of the methods used in the IP camera world incorporate lossy

compression. Lossy compression actually eliminates some of the data in the image and, therefore, provides greater compression ratios than lossless. Thus, the tradeoff is file size vs. image quality.

Lossless, on the other hand, consists of those techniques guaranteed to generate an exact duplicate of the input data stream after a compress/expand cycle. The difference is basic lossy compression takes advantage of two principles: *irrelevant reduction* and *redundancy reduction*.

Lossless compression algorithms reduce file size with no loss in image quality, although compression ratios are generally weak. Most images destined for print, or when image quality is valued above file size, are compressed using lossless algorithms. File formats with an extension of .tiff or .gif are usually listed as lossless compression methods.

### How Reduction Processes Work

Irrelevancy reduction omits parts of the video signal that is not noticed or perceived by the signal receiver, which in this case is the human eye. Research has shown that small color changes are perceived less accurately

more noticeable conditions of a video image, whereas the higher frequencies are usually related to the finer details of a video image.) File formats with an extension of .bmp, .jpg or .mpg are usually listed as lossy compression methods.

Redundancy reduction is accomplished by removing duplication from the signal source, which is found either within a single image or between multiple images of a video stream.

The first of three redundancy reduction methods is labeled *spatial reduction*. This is the reduction of the correlation between neighboring pixel values.

As seen in the simplified diagram to the left, the data stream can be reduced to single values for each of the four quadrants.

The next reduction method is spectral redundancy. This is the correlation between color planes or bands within an image. As an example, look at the blue sky in the diagram below. Many areas of that sky have the same numeric value and, therefore, the amount of stored information can be reduced to reproduce that same image in the decompression mode of operation.

## Spectral Reduction



**Spectral reduction refers to the correlation between color planes or bands within an image. In the blue sky shown to the left, many areas have the same numeric value and, therefore, the amount of stored information can be reduced to reproduce that same image in a decompressed mode.**

than slight alterations in brightness, and since it is less noticeable, why bother saving this information? In the areas of low and high frequencies, it is also known that low frequency changes are more noticeable to the human eye than high frequency. (Low frequency controls the coarser or

The last area is known as *temporal reduction*. This is the correlation between adjacent frames in a sequence. This information is the basis for MPEG as well as the H.263 and H.264 series of compression methods.

In temporal reduction, two types of image arrangements are analyzed.

The first is a full representation of the viewed image. This is known as the I-frame and is encoded as a single

compression standards presently incorporated by mainstream IP camera manufacturers.

be achieved with only a small amount of image degrading.

JPEG divides up the image into 8 X 8 pixel blocks, and then calculates the discrete cosine transform (DCT) of each block. A quantizer rounds off the DCT coefficients, according to the quantization matrix. This step produces the lossy nature of JPEG, but allows for modest compression ratios. JPEG's compression technique uses a variable length code on these coefficients, and then writes the compressed data stream to an output file (\*.jpg).

## Temporal Reduction



In temporal reduction, two types of image arrangements are analyzed. The first is a full representation of the viewed image, known as the I-frame, and is encoded as a single image, with no reference to any past or future images. In the second method of redundancy reduction, any detected movement triggers the compression process.

image, with no reference to any past or future images. In some circles it is also referred to as the Key-frame. The second concept of temporal is based on the question if there is no movement, why bother saving the information? Any detected movement triggers the compression process.

### More Than 1 Way to Compress

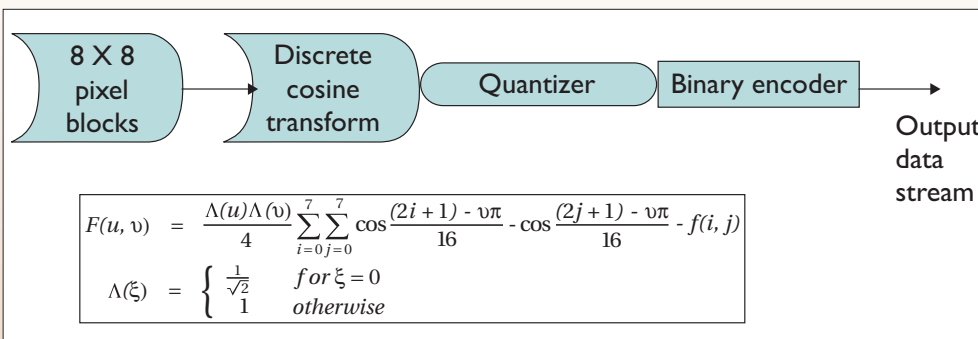
With a basic background about the many different compression theories as well as the different ways video information is reduced, we can now apply this knowledge to the different compression standards available throughout the IP camera industry. This article will only cover the major

**JPEG (Joint Photographic Experts Group)** — JPEG can be a lossless or lossy compression; however, most all IP camera manufacturers incorporate the lossy compression method, meaning that the decompressed image isn't quite the same as the one with which you started. JPEG is designed to exploit known limitations of the human eye, notably the fact that small color changes are perceived less accurately than minor fluctuations in brightness. Thus, it is intended for compressing still images that will be viewed by humans. Data compression is achieved by concentrating on the lower spatial frequencies. According to the standard, modest compression of 20:1 can

**MPEG (Moving Picture Experts Group)** — MPEG compression standards have many facets. Each has its own special features and, as always, improvements are consistently being incorporated that add to the already existing components of the MPEG standard. However, the basics are similar for all versions. MPEG incorporates the same compression method as JPEG (DCT). However, MPEG is based on the group of images concept. The groups of images are defined as the I-Frame, P-Frame and B-Frame.

The I-Frame (intra) provides the starting or access point and offers only a small amount of compression. P-Frames (predicted) are coded with reference to a previous picture, which can be either an I-Frame or another

## JPEG Image Formula



JPEG divides up the image into 8 X 8 pixel blocks, and then calculates the discrete cosine transform (DCT) of each block. A quantizer rounds off the DCT coefficients according to the quantization matrix. This step produces the lossy nature of JPEG, but allows for modest compression ratios.

P-Frame. B-Frames (bi-directional) are intended to be compressed with a low bit rates, using both the previous and future references. The B-Frame is never used as the references. The relationship between the three frame

compress interlaced video. It also is capable of coding standard-definition television at bit rates from about 3-15Mbit/sec and high-definition television. Compression ratios for MPEG-2 vary. The ration depends on the type

found in the MPEG series of compression, but does not use the interframe prediction part of the same series.

**Intraframe/Interframe** — One of the most powerful techniques for compressing video is interframe compression. This method uses one or more earlier or later frames in a sequence to compress the current frame, while intraframe compression uses only the current frame, which is effectively basic image compression.

The most commonly used interframe method works by comparing each frame in the video with the previous one. If the frame contains areas where nothing has moved, the system simply issues a short command that copies that part of the previous frame. If sections of the frame move in a simple manner, the compressor emits a (slightly longer) command that tells the de-compressor to shift, rotate, lighten or darken the copy.

Interframe compression works well for programs that will simply be played back by the viewer, but can cause problems if the video sequence needs to be edited.

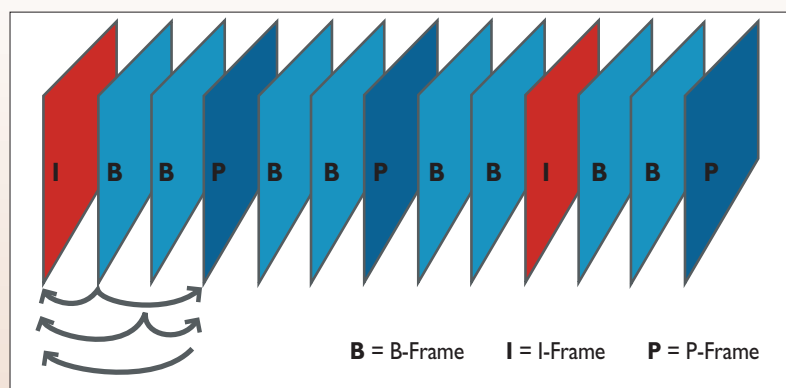
Since interframe compression copies data from one frame to another, if the original frame is lost in transmission the following frames cannot be reconstructed properly. Another difference between intraframe and interframe compression is that with intraframe systems, each frame uses a similar amount of data. In most interframe systems, certain frames (such as "I-Frames" in MPEG-2) aren't allowed to copy data from other frames, and so they require much more data than other nearby frames.

MJPEG is incorporated in many IP cameras because it prioritizes image quality over frame rates.

### Frame Rates Play 'Tricks' on Eyes

According to the theory of persistence of vision, the perceptual process of the retina of the human eye retains an image for a brief moment. Persis-

## MPEG-4 Block Diagram



**MPEG compression standards have many facets. Each has its own special features; however, the basics are similar for all versions. MPEG incorporates the same compression method as JPEG (DCT); however, MPEG is based on the group of images concept. The groups of images are defined as the I-Frame, P-Frame and B-Frame.**

types is described in the MPEG standard; however, it does not restrict the limit to the number of B-Frames between the two references, or the number of images between two I-Frames.

There are many previous MPEG standards that have evolved the past few years. The method currently in use is MPEG-4. The earlier forms of MPEG standards included MPEG-1 and MPEG-2.

The MPEG-1 standard has a resolution of 352 X 240 (CIF) at 30 images a second and incorporates progressive scanning. It is designed for up to 1.5Mbits/sec with compression ratios listed as 27:1.

MPEG-2 is a standard that was introduced in 1994 and has a resolution of 720 X 480 (4CIF), and incorporates both progressive and interlaced scanning. (Interlace scanning is the method used in the CCTV industry to produce images on monitors.) The most significant improvement over MPEG-1 is its ability to efficiently

of signal and number of B, P and I frames. On average, this ration can vary from 50:1 to 100:1.

**MPEG-4** — MPEG-4 is a popular standard for multimedia and Web compression because it is designed for low bit rate transmission. It is based on object-based compression, in which individual objects within a scene are tracked separately and compressed together. This method offers a very efficient compression ratio that is scalable from 20:1 up to 300:1. In today's CCTV industry, more and more manufacturers are turning to MPEG-4 for remote viewing of compressed video images.

**MJPEG (Motion JPEG)** — MJPEG is an informal name for multimedia formats where each video frame or field of a digital sequence is separately compressed as a JPEG image. However, it also uses intraframe coding technology that is very similar to the I-Frame

tence of vision is said to account for the illusion of motion that results when a series of film images are displayed in quick succession, rather than the perception of the individual frames in the series. Humans are not cameras, but the combination of motion, detail and brightness will create a constant movement of images to the individual. The frequency at which flicker becomes invisible is called the *flicker fusion threshold*.

As a refresher, let's review image rates and what they represent. Real-time video is listed in the analog world as 30 frames or 60 fields per second. In the digital world, they are usually referred to as images per second. At these rates, the video produces a steady flow with no loss of movement or information. However, in the IP world, transmitting 60 images per second (ips) per each camera requires a large amount of network bandwidth.

As a result, a compromise was reached and the term used to describe it was referred to as *real-motion video*. This is best conceptualized by considering that if you flash images before your eyes, there is a point in which the flashing rate appears in real-time with apparent missing information. The rate is usually between 15 to 20 images per second (again, this will vary due to the parameters previously mentioned).

As an example, let us compare the bandwidth requirement for two cameras. The first IP camera uses a high quality image size of 4CIF in a real-time color application with no video compression. The bandwidth requirement would calculate as:

$$4\text{CIF} @ 30 \text{ ips} = 146\text{Mbps}$$

With an image size of CIF @ 15 ips = 18.25Mbps, or an eight times reduction in bandwidth requirements.

## Putting It All Together, What's Next

Image compression and images per second, as well as the CIF size of an image, plays a very important part in the transmission of IP images. Most of the equipment offered today gives operators the capability to set up and/or change these parameters in order to meet their networking needs.

You may not see the actual compression ratio settings on any of the setup screens. However, image quality is related to this function. A rating of the highest quality will relate to very small amounts of compression, while the lowest quality setting will pertain to the best of the compression algorithms.

With all of the different forms of reducing image sizes to ensure proper network applications and the many different compression methods, it is no wonder many people get frustrated about IP cameras and networking. The crossover between analog video and IP cameras can be very confusing.

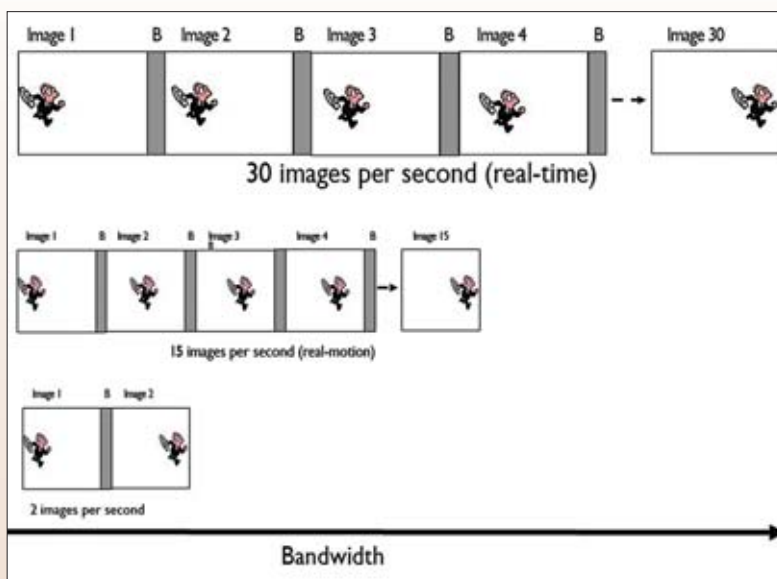
With the variety of image reduction or compression methods, there is one single item to keep in mind: The quality of the reproduced image — whether from a storage device such as a network video recorder (NVR) or a remote location — will depend on the application of that system. Not every method is designed to match all requirements.

When designing and setting up your system, keep in mind if the image quality, speed and storage requirement is what you had expected, then you have made the right choice. The days of trying to fully understand all of the theories, relationships and standards can be long and quite frustrating.

The final part of this series will address the array of different applications, both good and bad, for deploying IP cameras over networks. ■

*Robert (Bob) Wimmer is president of Video Security Consultants and has more than 35 years of experience in CCTV. His consulting firm is noted for technical training, system design, technical support and overall system troubleshooting.*

## Frame Rate Comparison



**Real-time video is listed in the analog world as 30 frames or 60 fields per second. In the digital world, they are usually referred to as images per second. At these rates, the video produces a steady flow with no loss of movement or information. However, in the IP world, transmitting 60 images per second (ips) per each camera requires a large amount of network bandwidth.**