



H.264: Facts and Fiction

It seems everyone in the security industry is talking about the H.264 compression standard for digital video, which produces high-quality video using less bandwidth than commonly used JPEG compression. But how does H.264 differ from JPEG, and are the proposed benefits of H.264 compression too good to be true? Are there any hidden costs to using H.264 in security applications? The industry must focus on the basics of the H.264 compression technology to separate facts from fiction and dispel a few myths and misconceptions.

THE SIMILARITIES

H.264 and JPEG are two closely related standards: computationally they belong to the same family of compression methods. Both use similar or identical techniques to compress the video, such as transforming the video signal into frequency domain, applying quantization to the frequency-transformed signal and using variable length coding. Because the compression methods are similar, the distortion introduced into the video in the process of compression also is similar. The degree of video distortion is proportional to the degree of compression: both standards support a wide range of compression levels and, accordingly, a wide range of achievable video quality (the inverse of video distortion).

There are many metrics of video quality, some objective and some subjective. Using any measure, one can precisely demonstrate that when the compression parameters of the two standards are matched, the video quality of the same scene under like conditions is indistinguishable across a wide range of settings, with the possible exception of the extreme high-compression limit. In particular, this is easy to demonstrate using Arecont Vision's megapixel IP cameras that feature instant switching of the on-camera encoder between JPEG and H.264. In fact, if video quality was the only measure for choosing one compression standard over another, it would be difficult to make the choice.

So, if the video quality of the two standards is very much alike, then how are they different?

THE DIFFERENCES

The main difference between H.264 and JPEG is the consumed bandwidth per given video quality—H.264 offers a major reduction in bandwidth relative to JPEG. Bandwidth reduction translates to a major reduction in cost of security installations: the requirements for networking equipment and disk storage are accordingly reduced.

Reduction of bandwidth is achieved at the cost of high computational complexity of the H.264 encoder. Put simply, the more computation there is, the more efficiently the data is organized and packed. Decoding the compressed video stream is an entirely different matter. The H.264 standard is asymmetrical—all of its computational complexity is on the encoder side—while the H.264 decoder is similar in complexity to a JPEG decoder.

Arecont's megapixel IP cameras use a patent-pending, massively parallel H.264 hardware encoder that achieves 80 billion operations per second. The high computational capacity is needed to process a large number of computational add-ons used in H.264 relative to JPEG, some of which were introduced in the earlier standards of the MPEG family to which H.264 belongs. A major departure from JPEG is that instead of encoding the video signal itself, only the inter-frame signal differences are encoded. The smaller the difference, the more economically it can be encoded into the video stream.

There are two sources of inter-frame signal differences: motion in the scene and random noise.

Noise is always present, and it is notoriously difficult to compress due to its random nature. High levels of noise are typically caused by low-light conditions—they require larger bandwidth and larger disk storage space to archive.

Signal differences due to motion are much easier to compress—the majority of computational effort is typically concentrated in estimating motion. The goal of motion estimation is to locate blocks of pixels in the current video frame that closely match blocks of pixels in the previous frame corresponding to the portions of the scene that may have moved during the interval between frames.

Because the direction and the distance of such movement are unknown in advance, the motion estimator must search hundreds of possible positions to find the best match. The closer the match, the smaller the signal difference to be encoded and, accordingly, the smaller the resultant video stream. Computational power of the motion estimator often determines the quality of the entire H.264 encoder: the larger the search area, the higher the chance to find the best possible match. While many motion estimators conduct only an approximate non-exhaustive search to reduce the amount of computation, other motion estimators conduct an exhaustive search over a large search area to find the best possible match.

Motion estimation and other computational components of H.264 compression explain its amazing ability to compress video into a low-bandwidth stream while maintaining high video quality. It also is the reason why H.264 is being embraced by broadcast television, DVD distributors and other industries, including the professional security and surveillance market.

NO HIDDEN COST

A common myth about H.264 is its so-called hidden cost—an erroneous belief that because the computational complexity of the H.264 encoder is high, the required decoder resources must be high as well. The hidden cost, as the theory goes, is in the additional computer server power needed to decompress multiple H.264 video streams in a multi-camera security installation to display live video from multiple cameras. This hidden cost is alleged to be especially high for megapixel cameras.