White Paper

H.265 High Efficiency Coding

Video Compression for Security Applications

Introduction

H.265 (also known as High Efficiency Video Coding, or HEVC) was jointly developed by the ITU-T Video Coding Experts Group and the ISO/IEC Moving Pictures Experts Group, the same groups that worked together on the H.264 (or Advanced Video Coding, AVC) video compression standard. H.264 has been a very popular and successful compression standard in several markets including security; it was used as the coding and playback engine in video applications including QuickTime, Vimeo, YouTube, and Blu-ray disc players, as well as a wide range of digital security cameras, transmission, and recording devices.

Hanwha Techwin America

As good as H.264 is, the expert groups that created it had proposed and considered a number of enhancements to the standard that were just not feasible at that time due to processing power limitations. Today, processing power has increased significantly at ever-decreasing costs, and it is now possible to implement the more advanced ideas economically. H.265 is, thus, the natural evolution from H.264 and represents another leap forward in video compression technology that has direct applications in the security market.

This video standard is critically important for video security systems because of the increasing deployment of megapixel, HD and UHD (4K) cameras and recorders. These devices capture a much larger amount of high-quality video data than previous lower-resolution models, putting a strain on the transmission and storage network that can be reduced with improved compression. This paper describes some of the new features of H.265 video compression, and how they apply to the security market.

Why and how was H.265 developed?

Technology drivers

Several market forces – consumer electronics in particular – have fueled the demand for ever-increasing display screen sizes and resolution. Many imaging technologies have pushed to increase their resolution, including smart phones and tablets, computer monitors, televisions, and cameras. Security equipment providers have also pushed for higher resolution cameras and more detailed images to support improvements in monitoring, analytics, and forensic review. Supporting the higher resolutions of these devices has been an ongoing challenge for network infrastructure, content providers and distribution systems, who need to support the large quantities of information being captured, transmitted and stored.

The most recent previous video compression standard that supported many of these applications was H.264, or Advanced Video Coding (AVC). The initial standard for H.264 was completed in May of 2003, and was the result of intense effort from both the Moving Pictures Experts Group (MPEG), who were focused on compressing and transmitting consumer video content efficiently, and the ITU's Video Coding Experts Group (VCEG) who were focused on video applications for mobile devices. Together, they tackled the challenge of doubling the performance of the best previous video compression scheme.

After developing H.264, the expert team knew that a more advanced version would be needed in the future, and began work on H.265 soon thereafter. Several companies have proposed and/or implemented enhancements to the H.264 protocol in an effort to extend its life, but none of these initiatives were sufficient to achieve the next level of performance.

Intent of the new H.265 protocol

The expert team set an ambitious goal for themselves and the new codec project: to achieve twice the compression efficiency of the previous standard, H.264/AVC, and support for resolutions up to 8192x4320.
Based on their work on H.264, they knew that significant improvements could be achieved, but that the high cost of the necessary computing power made those improvements not economical for most devices. They reasoned that as computing power continued to improve, the next generation protocol could take advantage of these improvements to achieve the challenging goal.

Although the specification was primarily developed for consumer and business video applications, the system performance enabled by H.265 is ideal for professional surveillance applications – and could even be an early driver for manufacturing and chipset development.

What is H.265, and how is it better?

In this section, H.264 video compression is briefly reviewed, and the key enhancements of the new H.265 compression standard are highlighted.

Sidebar: Compared with the text of a novel or the sound of a favorite song, video content contains enormous amounts of information. Storing the text of a book is relatively easy: 500 pages of text only takes about 1.5MB. CD-quality sound (16 bit stereo, at 44KHz) requires 10.6MB per minute, so a three-minute radio edit takes almost 33MB. (CDs themselves were sized to comfortably hold an hour of music) A single color image 1280 pixels wide and 1024 pixels high (1.3 Megapixel) takes about 4MB. A 1.3 MP video camera might generate 30 such frames per second, for an uncompressed data rate of 120MB per second – almost 700 times higher than the music file. This volume of data is far higher than can be reasonably transmitted and stored for most applications, including security. Thus, all current video formats make use of compression schemes to reduce the amount of data transmitted and stored.

Video compression concepts

To illustrate the enhancements that H.265 brings to video compression, we must first briefly review the basics of video compression, and how the previous state-of-the-art scheme, H.264, works.

To begin, it is important to note that video compression is a complex process that is made up of a combination of techniques. Each technique will have a range of compression results that depend on the actual nature of the video being compressed. This is why each particular compression scheme has a range of overall performance results that will depend on the size and complexity of the images, the speed and nature of the moving elements within the frame, and other factors. By adjusting key settings, the process can be tailored for the best match to the needs of each particular application. For this example, we will describe a simplified version of the actual steps and calculations.

Consider the following two image frames in Figure 1, showing a car moving away from the camera along a roadway:

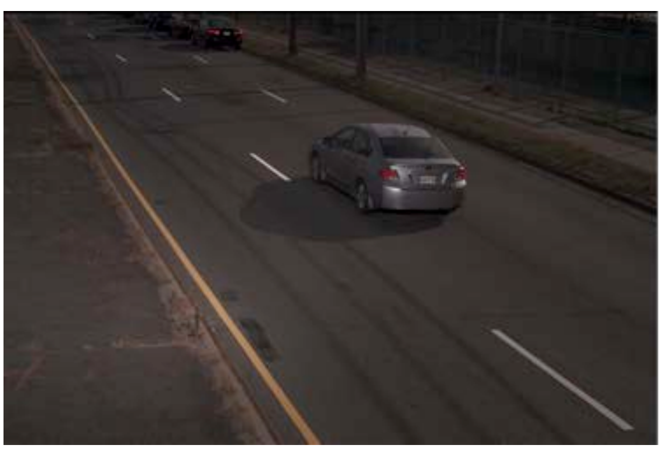
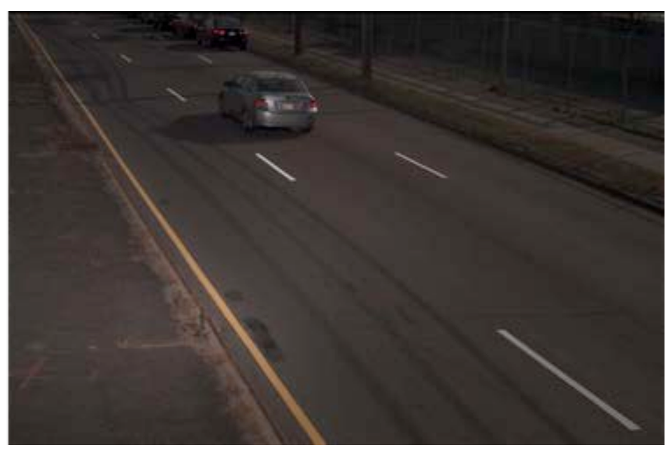



Figure 1: Two complete image frames as captured by a camera

Without compression, we would need to send complete descriptions of both frames. Each frame could be compressed using basic data compression techniques, but even in that case, there would still be a lot of data to send. This process would be wasteful, because most of the second image is just a repeat of information contained in the first frame. H.264 compression provides a great improvement by significantly reducing the amount of data needed.

With H.264 compression, the first of the two frames would be designated as the primary reference frame, or "I-Frame", and it would be encoded in such a way that it could be decoded alone; that is, without any information about previous or following frames. I-Frames can serve as a starting point for video playback, because they can be decoded and displayed as a complete image. To serve as a reference frame, the image is divided into regions called "macroblocks" as shown in Figure 2, with each block representing a square grouping of 16x16 pixels. (The regions in the figure are larger than 16x16 pixels for illustration.) These groupings can be subdivided into a limited set of sub-blocks to capture image details.

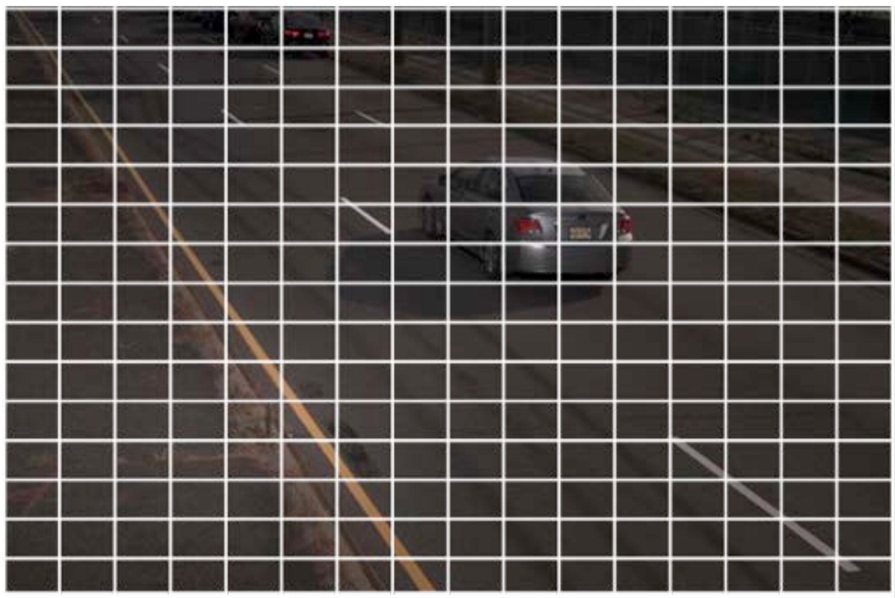


Figure 2: I-Frame divided into blocks using H.264

The second image would be designated as a "P-Frame", which indicates that it depends on prediction from a previous frame for its full description. Rather than resending the entire image, the frame is examined to see which areas changed from the previous frame. If a block did not change, then it is not resent or stored. If it did change, then it is compressed and sent. Figure 3 shows what the processed second frame would look like for the moving car image. Note that the changed blocks not only include the car, but also the part of the road that was revealed by the car's movement, including its shadow.

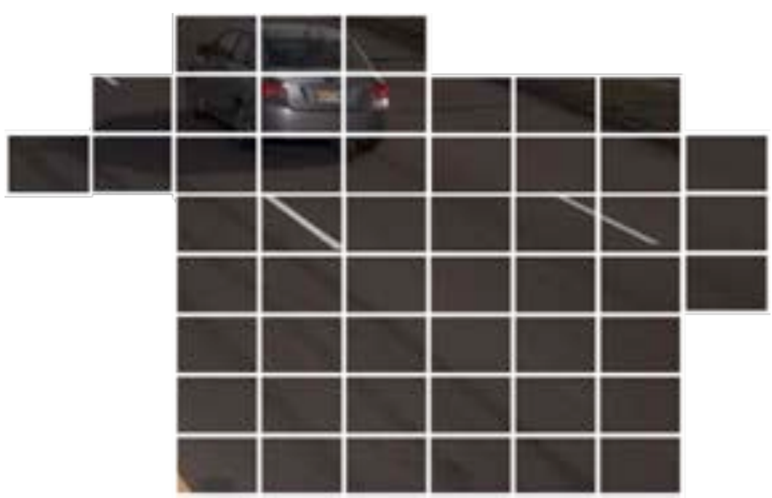


Figure 3: P-Frame showing only changed blocks from reference I-Frame

By using this scheme, it is clear that a significant data volume savings has been achieved. Rather than sending two full images, the second image has been reduced to only 20% of the original size, and basic compression can be applied to both images. The result in this simple example, assuming the basic compression could achieve a savings of just 20%, would be approximately half the data of the full two frames. If the process continued to a third frame, where again only the changing blocks would be compressed and sent, the savings would be even larger. In practice, it is usually possible to send between 15 and 30 P-frames before sending another I-Frame, and in some security applications where little movement occurs, even more.

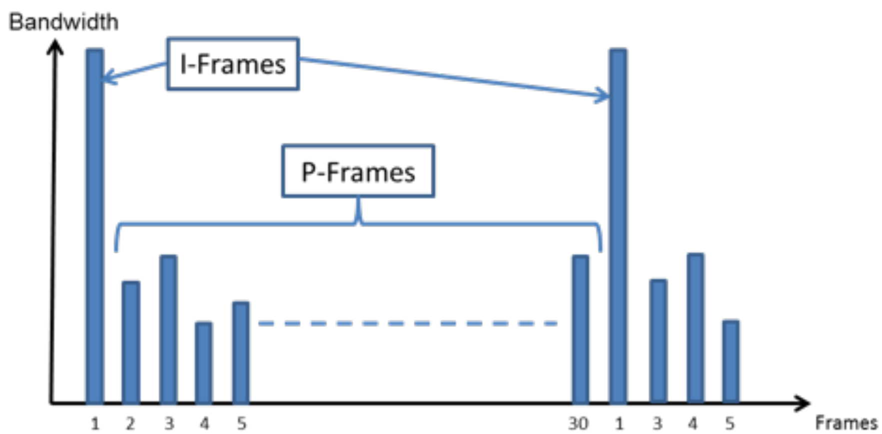


Figure 4: An I-Frame (or reference frame) is sent as frame 1, followed by 29 P-Frames.

The bandwidth required to transmit and store such encoded video is illustrated in Figure 4. In this illustration, an I-Frame (or reference frame) is sent as frame 1, followed by 29 P-Frames. In this case, an I-Frame would be sent once per second based on 30 frame per second video stream. The bandwidth required for the I-Frame is higher because of the larger amount of information sent. The bandwidth required for the P-Frames is much lower, particularly for applications where the amount of movement in the frame is small. The number of P-frames that will be sent between I-Frames can be specified with a system setting to control the overall average bandwidth for each camera.

H.265 builds and improves upon, H.264

There is no question that H.264 video compression was a leap forward in video compression technology. It enabled the widespread adoption of HD video, not only in its use in BlueRay disc systems, but also in video playback systems such as Vimeo and YouTube, web software including the Adobe Flash player, and video distribution and streaming systems including on-demand video and high-definition television, cable and satellite systems.

As good as it was, however, the truth is that the video experts that developed the standard had hoped that it could be even better. They proposed and considered a number of enhancements to the standard, but the trade-offs in terms of computational load were just too high to meet the goals of the program. The engineering challenge was this: with a little processing, you can compress the video a little bit. To compress the video a lot, requires a lot more processing and computations. Since the goal of the effort was to achieve a balance of compression while not overloading the processors available at that time, the decision was made to limit the compression and avoid the need for costly processing power. It was the right decision at the time, and the standard has been extremely successful.

Today, several years later, processing power has increased significantly at ever-decreasing costs, and it is now possible to implement the more advanced ideas economically. H.265 the natural evolution from H.264 represents another leap forward in video compression technology. With H.265, users can expect up to 50% reduction in bandwidth while maintaining the same video quality for professional surveillance applications, or with the same bandwidth, users can expect a doubling of the possible transmission data volume – enabling the use of higher-resolution video including 4K.

H.265 enhancements over H.264

The H.265 standard includes two major enhancements over H.264, plus a number of other improvements.

More efficient/flexible macroblocking

The most significant is an increase in the size and flexibility of the macroblocks. As was discussed earlier, splitting the image into blocks allows the system to send a subset of blocks for the P-Frames. In H.264, the blocks had a maximum size of 16x16 pixels, and a limited number of sub-blocking patterns. The possibilities are shown in Figure 5 below.

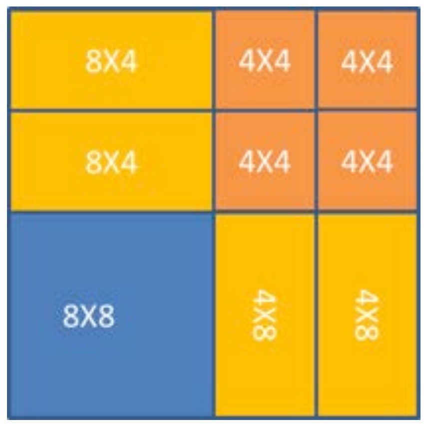


Figure 5: H.264 Block options

In the enhanced H.265 system, the macroblocks can be set to 16x16, 32x32, or up to a maximum size of 64x64, with the larger sizes usually increasing the coding efficiency. As was noted previously, the greater efficiency of larger block size options was well understood when H.264 was developed, but could not be supported by low-cost processors at that time. Now that processing power has been improved even for low-cost devices, the new standard takes advantage of this capability to improve performance. The size difference between the maximum block sizes is illustrated in Figure 6.

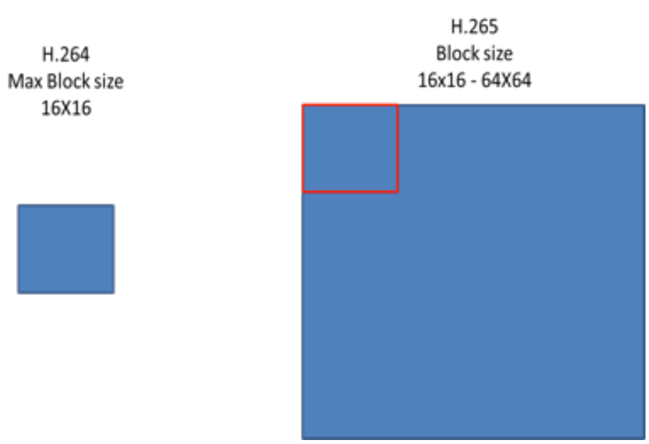


Figure 6: New range of macroblock sizes in H.265

Rather than using the fixed size blocks as in H.264, the new H.265 system implements a structure called "coding tree units", or CTUs, which defines the sub-blocks of each macroblock efficiently for processing. The new set of possibilities includes all the previous block sizes from H.264 plus a number of new shapes as shown in Figure 7.

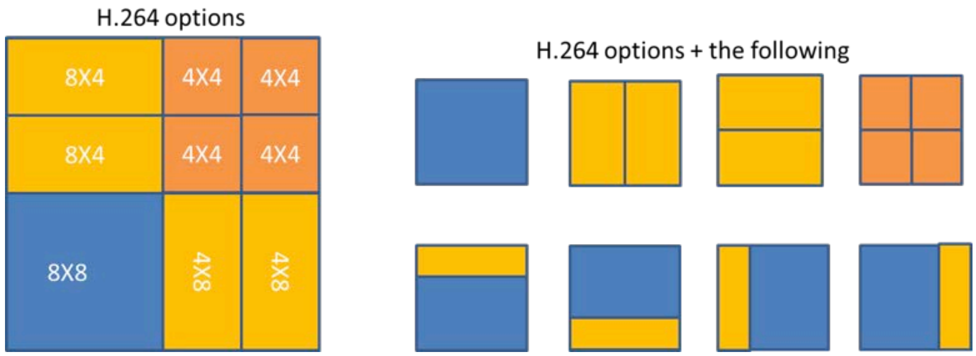


Figure 7: Enhanced block sizes and shapes in H.265

With the new sizes and shapes available, the encoding system can break an image into much more efficient large blocks while also applying smaller blocks wherever needed to efficiently capture the range of necessary detail. Figure 8 is a clear illustration of how H.265 can apply the range of block sizes and shapes to an actual security camera image. Note that the lower detail areas are captured in larger, more efficient blocks, while smaller blocks are assigned to the detail and activity areas.

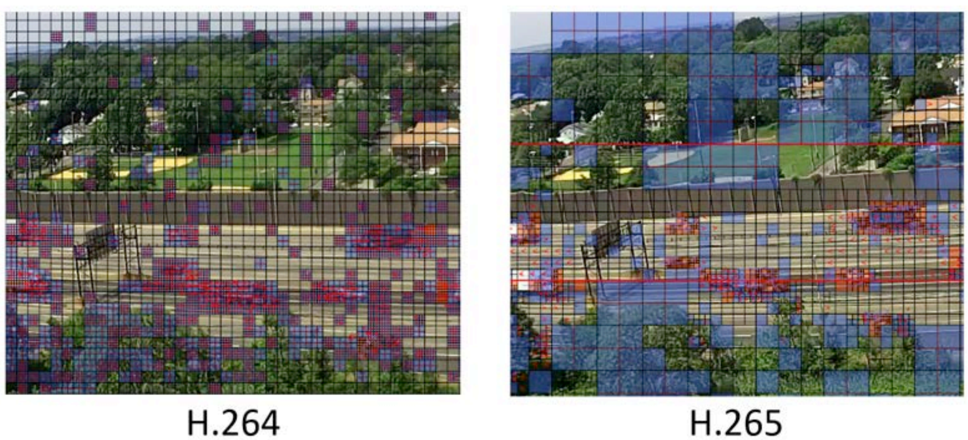


Figure 8: Comparison of H.264 and H.265 Macroblocking

As a further example of H.265 flexibility, consider the upper right block of one grouping in the security image we reviewed above, shown here in Figure 9:

Figure 9: Blocked image showing edge of vehicle

In this instance, as the vehicle continues to move, only a portion of the block marked with the red box will change. H.264 will recode the entire block in the subsequent frame, but as shown in Figure 10, the additional flexibility of H.265 lets the system select only the portion of the block that has changed – further reducing the amount of information that needs to be compressed, transmitted, and stored while continuing to capture all the important information from the image.

Figure 10: H.264 coding block with reduced size H.265 block

Improved inter-prediction using motion vectors

The other major enhancement in H.265 is in how it handles motion and moving objects. We did not cover this in the simple process description above, but it is another way that compression systems can achieve large bandwidth savings. Here's how it works: when objects move within a video, it is often the case that portions of their image look the same in subsequent frames, but have moved to a new position. If it is possible to locate blocks that look the same, but have just moved to new positions, then it is more efficient to transmit that information rather than re-transmit the image itself. This situation is illustrated in Figure 11.

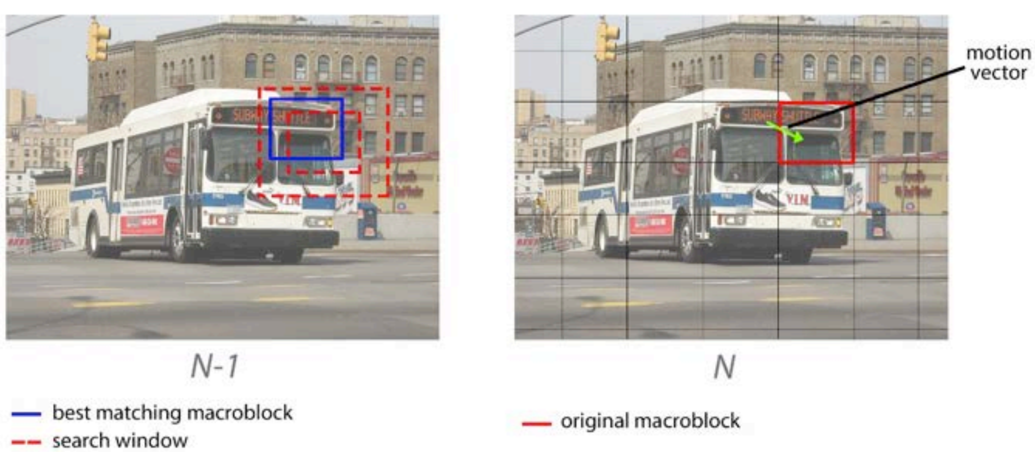


Figure 11: Use of motion vector to save bandwidth

In practice, it is rare for these matching blocks to match exactly, so in addition to transmitting the block specification and the movement to the new position, it is also necessary to transmit the description of any small changes to the block – a smaller version of the similar process of just describing the blocks that have changed from the original I-Frame. In H.264, the protocol has 9 defined directions to describe the movement of a block, as shown in Figure 12. Because the 9 directions cannot exactly match the movement in many cases, there are usually small adjustments that need to be transmitted and implemented for the translated blocks. In contrast, the new H.265 protocol uses 35 directions to describe the movement of video blocks much more accurately, which results in a corresponding decrease in the amount of information needed to describe the discrepancies or changes. This additional refinement of the process is another example of what has been made possible by the advances in computational power available for this purpose.

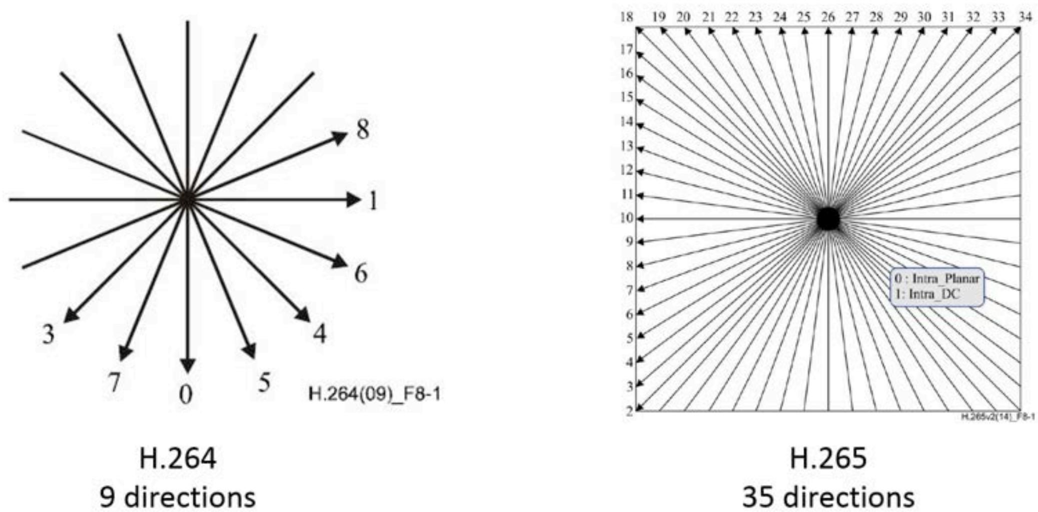


Figure 12: Greatly refined directional information used in H.265
(images from ITU-T standards documents describing H.264 and H.265)

Figure 13 below illustrates the effect of the improved directional descriptions on the amount of information that needs to be transmitted and stored for accurately capturing the images of a moving object in the frame. In the figure, the top section shows an image of a large moving truck. On the left side, one area of the image is shown zoomed in for clarity, and the directional information has been superimposed on the blocks to show how the system captured the movement of the video blocks. In this case, the closest direction to the actual detected movement was a horizontal vector. (Note that although the directional vector is pointing to the right, this does not mean that the truck is moving to the right; it only captures the reference direction, and the value in this case would be negative, indicating movement horizontally to the left. If the truck were moving to the right, the same directional vector would describe the motion, but with a positive value.)

As shown in Figure 13, the H.265 system is able to locate many more blocks that can be reused with translation and minor corrections (as indicated by blue blocks in the images) to significantly increase the compression of the video stream with no loss of image quality. A close look at the magnified images on the left side of the figure reveals the more refined directional information that improves this process in concert with the higher processing power that lets the algorithms locate the matching tiles quickly.

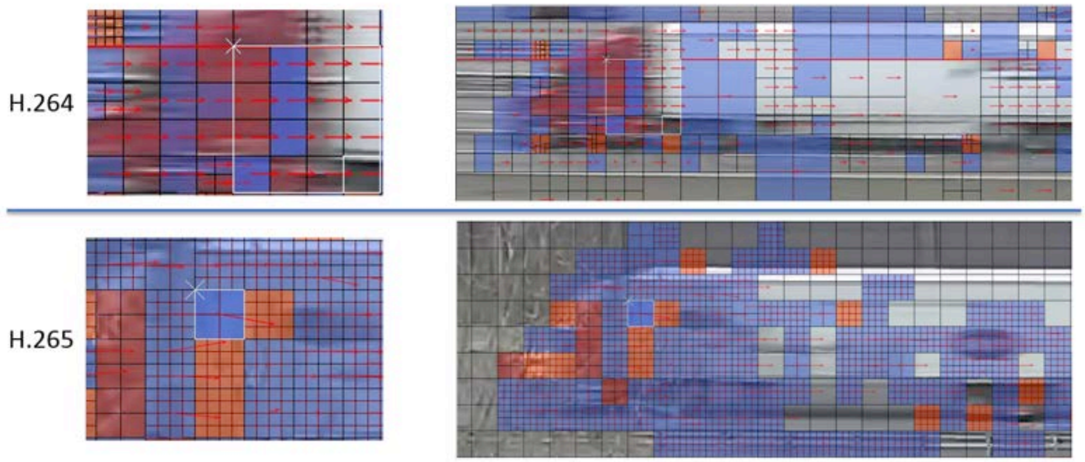


Figure 13: Improved block reuse in H.265 over H.264

Other important enhancements implemented in H.265

There are a number of other important enhancements in the new H.265 video compression protocol that improved its performance over the older H.264 protocols. Here are short descriptions of some of these key improvements:

Adaptive motion vector prediction
As described above, H.265 makes use of an increased number of motion-prediction directions for higher accuracy over H.264. In addition to a greater number of directions, the motion prediction algorithm itself has been improved to let the codec find more inter-frame redundancies, further increasing compression and quality performance.

Superior parallelization
The new H.265 protocols are designed to make better use of parallel processing – that is, the steps in the encoding and compression process can be more easily split up and handled by separate processors, rather than needing to be performed in a specific linear sequence. In this way, H.265 can make more efficient use of multi-core computing environments to further speed processing.

Improved entropy coding
"Entropy coding" is a shorthand referring to a very effective numerical compression technique that is applied in the final stages before transmission of the compressed video. H.264 used two techniques, one for the lower profiles and one called CABAC for the higher profiles, because CABAC was better but more demanding computationally. H.265 uses only CABAC for all profiles, and has implemented several adjustments to maximize throughput beyond H.264 speeds.

Improved de-blocking filters to reduce block edge artifacts
In the decoding and image reconstruction process, one of the final steps is to smooth any block edge effects and artifacts (features that were not in the original image before compression) that might be a byproduct of compression and restoration. At the same time, any such process should not reduce the resolution of actual image textures or features. H.265 incorporates improved filters that maintain higher image resolution without passing unintended block edge artifacts to deliver improved replay image quality.

Together, these H.265 enhancements work together with the new macroblocking and motion vector methodologies and the increased computational processing power of modern devices to deliver significant increases in compression while maintaining the highest video quality. The new H.265 standard has been developed and implemented to maximize compression, and thereby usability and market acceptance for the newest high-resolution formats (4k) as well as the pending/upcoming formats up to 8k.

Performance indicators

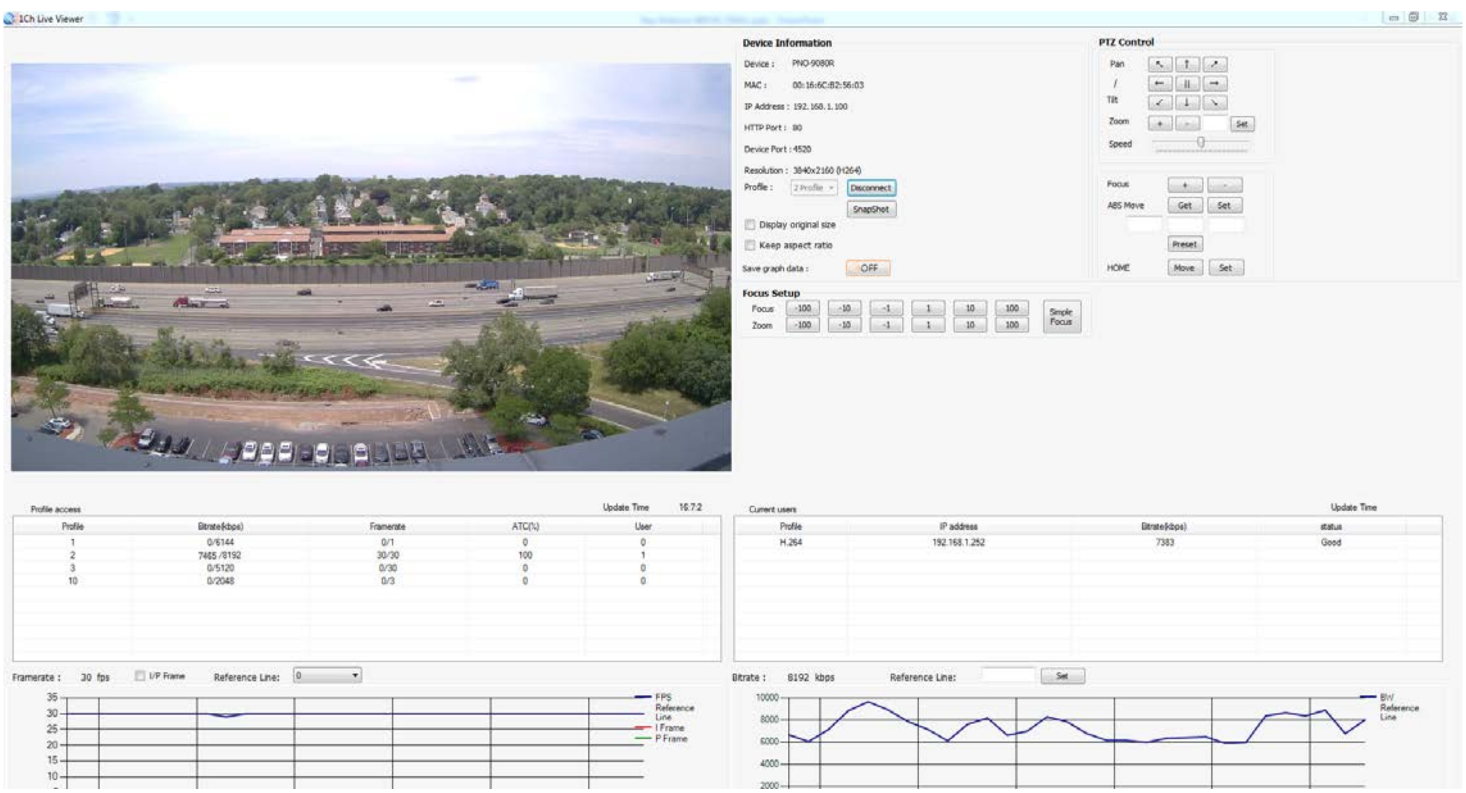
Recall that the goal of H.265 was to halve the bandwidth required to transmit and store current video material versus the use of H.264, or alternatively, to be able to transmit video streams with twice the image content over an existing data infrastructure. In either case, the new system should not cause any degradation of the video quality.

Experts agree that H.265 delivers an enormous improvement in video compression – in many ways, exceeding their targets, as shown below. To evaluate the success of meeting the goal, however, a single measurement is not sufficient. Recall that the goal included "no decrease in video quality," which is not a completely objective measurement – the fact is, the image must be interpreted by human eyes in order to assess whether quality has remained constant, increased, or decreased.

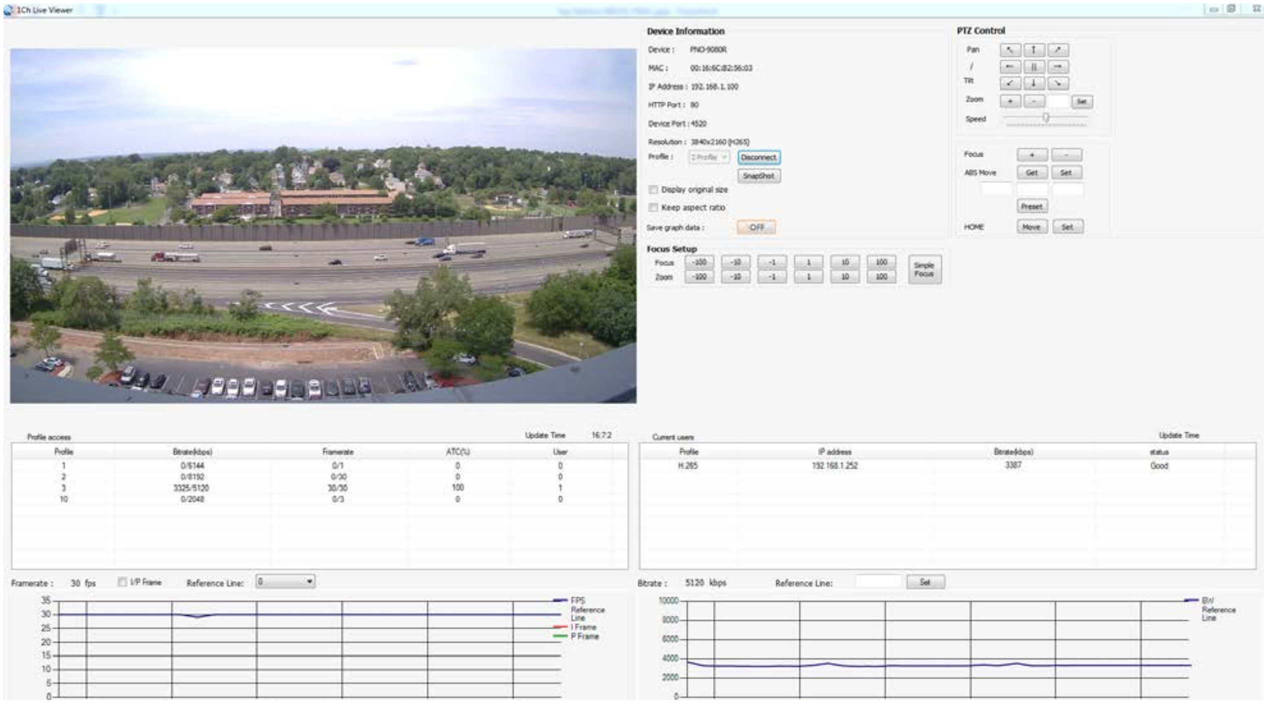
For Security applications, the most important measure is the availability of adequate resolution images to support surveillance and forensic evidentiary needs. So, to evaluate the performance of H.265 versus H.264, we can look at actual video transmission bit rates (to evaluate the compression performance) and also look at the quality of the images (to evaluate their suitability to capture security events and serve as evidence if needed).

Image 1 & 2 are example comparisons and measurements for sample security images, comparing the performance of H.264 with H.265 while using the same camera, at the same resolution and with otherwise the same settings, including 30 frames per second. For these images, camera model Hanwha PNO-9080R was used.

Example Image 1: H.264 protocol



Example Image 2: H.265 protocol



As shown in the example above, the average bitrate for H.264 was more than 7.6 Mbps, while the same video encoded using H.265 had an average bitrate closer to 3.4 Mbps. Both images display excellent image qualities in terms of color reproduction and video resolution, yet the video encoded in H.265 achieves over a 50% reduction in bandwidth usage.

Market impact of H.265 introduction

Drivers and obstacles for adoption

The most significant driver in the security industry for rapid adoption and implementation of H.265 based systems is the increased demand for higher-resolution cameras. There are at least two distinct reasons for this. First, the quality of the images from older analog and digital cameras was not sufficient to fully meet user needs. It is no longer acceptable to get a vague, fuzzy image of an intruder or license plate; users demand clear, detailed images that allow for identification beyond a reasonable doubt. Second, system designers are specifying fewer cameras that cover larger, more complex areas in order to save on infrastructure, wiring, installation, and maintenance costs. Capturing these complex scenes requires that each camera is capable of capturing and transmitting higher resolution images and higher bandwidth information.

This demand for higher resolution must be balanced to some extent by a concern over the cost and capabilities of the infrastructure needed to support these higher bandwidths. An interesting note is how this factor effectively bifurcates the market into two groups. The first group has existing installations and would like to upgrade their cameras to higher resolution; thus, this group wants to send more information through their existing wired infrastructure. The second group is installing new systems and wants to minimize the cost of the necessary infrastructure. If this set of two goals sounds familiar, it should – it perfectly matches the complementary goals and purpose of the H.265 development effort. Fortunately for both groups, advances in computer processing speed and data storage systems have continued to follow Moore's law (doubling of chip performance every 18 months), even if network/Internet bandwidth has not matched that pace. And note that implementing H.265 is not just a software update – it will require new hardware. But, the availability of processing power and storage technologies at current price points will allow suppliers to deliver this superior H.265 compression and satisfy the needs of both of these customer types.

To implement an H.265 system requires cameras with H.265 encoders and NVRs and VMS systems that support H.265. At the time of this publishing, Hanwha Techwin America is one of the only video surveillance manufacturers that has adopted H.265 codec in their next generation cameras. Additionally, their new NVR line and management software support H.265, making them one of the only security manufacturers that offer a full H.265 security solution.

Possible courses of action

It is widely believed that H.265 will replace H.264 as the standard codec for IP video security; its advantages in terms of image quality and bandwidth reduction are clear and compelling. The only question is the timing. Some manufacturers are moving and speaking cautiously – perhaps to protect their existing product lineup as long as they can. Others, however, are looking forward and moving briskly to this powerful new format, because it matches the needs of the security market so well.

In terms of technology, here are some sensible recommendations to help installers and end users that are thinking about new systems, or considering upgrades to their existing systems:

1. **Completely new systems**
   If a new business, building or campus under construction needs a completely new system, push to specify a fully H.265 system. There is no reason to install a waning technology now when the advantages of the new system are so clear, and the time to install the appropriate infrastructure (specifically, signal wiring and power taps) is during construction, not afterwards. Specifying the H.265 end point equipment (cameras and recorders) may cost more than H.264 equipment today, but the infrastructure savings should more than offset that difference for all but the very smallest systems.

2. **Upgrading an existing system**
   Next to a completely new system, a full upgrade of an existing system is the largest possible project – and if the wiring needs to be replaced, the project can easily cost more than an entirely new system described in #1 above. It is critical to plan carefully, and work with an installer/integrator experienced with upgrade projects. If the entire system must be pulled and replaced, then follow the recommendation above, knowing that the new system will be an investment with a long planned life. If the existing wired infrastructure is in good condition, then you are in luck: it is highly likely that you can upgrade the endpoints to H.265-capable devices (cameras, recorders and VMS) and gain all the benefits of the latest protocols with only a fraction of the cost of a new system.

3. **Upgrading a part of a system, or adding a new section onto an existing system**
   This is where planning, execution, and the ongoing management of the system get tricky. In general, new H.265 devices such as cameras will also be capable of using the older H.264 protocols for the first few years, if not longer, depending on market trends. For this reason, it is quite possible to continue to use existing H.264 NVRs, for example, while replacing old cameras with new H.265 models and setting them up to use the older H.264 protocol temporarily. Once budget is available, the NVRs can be upgraded and the cameras reset to operate using H.265. In a similar vein, a new extension of the system can be installed with H.265 capability (in a newly constructed wing, for example) but set to operate using the older protocol for a time to integrate into the existing system.

Summary

The H.265 protocol will help usher in the next evolution of superior video surveillance performance by implementing many of the best ideas of the expert groups that developed them, and by taking advantage of the latest processing and storage technologies. With the H.265 compression protocol, the professional security market will continue to develop and deploy more advanced high-resolution, high-performance cameras, and users will be able to send more video data over their existing infrastructure, or achieve the same information content with less infrastructure. Either way, the professional security market benefits from this important technical achievement.

References:

- Recommendation ITU-T H.264 from the International Telecommunications Union (02/2016)
- Recommendation ITU-T H.265 from the International Telecommunications Union (04/2015)
- Both are available online from the ITU-T Publications site; https://www.itu.int/rec/T-REC-H/en

SAMSUNG
WISENET

Hanwha Techwin America
100 Challenger Rd. Suite 700 Ridgefield Park, NJ 07660
Toll Free : 877.213.1222
www.HanwhaSecurity.com

Hanwha
Techwin America

I.H-1607