HIKVISION
H.265+ Encoding Technology

Halve Your Bandwidth and Storage
Enjoy the Ultra HD and Fluency

WHITE PAPER

Table of Contents

1. BACKGROUND

Several years ago, the ultra HD surveillance camera was developed. However it has not been widely applied so far because it requires wide transmission bandwidth and massive storage capacity. Thus, how to limit the bitrate of the ultra HD video stream becomes a critical problem that decides the ultra-HD camera popularity.

H.265+ technology is an intelligent algorithm developed by Hikvision. It is a remarkable encoding technology based on the H.265/HEVC (High Efficiency Video Coding) standard and is optimized while fully considering the following features of the surveillance video.

- Background information stays stable and rarely changes.

- The moving objects may appear only several times in a substantial portion of surveillance videos.

- The viewer mainly focuses on the moving objects.

- 24 hours non-stop surveillance and the noise on the video has a relatively great impact on image quality.

H.265+ is able to greatly decrease the bitrate of the surveillance video so as to reduce the bandwidth and storage cost.

2. KEY TECHNOLOGIES

H.265+ improves the compression ratio based on three key technologies: prediction encoding technology based on the background model, background noise suppression technology, and long-term bitrate control technology.

2.1. PREDICTION ENCODING

The current mainstream compression algorithms, such as MPEG2, MPEG4, H.264/AVC, and the latest H.265/HEVC, are all based on the frame of the hybrid encoding. Prediction encoding is one of the core technologies that influences the compression performance. It can be divided into inter-frame prediction encoding and intra-frame prediction.

- Inter-frame prediction creates a prediction model from one or more previously encoded video frames or fields using block-based motion compensation.

- Intra-frame prediction means that the samples of a macroblock (processing unit) are predicted by using only information of already transmitted macroblocks of the same frame.

For different frames in the video stream, different encoding methods are adopted. I-frame can be encoded separately and it adopts the intra-frame prediction encoding technology. P-frame encoding relies on the encoding of I-frame or P-frame, and it adopts the inter-

frame prediction encoding.

2.1.1. P-FRAME ENCODING

With inter-prediction, you can lower the bitrate by compressing only the difference between the current frame and the reference frame. Therefore, selecting an appropriate reference frame is the key to improving the stream compression ratio.

For a surveillance video, the background information is usually stable. Thus, we can extract a background frame as a reference frame to encode. The background frame should contain as few moving objects as possible.

As shown in Figure 1, among the three frames, T0 and T1 are the encoded images. You can take a background frame as a reference frame, and then encode T2 based on the similarities and differences between T1 (reference frame) and T0 (background frame). Since T0 contains fewer moving objects, it's a good choice as the background frame.

T0　　　　　　　　　　　T1　　　　　　　　　　　T2

Figure 1　Choose Best Background Frame

We can extract the moving object from T1 and T2, as shown in Figure 2. The object moves B to A, when encoding the T2 frame, the B area is the newly exposed area.

Figure 2　The Object Moves from B to A

Example 1:

If we take only T1 as the reference frame, it will cost more bits when encoding the B area since we cannot find an area in T1 that is similar to the B area.

Reference frame   Residual part   Final image

Figure 3   Ordinary Reference Type

Example 2:

However, if we take only the background frame (T0) as the reference frame, we can find a block that is almost the same as the B area, while the A area is a new area. Then it will cost more bits when encoding the A area.

Example 3:

Moreover, if we take the background frame and T1 as the reference frames, we can simply find the areas that are similar with the A area and B area. Then the bit cost can be reduced to the lowest value.
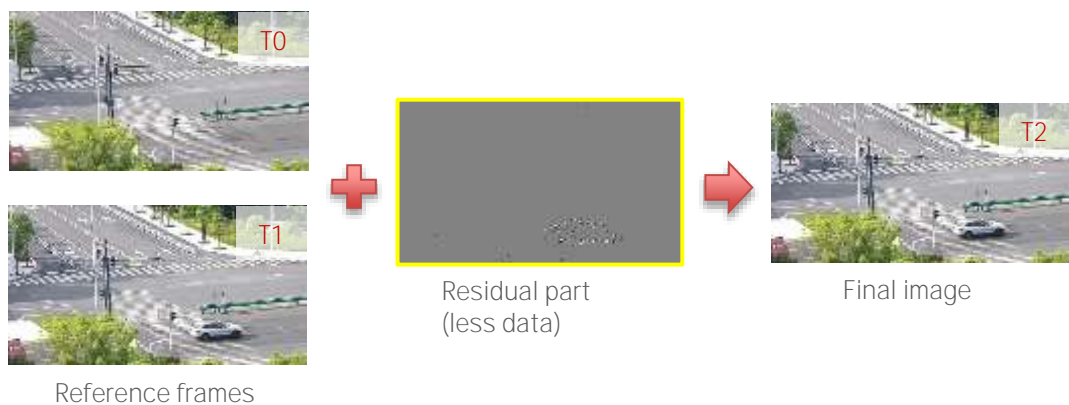

Residual part
(less data)   Final image

Reference frames

Figure 4   Background Reference Type

2.1.2. I-FRAME AND R-FRAME ENCODING

I-frame pops up every few seconds when encoding the surveillance video. As a result, the I-frame bitrate takes quite a high percentage during encoding, especially for environment that are relatively motionless. In some cases, I-frame may take up fifty percent of the encoding resources. Moreover, the information displayed by I-frame is repetitive when the background is stable.

In order to reduce the bitrate cost of the repetitive I-frame, H.265+ designs a prediction encoding reference relationship (based on background model) shown as Figure 5.
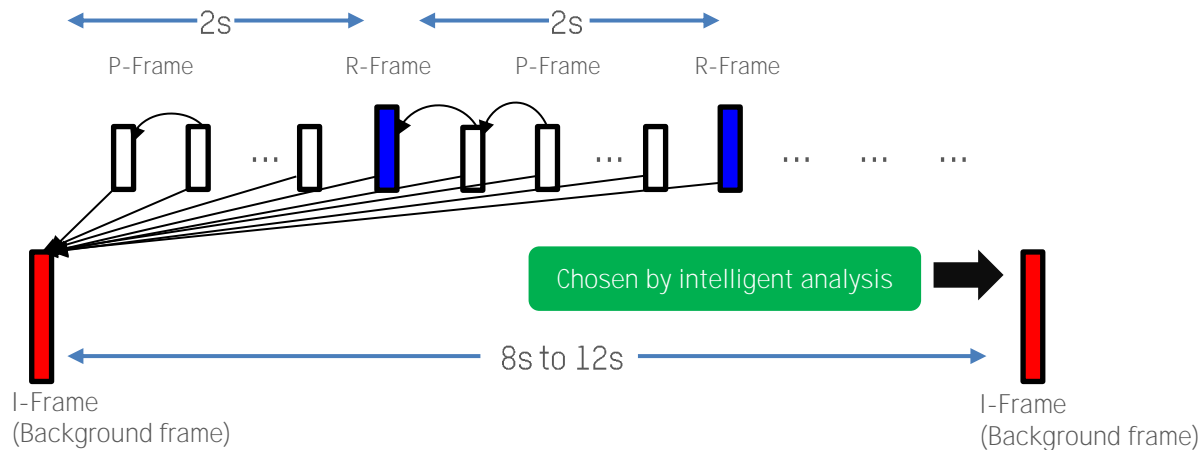
Figure 5   Prediction Encoding Reference Relationship

Frame Type	Interval	Description
I-Frame (Background Frame)	8s to 12s	I-frame is completely encoded based on the current image adopting the intra-frame prediction encoding. It's a frame that is chosen by the intelligent algorithm and contains as few moving objects as possible.
R (Refresh)-Frame	2s	The R-frame adopts intra-frame prediction encoding (for moving objects) and inter-frame prediction encoding based on I-frame (for motionless objects). The R-frame works as I-frame in the video stream during the random access to guarantee the playback user experience.
P-Frame	Same as the frame rate	The frame adopts the inter-frame prediction encoding that is based on the previous frame (P-frame or R-frame) and I-frame.

With the R-frame, bitrate cost can be lowered with the playback user experience is guaranteed as well. See the figure below for the R-frame encoding process. The moving objects, which are marked with red rectangles are encoded with the intra-frame prediction encoding and have a good quality. The background adopts the inter-frame prediction encoding.
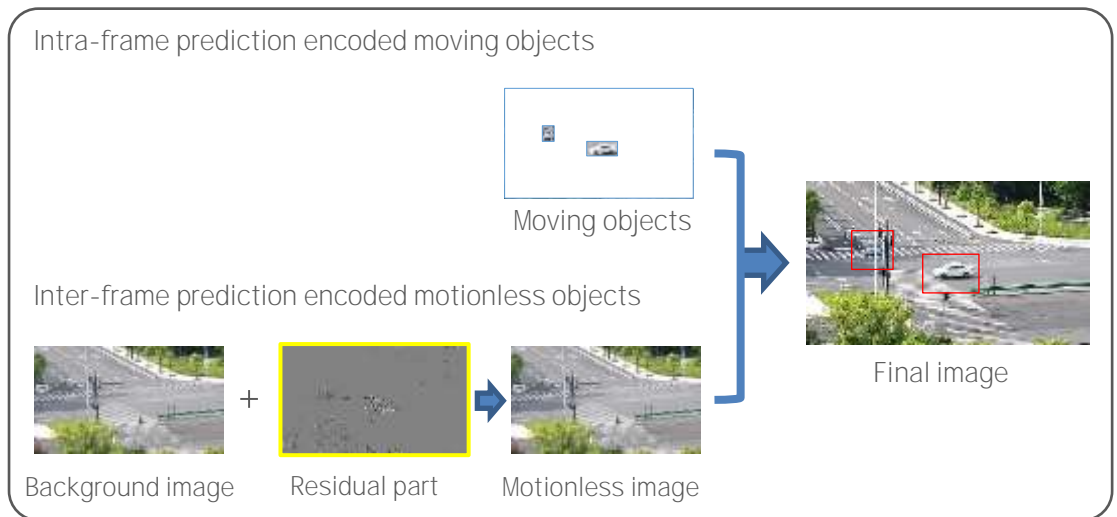
Figure 6   R-Frame's Mixed Prediction Encoding

2.2. NOISE SUPPRESSION

Normally, in order to guarantee the quality of the moving objects, the encoding module also encodes the noise in the environment. But now, with the intelligent analysis algorithm that can distinguish between the background image and the moving objects, the moving objects and the background can be encoded with different encoding strategies.

As shown in Figure 7, the intelligent analysis algorithm extracts the background image and the moving vehicle. The background image is encoded with high compression in order to suppress the noise and to lower the bitrate.
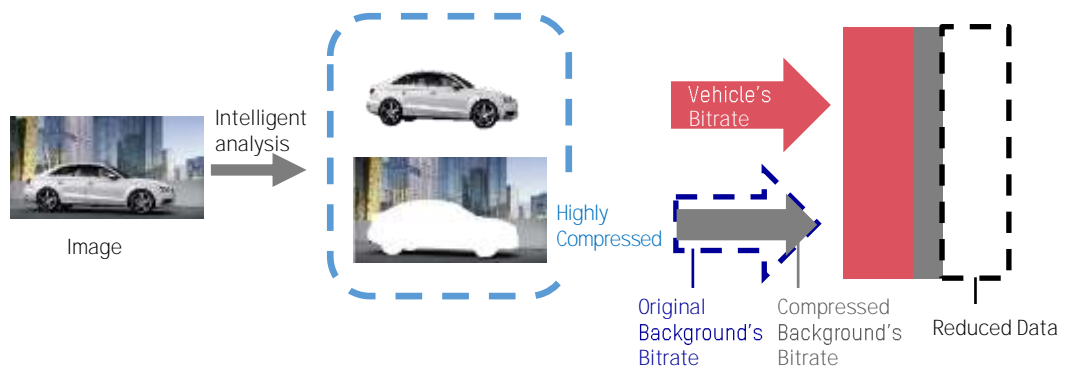

Figure 7   Noise Suppression

2.3. LONG-TERM BITRATE CONTROL

Hikvision introduces a new kind of bitrate concept named "Long-Term Average Bitrate" to make full use of the bitrate. The Long-Term Average Bitrate means the average bitrate during various time periods (usually 24 hours). With the average bitrate control, the camera can assign more bitrate to the busy hours while reduce the bitrate in the idle hours (such

as 0:00 to 9:00 and 20:00 to 24:00 in the office), as shown in the below diagram (we take the constant bitrate control as the example).
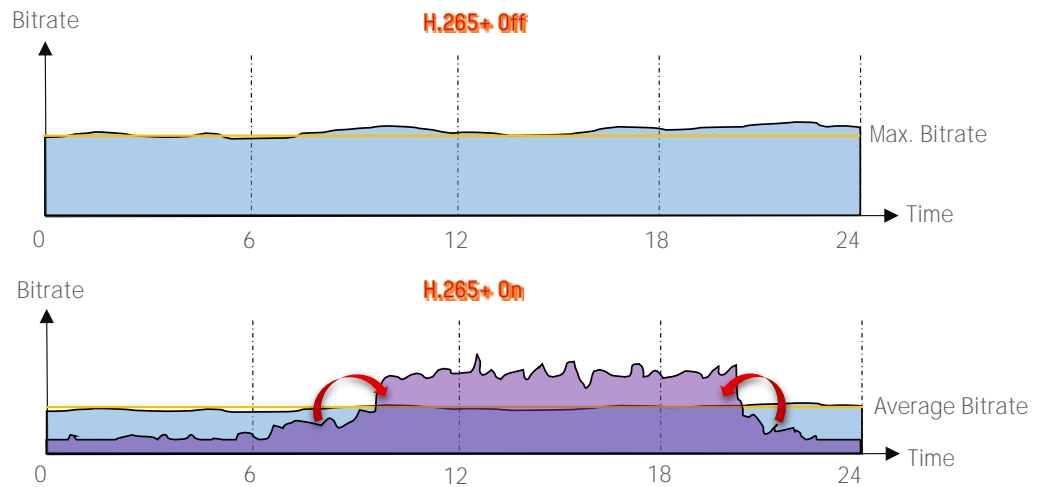

Figure 8   The Bitrate Control of the Constant Bitrate

- Constant Bitrate Control

For H.265 encoding with the bitrate configured as constant, the bitrate slightly varies near the predefined Max. bitrate value. With H.265+ on, the average bitrate can be kept as the half value of the Max. bitrate (take the office surveillance as the example, the actual bitrate reduction rate may vary with different surveillance scenes) and the image quality can be optimized since the H.265+ technology makes full use of every bit.

- Variable Bitrate Control

In the variable bitrate mode, the instant bitrate varies according to the scene busyness, while the image quality is steady. When H.265+ is on, the bitrate change can be divided into two cases.

o If the set average bitrate value is limited, the H.265+ encoding can provide a better image quality with the limited bitrate.
o If the set average bitrate value is high for the monitoring scene, the actual average bitrate can be lower than the predefined value, thus the storage can be saved.

3. BITRATE REDUCTION TEST RESULT

The bitrate reduction test is based on the 1080p@25fps cameras. The result can be divided into two parts: the instant bitrate comparison under different circumstances and the 24-hour file size resulted in the different encoding standards.

3.1. INSTANT BITRATE OF DIFFERENT SCENES

Table 1 Instant Bitrate Comparison between H.264 and Hikvision H.265+

No.	Scene Description	Bitrate (kbps)		Decreased
		H.264	H.265+	Rate
1	Cafe, Sufficient Illumination, Many Moving Objects	3,481	650	81.3%
2	Cafe, Sufficient Illumination, A Few Moving Objects	2,253	340	84.9%
3	Cafe, Low Illuminated (IR On), Motionless	930	108	88.4%
4	Street, Sufficient Illumination, Many Moving Objects	4,403	970	78.0%
5	Street, Sufficient Illumination, A Few Moving Objects	4,096	518	87.4%
6	Street, Low Illuminated, Motionless	2,662	480	82.0%
			Average Decreased Rate	83.7%

Table 2 Instant Bitrate Comparison between H.265 and Hikvision H.265+

No.	Scene Description	Bitrate (kbps)		Decreased
		H.265	H.265+	Rate
1	Cafe, Sufficient Illumination, Many Moving Objects	1,843	650	64.7%
2	Cafe, Sufficient Illumination, A Few Moving Objects	1,289	340	73.6%
3	Cafe, Low Illuminated (IR On), Motionless	453	108	76.2%
4	Street, Sufficient Illumination, Many Moving Objects	2,154	970	55.0%
5	Street, Sufficient Illumination, A Few Moving Objects	1,331	518	61.1%
6	Street, Low Illuminated, Motionless	1,946	480	75.3%
			Average Decreased Rate	67.6%

Conclusion:

1. The average decreased rate between H.264 and Hikvision H.265+ is 83.7 percent, and the rate between H.265 and Hikvision H.265+ is 67.6 percent. Hikvision H.265+ can greatly decrease the bitrate in the same scene.

2. The decreased rate falls as the number of moving objects in the scene increases.

3.2. 24-HOUR FILE SIZE OF DIFFERENT SCENES

- Scene 1: Café

Table 3 24-Hour File Size Comparison - Café

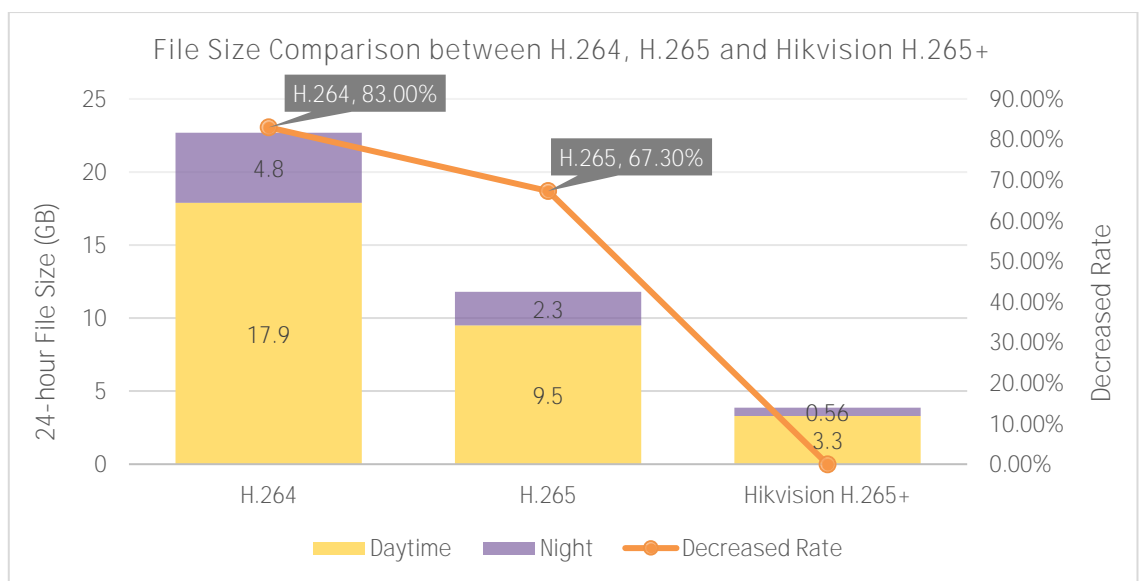| Time	Average Bitrate (kbps)		
	H.264	H.265	H.265+
09:00-21:00 (Daytime)	3,482	1,843	650
21:00-Next Day 09:00 (Night)	930	453	108
Theoretical File Size in 24 Hours (GB)	22.7	11.8	3.9

Figure 9   Comparison Chart - Café

- Scene 2: Intersection

Table 4 24-Hour File Size Comparison - Intersection

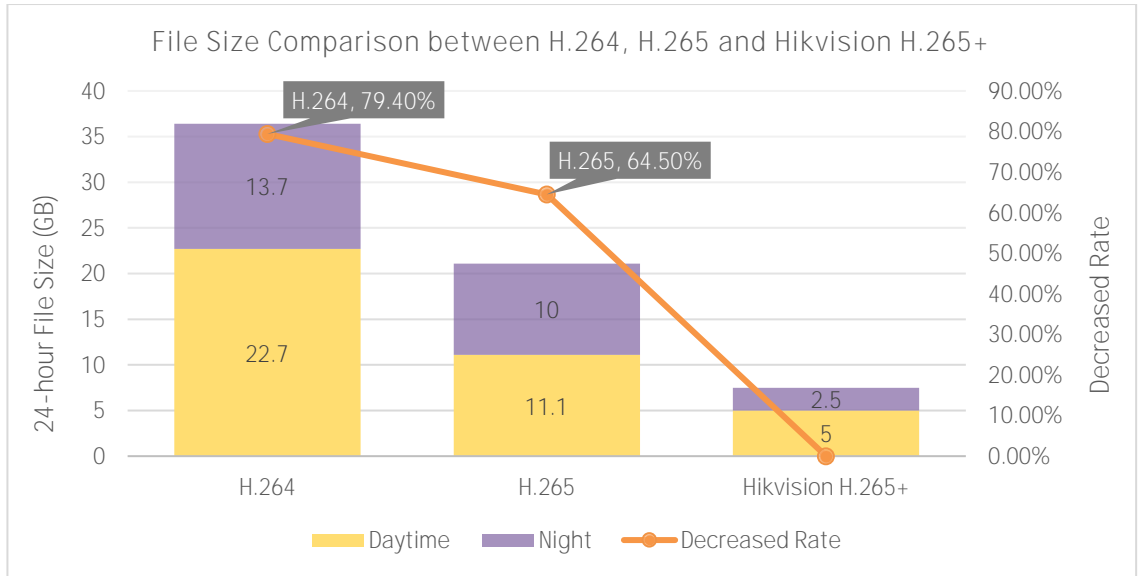| Time	Average Bitrate (kbps)		
	H.264	H.265	H.265+
09:00-21:00 (Daytime)	4,403	2,150	970
21:00-Next Day 09:00 (Night)	2,662	1,945	480
Theoretical File Size in 24 Hours (GB)	36.4	21.1	7.5

Figure 10   Comparison Chart - Intersection

Conclusion:

For the café monitoring, the decreased rate of the 24-hour file size between H.264 and Hikvision H.265+ is 83.0 percent, and the rate between H.265 and Hikvision H.265+ is 67.3 percent.

For the intersection monitoring, the decreased rate of the 24-hour file size between H.264 and Hikvision H.265+ is 79.4 percent, and the rate between H.265 and Hikvision H.265+ is 64.5 percent.

Hikvision H.265+ can greatly decrease the file size in the all-day monitoring and decrease the storage cost in the end.

4.   SUMMARY

Hikvision H.265+ is an optimized encoding technology based on the standard H.265/HEVC compression. With H.265+, the video quality is almost the same as that of H.265/HEVC but with less transmission bandwidth and storage capacity required. It extends the application of the ultra HD video in video surveillance, such as 8MP and 12MP devices.

Hikvision H.265+ meets the H.265/HEVC standard, and it is compatible with most of the software/hardware supporting H.265. It plays a significant role in cutting the storage cost and promoting the ultra-HD video popularity.

HIKVISION®