WHITE PAPER

HIKVISION

Hikvision H.264+ Encoding Technology

Encoding Improvement / Higher Transmission / Efficiency Storage Savings

Contents

1. INTRODUCTION

As the global market leader in video surveillance products, Hikvision Digital Technology Co., Ltd., continues to strive for enhancement of its products through application of the latest in technology. H.264+ Advanced Video Coding (AVC) optimizes compression beyond the current H.264 standard. Through the combination of intelligent analysis technology with predictive encoding, noise suppression, and long-term bitrate control, Hikvision is meeting the demand for higher resolution at reduced bandwidths. Our customers will benefit from the lower bitrate provided by H.264+, which plays a significant role in reducing bandwidth requirements, improving network throughput, and savings in security system costs.

2. BACKGROUND

Massive video data requires increased storage capacity, and the popularity of high-definition video with growing bitrates and resolutions demands higher capacity devices and increased system costs.

To resolve this problem, surveillance manufacturers use the latest video compression technologies, including MPEG2, MPEG4, H.264/AVC, etc., among which H.264/AVC, as the most popular compression standard, has become the main implemented compression technology.

Compared to normal videos, surveillance video has the following characteristics:

- Background information stays mostly constant and rarely changes
- Moving objects may appear only sporadically for a substantial portion of time
- The viewer focuses primarily on the moving objects in the scene
- 24-hour non-stop surveillance has an impact on storage requirements
- Video noise has a relatively big impact on quality and storage requirements

Hikvision’s enhancement of H.264 is an innovative encoding technology aimed at surveillance video. It greatly lowers the video bitrate to save implementation costs.

3. KEY TECHNOLOGIES

H.264+ improves compression performance based on three key technologies: predictive encoding based on the background model, background noise suppression, and long-period bitrate control technology.

3.1 PREDICTIVE ENCODING

Current mainstream compression algorithms such as MPEG2, MPEG4, H.264/AVC, and the latest HEVC, are all based on hybrid encoded frames. Predictive encoding is one of the core technologies of compression performance, and it can be divided into spatial-domain predictive encoding and time-domain predictive encoding.

I-frames can be encoded separately and adopt spatial-domain predictive encoding technology. P-frame encoding relies on encoded I-frames or P-frames and adopts time-domain predictive encoding.

For time-domain predictive encoding, lower bitrate is accomplished by compressing the difference between the current frame and a reference frame. Therefore, selecting the appropriate reference block is crucial.

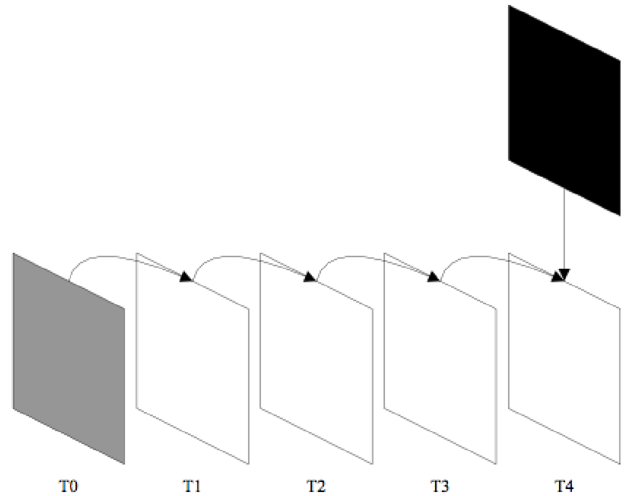


Figure 1 *Predictive Encoding*

For surveillance video, the background information is usually constant, and thus we can extract a background image as a reference frame to encode. Steadiness should be taken into consideration when choosing a background image. As shown in ***Figure 1*** *Predictive Encoding*, among the five frames, T0 to T3 are the encoded images. The black background image can be used as the reference frame and T4 can be encoded based on the similarities and differences between T3 and the black background image.

Take ***Figure 2*** *Object Moves from B to A* as an example: An object is moving from B to A, so when encoding the T4 frame, the B area is the newly exposed area.

Figure 2 *Object Moves from B to A*

If only the T3 frame is taken as a reference frame, there won't be an optimized matching block for the B area. Thus, more frames are required. If, in this case, the black background image is used as a reference frame, in most cases there will be an optimized matching block for the B area.

However, if we copy the same information in the black image, and then take T3 as the reference image, we can simply find the ideal blocks to encode T4, which guarantees image quality and reduces the bitrate. Taking the background image as the reference frame not only improves the encoding compression performance of the motionless objects, but also lowers the I-frame bitrate.

I-frames are required every few seconds (typically 1 to 4 seconds) when encoding surveillance video. As a result, the I-frame bitrate takes a large percentage of encoding resources when being encoded, especially for environments that have considerably detailed scenes that are relatively motionless; I-frames may take up to 50 percent of encoding resources. Moreover, the information displayed by I-frames is repetitive when the background is unchanged.

In order to reduce the bitrate cost of repetitive I-frames, H.264+ implements a predictive encoding reference relationship (based on the background model) shown in ***Figure 3*** *Predictive Encoding Reference Relationship*. It lowers storage requirements while optimizing the user playback experience.

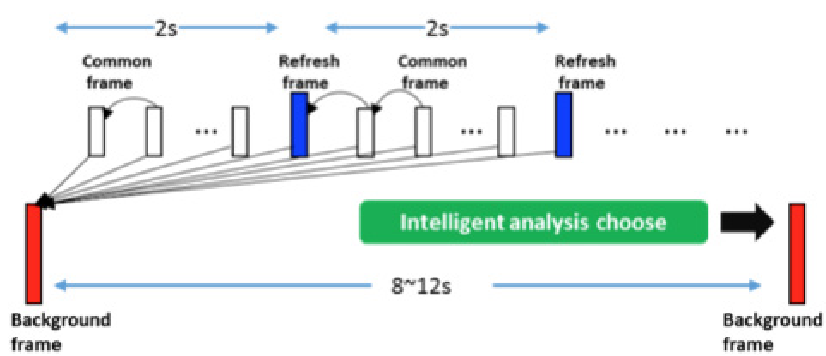


Figure 3 *Predictive Encoding Reference Relationship*

In ***Figure 3*** *Predictive Encoding Reference Relationship*, the red frame is the background frame, using intra-frame predictive encoding; the blue blocks are the refresh frames, adopting intra-frame predictive encoding (for moving objects, marked with red boxes in ***Figure 3*** *Predictive Encoding Reference Relationship*) and inter-frame predictive encoding (for motionless objects). The white blocks are the normal frames, using intra-frame predictive encoding.

The intelligent analysis algorithm selects the red block, which has fewer moving objects. The background frame data size is almost the same as that of the I-frame, and the time interval between the background frames is much longer than the interval between the I-frames. The refresh frame data size is much smaller than that of the I-frame. The time interval between two refresh frames is the same with that of the I-frame. The refresh frame can be used as the I-frame.

The background frame size is almost the same as that of the I-frame, and the time interval between the background frames is much longer than the interval between the I-frames. Also, the refresh frame data size is much smaller than that of the I-frame. The time interval between two refresh frames is the same as that of the I-frame. The refresh frame can be used as the I-frame.

3.2 NOISE SUPPRESSION

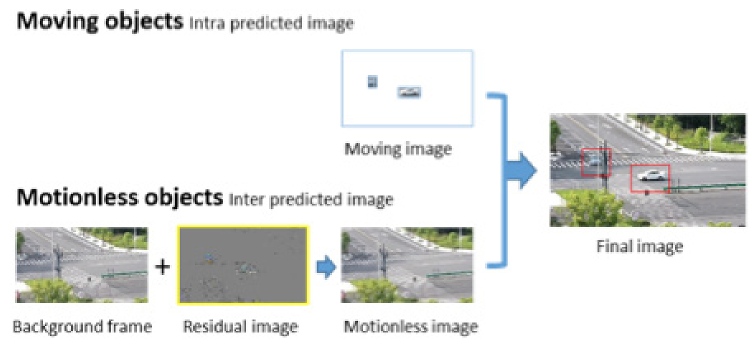


Figure 4 *Refresh Frame Predictive Encoding*

Since a typical surveillance environment's background remains relatively unchanged, the intelligent analysis algorithm can extract the background image and the motion objects (marked with red boxes in ***Figure 4*** *Refresh Frame Predictive Encoding*). Normally, in order to guarantee the quality of the moving objects, the encoder has to encode the noise in the scene. However, with the intelligent analysis, the encoder can process the moving objects and background information using different encoding strategies (e.g., on the premise that the video quality is guaranteed, the background image can be encoded with a high compression ratio, to suppress the noise to some degree and lower the bitrate). See ***Figure 5*** *Noise Suppression.*

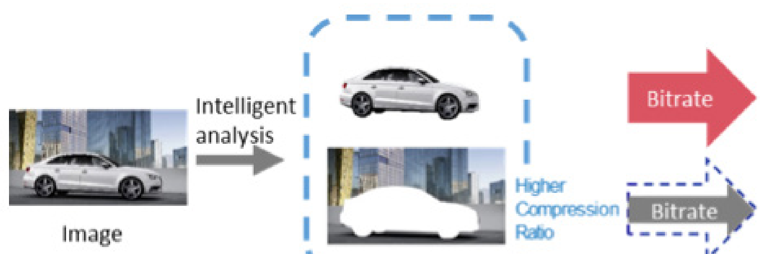


Figure 5 *Noise Suppression*

3.3 LONG-TERM BITRATE CONTROL

With noise suppression on the background image, the bitrate fluctuates according to the background area size (***Figure 6*** *Bitrate Fluctuation*). On a street, for example, the background area is relatively small because there are many people and vehicles during the day, which leads to a high bitrate. Conversely, the same background area will be relatively larger with fewer people and vehicles at night, which leads to a low bitrate. Allocating different bitrates according to different time periods not only guarantees the moving objects' image quality, but also lowers the storage requirement.

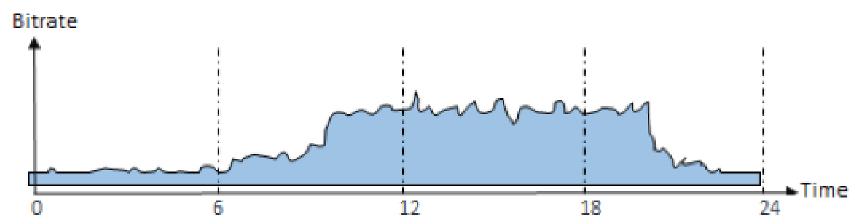


Figure 6 *Bitrate Fluctuation*

Long-term bitrate control is an effective means that can be self-adaptive to different bitrate demands of different time periods and can ensure the average bitrate reaches its target value.

To make full use of the bitrate, Hikvision introduces a bitrate concept called "Average Bitrate," which is the average bitrate during various time periods (usually 24 hours). To maintain encoding quality for moving targets and reduce storage space, H.264+ analyzes the bitrate of different periods, self-adjusts the bitrate, and allocates spare bitrate to time periods that need more. At the same time, the H.264+ long-period bitrate control technology maintains the average bitrate as the set value.

In ***Figure 7*** *Average Bitrate*, periods A and B have spare bitrates, and period needs more bitrate, H.264+ self adjusts the bitrate and allocates the spare bitrate to period C.

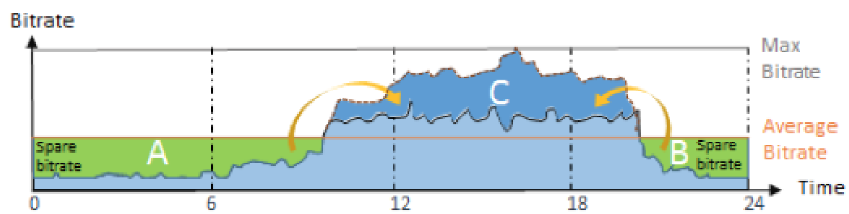
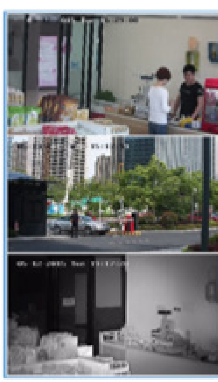


Figure 7 *Average Bitrate*

When using H.264+, the Bitrate Type must be set to be "variable" since H.264+ works under variable bitrate types. When H.264+ is enabled, average Bitrate is enabled. The default average bitrate is calculated by an intelligent algorithm based on Max Bitrate.

In most situations, the average bitrate doesn't need to be set. According to the different monitoring scenarios, the value can be adjusted based on the environment such as being decreased in environments with few moving targets and increased when there are lots of moving targets.

Figure 8 *Scenes with Bitrate Reduction* shows examples of surveillance scenes where Hikvision H.264+ can reduce storage needs. The bitrate reduction is related to background size and amount of movement in the scene. The table shows the total bitrate reduction.



Indoor		
Background	Movements	Bitrate Reduction
Small	Sparse	55%
Outdoor		
Background	Movements	Bitrate Reduction
Medium	Many	60%
Nighttime		
Background	Movements	Bitrate Reduction
Large	Few	95%

Figure 8 *Scenes with Bitrate Reduction*

4. APPLICATIONS

H.264+, when applied to a high definition or megapixel surveillance camera, provides image quality equivalent to H.264/AVC while requiring less storage space. For surveillance environments with few moving objects, using H.264+ can reduce storage space by 75 percent. For scenes that have many moving objects at certain times, H.264+ can save up to 50 percent of the storage required. For scenes with many moving objects all the time, the H.264+ storage requirement is similar to that of H.264/AVC.

Moreover, the H.264+ peak bitrate may be higher than the average bitrate in order to guarantee the encoding quality of moving objects because the larger the moving objects, the higher the bitrate cost (but no higher than that of H.264/AVC).

H.264+ is fully compliant with the H.264/AVC standard and is compatible with most software/hardware that supports H.264 without requiring additional plug-ins, etc. H.264+ offers H.264/AVC image quality with higher compression and seamless integration with existing systems. Slight adjustments to the decoding settings may be needed to optimize viewing when using H.264+ on systems that decode regular H.264/AVC.

2. CONCLUSION

Based on general encoding technology, H.264+ combines intelligent analysis technology with predictive encoding, noise suppression, and long-term bitrate control to realize a lower bitrate, which plays a significant role in cutting storage costs and provides a higher return value for the investment. ■

HIKVISION®

Hikvision USA Inc.
908 Canada Court
City of Industry, CA 91748
Email: sales.usa@hikvision.com

Hikvision Canada Inc.
4485 rue Dobrin
St. Laurent, QC H4R 2L8
Email: sales.ca@hikvision.com

Contact Information:

Tel: +1 909-895-0400
Toll Free: +1 866-200-6690 (U.S. and Canada)
Fax: +1 909-595-2788