IPC Smart Code Technology White Paper

White Paper by Dahua Technology

Release 1.0

Table of Contents

1 Foreword

In recent years, with the promotion in construction of safe cities and smart cities, the demand for video monitoring sites has increased exponentially, which also caused explosive increase in video data. Both network bandwidth and storage space are likely to become bottlenecks. At the same time, with the rapid development of image sensor, optical device and embedded image processing technology in IPC (Internet Protocol Camera), the resolution, frame rate and dynamic range of videos are increasingly high, resulting in the need for more video storage space.

Due to the bandwidth and storage cost during business expansion, users are becoming concerned about the issues of transmitting videos efficiently in limited network bandwidth and storing as many effective videos as possible in expensive storage space.

2 Existing Technologies and Countermeasures in the Industry

The solution to limited bandwidth and storage space is to accelerate the development of bandwidth and storage space as well as the update of video coding standards. However:

(1) Network bandwidth, its high cost, and network stability have always been the pain points for users.

(2) The development of hard drive storage technology lags far behind the explosive increase of video data.

(3) Video coding standards (for example: H.264/H.265/H.266, AVS1/AVS2/AVS3, and VP8/VP9) aim to compress video data to a larger extent on the premise of ensuring video quality, but they require a long development period and various patent fees.

Industry manufacturers also reduce the amount of video data from the perspectives of dynamic GOP (Group of Pictures), dynamic ROI (Region of Interest) or dynamic frame rate, but the decline and effects are not satisfactory.

- With GOP, the interval between adjacent I-frames in video streaming is increased to

reduce the amount of I-frames, thus saving bitrate. However, GOP causes poor random access of video streaming. For example, clients cannot operate flexibly in some typical applications (fast-forward or fast-reverse).

- Dynamic ROI generally improves the coding quality of ROI, resulting in the increase of overall bitrate.
- Dynamic frame rate is a technology that dynamically adjusts the frame rate of video images. It obviously does not work in scenarios with real time requirements and rigorous frame rate standards. Generally, this technology is not well recognized by clients.

3 Dahua Smart Code Technology

With the aforementioned background, Dahua Smart Code technology was launched. It aims to provide secondary improvement on the compression efficiency based on the existing video coding standards on the premise of limited bandwidth and storage space.

3.1 Framework

Dahua smart code technology mainly includes three engines: Analysis Engine, Bit-allocation Engine and Stream Compression Engine (see Figure 1).

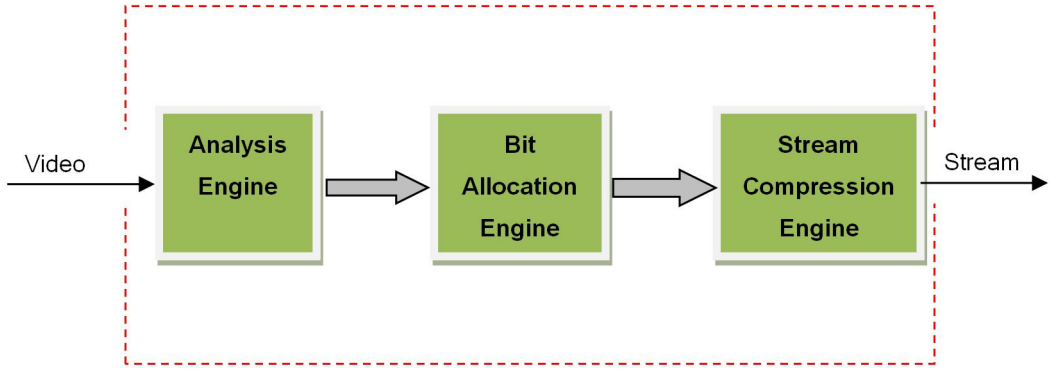


Figure 1 Flow chart of Dahua smart code technology

Analysis Engine: Mainly for complete judgment and analysis of scenarios (such as determining whether it is day or night, sunny, rainy, snowy, or foggy). The output of the user's areas of interest and the analysis of the target's motion status are used to guide the

Bit Allocation Engine;

Bit Allocation Engine: Bit is a representation unit of data quantity. The more bits allocated and the greater amount of encoded data, and the higher the image quality will be. The bit allocation engine accurately allocates bits to the best position in the image according to the results of the Analysis Engine. For example, the Analysis Engine finds a face area in a scene that is of great interest to users, and an area that they are not interested in. The Bit Allocation Engine will allocate more bits to the face area and less bits to other less important areas. Because the less important area is much larger than the area of interests in common security scenario, this engine ensures the quality of the most important areas in a scene, and saves overall bitrate as well.

Stream Compression Engine: According to the bit allocation results, the compression engine with the best configuration of the existing encoder is adopted for encoding and compression.

The Analysis Engine outputs the analysis results in real time and guides the Bit Allocation Engine. This technology then adopts the Stream Compression Engine with the best configuration for encoding and compression, and finally outputs the high-quality compressed stream as required by users. The three engines work in a collaborative way. The highly efficient and high-quality standard compressed stream is outputted by reasonably and efficiently removing the two-dimensional, three-dimensional, and human eyes perception redundancies.

3.2 The Principle of Removing Two-dimensional and Three-dimensional Redundancy

The surveillance video is composed of continuous frames, causing huge two-dimensional redundancy between image frames (there are large areas of similar or identical areas between two consecutive frames, resulting in useless data — redundant data in time sequence). In addition, due to the distinctiveness of a security monitoring scene, there are large continuous

areas (such as sky, roads) most of the time. The data in these areas are almost the same or similar, resulting in more three-dimensional redundancy in the same image.

During the development of video coding standards, block-based hybrid coding technology continuously increases the depth of image block and the direction of prediction, in order to improve the compression efficiency of images with high resolution. On this basis, the Dahua Smart Code Technology analyzes more thoroughly and reasonably the correlation between image frames and the correlation within the images through the analysis engine. The Bit Allocation Engine allocates bits to the best position through a robust bit rate allocation mechanism, reasonably removing a larger two-dimensional and three-dimensional redundancy to achieve win-win quality and bitrate.

3.3 The Principle of Removing Redundancy in Human Eye Perception

Due to the visual masking effect of human eyes, users pay more attention to key areas and ignore the image quality of non-key areas. Appropriate and reasonable reduction in the image quality of non-key areas is almost imperceptible to human eyes. The technology of removing redundancy in human eye perception ensures the quality of targets or areas of focus by reducing the quality of the non-focus area to a certain extent (see Figure 2 and Figure 3).

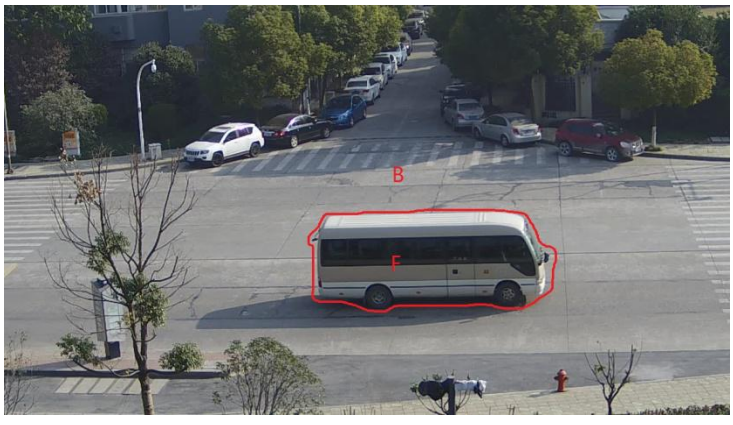


Figure 2 The Analysis Engine outputs the target area of focus.

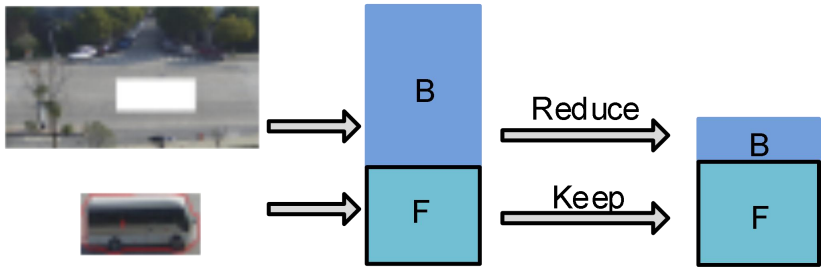


Figure 3 The Bit Allocation Engine reallocates bits.

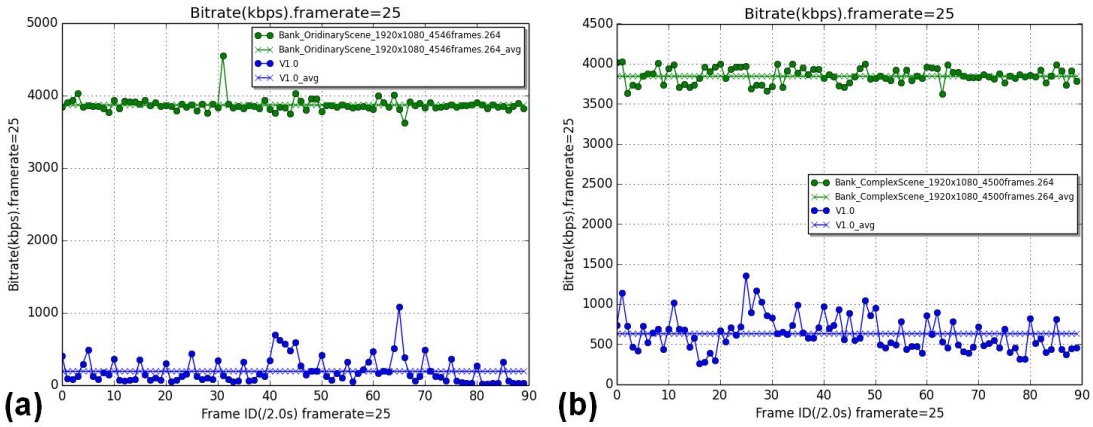
As shown in Figure 2, the Analysis Engine outputs the result of the areas of interest: The area surrounded by the red frame is the area of focus, marked as F; the rest of the area is the background, marked as B. Since users usually pay more attention to the image quality of Area F, the encoding (shown in Figure 3) ensures the image quality of this area (the number of bits consumed in encoding is unchanged), and appropriately reduce the quality of Area B (the number of bits consumed in encoding in this area is appropriately reduced), thus reducing the bitrate without affecting the viewing effect for users.

3.4 Advantages and Effects

Dahua Smart Code technology integrates two-dimensional and three-dimensional compression effect of the original video coding standards, and adopts self-developed algorithms to further reduce the amount of video data. Because this technical solution does not rely on hardware, users only need to upgrade firmware as long as the device supports the technology. In addition, as a secondary development solution based on the existing coding standard protocols, this technology has no requirements on back-end devices, posing no problem for decoder adaptability.

Figure 4, divided into a and b, takes H.264 as an example (H.265 works better). It compares the effects of one scenario after Dahua smart code is enabled and another scenario without any smart code solution. The green curve represents the stream curve without enabling smart code, with the average stream of about 3,800 Kbps. The blue curve represents the amount of the stream after smart code is enabled under the same condition, with the

average stream of about 250 Kbps and 700 Kbps respectively, which are far lower than the data without enabling smart code. Therefore, customers can use smaller bandwidth and lower costs to realize single-channel or multi-channel IPC access. A storage device with lower capacity can also be used to store monitoring videos of the same duration, or to prolong the time for storing videos and the service life of the storage device.



(a) Common scenarios (b) Complex scenarios

Figure 4 Comparison of data volume when Dahua smart code is on (blue curve)/off (green curve)

Compared with the traditional Smart Code technology in the industry, Dahua Smart Code technology also has obvious advantages. Figure 5 takes H.264 as an example (H.265 works better). In an ultra-complex business scenario, the green curve represents the effect after the traditional smart code is enabled, with the average stream of about 2,000 Kbps. The blue curve represents the effect after Dahua Smart Code is enabled, with the average stream of about 1,400 Kbps.

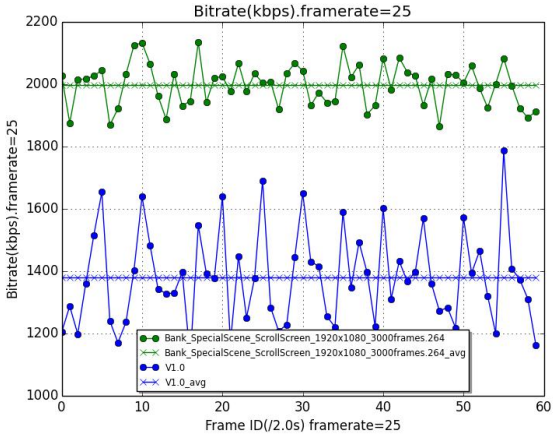


Figure 5 Comparison of data volume of Dahua smart code (blue curve) and traditional smart code

(green curve)

Dahua Smart Code technology can also guarantee clear images while greatly reducing the stream of video transmission. Comparing Figure 6 and Figure 7, in the same scenario, the target snapshots after smart code is enabled can have the same quality as the normal mode, and the overall picture quality is basically unchanged.

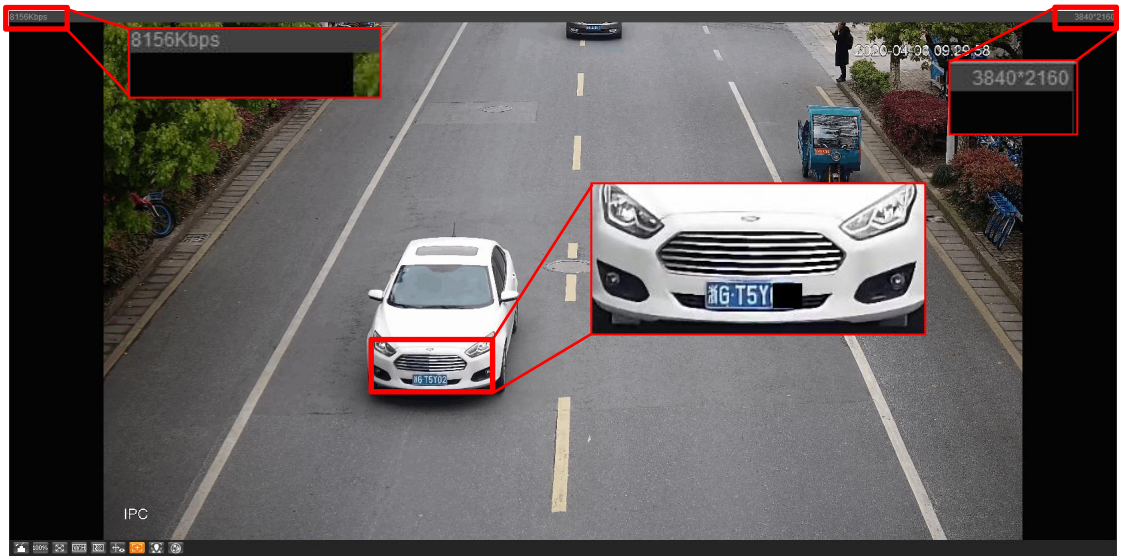


Figure 6 8M stream, smart code is disabled in normal mode

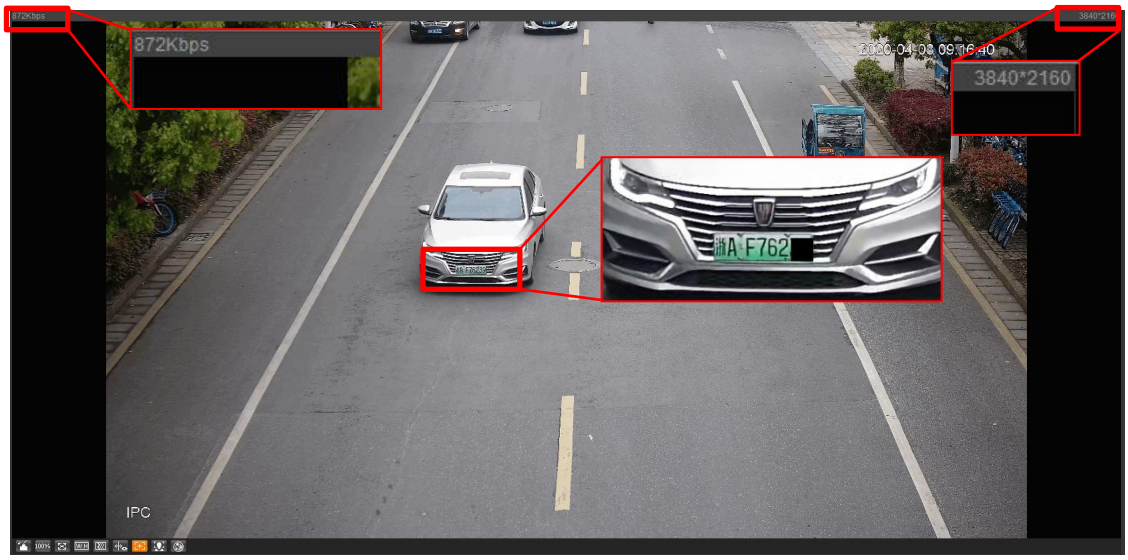


Figure 7 The license plate snapshot is still clear after the smart code is enabled and the stream is reduced to about 1M.

4 Summary

Dahua Smart Code technology is an independently-developed solution for reducing video stream based on existing video coding standards and particular scenarios of security surveillance. The technology can effectively reduce the amount of video transmission data in security scenarios, occupying less network bandwidth and less storage space. In this way, it significantly resolves the conflict between explosive growth of video monitoring points, video data volume and slow-developing network and storage space, facilitating the construction of safe and smart cities.