



WHITE PAPER

AI Smart Coding Technology

i-PRO Video surveillance systems



Table of contents

1. Introduction	3
2. i-PRO's network camera stream	3
3. Smart Coding Technology	4
3.1. ROI	5
3.1.1 Smart VIQS	5
3.1.2 Smart P-picture control	5
3.2. GOP Control	6
3.2.1 Variable GOP Low/Mid	6
3.2.2 Frame rate Control	7
3.2.3 Advanced	7
4. Conclusion	9

1. Introduction

Advancements in video surveillance techniques since the launch of IP cameras have seen many technological advancements where better lenses, sensors and image processors deliver higher resolution and improved image quality. These advancements are capable of bringing much needed improvement in quality and clarity of evidential images obtained from the video surveillance system.

Delivering these advancements require redesigning of bandwidth realm. The higher the resolution become, such as 4K, the greater the impact on the design and cost for both the network streaming and the recording hard drive. Thus, system designers must consider not only the angle of view or depth of field captured by the camera but also consider the purpose of the camera and what is useful and what is not for the particular user. For example, recording background image, such as walls or roads without important objects by high resolution only leaves useless image data consuming large bandwidth as well as storage capacity.

i-PRO is now introducing a viable solution named Smart Coding, that allows users to capture the most important object while taking out the less important objects from the target viewing area using the available bandwidth efficiently. This solution simply reduces the total cost of surveillance system ownership and makes it possible for potential owners to investment in high-end items.

2. i-PRO's network camera stream

The video stream is comprised of I-picture which is encoded solely from its own picture data and P-picture data which is an encoded data of difference information from the previous picture. If refresh interval (interval between I-pictures) is set to 1 second at a frame rate of 30 fps, stream data is generated in which one I-picture is generated followed by 29 P-pictures. From this I-picture to the picture one before the next I-picture is called Group of Picture (GOP). To play the stream data generated in this pattern, decoding must be done from the first I-picture. This first I-picture is also called a key frame. More decoding will be required until the last picture of the GOP can be played if the refresh interval is set longer.

The data size of I-picture is relatively large as it is encoded from the entire picture. On the other hand, the data size of P-picture is smaller as it is encoded from the difference of current picture and previous picture.

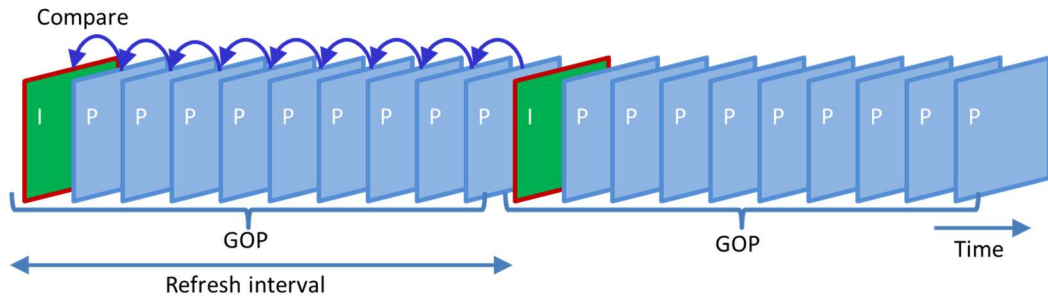


Figure 1 : The image of video stream

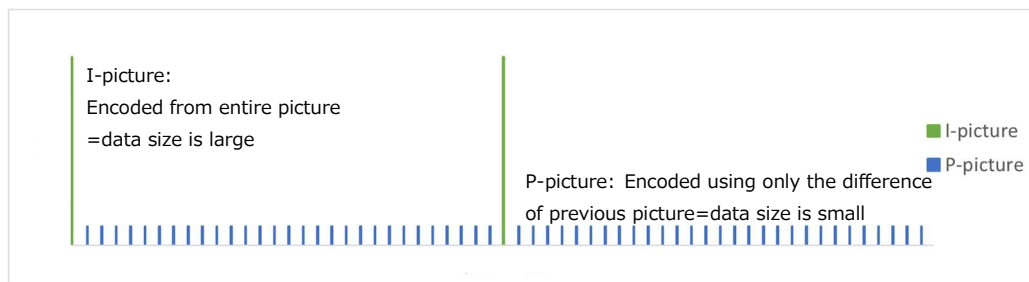


Figure 2 : The data size of I-picture and P-picture

3. Smart Coding Technology

i-PRO's camera is equipped with AI (Artificial Intelligence) processor which detects people, cars and motorcycles and automatically adjust the image property . Smart Coding is a license-free feature using this AI capability.

Using Smart Coding Technology, sufficient data capacity is assigned to important objects such as people or cars and those data are streamed and recorded in high quality while background objects such as walls and roads are assigned much less data or none reducing the total bit rate for streaming and recording.

Smart Coding with two main technologies that utilize bit rate effectively are indicated below.

1. ROI (regions of interest)

Technology that controls the amount of data assigned to object detection areas and object non-detection areas.

2. GOP Control

Technology that changes the GOP length or frame rate depending on with or without object detection areas.

3.1. ROI

ROI reduces bit rate via AI object detection assigning different amount of data to areas identified as important and non-important. There are two types of ROI, “Smart VIQS” and “Smart P Picture Control”.

3.1.1 Smart VIQS

Smart VIQS with AI object detection can reduce data size by compressing data of non-detection area.

The yellow area and red area in *Figure 3* indicate low compression blocks and high compression blocks respectively.

Smart VIQS has two levels, “Low” and “High” where “High” is higher compression rate for the non-detection area.



Figure 3: Left: Original image. Middle: Smart VIQS Low. Right: Smart VIQS High.

3.1.2 Smart P-picture control

Smart P-picture Control with AI object detection can reduce data size of p-picture by skip-blocks control, meaning to skip encoding of non-detection area. *Figure 4* shows appearance tendency of skip blocks in green. The center of *Figure 4* indicates that areas where leaves and people exist are both encoded when Smart P-picture Control is turned off. On the right side, it indicates only the area of people is encoded when Smart P-picture Control is turned on. Image quality of object detected area which is important for the surveillance is maintained even if there happen to be moving objects in non-detection area. *Figure 5* shows difference of data size in a GOP in the cases when Smart P-picture Control is enabled and disabled.

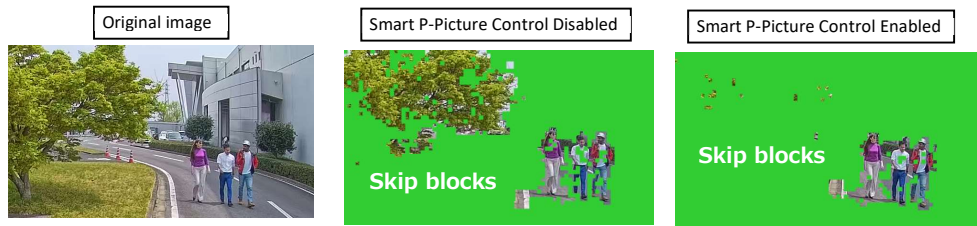


Figure 4: Left: Original image. Middle: Smart P Picture Control Disabled. Right: Smart P Picture Control Enabled.

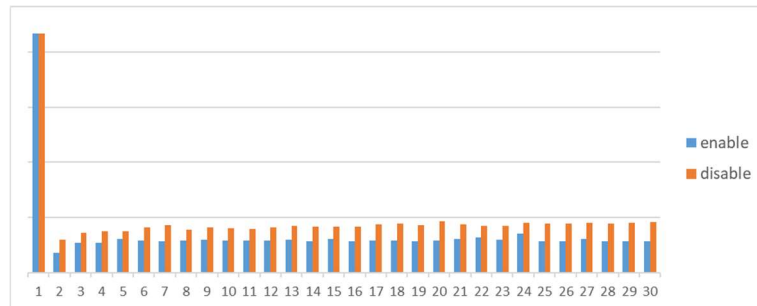


Figure 5: the data size difference between with and without Smart P-picture Control.

3.2. GOP Control

GOP control shortens the GOP length in scenes where object detection areas exist to secure random accessibility. On the contrary, it extends GOP length in scenes where object detection area does not exist to reduce the bit rate. There are three types of GOP control, “Variable GOP Low/Mid”, “Frame rate control” and “Advanced Mode”.

3.2.1 Variable GOP Low/Mid

Variable GOP Low/Mid with AI object detection makes refresh interval shorter when there are object detection areas or longer otherwise. By making the refresh interval longer in the scene without object detection areas, the number of I-pictures with large data size can be reduced, consequently reducing the total bit rate considerably. In the scene with object detection areas, refresh interval is made shorter to ensure random accessibility to scenes the user wants to view. Refresh interval can be selected from 1-8 seconds for Low or 4-16 seconds for Mid.



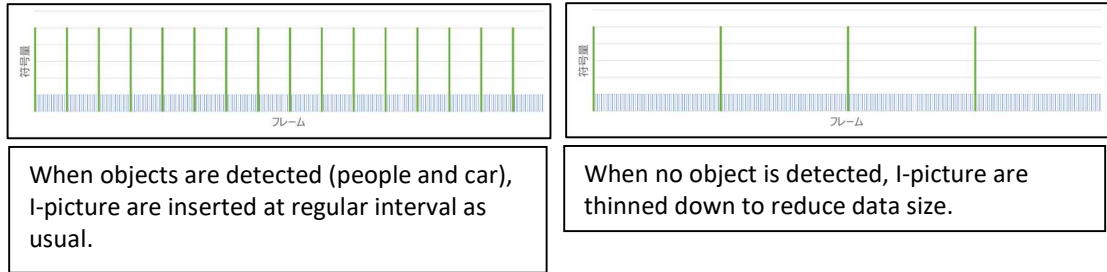


Figure 6 : The difference of I-picture insertion interval

3.2.2 Frame rate Control

Along with Variable GOP Mid feature, frame rate Control reduces frame rate down to 1fps when there are no object detection areas. As the frame rate changes, refresh interval also changes. Figure 7 shows an example of how Frame rate Control works when frame rate setting is 30fps.

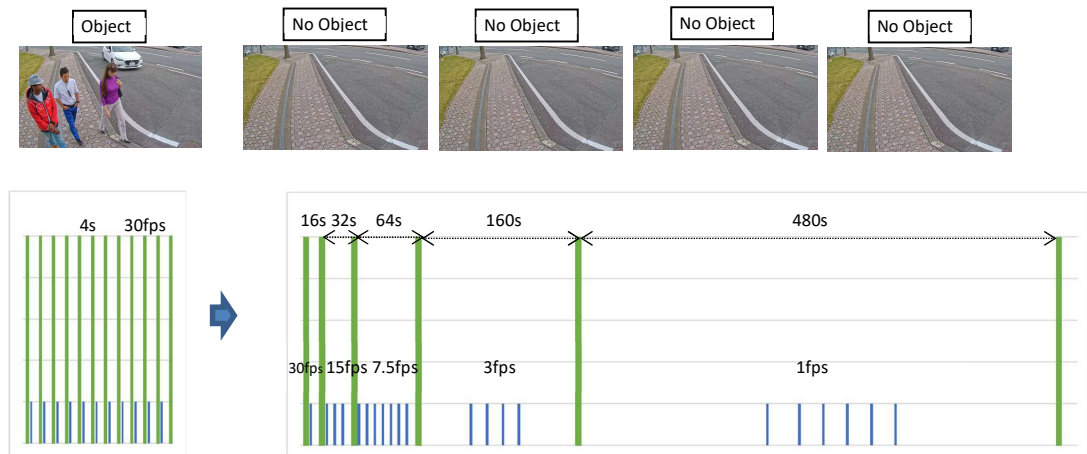


Figure 7: Transition of frame rate and refresh interval

3.2.3 Advanced

i-PRO H.265 cameras have Advanced mode in addition to the conventional GOP control described in the previous page.

The reference frame structure of Advanced mode is shown in *Figure 88*.

Advanced mode utilizes the concept of long term prediction frames. The reference frame structure is consisted of I-pictures, P-pictures that reference two frames (I-picture and previous P-picture), and key frames (P#30, #60,) that only reference I-picture.

Figure 88 shows the reference relationship at 30 fps with 60 seconds I-picture interval and 1 second key frame interval.

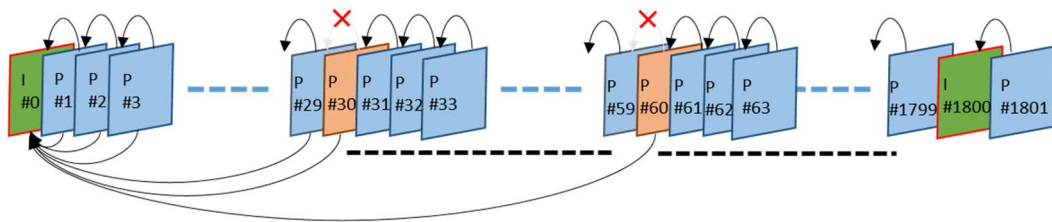


Figure 8 : reference frame structure of Advanced mode

Decoding starts sequentially from the first I-picture in the GOP if there are no key frames in the GOP. In fact, key frames can be decoded using I-pictures only. For example, if key frames exist in the GOP, #61 can be displayed by decoding only IDR#0, P#60, and P#61.

Additionally, random accessibility can be secured by inserting key frames.

The Advanced mode improves prediction efficiency by using two kinds of reference P-pictures and reduces bit rate by expanding the I-picture interval whilst maintaining random accessibility.

4. Conclusion

i-PRO's Smart Coding technology and AI together contribute to suppress cost of video surveillance system. They efficiently allocate data capacity by determining the priority level of target scenes and objects. Thus the required bandwidth is successfully reduced.



Mari Yasuda

SoC Module Development