Intelligent streaming

Bitrate-optimized high-quality video – White Paper

Table of contents

1 Introduction

The main drivers of a video security system are high quality video images and reliable recording. Over the years, resolution and framerates of security cameras have increased, now providing high quality images with more details than ever before. At the same time, the amount of cameras that are recorded, and the retention time of their recorded images has grown rapidly. The storage capacity required to store all this video data is a major portion of the total costs of a video system.

Technologies to limit the bitrate of the cameras while maintaining the necessary video quality can help reduce storage costs. Intelligent Dynamic Noise Reduction, introduced several years back, was a great step forward in reducing bitrates. Now, a new technology takes this to the next level.

Intelligent streaming is a well-tailored solution that helps achieve the best possible trade-off between image quality and video bitrate.

The main concept behind **Intelligent streaming** is to reduce redundant and irrelevant information in video data. Therefore, a well-defined combination of technologies is merged into our IP cameras, ranging from customized noise reduction algorithms to content-dependent bitrate regulation.

Enhanced encoding features such as B-frames are used to further optimize the video encoding according to the human visual system. Irrelevant image regions can be selected and encoded with lower quality. A long-term bitrate regulation has been implemented to bring precise control over the amount of data the cameras produce.

These are just some examples. This white paper provides detailed insights in the technologies behind **Intelligent streaming** and explains how it can be configured in an optimal way.

Chapter 2 introduces the basic structures and processes behind image recording and processing. Chapter 3 deals with the main concepts and components of our **Intelligent streaming** solution.
Chapter 4 shows the ease of configuration, and Chapter 5 then introduces the modes of application and shows the benefits of **Intelligent streaming** in real examples.

2 Basic structures and processes

In order to fully understand the technical concepts behind **Intelligent streaming**, some of the fundamental principles of image recording need to be understood.

The following sections thus give insight to where image noise comes from and how image noise differs from real image content. Further, image content is described that is typical for video surveillance scenes, and a short introduction to human visual perception is given. Based on these findings, an introduction to the topic of video encoding follows.

2.1 Image capture process

Digital images generally contain image noise from different sources. The most important and dominant noise types in digital images are the so-called photon shot noise and read noise.

Shot noise is generated during the capturing process of an image and is caused by statistical fluctuations in the number of photons that reach the image sensor. A perfect sensor counts the number of incoming photons per pixel, so even a perfect sensor would include noise in the captured image. Over the past decades, image sensors have begun approaching the behavior of the ideal sensor.

The other main source of noise is the so-called read noise, which consists of the amplifier noise, fix pattern noise, reset noise, and the analogue-to-digital converter noise. The noise level of the sensor increases by the applied gain, often referred to as ISO value. [1]

2.2 Image content

Natural scenes typically contain highly structured image regions such as trees as well as smooth regions such as the sky. Those natural scenes typically have an amplitude that follows a 1/f amplitude distribution, which means that the main energy can be found in the low-frequency components of an image. The main for reason for this result is that neighboring pixels are quite similar. This is an example for a high correlation.

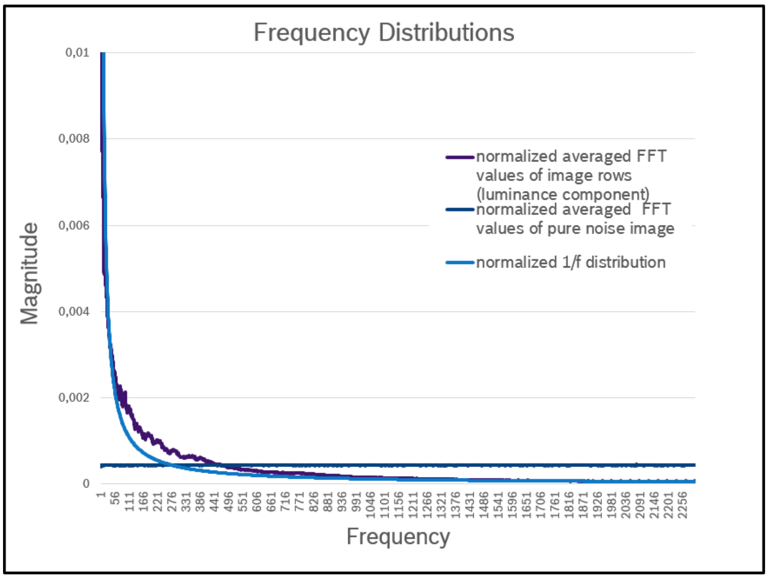

Figure 1: Frequency amplitude distribution

The correlation of the pixels is a basic requirement that allows strong compression of image content.
Image noise (white noise) breaks up this correlation of pixel values as it is distributed randomly across the image. As a result, the more noise an image contains the harder it is to compress. [2]
We thus require images with as little noise as possible.

2.3 Scene content

The content that is recorded with our cameras might vary widely. However, there are some scenes that are characteristic for surveillance cameras. These so-called "*security scenes*" show typical conditions:

- Mostly 24 hours, day/night, during nighttime the sensor creates a high noise level.
- Very often scenes are static for most of the time, e.g. company floor, parking lot, etc.
- There is special image content of interest, e.g. faces, people, cars, etc.
- Outdoor scenes sometimes include irrelevant information, like moving trees, rain, snowfall, etc.

Figure 2: A typical security scene, matching most of the mentioned conditions

2.4 Human vision

The human visual system is based on a complex circuitry of biological and psychological mechanisms, which is still not fully understood. A basic understanding of visual perception however is useful when it comes to topics like image compression.

Contrast sensitivity measures how well humans can distinguish between different levels of luminance in a static image. Research shows, that this ability is a function of the spatial image frequency.[3] In regions with higher frequency content, humans tend to perceive less noise than in smooth image regions. This effect also leads to encoding errors being more visible in areas of low contrast than in areas of high contrast.

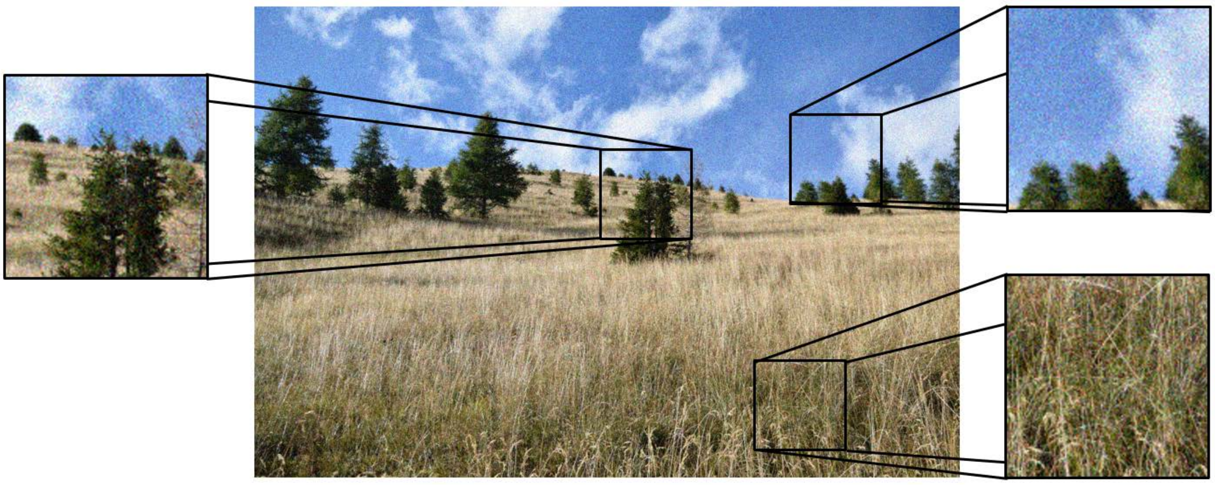

Figure 3: This image contains artificially added noise to showcase the human vision impact.
The noise is equally pronounced in all image areas, but is perceived as more disturbing in homogeneous image areas.
As an illustration, the three cut-out images are from sections with different structural content.

2.5 Video compression

The basic goal of video compression is to reduce the amount of image/video data without throwing away any relevant information.

Many different encoding algorithms exist on the market, such as H.264 or the newer H.265 algorithm. The main idea behind video compression is using pixel correlations in a single image or between following image frames. In general, there are three different frame-types in video encoding algorithms:

► I-frames only use the correlation within a single image; they can be compared to a JPEG-image
► P-frames can use data from other frames as a reference, whereas
► B-frames can use blended data from two different frames (cf. section 3.3.3)

The video compression process mainly consists of two major steps. The first step is to use correlations in the image (used in I/P/B-frames) or between images (used in P/B-frames). In a still scene without noise, for example, the next picture in the case of P/B-frames can be predicted perfectly from the preceding image. In this case, bitrate is only spent on the I-frames, because there the preceding frames cannot be used in the prediction process.

However, as already mentioned, this compression is impossible, when there is white noise, because this kind of noise is temporally and spatially independent for every pixel. So, in case of P/B-Frames, this component cannot be predicted from the preceding frame.

The next compression step really removes information. This is typically done by a quantization process that reduces the resolution of the DCT2 (2D discrete cosine transform) components by a division with a certain factor. The DCT2 is generally used to remove local correlations (periodic patterns) on a block base. This process is often used to achieve a certain bitrate, and allows for trade-off between encoding quality and bitrate.

3 Main concepts and components

Bosch **Intelligent streaming** focuses on three main concepts:

- avoiding the encoding of noise,
- optimizing encoding according to human vision,
- and trying to avoid spending too much bitrate on irrelevant regions.

There are three main components involved in the optimization of images captured by the sensor:

- Image Processing
- Smart Encoder
- Video Content Analytics (VCA)

Figure 4 gives an overview of the relation between these components.

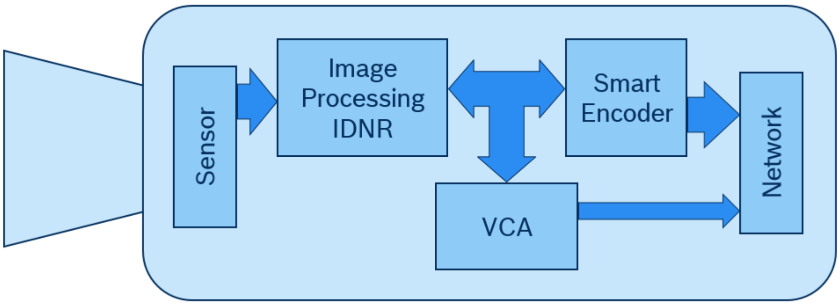

Figure 4: Illustration of the components related to Intelligent streaming.

3.1 Image processing

Surveillance cameras can be used for very different purposes, and the content of the captured scenes may vary widely depending on the use-case. To adapt perfectly to this variety of conditions, the image processing unit makes use of content-based imaging technologies and Intelligent Digital Noise Reduction (IDNR). Content-based imaging technologies dynamically adapt camera settings based on scene movement and changing lighting conditions whereas IDNR reduces image noise based on light levels and data created out of content-based imaging technologies.
Advancements in both technologies support the successive encoder to become smarter.

Overall, our camera sequentially applies different filters and functions to the captured images. The applied filters might vary depending on the various types of scene regions. For instance, our image processing pipeline is generally tuned to produce lowest possible noise in static regions of a scene while it may leave more noise in regions with moving objects to avoid a so-called "smear effect". Further, we apply motion-adaptive/compensated temporal filters to reduce motion artifacts.

"Scene modes", image setting templates predefined for certain scene types, control how these filters are applied. Our default scene mode configuration is compatible with the **Intelligent streaming** approach.

3.2 VCA

Our cameras contain a special unit that is able to analyze video content and provide recommendations for further image processing steps. This processing unit is called video content analytics (VCA). The data from VCA is sent to the image processing unit of the camera. The image processing unit uses this information and applies it for tasks like exposure control and image noise reduction.

3.3 Smart Encoder

3.3.1 Metadata exchange

The encoder receives metadata from both, the image processing unit and our VCA unit. The received metadata is then used to optimize the encoding settings. As an example, if the VCA does not detect any moving objects in the scene, the encoding quality is lowered automatically, reducing the resulting bitrate.

With the additional Intelligent streaming “Dynamic sharpness & noise filtering” mode, we also use information from the encoder to optimize the image processing according to the encoder requirements. There is always a trade-off between noise filtering, sharpness and bitrate. Depending on the encoder settings (e.g. bitrate reduction mode, bitrate, frame rate, resolution, etc.), the image processing pipeline is tuned in a way to achieve the best image quality for a given bandwidth budget.

3.3.2 Encoder quality regions

Some image regions might constantly cause a very high bitrate without containing any relevant information. The best example for such is a region with leaves that move in the wind. Our encoder allows manual selection of irrelevant regions that are then encoded with a different, typically lower quality. This further reduces bitrate while no relevant information is removed from the image. With this strategy it is possible to spend more of the available bitrate on highly relevant regions, such as facial details or number-plates.

3.3.3 Enhanced prediction modes

As already mentioned, the basic idea behind video encoding is to use correlation within a single image or between different video frames. These correlations are used to predict the content of a frame based on leading/trailing frames. The H.264 and H.265 encoding algorithms provide various tools for prediction, dependent on the frame type. For I-, P- or B-frames, different tools are available. The full toolset is only available with B-frames. There are multiple Bosch cameras available, which are capable of making use of B-Frames and other enhanced prediction modes.

- B-frames can help to further reduce the bitrate because they additionally have the option to blend data from two different frames to predict the image content.
- Long-term references enable the encoder to use content from the last I-frame as prediction. This can help, for example, if the background is hidden for a short time.
- Perceptual encoding allows the encoder to adapt the encoding quality according to the human vision.
- As B-frames typically do not contain references for other frames, they can be skipped during decoding, which leads to a better seek and time lapse performance. This is true even if they are used as reference frames (B*) in a hierarchical manner like in **Figure 5**.

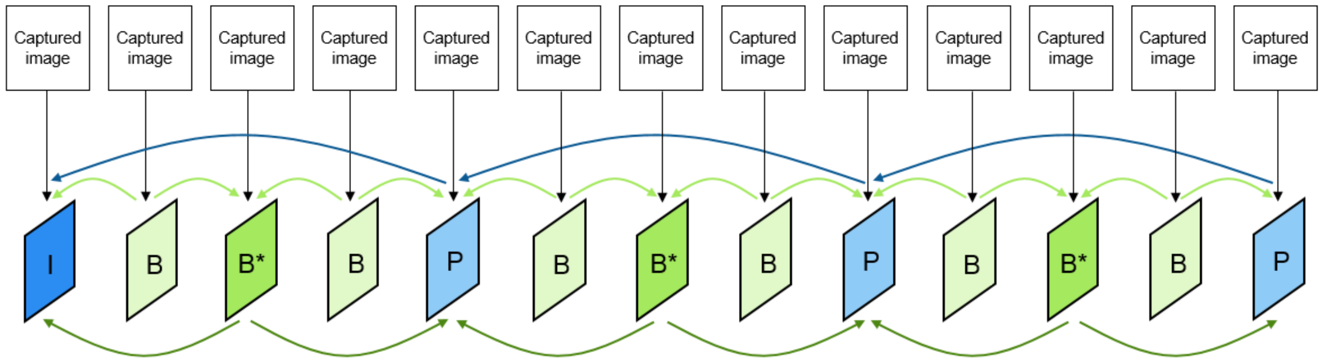

Figure 5: Encoded frames in a group of pictures (GOP) and their references to related pictures

3.3.4 Extended GOP structure

An I-frame naturally holds the largest amount of data as it provides full image information, thus it is typically much larger than P- and B-frames. From a bitrate-point of view, the goal is to send I-frames as rarely as possible. However, frequent I-frames are required for usable playback of video streams. As the decoder needs an I-frame to start decoding of a stream, the I-frame distance defines the time granularity for fast seeking into recordings.

With shorter I-frame distance at constant video quality the bitrate would increase. With too long of a distance between I-frames, the decoder would be required to read a lot of data and decode it before it could display the wanted image, likely resulting in annoying latency. Hence, the goal is to find an I-frame distance, which is a good trade-off between bitrate optimization and usability. We accounted for this in our **Intelligent streaming** solution.

3.4 Rate control

Overall, bitrate optimization technologies lead to a strongly scene-dependent bitrate. For example, if there is no movement in the scene the bitrate can be very low, whereas the bitrate is much higher if there is a lot of movement in the scene. To keep the bitrate predictable, **Intelligent streaming** provides a long-term rate control, which has two settings:

- an *Average bitrate*, which is realized by the long-term rate control, and
- a *Maximum bitrate*, which is the maximally allowed bitrate for the encoder.

The long-term rate control can run with an averaging interval of one day creating a predictable average bandwidth making it possible to precisely calculate storage requirements.

3.5 Encoding algorithm

The encoding algorithm itself, typically H.264 or H.265, is not part of **Intelligent streaming**. The smartness of the encoder is independent of the encoding algorithm in use. However, the encoding algorithm can further support bitrate optimization.

Scene content and complexity as well as resolution define how much encoding algorithms may help in saving bitrate. The industry acknowledges that the new H.265 encoding algorithm (also known as High Efficiency Video Codec or HEVC can deliver a bitrate reduction of up to 50% compared to H.264 without impact on image quality.

4 Configuration of Intelligent streaming

Among a camera's configuration pages, two of them allow the user to review and adapt the default Intelligent streaming settings.

4.1 Intelligent streaming encoder configuration

Intelligent streaming is configured via four parameters within the *Encoder Profile* pages. Each of the profiles contains the same set of parameters with different default values that match the intended use case of the profile.

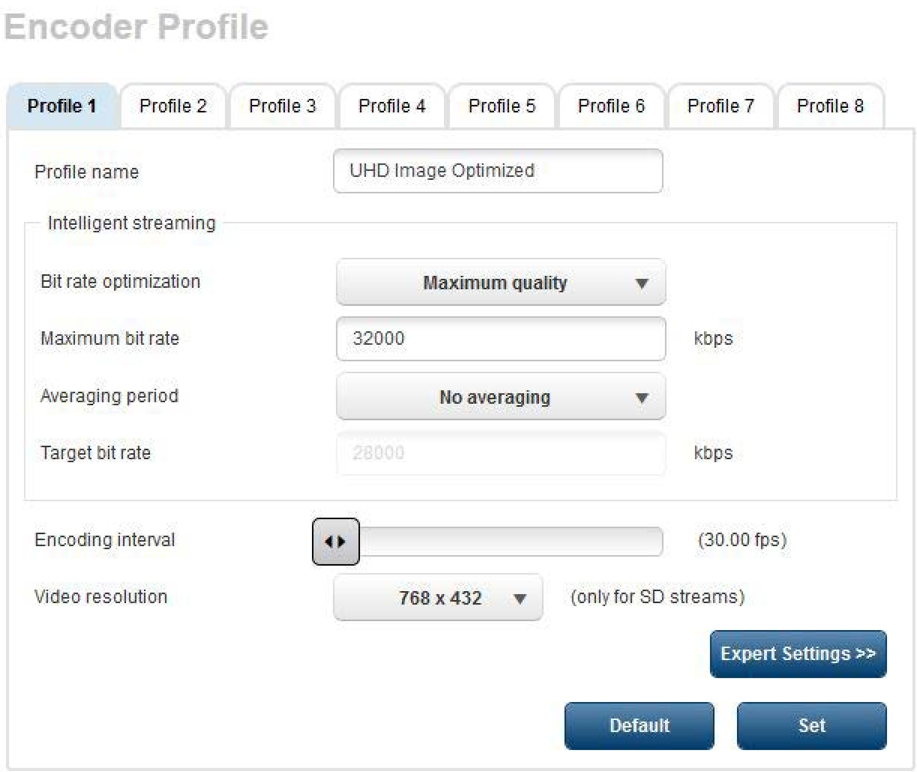

Figure 6: The parameters to configure Intelligent streaming within an encoder profile

The first parameter generally defines the bitrate optimization strength, the second sets a maximum allowed bitrate. These two parameters already define the working range of **Intelligent streaming**.

The bitrate optimization strength levels are defined as:

- Maximum quality
- High quality
- Medium
- Low bitrate
- Minimum bitrate

Please note that camera scene mode settings optimize image preprocessing in a certain way. These settings are then merged in a smart way with intelligent streaming and thus result in different tradeoffs depending on the scene mode.
A change in image resolution does not automatically trigger a change in the bitrate settings. For such cases, the bitrate settings need to be manually adapted to the image resolution. The same holds for switching to H.265. In this case the bitrate settings should be reduced by 25% compared to H.264.

In the "*Image-Optimized*" profile, the bitrate optimization strength is set by default to "*High Quality*", putting less focus on bitrate reduction. In the "*Balanced*" profile, typically used for live viewing (default live profile), the bitrate optimization strength is set by default to "*Medium*", with a higher maximum bitrate than the *"Bit Rate-Optimized"* profile. The *"Bit Rate-Optimized"* profile is the default recording profile.

The next two parameters define an averaging period and a target bitrate. The averaging period specifies the time interval that is used for future bitrate prediction, whereas the target bitrate describes the overall desired long term bitrate. The impact and benefit of both settings will become clearer in the following chapter.

4.2 Dynamic sharpness and noise filtering

The latest "Dynamic sharpness and noise filtering" extension of Intelligent streaming can be enabled or disabled on the *Enhance* page. If enabled, sharpness and noise filtering will be adapted depending on the encoder configuration of stream 1 and 2, where, for example, bitrate optimization strength level, bitrate, frame rate and resolution are taken into account. The resulting offsets are displayed on the right-hand side of the three corresponding parameters, highlighted by a box, which is titled "Intelligent streaming dynamic offsets".
In addition, there is a drop-down list ("Priority encoder stream"), where priority with regard to image processing optimization can be assigned to a certain stream. Default is the "Auto" approach. In this case the tuning is based on a balance between the two encoders of stream 1 and 2. Moreover, the requirements of the encoder requesting a higher level of sharpness are used to internally define a limit. This limit is applied in order to avoid that one stream, configured for a very low bitrate, introduces too little sharpness compared to the other encoder, configured for higher bitrates.
If a certain stream is selected from the drop-down list, thus not using the "Auto" mode, only the requirements of this encoder are taken into account for the optimization. Furthermore, an information box is provided (the "I" icon inside the Intelligent streaming box), which provides an estimation of the expected bitrate reduction, based on all the current Intelligent streaming settings.

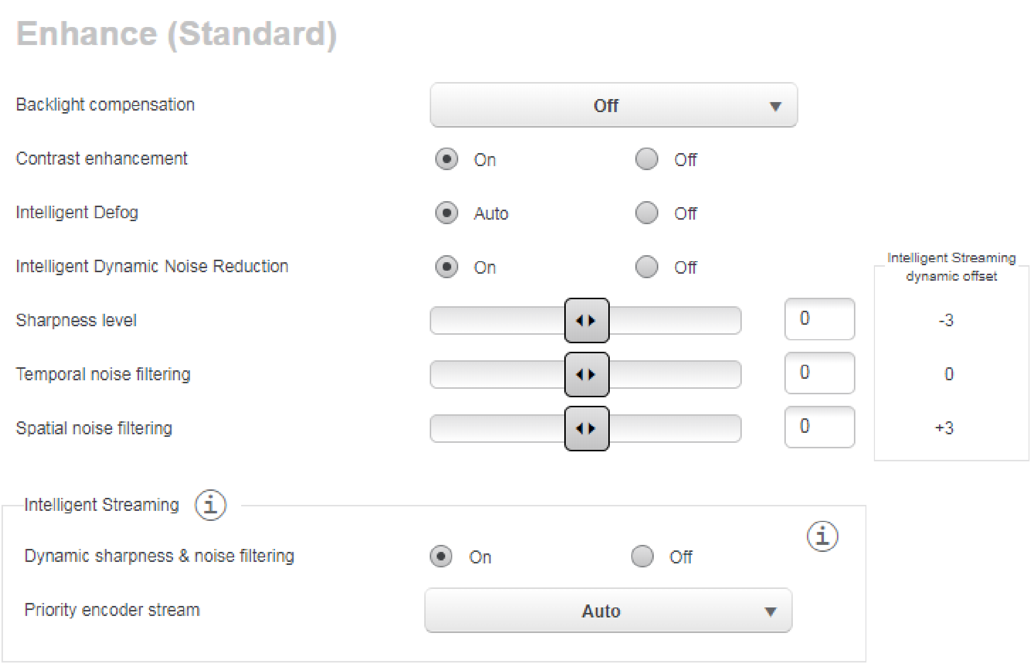

Figure 7: Activation and prioritization of "Dynamic sharpness and noise filtering"

5 Application Modes

There are two principal modes when utilizing **Intelligent streaming**. The first mode is quality driven, the second tries to achieve an average bitrate over a specific timeframe. In the previous chapter, we already learned about configuration of **Intelligent streaming**. We speak of quality driven mode if only the first two settings (bitrate optimization strength and maximum bitrate) are used and the averaging period is set to inactive. As soon as the averaging period is set active, we speak of a bitrate driven application mode. Both modes are explained in the following chapters 5.1 and 5.2. In all following tests, H.264 was used, as it is currently the default setting on our cameras. H.265 would provide an additional bitrate reduction, but also typically reduces the maximum encoding performance.

5.1 Quality driven

In the quality driven mode, the encoder generates a bitrate up to the maximum setting if the scene requires it. The actual bitrate strongly depends on the current scene conditions and the bitrate optimization setting. It is recommended to watch the scene activity and the resulting bitrate to tune the settings in order to to meet expectations, especially when this stream is intended to be recorded somewhere.

Cameras equipped with firmware version 6.5 or newer, a bitrate monitoring tool is available which provides bitrate statistics for seconds, minutes, hours, days and weeks. This allows for a good prediction of the average bitrate and the resulting recording time if the scene is periodic, even when no long-term rate control is used. The upcoming two sections (5.1.1 & 5.1.2) give insight to the bitrate distribution of a typical indoor and outdoor scene.

The following two figures show a comparison of default recording settings with the additional Intelligent streaming “Dynamic sharpness & noise filtering” enabled versus disabled. This enhancement was first introduced in firmware version 7.10.
In the comparison, two DINION IP starlight 6000i IR cameras were used, both recording the same scene. Stream 1 was recorded with a resolution of 1920x1080 @ 60fps, using the default recording profile “HD Bit Rate Optimized” (max.

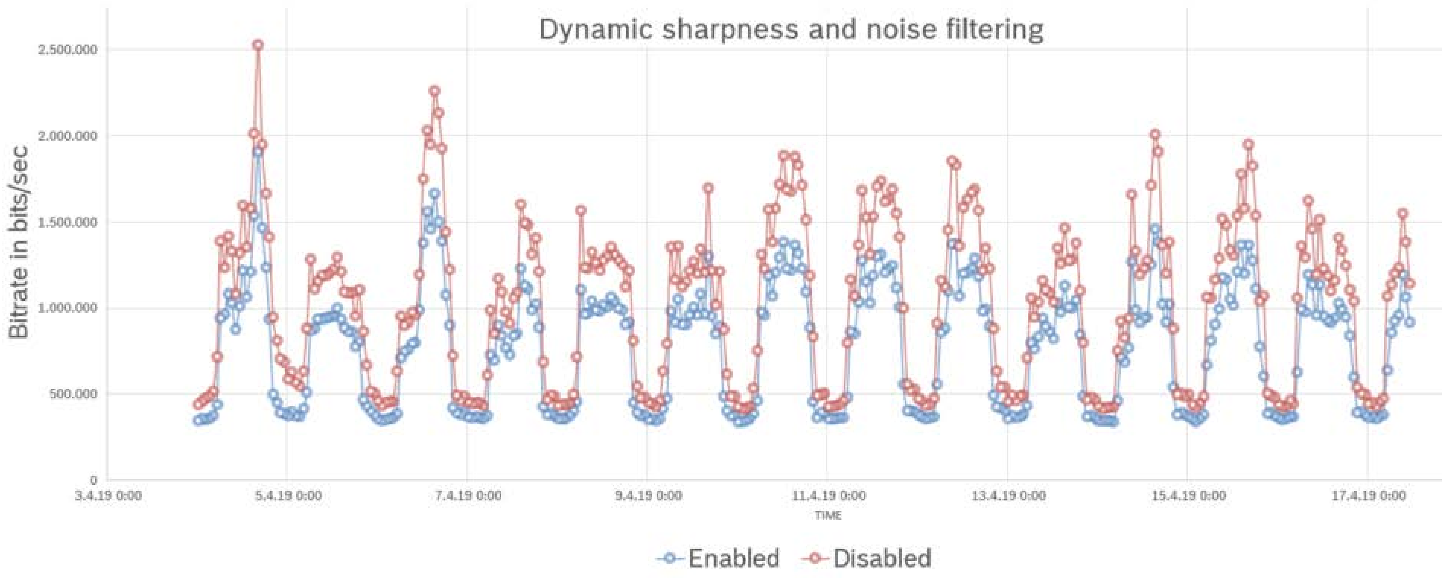

Figure 8: Comparison of Intelligent streaming with and without "Dynamic sharpness and noise filtering"

3400/kBits). Stream 1 is displayed in figure 8 for both cameras. Stream 2 was running in SD resolution using the “SD Balanced” profile. The only difference is that for one camera “Dynamic sharpness & noise filtering” is enabled and for the other camera, it is disabled. With this Intelligent streaming enhancement, an additional reduction of 25% bitrate was achieved on average in this example, while in general using default recording settings.

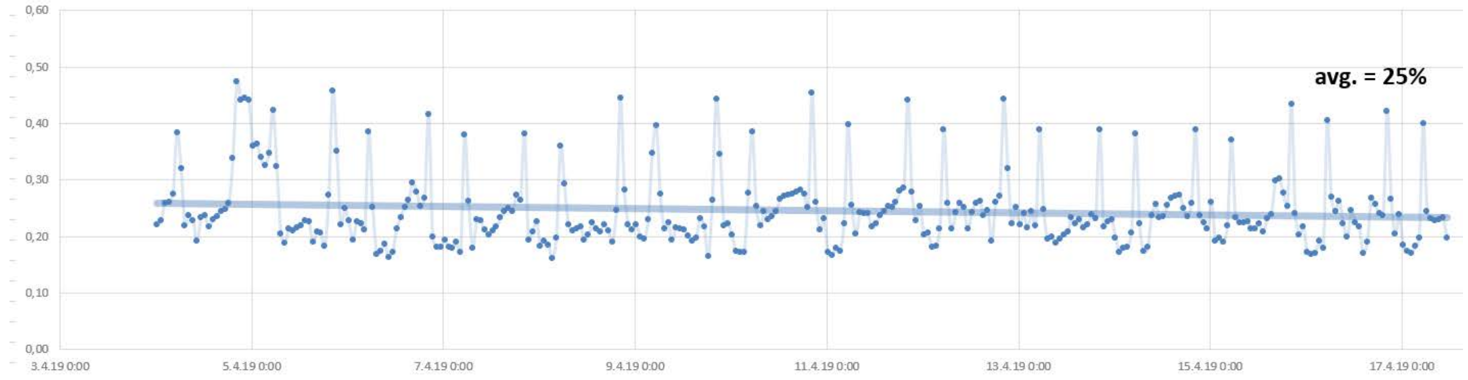

Figure 9: Relative bitrate savings of Intelligent streaming with "Dynamic sharpness and noise filtering" enabled vs. disabled

The quality driven mode is overall the mode of choice for all applications where the focus lies more on image quality and less on predictable bitrates and predictable recording time. However, the maximum bitrate setting also indirectly defines a minimum available recording time. In reality, recording time is much longer as the maximum bitrate is not permanently required.

The following chapters 5.1.1 and 5.1.2 visualize how changing scene content generally might affect the produced bitrates and chapter 5.1.3 emphasizes the benefits of our **Intelligent streaming** solution.

5.1.1 Bitrate monitoring of an outdoor scene

Outdoor scenes usually have a strong periodic structure with an interval of a day, caused by reoccurring events such as sunset and sunrise. Typically, variance can be caused by random events like snowfall, rainfall and strong wind causing movement in leaves.

Figure 10 and **Figure 11** show the bitrates of the same typical outdoor scene over one week. The figures were generated with different averaging periods, which means the temporal resolution of the captured averaging samples. **Figure 10** was sampled with an averaging period of one hour. A strong periodic structure can be recognized. During night, very low bitrates are generated, whereas the bitrate is much higher during the day. **Figure 11** was sampled with an averaging period of a day. The periodic structure leads to less variation when averaging over one day.

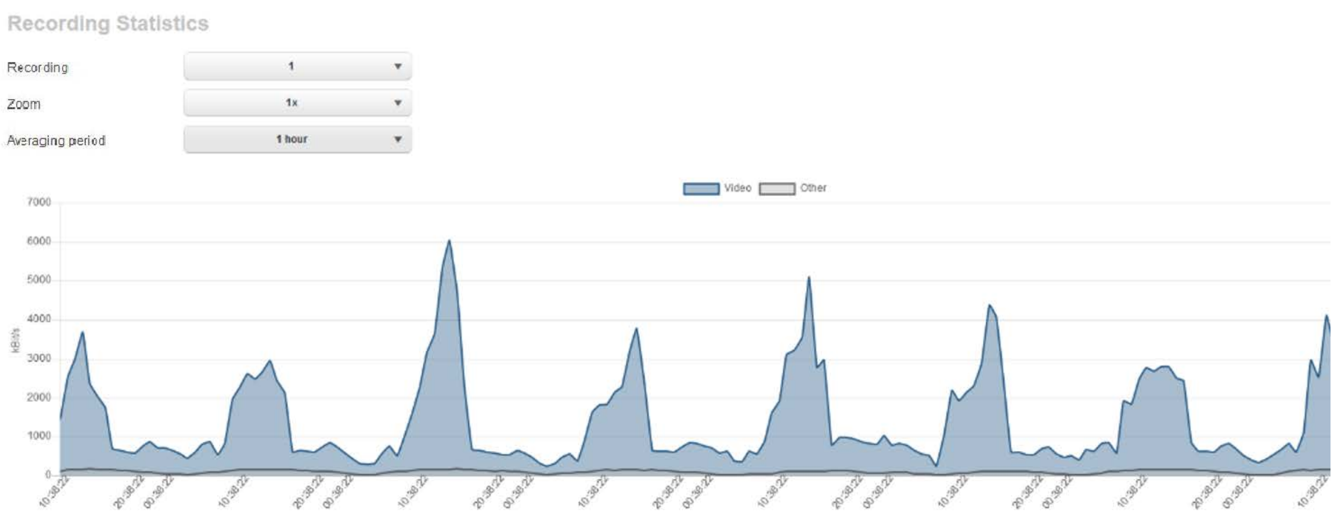

Figure 10: Bitrates with an averaging period of an hour captured with DINION IP starlight 7000 HD

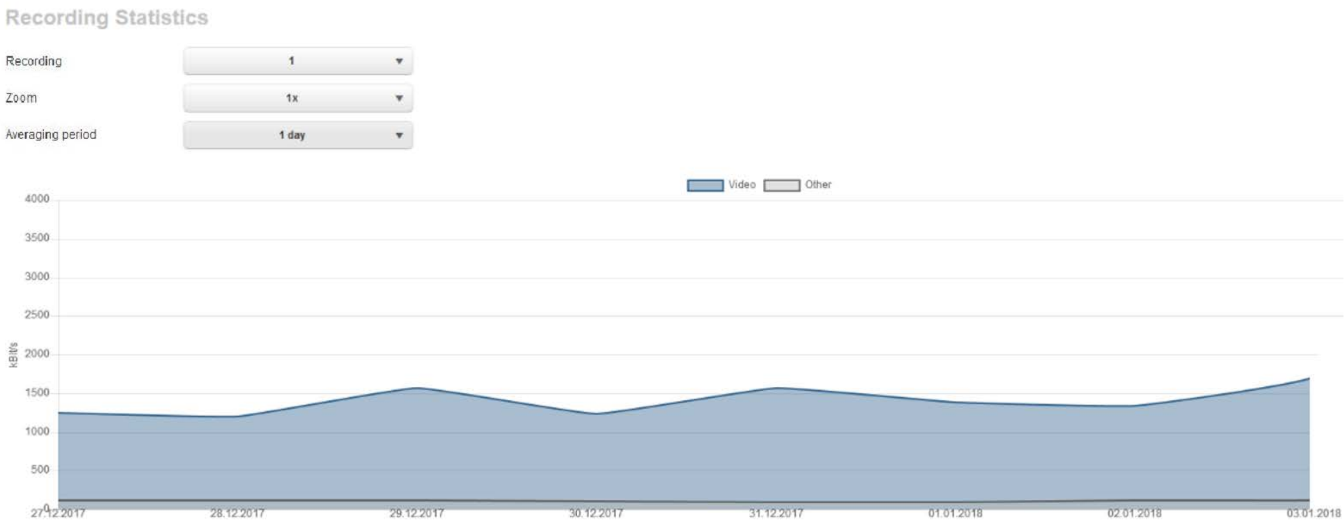

Figure 11: Bitrates with an averaging period of a day captured with DINION IP starlight 7000 HD

5.1.2 Bitrate monitoring of an indoor scene

Indoor scenes typically also have a strong periodic structure with an interval of a day, caused by reoccurring events during one day. In addition, indoor scenes contain fewer random events like snowfall, rainfall and strong wind. This often leads to a less variable bitrate when compared to outdoor scenes. However there may also be periodic structures on a weekly basis, for example, when looking at a business which is only open during week days (e.g. closed nights and weekends).

The following graphs **Figure 12** and **Figure 13** show a simple indoor scene captured with DINION IP 4000i IR using H.265 compression over 5 days. The figures were again generated with different averaging periods. **Figure 12** was sampled with an averaging period of one hour. A strong periodic structure can be recognized again. During night, very low bitrates are generated, whereas the bitrate is much higher during the day. **Figure 13** was sampled with an averaging period of a day. The periodic structure leads to much less variation when averaging over one day.

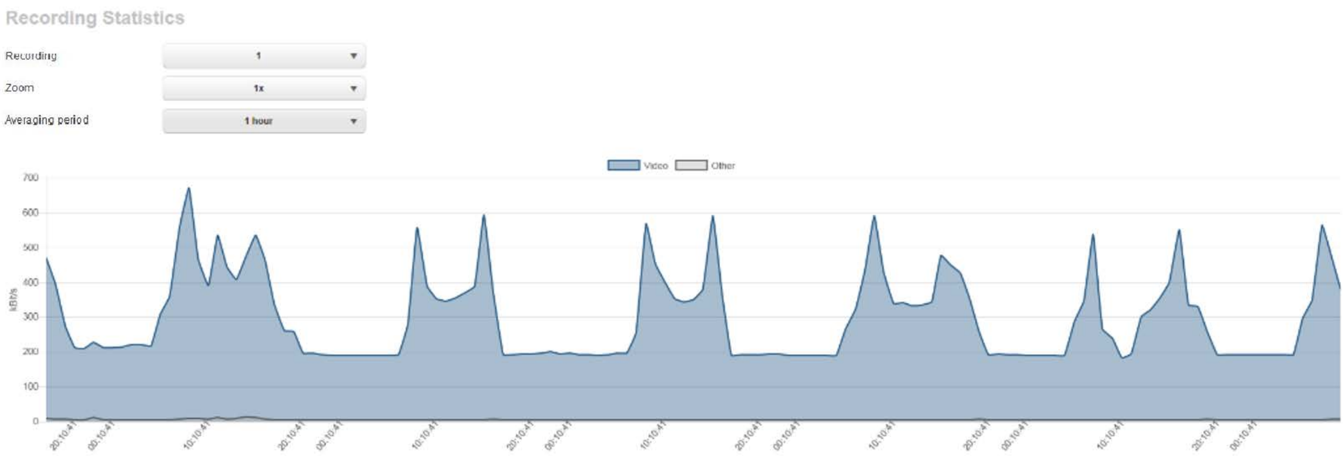

Figure 12: Bitrates with an averaging period of an hour captured with DINION IP 4000i IR and H.265

Figure 13: Bitrates with an averaging period of a day captured with DINION IP 4000i IR and H.265

5.1.3 Proving the benefit: bitrate monitoring of a full-day cycle

In this chapter, we will illustrate the benefits of **Intelligent streaming**. The following **Figure 14** shows the recording statistics over a full day circle with **Intelligent streaming** set active. The camera used in this example is a DINION IP ultra 8000 MP that is set to 4K resolution at 30 fps. The camera was pointed at a street with moderate traffic, mainly before and after business hours. The typical periodic structure in the produced bitrate is easily recognized in the graph.

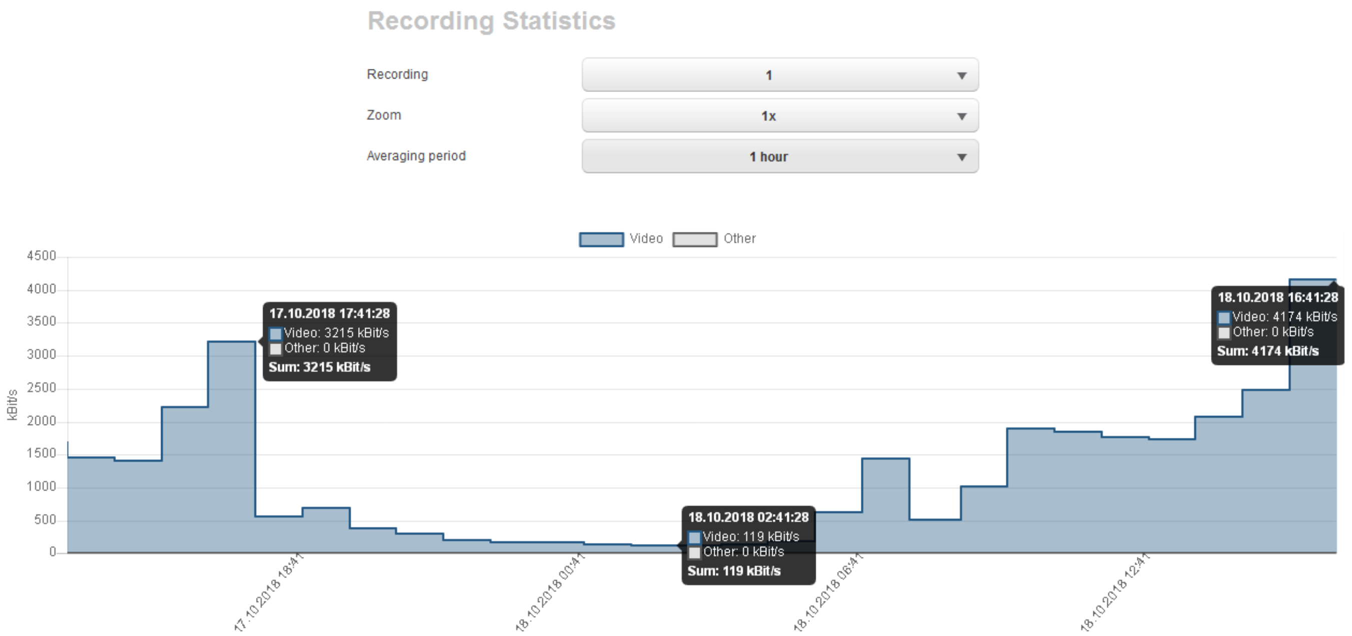

Figure 14: Example full-day statistics from afternoon to afternoon with Intelligent streaming.

Figure 15 shows the encoder configuration that was used to generate the data in **Figure 14.** The bitrate optimization parameter was set to the default-setting medium and the camera records only stream one in continuous mode without scheduling.

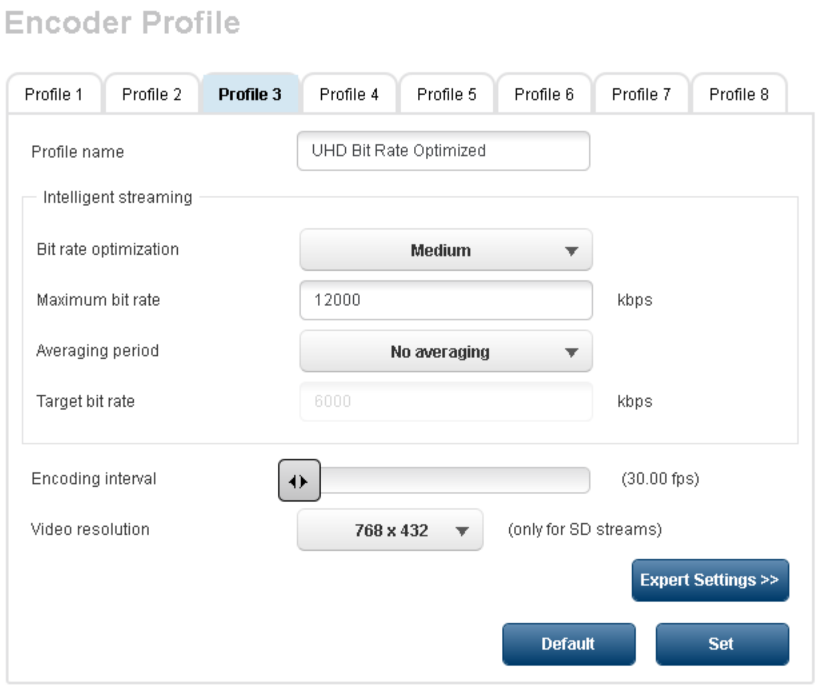

Figure 15: Profile settings of the recorded UHD stream with bitrate optimization *Medium* and 12 Mbps maximum bitrate.

As **Figure 14** shows the averaged bitrate within an hour, it does not give an indication of the real bitrate. Reducing the averaging period to 1 minute gives an impression in **Figure 16** that the encoder can still use the full dynamic range of up to 12 Mbps when the scene and motion in there require it. Reducing the averaging period further to 1 second would show the real bitrate values, but as the number of values for the history is limited, it would cover only the last few minutes. It requires full motion happening in the scene in that period to see the bitrate going close to the limit.

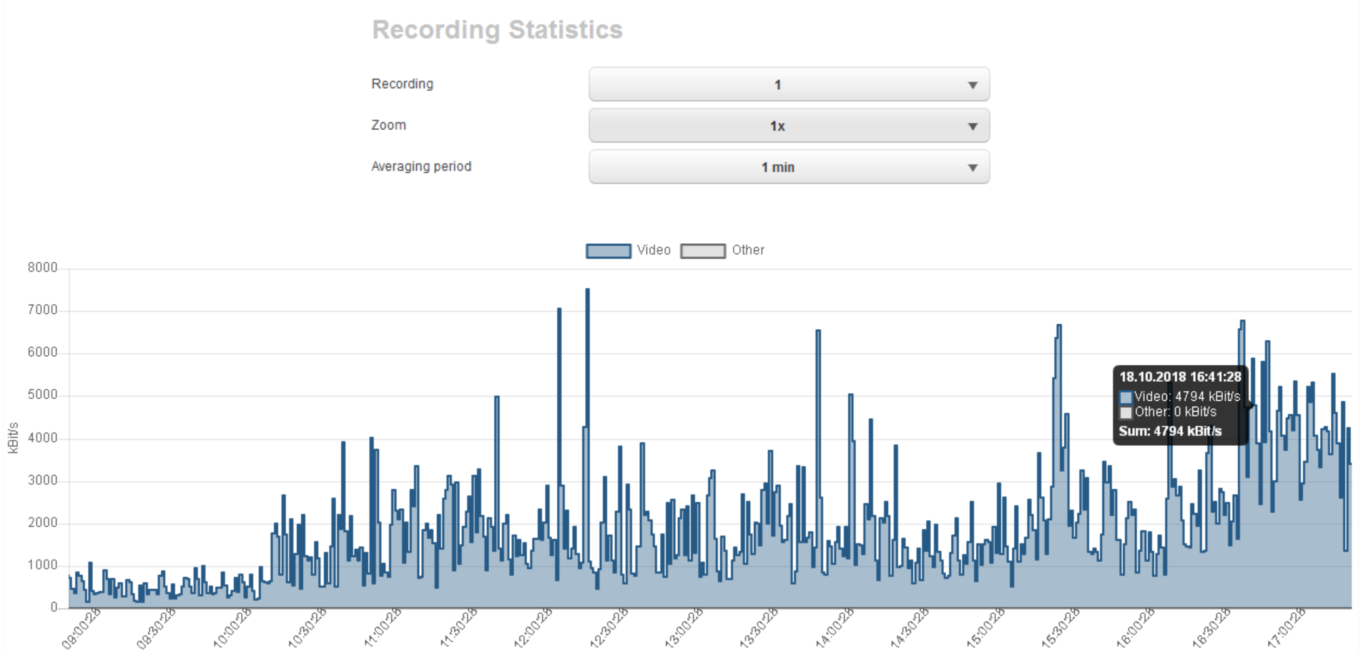

Figure 16: Impression of the bitrate dynamics in the recorded stream at a shorter averaging period.

Then, the same camera setup as above is used, however, bitrate optimization was switched off so that **Intelligent streaming** is inactive. The camera watched the same scene during a working day and under similar weather conditions. Its encoder profile settings were unchanged with 12 Mbps maximum bitrate, and the encoder will try to achieve a target bitrate of 6 Mbps where possible. **Figure 17** shows the corresponding bitrate distribution.

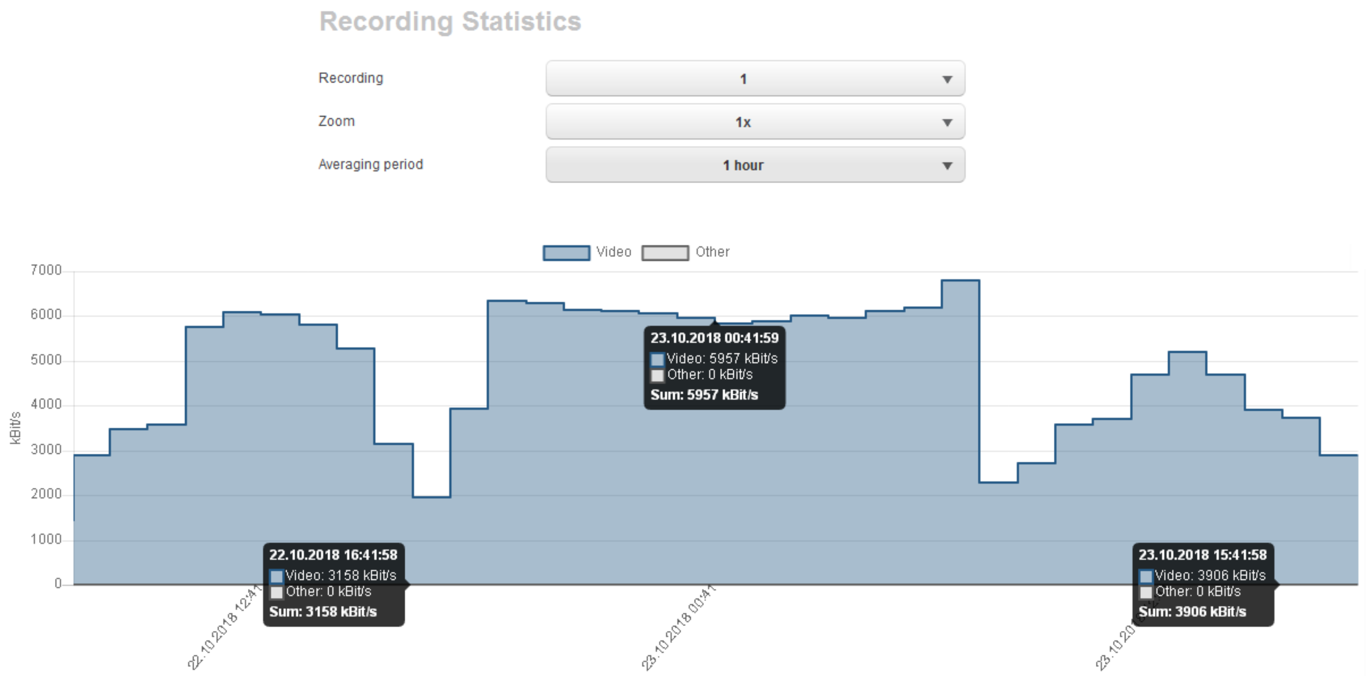

Figure 17: Example full-day statistics from afternoon to afternoon without Intelligent streaming

During daytime, with lighting conditions producing sufficient signal to noise ratio, the encoded bitrate follows the ratio of motion in the scene, creating a shape similar to **Figure 14**, but with noticeably higher bitrate overall. During night, the camera switches to night mode and increases the gain for the sensor signal. This leads to a larger amount of noise being encoded, resulting in an overall higher bitrate even in the absence of motion in the scene. The overall bitrate is around 6 to 7 Mbps, adequately matching the configured target bitrate.

The previous examples show how **Intelligent streaming** helps save bitrate during the course of a day. **Figure 18** now illustrates the impact of **Intelligent streaming** on the produced bitrate over a longer period of six days. The first day was recorded with **Intelligent streaming** set active, which results in a bitrate of 1.4 Mbps. However, the encoder can still use the full bitrate dynamic of up to 12 Mbps to ensure a good image quality in scenes with high activity.

The days after, **Intelligent streaming** was switched off for comparison, resulting in bitrates around 5 Mbps. The benefit of **Intelligent streaming** with significantly lower bitrates at comparable image quality is clearly visible.

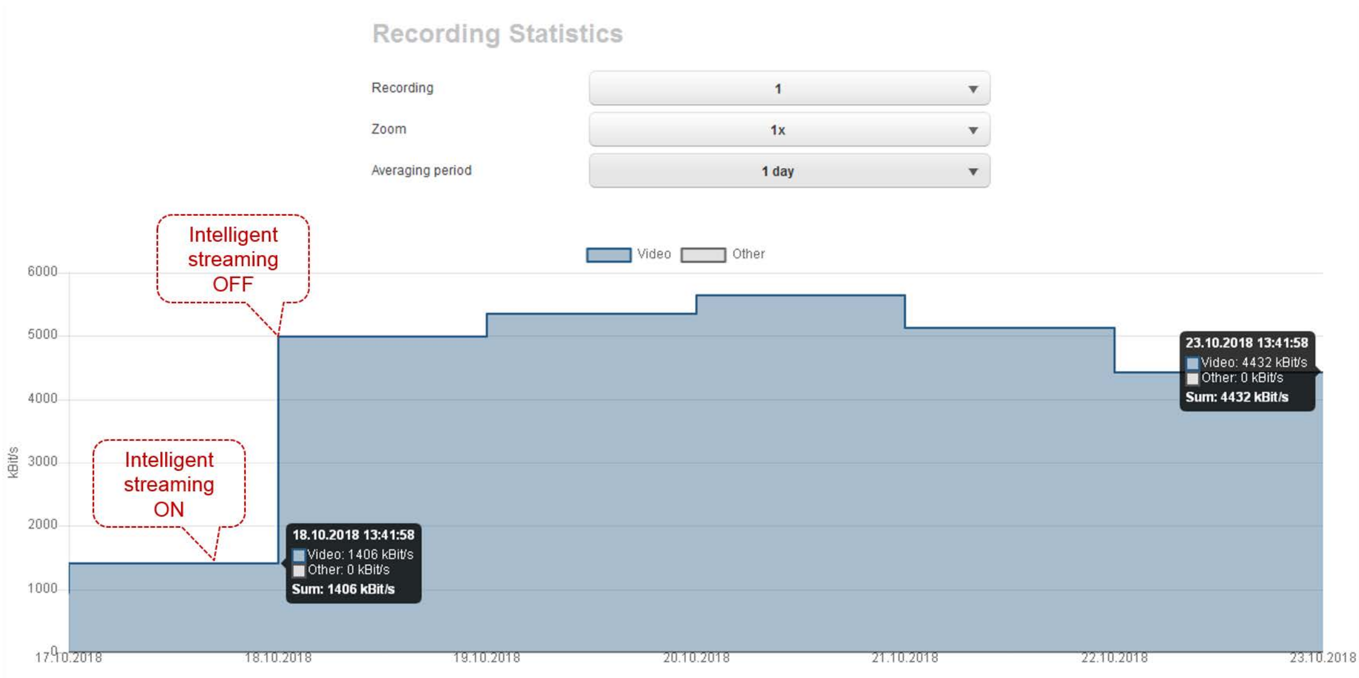

Figure 18: Comparison between Intelligent streaming on and off on a daily recording average

5.2 Average bitrate driven

The average-bitrate driven mode makes use of an internal loop-back of the long-term rate control to the encoder. This loop-back sets the priority on the resulting bitrate instead of on the configured video quality.

In the average bitrate driven mode, **Intelligent streaming** tries to achieve a target encoding bitrate. In addition to this target bitrate, an averaging period must be configured (cf. figure 15). The averaging period specifies the time interval that is used for future bitrate prediction. Most of the time an averaging period of a day is a good value because of the periodic sunrise. In other scenarios, like when looking at a business which is only open during week days, a week can be a good averaging period. Then, the long-term rate control continuously fine-tunes the encoder settings, based on its predictions, to meet the configured target bitrate.

This mode is especially helpful if recording space must be managed to ensure minimum retention times in systems that are at the limit of storage capacity, or where remaining space should be prioritized for more important or more dynamic cameras.

It is recommended to watch the scene activity and the resulting bitrate to tune the settings to meet expectations. If the average activity in the scene exceeds the matching level, the encoder is forced to reduce the image quality to stay within the budget, maybe more than what is considered acceptable. If the average activity is lower, recording space is left unused. Such a tuning should consider sufficient room for dynamic changes in the scene, and should be repeated after a number of averaging cycles. The recording statistics with its flexible averaging period option visualizes the cumulated recorded bitrate and thus perfectly supports such a tuning process.

The following two figures show the bitrates of a scene where a toy train passes the camera every 30 sec. The figures were generated with different averaging periods. **Figure 19** was sampled with an averaging period of one second and **Figure 20** was sampled with an averaging period of 1 minute. The spikes in **Figure 19** are caused by the comparatively large I-frames. In **Figure 20,** those spikes are balanced out due to the larger averaging period.

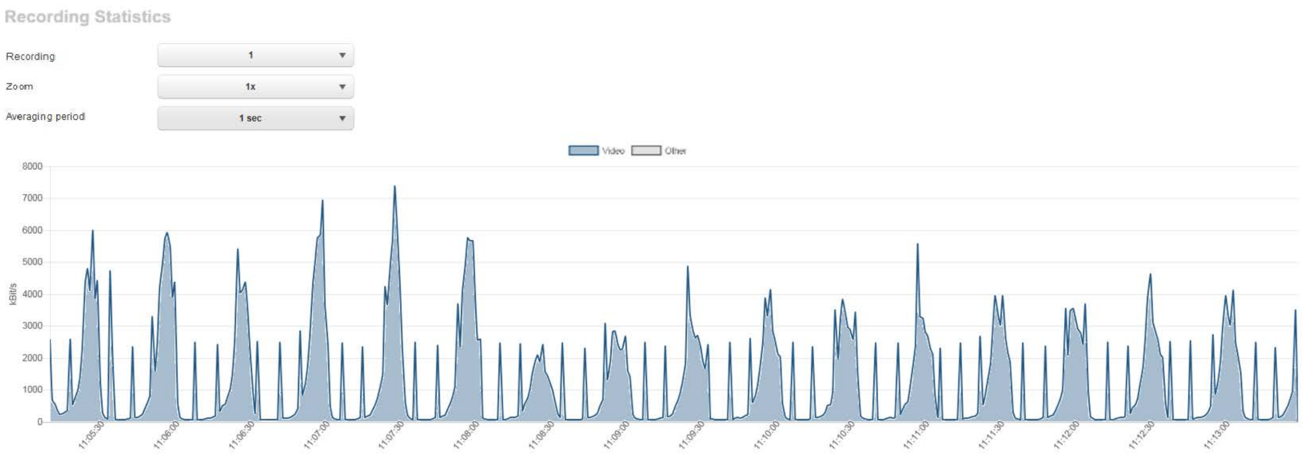

Figure 9: Bitrate sampling with an averaging period of 1 second

Figure 10: Bitrate sampling with an averaging period of 1 minute

Until 11:08:00, the maximum bitrate is configured to 6.6 Mbit/s, and bitrate reduction mode is set to medium. **Figure 20** shows that during this time the bitrate averaged over one minute is around 1.6 Mbit/s. At 11:08:00 long-term rate control with a target bitrate of 1.2 Mbit/s is activated. It can be seen now in **Figure 20** that the rate control starts to kick in and achieves an average bit rate of 1.2 Mbit/s. However, **Figure 19** shows that peaks up to 4Mbit/s still occur and thus allow for sufficient image quality.

6 Conclusion

As a camera manufacturer, our goal is to provide a proper video quality while reducing video storage costs as much as possible. Video storage cost directly correlates to the produced bitrate, and the produced bitrate heavily depends on the specific video content such as moving objects or image noise. Image noise generally causes a higher bitrate without contributing any relevant information. With **Intelligent streaming**, we strongly reduce the bitrate wasted by image noise using a well-harmonized setup of noise reduction algorithms and encoder tuning. This setup allows us to spend the available bitrate only for relevant image content.

In **Figure 21** there is a comparison of full-day bitrate statistics of a camera with and without **Intelligent streaming** enabled. The bitrate graph of **Figure 14** is laid over the graph of **Figure 17** as a red-dotted line for direct comparison and thus, on an hourly basis, gives an indication where **Intelligent streaming** provides the biggest benefit. As the recordings happened at different days with different motion, hours cannot be compared one-to-one. But as both were working days with similar characteristics, a rough comparison is nevertheless possible. As **Figure 21** shows, considerable savings are achieved during daylight, as only true motion leads to encoded bitrate. During night, where noise increases but true motion decreases, even more bitrate can be saved.

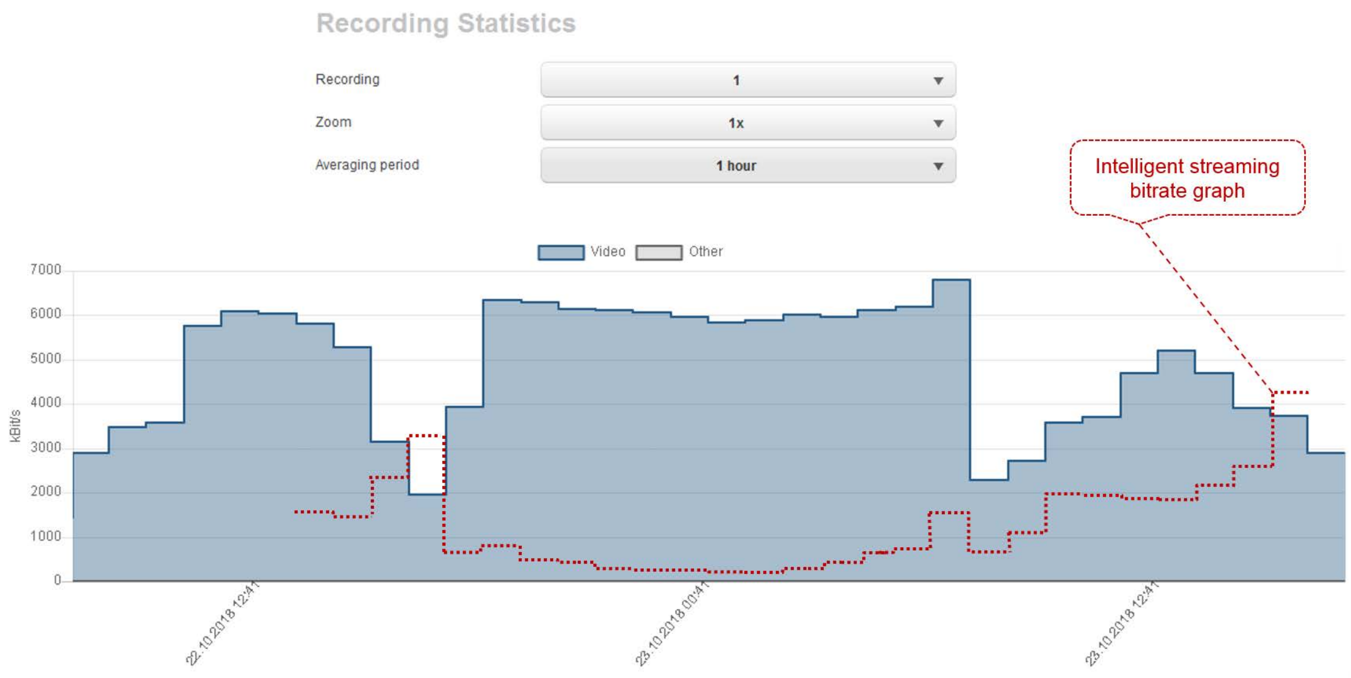

Figure 11: Comparison of bitrates for a similar scene with Intelligent streaming versus without Intelligent streaming

The content that is captured with our cameras might vary widely among customers, however most surveillance scenes show a strong periodic structure. This periodic structure mostly depends on the sunrise and sunset as well as on situational conditions (e.g. there are generally less cars on the street during night, and more people in shops during opening hours). In our well-harmonized setup, the periodic structure of scene content also causes a periodic structure in the produced bitrate.

Intelligent streaming offers functionality with a simple-to-use interface that enables our customers to monitor, analyze and control the produced bitrate based on their very specific use cases. A long-term rate control ensures predictable bitrates, and the use of enhanced encoding features ensures proper video quality even for lower bitrates.

In summary, **Intelligent streaming** provides the possibility to precisely manage the produced bitrate while automatically using the available bitrate budget only on relevant image content. Without **Intelligent streaming**, either available storage space might limit image quality in relevant scenes or expensive storage space might be wasted.

7 Authors and References

Konrad Simon, Product Manager Platform Firmware and Software
Jan Brandstädter, Engineering Technologies and Platforms
Florian Kuhn, Engineering Technologies and Platforms
Robert Haase, Product Manager Platform Firmware and Software

---

[1] https://ece.uwaterloo.ca/~ece434/Winter2008/Noise.pdf

[2] https://www.cs.unm.edu/~williams/cs591/ne940406

[3] https://web.stanford.edu/class/ee368b/Handouts/09-HumanPerception.pdf

Bosch Sicherheitssysteme GmbH
Robert-Bosch-Ring 5
85630 Grasbrunn
Germany
www.boschsecurity.com