1. Preface

2. Hanwha Techwin's Analysis Technology
   2.1. Tampering Detection
   2.2. Face Detection
   2.3. IVA (Intelligent Video Analysis)
   2.4. Audio Detection
   2.5. Audio Source Classification
      2.5.1. Category
      2.5.2. Noise Cancellation
      2.5.3. Detection Level
   2.6. Image Stabilization

3. Conclusion
1. Preface

To protect one's life and assets against various accidents, incidents and terror attacks, the number of video security installations is increasing at a high rate. However, the number of security personnel has not increased at the same rate, resulting in an average of 20 to 50 units being monitored by one security personnel. Monitoring more than 2 video security units simultaneously prevents security personnel from focusing on monitoring duties for more than several minutes which can lead to an increased probability of missing critical situations.

For these reasons, interest in intelligent audio and video source analysis technology for overcoming such limitations and for efficient monitoring is expanding, and research is striving to advance. Intelligent audio and video analysis is a technology which alerts the operator of abnormal activities detected through analyzing video and audio information designed to prevent accidents and to minimize damage through prompt response in case of an accident.

Hanwha Techwin is continuing to invest in the research and development of intelligent analysis technology, and this white paper is designed to provide information on the intelligent audio and video source analysis technology featured in Hanwha Techwin network cameras.
2. Hanwha Techwin's Analysis Technology

2.1. Tampering Detection

Tampering detection is a technology which detects events that disturb normal monitoring, and it is a crucial technology which all monitoring systems must provide. In case of sudden changes such as the following, the camera may not be able to perform normal monitoring. If the following changes occur, check the device on site and implement suitable measures.

- Camera direction changed due to impact
- Camera focus significantly impaired
- Camera vision lost due to object being covered or the camera being spray painted on
- Camera video lost due to intentional blockage

In normal monitoring environments, insignificant sudden or gradual lighting changes may be present, or the camera may be subject to repeated vibration due to wind or vibration from the installation location. Furthermore, an object temporarily appearing on the screen or repeated changes on portions of the screen may be detected.

Hanwha Techwin's tampering technology effectively excludes such elements in normal monitoring environments and is designed to detect significant events. Furthermore, it provides real-time, visual screen change levels which allow optimization according to various monitoring environments for defining minimum detection time until sounding an alarm.

2.2. Face Detection

Face detection is a technology which identifies human faces and defines facial ranges from video images by utilizing the key features of human faces. There are a variety of methods used to detect faces, and here are the three most noted face detection methods:
**Template Matching Method**

This method develops templates based on facial information extracted and registers the relationship in the system. Then it calculates the similarity between faces in video images and the templates.

![Image 1. Sample template with facial and relation information](image)

**Feature Invariant Approach**

This face detection method utilizes facial features which are less influenced by rotation, size and lighting changes. It combines information about eyes, noses and mouths to determine the presence of a person's face.

![Image 2. Facial feature information](image)

(a) eyebrow  (E) eye  (C) nose  (d) mouth

**Boosting Approach**

This method utilizes basic patterns of faces which are compared to a classifier containing facial feature information for determining an individual's face.

![Image 3. Basic facial pattern sample](image)

- Image 1, 2, 3 source: “Real-time face detection technology research trend” by Prof. Dongil Han of Sejong Univ.
Facial detection requires a certain amount of video information, and products equipped with Hanwha Techwin Wisenet5 SoC significantly improves its face detection performance by collecting more detailed video information from user designated areas.

![Image 4. Face detection size of previous product (L) and X series (R)](image)

2.3. IVA (Intelligent Video Analysis)

The system can be set to generate an event signal in cases where it detects movement or a situation that satisfies defined event rules.

**Crossing Line Detection**

Objects crossing a designated virtual line can be detected.

**Enter / Exit Detection**

Objects entering/exitng a designated virtual area can be detected.

![Image 5. Crossing line detection & Enter / Exit detection samples](image)
**Appearing Detection**

Objects appearing in a designated virtual area and holding their position for more than the set observation time are detected.

![Image 6. Appearing detection sample](image)

**Disappearing Detection**

Objects disappearing in a designated virtual area and remaining absent for more than the set observation time are detected.

![Image 7. Disappearing detection sample](image)
**Loitering Detection**

Objects loitering in a designated virtual area for more than the set observation time are detected.

![Image 8. Loitering detection sample](image)

It does not detect movements that are smaller than the user-defined minimum movement and larger than the user-defined maximum movement. To avoid detection errors due to various noise, set suitable minimum/maximum detection size for the installation environment. However, as identical movement from identical locations may be detected differently, be sure to include margins in the minimum/maximum size limitations.

**2.4. Audio Detection**

Audio detection is a technology that detects audio levels which exceed the user-defined levels. As audio levels are greater in abnormal situations than in normal situations, audio levels exceeding set levels are detected as being an abnormal situation. Through audio detection technology, the camera is able to detect abnormal situations, then notify the operator via event signals allowing the operator to implement suitable measures.
2.5. Audio Source Classification

Audio source classification is a technology classifying audio types of the detected audio. As audio detection technology generates alarms based on simple audio size, it is possible to generate events even under normal situations. To overcome such limitations, technologies to classify audio source types are being developed.

Audio source classification technology classifies audio types detected during camera recording, and it can generate events when an audio source satisfying the criteria the operator defined is detected. The camera detects the audio source type, then notifies the operator via event signals allowing the operator to implement suitable measures.

Hanwha Techwin’s audio source classification technology X series camera features three customizable settings for category, noise cancellation and detection level for optimum performance in a variety of installation environments. It also provides a graph which visualizes audio source levels to allow for the intuitive checking of noise cancellation and detection levels setup.

2.5.1. Category

Generates events based on audio source type detection. An operator can select the type of audio source for detection, and multiple audio sources can be selected.

![Image 9. Audio source detection category setup](image)

- **Scream**: Generates events based on detections of large voices made by humans such as screams and the yells of adult males/females and children.
- **Gunshot**: Generates events based on the detection of non-continuous gunshot sounds.
2.5.2. Noise Cancellation

If surrounding noise of the camera is greater than 55dB-65dB, the system can enable the noise cancellation function.

If an audio source is generated within an environment with noise, disabling noise cancellation functions can show surrounding noise in the form of a graph, and enabling the function can show canceled surrounding noise in the form of a graph.

• With noise cancellation enabled, the system analyzes the attenuated audio source. As such, the audio source classification performance may be hindered or generate errors.

![Image 10. Audio source energy graph with noise cancellation disabled (top) and enabled (bottom)]
2.5.3. Detection Level

Sets audio source energy levels to perform audio source classification. Energy levels of the input audio source are periodically updated from right to left forming a graph. Audio source classification is performed only on audio sources exceeding the set level. In other words, only input audio sources with energy exceeding the threshold undergo audio source classification.

Lower thresholds result in greater audio source classification data and greater misdetection probability. Higher thresholds result in less audio source classification data and greater non-detection probability. The threshold must be set appropriate to the surrounding noise level of the camera.

Image 11. Audio source classification targets for level 33 (top) and level 50 (bottom)
2.6. Image Stabilization

Image stabilization is a technology that compensates image shaking due to vibrations from the environment to produce a stable image. In general, image stabilization technology is classified as a hardware method which utilizes the camera lens or image sensor to compensate for shaking, and DIS (Digital Image Stabilization) which utilizes a software analysis of shaking based on the image. As DIS compensates for shaking with software unlike hardware compensation methods, it can reduce a product's price by reducing the amount of hardware in the product.

As more cameras are installed in public organizations, parking lots, city outskirts as well as various environments and fields, and as the technology for image analysis advances, image stabilization technology is being regarded as a crucial feature for allowing different technologies to perform at their maximum capacity beyond simply recording stable images.

Hanwha Techwin's image stabilization is based on the software compensation method, DIS. This technology does not require any additional hardware giving it an advantage in cost, but as it is highly dependent on the image, it has the disadvantage of malfunctioning due to movement vector estimation errors.

For this reason, Hanwha Techwin developed a gyroscope sensor integrated DIS technology to reduce malfunctioning and improve DIS accuracy. Independently operating gyroscope sensors collect camera shake information aside from the movement vector information collected through image analysis reducing the probability of malfunctioning.
3. Conclusion

The intelligent audio and video analysis technology featured in Hanwha Techwin's network cameras notifies the operator of pre-defined situations detected through audio and video analysis. Through these technologies, they are able to not only monitor all videos 24/7, but they can also ensure operation efficiency by confirming and determining the circumstances of an event. With the use of intelligent audio and video analysis technology, a single individual can operate many video monitoring devices, reducing the amount of missed critical events and making prompt recognition and response possible in case of an accident which can result in minimized damage.