Hanwha Techwin

White Paper / Installation Guide

Wisenet AI Camera

22nd 04 2020

Contents

1. Introduction and Background

Since the development of CCTVs, humans have been responsible for simultaneous monitoring of multiple cameras and event detection from each of them. However, the number of video a human can review at the same time is limited, leading to fatal human errors and lowering the efficiency of operation.

To overcome such limitations, AI (Artificial Intelligence) technology is widely incorporated in video surveillance, offering a range of video analytics. AI-enabled Video analytics can identify persons and vehicles and alert meaningful events to operators through attributes extraction and Automatic Number Plate Detection (ANPD). This enabled effective monitoring of more number of cameras and maximized the efficiency of operation.

Hanwha Techwin is offering AI-enabled video analytics with its Wisenet AI cameras. The cameras with AI video analytics have "AI" menu in its configuration page. AI video analytics is used for various features of the cameras. The most notable features include "Object detection", events of "IVA (Intelligent Video Analytics)", "Statistics", and "Digital Autotracking." IVA feature defines virtual lines and areas for persons and vehicles to detect crossing, intrusion, enter, exit, or loitering. Statistics feature provides statistical data of observed people by adopting various conditions for People Counting, Queue management and Heatmap. Digital Autotracking can track detected movements of persons and vehicles in the video.

- Detect persons, faces, vehicles, and license plates from the video
- Extract attributes from detected persons, faces and vehicles

This document is designed to help users better understand and conveniently use our products and features while installing and operating Hanwha Techwin's AI cameras.

2. 1. Object detection

Object Detection technology classifies the types of objects (person, face, vehicle, license plate) and identifies their location in the video. Object Detection adopts deep learning algorithms to learn images of objects (person, face, vehicle, license plate) and detect similar objects within an image.
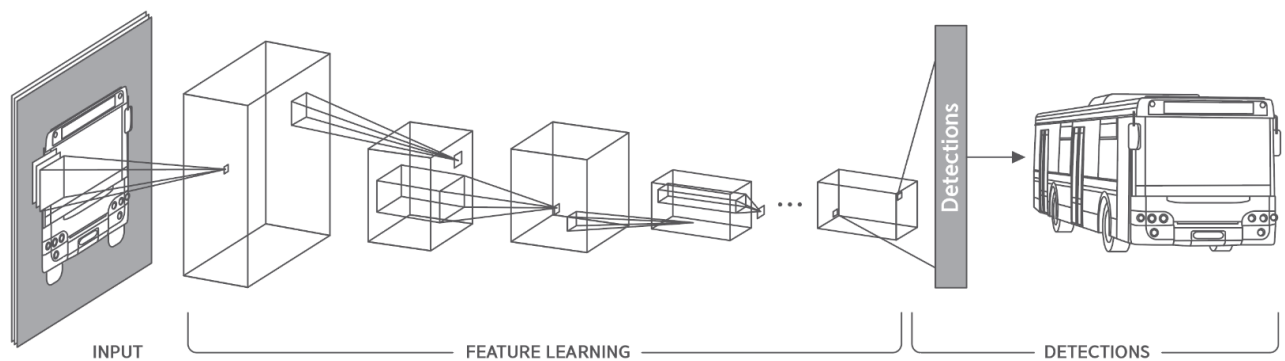

Image 1. Deep learning algorithm for detecting four types of AI objects

2. 2. Attributes extraction

Attributes extraction technology utilizes deep learning algorithms to classify detected objects (person, face, vehicle) into detailed and meaningful categories for each object type. For instance, persons can be categorized by their gender, clothing color or bags, while face attributes are by age or glasses. Vehicle attributes are classified by color or vehicle types such as trucks, buses, or bikes. Users can utilize such attributes information to increase efficiency in searching objects from large sized data in video recorders.

2. 3. BestShot capture

BestShot feature captures the best still cuts of the classified objects (persons, faces, vehicles, and license plates) which can be exported in maximum 4K resolutions. These still cuts also contain metadata of the objects' distinctive attributes, therefore, allowing operators to easily find specific objects in the video. BestShot search scenarios include facial recognition, license plate recognition, and re-identification of objects.

Users can go to "Object detection" > "BestShot" from the configuration tap and select BestShot options for persons, faces, vehicles, and license plates.
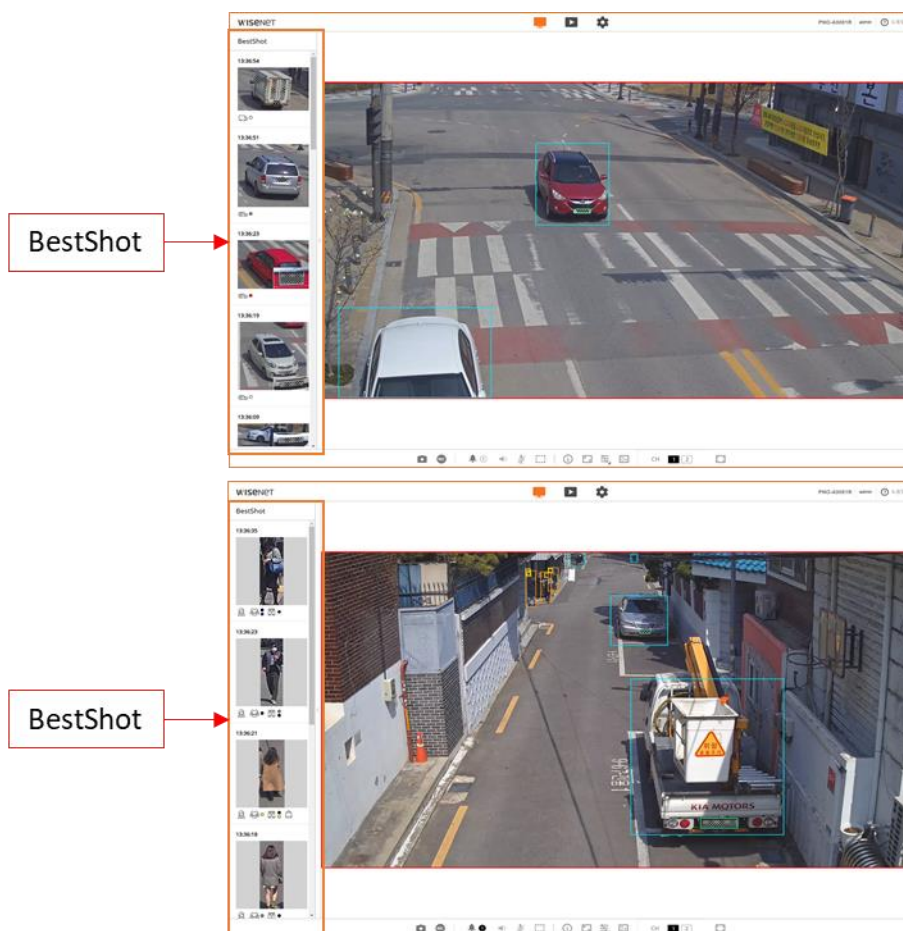

Image 2. BestShot images of AI camera

3. Camera installation guide

Hanwha Techwin's AI features can achieve the optimal performance when the cameras are installed and operated in accordance with this guide. For best results, please refer to the following instructions as the cameras' AI video analysis can be affected by lighting or camera field of view.

- This installation guide outlines minimum requirements and recommendations for using various AI features of the cameras. Performance of each feature cannot be guaranteed if the instructions are not followed properly.
- Video analysis can offer optimal performance under stable lighting conditions. Minimum 300 lux is suggested as suitable lighting level.

3. 1. Recommendations for AI camera installation

This section illustrates recommendations for installing Wisenet AI cameras to ensure reliable AI performance. If a camera is installed at a height of minimum 3 meters with a tilt of 45-degree angle, objects at least one meter away and higher than 170cm can be recognized. For reliable detection of AI cameras, it is recommended to install cameras at an angle of 45 degrees or larger, creating a side-view.
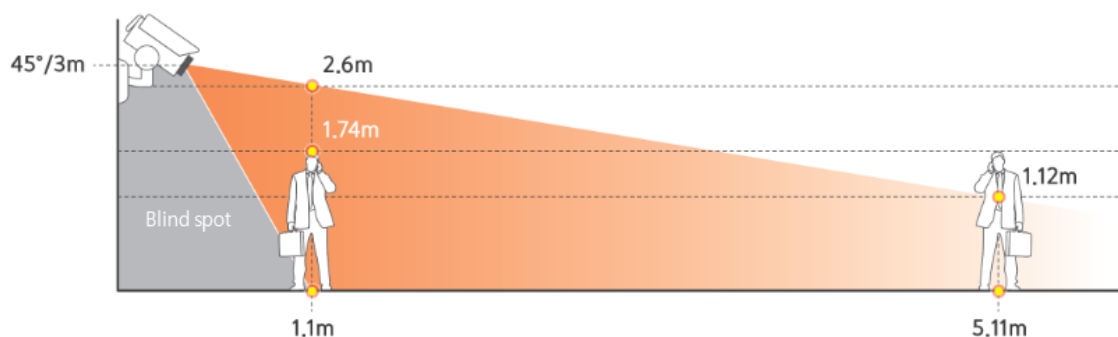

Image 3. Recommendations for AI camera installation (3m, 45 degrees)

- 4K (3840x2160) camera installed at 3m/45 degrees (lens focal length: 4.5mm)

   Resolution : 561 pixels per meter (171 pixels per foot) @ 1.10m

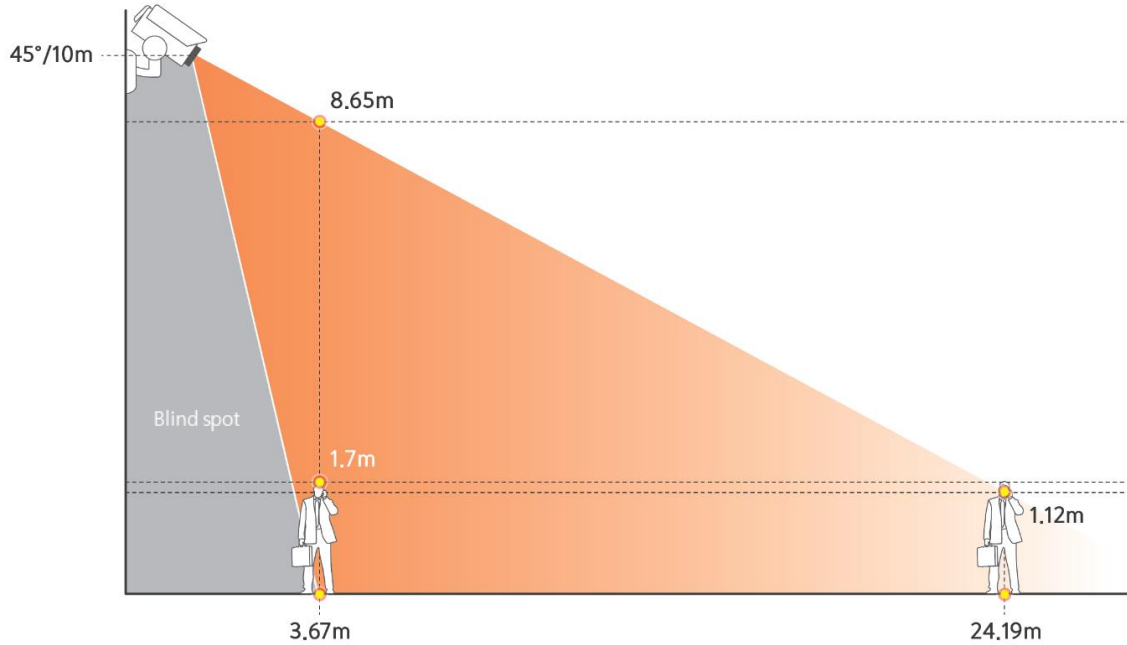   290 pixels per meter (88 pixels per foot) @ 5.11m (Over 2/3 of object)

Image 4. Recommendations for AI camera installation (10m, 45 degrees)

- 4K (3840x2160) camera installed at 10m/45 degrees (lens focal length: 4.5mm)

   Resolution: 168 pixels per meter (51 pixels per foot) @ 3.67m

   69 pixels per meter (20 pixels per foot) @ 24.19m (Over 2/3 of object)

Recomm-endations	Install height	Install angle	Max. resolution	Min. resolution
Min	Over 3 m	45° <	561 PPM @1.10m	290 PPM @5.11m
Max	Below 10m	45° <	168 PPM @3.67m	69 PPM @24.19m

Table 1. Minimal to maximum recommendations for AI camera installation

Criteria for PPM (Pixels per meter)

※ **Wisenet Toolbox: https://www.hanwha-security.com/wisenettoolbox/index.html#!/en/home**

- Pixels per meter is the number of pixels to represent one meter. With a greater number of PPM, the higher the resolution will be.

(1) 25 PPM or higher for general video surveillance: Detect an object's shape, color, rough size, or gender.

Not sufficient to recognize facial features or letters.

(2) 62 PPM or higher for object detection: Observe human faces or license plates from the video analysis module.

(3) 125 PPM or higher for object recognition: Recognize facial features or letters on a license plate.

(4) 250 PPM or higher for identifying details: Sufficient picture quality to identify details. Identify facial scars, eye colors, or tattoos.

3. 2. Recommended requirements for object detection and attributes

Wisenet AI cameras can detect persons, faces, vehicles and license plates. The number of objects that can be detected simultaneously is 256 (person, car, head, face, license plate information in total). This section illustrates recommended requirements for detecting each object type. The requirements can vary for each object. Even if all the requirements are met, the performance may vary according to different operating environments.

(1) Recommended requirements for person detection

- The minimum size supported for detecting a person in an image is 15 pixels for the shorter axis. The recommended size is 30 pixels or higher.
- Detection might not be possible if a full body (face/body/leg) is occluded more than 50% from left to right.

Image 5. (Undesirable condition) When a full body (face/body/leg) is occluded more than 50% from left to right.

- Detection might not be possible if upper or lower body is occluded more than

50%.

- The following conditions can reduce the accuracy of detection.

    ➢ Only parts of the body is captured in an image

    ➢ A person moving faster than 0.5m~1.5m/sec

    ➢ In high density crowds (e.g., standing shoulder to shoulder, large crowds)

(2) Recommended requirements for person attributes

- The minimum size supported for classifying person attributes in an image is 112 pixels for the longer axis. The recommended size is 352 pixels or higher.

- Types of person attributes

    ➢ Gender: Male, female

    ➢ Top/bottom wear color: Black, blue, gray, green, orange, red, purple, white, yellow (1~2 colors can be extracted from each wear at a time)

    ➢ Bag: Carrying, not carrying

- The following conditions can reduce the accuracy of attributes extraction.

    ➢ Only parts of the body, not full body (face/body/leg) is captured in an image

    ➢ In high density crowds (e.g., standing shoulder to shoulder, large crowds)

    ➢ Persons riding two wheelers (motorcycle/bicycle)

    ➢ Change of pose: Not standing straight (e.g., lying down, bending)

    ➢ Unidentifiable to the naked eye due to poor quality or blurry images

(3) Recommended requirements for head detection

- The minimum size supported for detecting head in an image is 12 pixels for the shorter axis. The recommended size is 20 pixels or higher.

- Detection might not be possible if the head is occluded more than 25% from left to right.

- Detection might not be possible if the face is occluded more than 15% from the eyes.

- Detection might not be possible if the face is occluded more than 50% from the mouth.

Image 6. (Undesirable condition) When a head is occluded more than 25% from left to right.

(4) Recommended requirements for face attributes

- The minimum size supported for classifying face attributes in an image is 80 pixels for the shorter axis. The recommended size is 120 pixels or higher.
- Types of face attributes
  - ➢ Gender: Male, female
  - ➢ Glasses: Wearing, not wearing
  - ➢ Age: Young (~19), adult (20~44), middle (45~64), old (65~)
- The following conditions can reduce the accuracy of attributes extraction.
  - ➢ Parts of face (eye, nose, mouth) is occluded
  - ➢ Change of pose: Non-frontal faces (e.g., profile view, keeping the head down)
  - ➢ Unidentifiable to the naked eye due to poor quality and blurry images or nighttime conditions

(5) Recommended requirements for vehicle detection

- If a vehicle is front-facing in an image, the minimum size supported for detection

is 15 pixels for the shorter axis. The recommended size is 40 pixels or higher.

- If a vehicle is side-facing in an image, the minimum size supported for detection is 15 pixels for the shorter axis. The recommended size is 35 pixels or higher.

- Detection might not be possible if the front of a vehicle is occluded more than 75% from left to right.
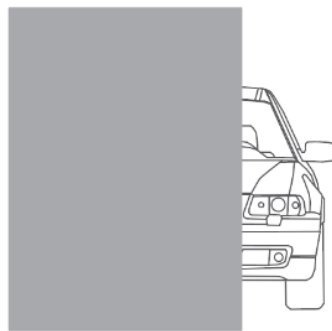

Image 7. (Undesirable condition) When the front of a vehicle is occluded more than 75% from left to right

- Detection might not be possible if the front of a vehicle is occluded more than 50% from top to bottom.

- Detection might not be possible if the side of a vehicle is occluded more than 25% from left to right.

- Detection might not be possible if the side of a vehicle is occluded more than 50% from top to bottom.
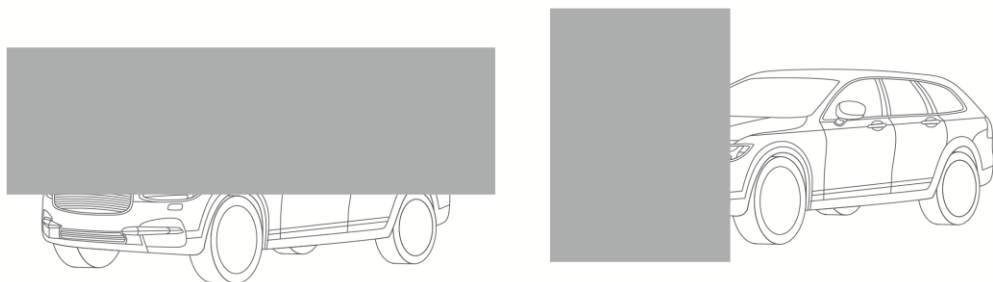

Image 8. (Undesirable condition) When the side of a vehicle is occluded more than 50% from top to bottom (left), 25% from left to right (right)

- Detection might not be possible if a vehicle is overturned.
- The following conditions can reduce the accuracy of object detection.
  - ➤ Only parts of a vehicle is captured in an image
  - ➤ If a camera is installed within 10m high at an angle of 45 degrees or higher or a vehicle is moving faster than 80Km/h
  - ➤ If a camera is installed at an angle of less than 30 degrees or a vehicle is moving faster than 40Km/h
  - ➤ In high density crowds (e.g., traffic congestion, waiting at red light)

(6) Recommended requirements for vehicle attributes

- The minimum size supported for classifying vehicle attributes in an image is 50 pixels for the shorter axis for two-wheelers and 80 pixels for four-wheelers. The recommend size is 100 pixels or higher.
- Types of vehicle attributes
  - ➤ Vehicle type: Car, bus, truck, bicycle, motorcycle
  - ➤ Vehicle color: Black, blue, gray, green, orange, red, purple, white, yellow (1~2 colors can be extracted from each vehicle)
- The following conditions can reduce the accuracy of attributes extraction.
  - ➤ If a vehicle is occluded more than 50% or on the edges of the screen
  - ➤ If vehicles are overlapped by others
  - ➤ Vehicle color unidentifiable due to reflection
  - ➤ Unidentifiable to the naked eye due to poor quality and blurry images or nighttime conditions

(7) Recommended requirements for license plate detection

- If a license plate is front-facing in an image, the minimum size supported for detection is 12 pixels for the shorter axis. The recommended size is 15 pixels or higher.

- Detection might not be possible if a license plate is not facing the front in the image.

- Detection might not be possible if a license plate is occluded more than 25% from left to right.
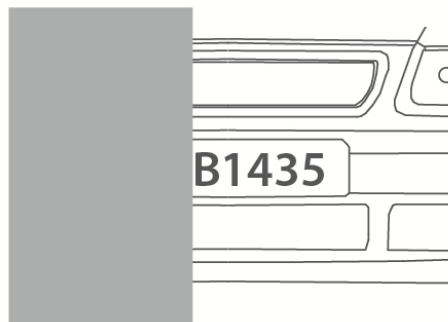

Image 9. (Undesirable condition) When a license plate is occluded more than 25% from left to right.

Object	Min. pixel for detection (Shorter axis) (@4K(3840x2160))	Recommended speed for detection (@ Within 100m)
Person	30 pixels or above	0.5m ~ 1.5m/sec (Cannot guarantee accuracy for objects walking faster than the normal walking speed)
Head	20 pixels or above	-
Vehicle	(Front) 40 pixels or above (Side) 35 pixels or above	Below 80km/h (Installed at 45° or larger) Below 40km/h (Installed at 30° or smaller)
License plate	15 pixels or higher	-

Table 2. Recommended requirements for AI object detection

Object	Min. pixel for attributes extraction (@4K (3840x2160))	Recommended pixels for attributes extraction (@4K (3840x2160))
Person	112 pixels or above (Longer axis)	352 pixels or above (Longer axis)
Face	Over 80 pixels or above (Shorter axis)	120 pixels or above (Shorter axis)
Vehicle	50 pixels or above (Shorter axis/two-wheeled) 80 pixels or above (Shorter axis/four-wheeled)	100 pixels or above (Shorter axis)

Table 3. Recommended requirements for AI attributes extraction

3. 3. Recommended requirements for BestShot

Wisenet AI cameras can detect best images (BestShot) for detected persons, faces, vehicles, and license plates. Users can go to "Object detection" > "BestShot" from the configuration tap and select BestShot options for person, face, vehicle, and license plate.

This section illustrates recommended requirements for detecting BestShot of each object type. The requirements can vary for each object to detect. Even if all the requirements are met, the performance may vary according to different operating environments. As the feature is enabled by object detection, its performance can be impacted if the performance of person/face/vehicle/license plate detection degrades.

(1) Recommended requirements for person BestShot
- The minimum size supported for a person BestShot in an image is 112 pixels in height and detection requires an aspect (width to height) ratio between 1.5 and 4.5. Detection is possible for objects existing in an image for more than one second.
- The following conditions can reduce the accuracy of person BestShot.
  ➤ Only parts of the body, not full body (Face/body/leg) is captured in an image
  ➤ In high density crowds (e.g., standing shoulder to shoulder, large crowds, or people standing close together)
  ➤ Fast movement of an object
  ➤ Change of pose: Not standing straight (e.g., sitting down, lying down, or bending)
  ➤ Unidentifiable to the naked eye due to poor quality or blurry images

(2) Recommended requirements for face BestShot

- The minimum size supported for face BestShot in an image is 80 x 80 pixels, while detection requires an aspect ratio of 1.5 or above. Detection is possible for objects existing in an image for more than one second.

  - The following conditions can fail or reduce the accuracy of face BestShot.

    ➢ Only parts of a face is captured in an image, not the entire face (including all of eyes, nose, mouth).

    ➢ Multiple faces are overlapped in high density crowds

    ➢ Fast movement of an object

    ➢ Unidentifiable to the naked eye due to poor quality or blurry images

(3) Recommended requirements for vehicle BestShot

- The minimum size supported for vehicle BestShot in an image is 80 x 80 pixels for four-wheelers and 50 x 50 pixels for two-wheelers. Detection is possible for objects existing in an image for more than 0.3 seconds.

  - The following conditions can fail or reduce the accuracy of vehicle BestShot.

    ➢ Only parts of a vehicle is captured in an image

    ➢ Vehicles are overlapped in high density crowds

    ➢ Fast movement of an object

    ➢ Unidentifiable to the naked eye due to poor quality or blurry images

(4) Recommended requirements for license plate BestShot

- The minimum size supported for license plate BestShot in an image is 80 pixels wide and 30 pixels high. Detection is possible for objects existing in an image for more than 0.3 seconds.

  - The following conditions can fail or reduce the accuracy of license plate BestShot.

    ➢ Only parts of a license plate is captured in the image

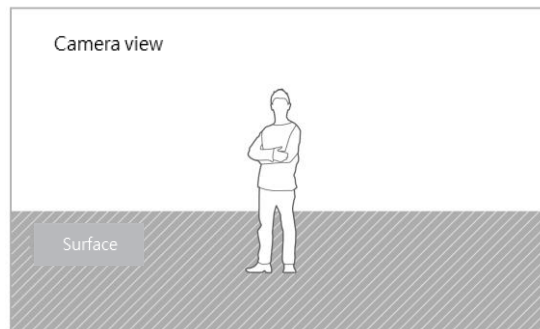   ➢ A license plate is occluded or overlapped in high density crowds

Image 10. Recommended Field of View

3. 6. Precautions for camera installation

False or failed detection could result if cameras are not properly installed according to recommended requirements. The following conditions can reduce the cameras' performance.

1. Camera installation angle at less than 30 degrees (Direct under)
2. Challenging lighting conditions with shadows or low-light environments
3. Scenes with strong background lighting such as direct sunlight, backlighting, or direct lighting
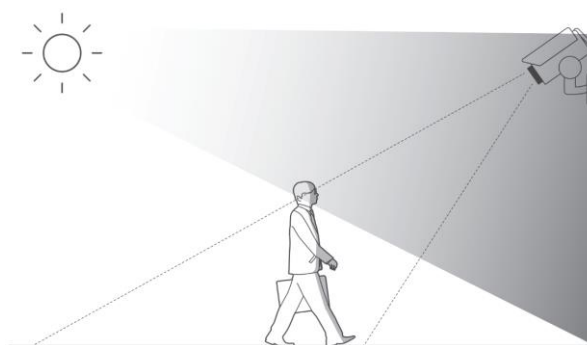

Image 11. (Undesirable condition) When cameras get affected by direct sunlight and light, or backlighting

4. Motion blur caused by slow shutter speed or low frame rate
5. Increased Dynamic Range[1]

   Dynamic range can be reduced with Wide Dynamic Range (WDR).

---

1 Dynamic Range: The difference between the darkest and the lightest tones in an image

6. Objects have similar illumination or color as the background.

7. Random objects occlude different objects

Image 12. (Undesirable condition) When random objects occlude different objects

8. Several objects moving together

9. Objects reflected in reflectors such as mirrors

Image 13. (Undesirable condition) When objects are reflected in reflectors

10. Objects moving fast

11. To detect fast moving objects, it is recommended to increase the field of view.

12. Designating object detection size as smaller than its actual size

13. Designating object detection size as bigger than its actual size

14. The ratio of the object's size to the whole image is too large.

15. Object occlusion or partial objects can still be detected in Object detection, but can cause false positives as the detection is enabled only by attributes from detected areas.

4. Conclusion

Hanwha Techwin's Wisenet AI cameras provide top notch video analysis technologies through AI. By accurately detecting and classifying attributes of meaningful objects (person, vehicle, face, license plate, etc) in video surveillance, users can utilize the information in various applications.

Hanwha Techwin offers a range of video analysis that can take full advantage of camera devices (systems). All video analysis features adopt intuitive Webviewer of cameras for easy configuration and operation. The features are expected to help users build efficient video surveillance systems and engage in systematic information gathering and analysis.

WISENET

Hanwha
Techwin