White papers

Similarity Search

June 2026

Contents

1. Introduction

This white paper focuses on Hanwha Vision's AI-based "*Similarity Search*" technology, designed to overcome the operational limitations caused by the explosive growth of video surveillance data and to significantly reduce the time required to investigate hours of recorded footage. Moving beyond conventional attribute-based filtering, this technology leverages high-dimensional similarity features extracted directly by edge AI to compare visual likeness. It analyzes visual similarity under varying lighting conditions and camera perspectives, enabling continuous search across wide-area surveillance environments.

Within this document, readers will gain comprehensive insights into the engineering principles of complex vector embedding, practical integration methods with global VMS platforms (such as Genetec and Milestone) and BLAZE, and strategic configuration guides for maximizing search accuracy while maintaining strict compliance with data privacy regulations.

Hanwha Vision's AI-powered *Similarity Search* technology was designed and developed under the framework of **ISO/IEC 42001**:2023, the world's first international standard for AI Management Systems (Certification No: AIMS-0012). This certification serves as a testament to Hanwha Vision's commitment to AI ethics and data management, ensuring that all core processes strictly adhere to established global standards.

Ultimately, this white paper provides a strategic roadmap for operators of smart cities and large-scale industrial facilities to minimize human resource expenditure and establish immediate response systems within the "golden time," supporting improvements in operational efficiency and a reduction in Total Cost of Ownership (TCO).[1]

[1] ※ This system does not perform identity recognition or identity verification of specific individuals. It functions as a decision-support tool that returns similar results based on visual similarity scores, and the final judgment remains the responsibility of the user. The generated feature vectors are stored as numerical embeddings and cannot be reverse-engineered into original images, nor are they used to uniquely identify any individual.

2. Background

Modern video surveillance infrastructure is expanding rapidly across diverse sectors, including urban safety, smart buildings, and large-scale industrial facilities. As thousands of cameras record high-resolution footage 24/7, the volume of data to be managed has already surpassed human cognitive limits, leading to a significant decline in monitoring efficiency. In particular, the traditional manual process—where security personnel must visually inspect numerous channels of recorded footage to track an person—is not only labor-intensive and time-consuming but also creates a critical bottleneck when an immediate response is required.

To overcome these limitations, the industry has continuously attempted to extract the visual attributes of people as metadata. However, simple classifications such as clothing color or gender may not be sufficient to reliably identify people in complex environments. This necessitates technology where Edge-based AI detects people in real-time and extracts "Visual Attributes"—including full-body appearance and facial characteristics—converting them into data. This approach clarifies that the technology is focused not on "identity recognition (Identification)" of a specific individual, but on "visual similarity comparison and search" against a user-selected target within vast data, in order to maximize efficiency.

Ultimately, the core competitiveness of next-generation Video Management Systems (VMS) lies in how efficiently they can identify targets using vast amounts of collected metadata. The integration of AI-based feature extraction and similarity search technology enables the comparison of visual similarities and precise visualization of movement paths over time, regardless of varying camera angles or lighting conditions.

These technical advancements provide substantial business benefits to customers. Security operators can improve operational efficiency by significantly reducing the time required for people tracking and path analysis. Furthermore, by identifying accurate information within the "golden time" for rapid response, organizations can maximize their incident prevention and post-response capabilities. This enables the construction of an innovative security environment that optimizes both human and material resources, ultimately reducing the Total Cost of Ownership (TCO).

3. AI-Based Feature Extraction and Matching Technology for Similarity Search

3.1. Definition and Process of Similarity Search

Similarity search is an intelligent retrieval technology designed to precisely identify people within a vast video database that most closely resemble a specific target. This technology operates through a structured framework to efficiently search for matching or similar candidates within large-scale datasets as follows:

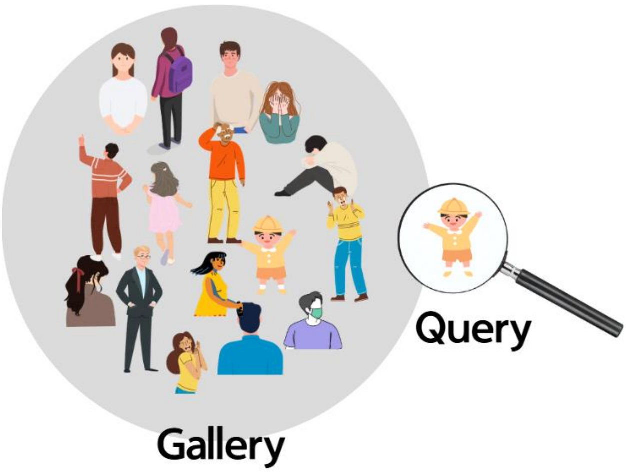


Figure 1. Concept of Gallery and Query

- **Gallery:** A comprehensive collection of person images and feature points collected from multiple cameras and stored in a database. It serves as the target pool for finding matches against a query image.
- **Query:** The input image or the specific visual feature points of a person that a user intends to identify or match within the database.

By ranking candidates with similar physical traits — such as body type or facial features — based on similarity scores, this technology supports efficient tracking of movement paths in wide-area security environments. This technology is intended for visual similarity comparison, not identity verification, and search results should be used for reference purposes only. The final determination of whether a search result matches the target individual rests solely with the operator.

3.2. From Pixel Matching to 'Vector embedding'

Traditional ‘Raw pixel matching’ requires a direct comparison of vast amounts of pixel data within a video, leading to extremely high computational costs and significantly degraded search performance. In large-scale dataset environments, this approach becomes a critical factor that severely hinders the efficiency of real-time searches.

To address this challenge, Hanwha Vision employs advanced Vector Embedding technology. Edge AI cameras use deep learning models to abstract visual information—such as color, texture, and shape—into a Feature Vector, which is a high-dimensional numerical array that cannot be directly interpreted by humans, thereby enhancing data management transparency. Key feature points within an image are hierarchically analyzed and essential visual characteristics are extracted. The extracted information is ultimately structured as a “Similarity Feature” data format that compresses and organizes only the most critical information.

Figure 2. Example of ‘Feature Vector’

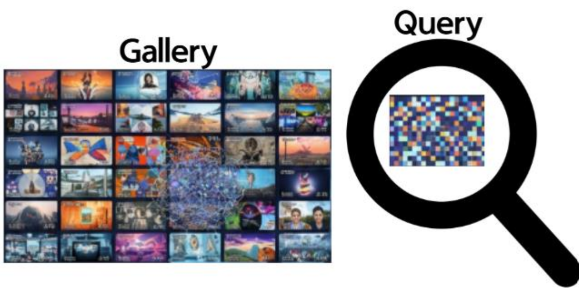
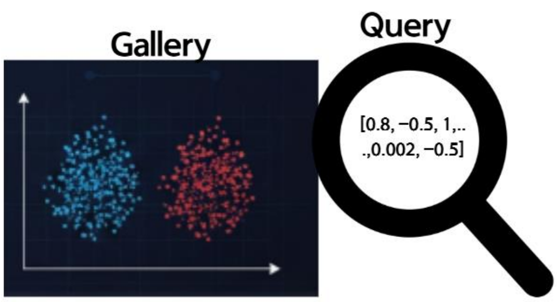
Image Comparison	Vector Comparison
	

Table 1. Differences between Image Comparison and Vector Similarity Comparison Methods

3.3. Vector Similarity Calculation Algorithm

Hanwha Vision's *Similarity Search* system consists of Edge AI cameras (which provide Feature Vectors for similarity search) and a VMS (providing Similarity Matching and search results). This feature can only be utilized when both the camera and the VMS support similarity search.

Figure 3. Similarity Search System Architecture

The Hanwha Vision's *Similarity Search* system is built on collaboration between Edge AI cameras and the VMS (Cloud/Server). Feature Vectors extracted directly at the camera level are transmitted to the VMS, which then calculates a Similarity Score by measuring the mathematical distance between people.

To determine these similarity scores, two primary methods are commonly used: measuring the straight-line distance between two points in a vector space to identify absolute differences, or comparing the directional alignment of two vectors to identify visual similarities.

As illustrated in Figure 4, the Feature Vectors extracted by Hanwha Vision's AI cameras using deep learning at the Edge are generated as numerical array values. This vector data is transmitted to the VMS along with the BestShot image, where it undergoes comparative computation on the server side to perform efficient searches.

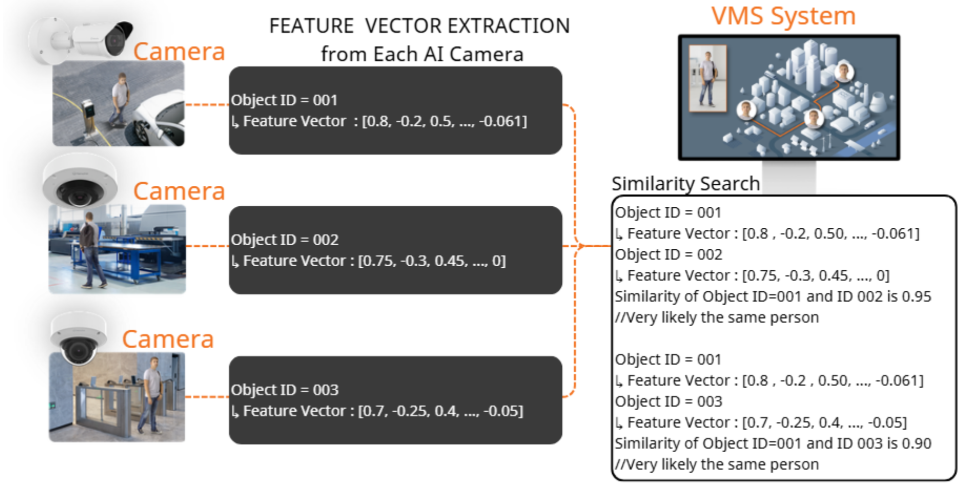


Figure 4. Example of a Feature Vector

4. System Architecture

The *Similarity Search* feature, introduced with Hanwha Vision's 2nd generation X and P series cameras released in 2025, is enabled by the cameras directly providing feature vectors dedicated to similarity matching. AI cameras supporting this feature extract feature values from an optimal frame (BestShot)—which is selected based on stringent criteria including sharpness, confidence, and occlusion (visibility > 70%). Through this process, the system calculates feature values for the same person.

The BestShots provided as metadata by the cameras include the Person BestShot and the Face BestShot. For each valid person and face detected, the camera transmits the corresponding feature vector values under the respective Object category directly to the VMS.

4.1. Data Flow for Similarity Search

The data flow consists of three stages: Generation, Transmission & Storage, and Comparison. Generation occurs at the camera edge, where similarity feature values for people are extracted from live video in real-time. Transmission & Storage refers to sending the extracted vector data and BestShot images to the VMS and storing them in the Gallery. Comparison is the stage where the data stored in the VMS Gallery is compared against the user's Query, similarity scores are calculated, and similar footage is displayed

4.2. Camera Configuration for Similarity Search

To utilize the similarity search function, users must first review the camera's datasheet to verify whether it supports the extraction of feature vectors. If the "Analytics" section of the datasheet indicates "*Similarity Search*: Support," the camera is capable of providing the vector data necessary for this feature.

However, hardware support alone is not sufficient; the BestShot feature must be actively enabled within the edge-based WiseAI App, and selecting the "Person" object is mandatory. Users intending to utilize combined facial data for enhanced matching must also select the "Face" object.

Since the performance of similarity search relies heavily on the quality of the BestShot, optimizing the AI's object detection accuracy will yield significantly more precise search results. These AI accuracy parameters can be fine-tuned in the Setup tab of the WiseAI App.

To enhance operational efficiency in large-scale environments, Person and Face BestShot settings can be remotely enabled and controlled directly through the VMS, eliminating the need to access the camera's web viewer. This allows for the seamless configuration of camera environments optimized for similarity searches.

4.3. Similarity Search in VMS

Because not all Video Management Systems (VMS) support similarity search natively, it is essential to consult the specific VMS datasheet to verify compatibility. Recently, major VMS and cloud service providers have been actively integrating new functionalities to support similarity searches based on the feature vectors extracted from cameras.

For major global Video Management Systems (VMS) such as Genetec and Milestone, feature vector-based similarity search capabilities can be implemented utilizing a dedicated plug-in provided by Hanwha Vision. Once a user designates a target subject from the available BestShots, the VMS receives the BestShot metadata and vector values transmitted from the camera to execute comparative computations. The scope of search may vary depending on the computational load and the system's operational policies.

Conversely, Hanwha Vision's in-house designed solution, BLAZE, natively integrates this functionality at the system level (built-in), offering distinct competitive advantages in terms of search flexibility and operational scope. BLAZE automatically performs searches across all cameras accessible to the user, and if required, can extend the search coverage to include the entire data retention period.

When the VMS aggregates BestShots and feature vectors from the cameras into a Gallery, users can select a specific Query image. The VMS then performs a similarity search by mathematically comparing the feature vectors of faces and persons in the Gallery against those of the Query. During this process, users can adjust the sensitivity threshold. If the calculated similarity exceeds the user-defined threshold, the system classifies the result as a similar candidate and displays it. A lower threshold means that matches will be returned even with a lower degree of vector similarity.

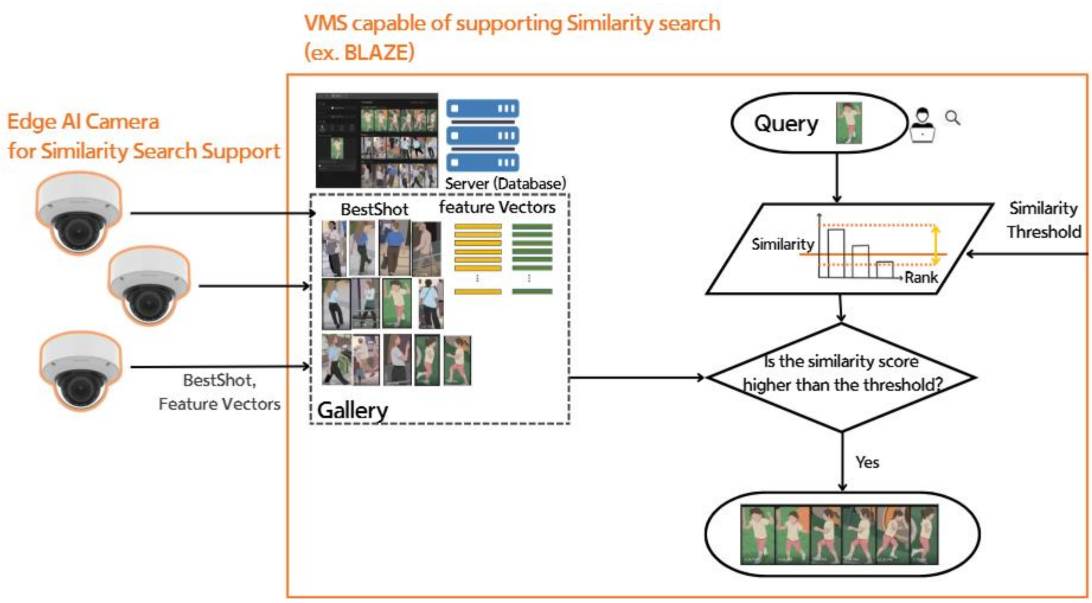


Figure 5. Sensitivity Setting During Similarity Search

To initiate a similarity search, the feature must first be enabled for compatible cameras within the VMS. When configuring a supported camera, users must navigate to the Objects tab and explicitly enable the option to store feature vectors in the VMS database for similarity search utilization.

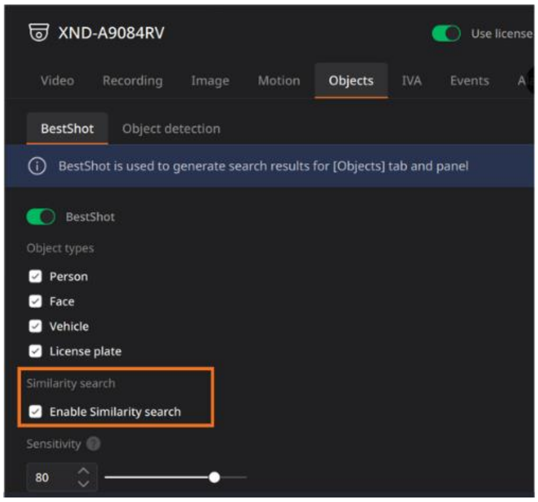


Figure 6. Enabling Similarity Search in the VMS

As previously detailed, similarity search utilizes two distinct types of feature vectors: full-body (Person) and facial (Face). To improve search performance by activating facial similarity feature values, users must navigate to the General tab and explicitly consent to the use of biometric information. If consent is not given, the camera will not transmit biometric information to the VMS.

Even if a feature vector is not generated, the camera will still transmit the BestShot to the VMS. However, when a BestShot is accompanied by similarity vector data, a distinct "intersection" icon is displayed in the bottom right corner of the BestShot image. To search for that specific person, the user simply clicks the BestShot.

Clicking the BestShot brings up "SIMILAR TO" in the lower right section of the interface, allowing the user to select the precise Query image to be used for the similarity search.

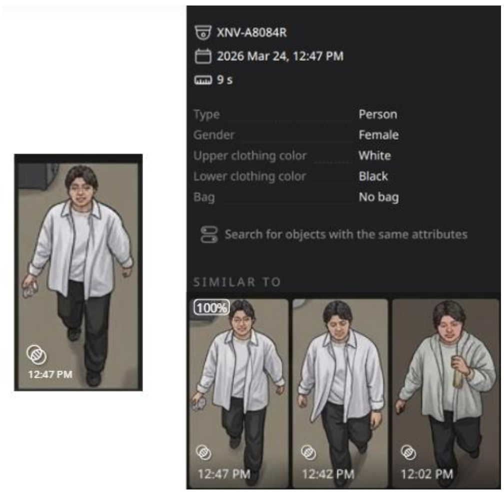


Figure 7. Example of a BestShot Containing a Feature Vector

Once a BestShot containing feature vector values is selected as the search query, the scope of the search can be controlled by adjusting the similarity sensitivity, as illustrated in the figure below. Configuring the sensitivity to ‘Low’ broadens the results to include people with relatively lower similarity scores. Conversely, setting the sensitivity to ‘High’ prioritizes results with high similarity scores. The similarity results are sorted by similarity score, and the time information is displayed alongside them.

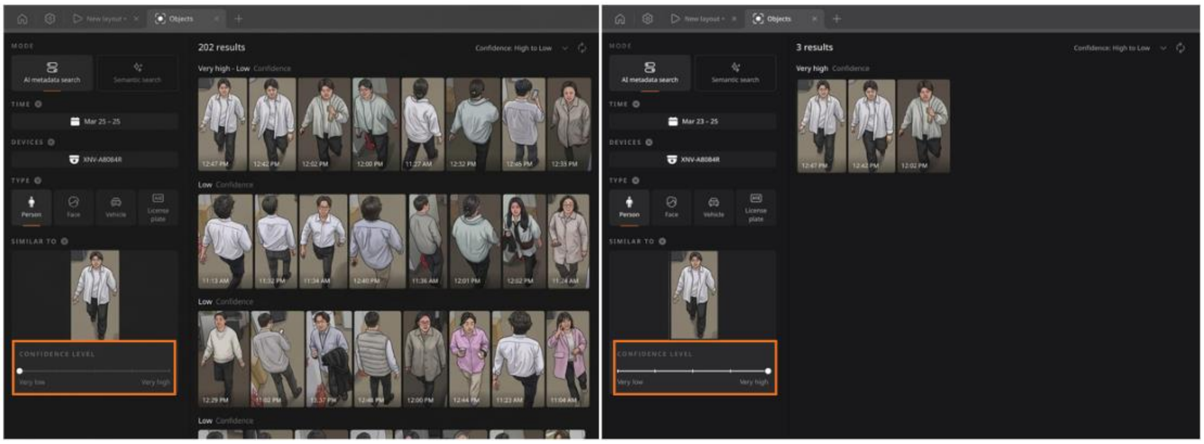


Figure 8. Example of Similarity Sensitivity Adjustment[2]

Through this advanced search system, operators can search for and reference visually similar individuals across multiple cameras, varying times, and diverse locations. By seamlessly integrating with Hanwha Vision’s VMS and cloud solutions, this technology enables more efficient search and monitoring even within large-scale surveillance environments.

2 ※ The UI screens and internal text shown in Figures 6–8 are subject to change depending on product updates.

4.4. Regulations & System Configurations to Consider Utilizing Biometric Data

The features of this system may be classified as a "Remote Biometric Identification System" under the EU AI Act and thus categorized as a High-risk AI System. In response, Hanwha Vision, as the Provider, has designed the system in compliance with the following strict obligations and management frameworks to proactively address regulatory requirements and mitigate legal risks.

- **Risk Management System and Quality Management System:** To identify and manage risks throughout the system lifecycle, Hanwha Vision has fully implemented and strictly adheres to the ISO/IEC 42001 (AI Management System) certification[3] framework—a globally recognized standard. The certified scope specifically includes Face Recognition and Re-ID, guaranteeing that biometric information and object attribute data are processed with transparency in accordance with international standards.
- **Forensic Search Design:** This system is explicitly designed for "post-event search" purposes—comparing similarities within recorded data after an incident—not for real-time control or real-time identification, thereby fundamentally eliminating the risk of prohibited AI practices under the EU AI Act.
- **Data Governance:** The representativeness of training datasets is secured and managed transparently to prevent bias.
- **Technical Documentation and Log Management:** Detailed technical documentation, including system design logic and algorithm specifications, is maintained, and all events (logs) are automatically recorded to ensure transparency.
- **Transparency and Human-in-the-loop:** Detailed guidelines are provided to help Deployers (customers) accurately interpret outputs, and all final determinations and decisions are restricted to be made by the operator (human).

The ultimate responsibility for compliance with personal data protection regulations — including the EU GDPR (General Data Protection Regulation) and the US BIPA (Biometric Information Privacy Act) — lies with the end-user operating the system. It is strongly recommended to obtain consent or a valid legal basis from data subjects for the collection, storage, and use of personally identifiable information such as facial data, strictly limit the purpose of processing (restricted to security and safety management), and set data retention periods to prevent unnecessary personal data collection in accordance with the principle of data minimization. Furthermore, VMS system administrators must acknowledge this end-user responsibility and explicitly consent to the use of biometric information before activating it to improve similarity search accuracy.[4]

---

[3] The scope of certification includes Face Recognition and Re-Identification (Re-ID), ensuring that biometric information and object attribute data used for similarity search are processed transparently in accordance with international security and management standards.

[4] The ultimate responsibility for compliance with applicable regulations, including the EU GDPR, lies with the end user operating the system.

Even without utilizing biometric data, Hanwha Vision's *'Similarity Search'* based full-body (Person) remains fully functional.

However, if the BestShot feature is disabled on the camera, the required metadata will not be transmitted to the VMS. Similarly, if the *Similarity Search* option is disabled in the VMS, the functionality becomes entirely unavailable. Note that similarity search features can only be enabled from the VMS by an administrator.

5. Application Guide

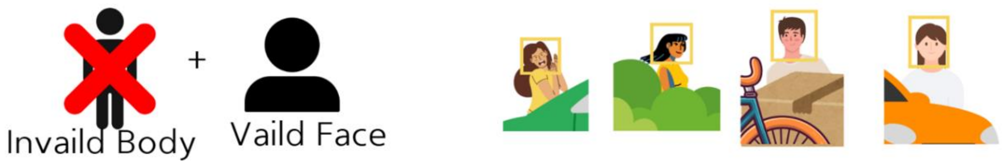
5.1. Constraints of the Similarity Search System

The performance of similarity search is most heavily influenced by the presence of a "Valid Body" and a "Valid Face." The following two search modes are available depending on which data is valid:

- **Valid Full-body Based Similarity Search:** The full-body appearance characteristics such as clothing color, body type, and shape are vectorized to search for similar subjects. Even when frontal facial data is unavailable due to a side or rear view, similarity search can proceed smoothly based on full-body information.
- **Valid Face-based Similarity Search:** Facial visual feature points are extracted and matched, focusing on deep learning-based visual similarity comparison rather than strict biometric measurement. Even if full-body information is insufficient due to a partial view or occlusion, similarity search can be performed based on facial data alone if valid frontal facial data is available.

Optimal capture of both the face and full body maximizes the precision of similarity search algorithms. However, environmental factors that prevent the camera from properly detecting the face or full body can impose constraints on the system. When both full-body and facial data are properly captured, conducting a similarity search based on this combined data yields the highest quality search results.

As such, if valid full-body information is available, similarity search can still be successfully executed based solely on the person's body—even if frontal facial data is missing (e.g., side profile or rear view). Alternatively, similarity search remains functional even if the user chooses not to utilize biometric information within the VMS.

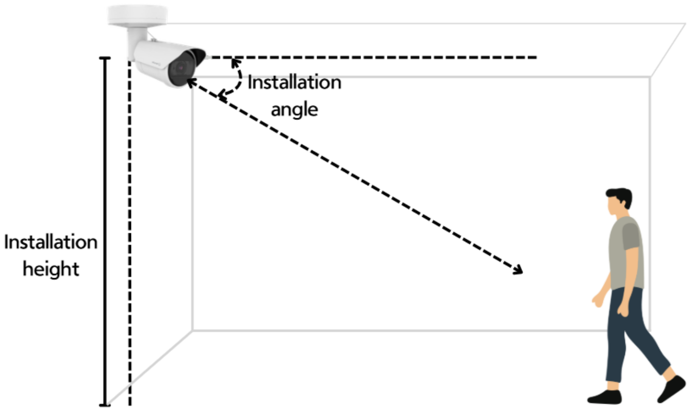


Conversely, if the full-body information is insufficient, searches based on body data will not function properly. This occurs when only the upper body is visible, the subject is seated, or the body is significantly obscured by objects. However, if sufficient frontal facial data is available, the similarity search can proceed based entirely on the face. Note that if the user has not consented to the use of biometric information, no facial data will be processed, and the individual cannot be searched using this method.

If neither the full-body data nor the facial data is sufficient, the similarity search will not operate. Similarly, if the camera fails to provide a BestShot, the search functionality will be unavailable.

5.2. Installation and Environmental Guidelines for Optimal Performance

To guarantee stable similarity search performance, the recommended minimum installation criteria require a camera height between 2m and 10m, with an installation angle of 45° or less from the ceiling. For reliable detection, a side-view installation angle is recommended. Additionally, strong light sources should be avoided from pointing into the camera's lens, as they can cause the image to wash out or blur even after focusing.



	Requirement
Installation Height	2m to 10m
Installation Angle	45° or less
Installation Background	Simple background recommended. Setting an 'exclude area' is recommended for zones with constant loitering.

Table 2: Recommended AI Camera Installation Conditions

To ensure that data input from multiple cameras yields consistent vector values, the installation conditions across the site should be as uniform as possible. Furthermore, since feature vectors are extracted from the BestShot, it is crucial to adjust the Setup settings in the WiseAI App according to the specific scene to maximize the AI's object recognition rate.

Regarding 'Person' objects, a BestShot is captured exclusively when the full body is detected. Furthermore, the full-body BestShot must exceed a resolution of 43x130 pixels to successfully extract the feature vector required for similarity searches. For facial data, the minimum resolution necessary to extract a feature vector is 60x60 pixels. Optimal search performance is achieved when both a frontal face BestShot and a full-body BestShot are concurrently available.

Person-based feature vectors (43x130 pixels) reflect not only similar physical traits (such as face, body type, hair, and gender) but also similar clothing. Therefore, information other than the face significantly impacts the results. Because the system might determine that objects wearing different clothing are not the same person, operators must actively utilize the “Confidence Level” adjustment.

5.3. Similarity Search Performance Limitations

Similarity search outcomes may vary depending on the camera's performance, resolution, installation angle, object detection accuracy, and the specific VMS utilized. Additionally, the searchable time frame may be limited by the VMS configuration.

Beyond these factors, performance constraints can arise from the environment. If the viewpoints of the capturing cameras are drastically different, the system may fail to extract consistent feature vectors. Other limitations include scenarios where a significant portion of the object is occluded, the AI misidentifies two or more people as a single person, or the object is not clearly recognized due to poor lighting or camera limitations.

While similarity searches can be executed relying solely on facial feature vectors—without requiring full-body feature vectors—the inherent characteristics of the algorithm dictate that candidates may be omitted from the search results if their face-based cosine similarity score falls below a predefined threshold. This mechanism is strictly implemented to mitigate false positives.

6. Conclusion

Hanwha Vision's *Similarity Search* is an advanced video search solution delivered through the seamless integration of its high-performance AI cameras with its native BLAZE VMS, or via dedicated plug-in for leading global VMS platforms such as Genetec and Milestone. Designed under the ISO/IEC 42001 certification framework—a global standard—it provides an intelligent video surveillance environment that supports regulatory compliance.

By utilizing feature vectors extracted directly from the BestShot—a core strength of Hanwha Vision's edge AI cameras—similarity feature values are extracted directly from the BestShot and distributed for transmission, efficiently optimizing server computational and data loads, enabling efficient search and monitoring with minimal latency even in large-scale enterprise environments.

Ultimately, this technology significantly improves upon the limitations of conventional metadata or natural language-based search methods currently prevalent in the security sector. By analyzing object similarities across multiple cameras, it enables operators to respond swiftly during integrated monitoring within the golden time, presenting a new standard for next-generation intelligent video surveillance.

7. Appendix

- **Applicable Cameras**

  All cameras indicating "*Similarity Search*: Support" within the Analytics section of their datasheet.

- **Applicable VMS**

  Genetec Plug-in (ver 1.08 or later)

  Milestone Plug-in (ver 1.09 or later)

  BLAZE (ver 1.2 or later)

Hanwha Vision Co.,Ltd
13488 Hanwha Vision R&D Center,
6 Pangyo-ro 319-gil, Bundang-gu, Seongnam-si, Gyeonggi-do
TEL 070.7247.8771 **FAX** 031.8018.3715
www.HanwhaVision.com