White Paper

Similarity Search

May 2026

Contents

1. Introduction

This white paper highlights Hanwha Vision's **AI-based Similarity Search** technology, a breakthrough that overcomes the limitations of modern video surveillance by transforming hours of manual footage review into seconds of high-precision data searching. Going beyond conventional attribute matching, this technology leverages high-dimensional Feature Vectors extracted directly by Edge AI. This ensures seamless person re-identification and persistent tracking across expansive security networks, even under extreme conditions with fluctuating lighting and diverse camera angles.

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," achieving a quantum leap in operational efficiency and a significant reduction in Total Cost of Ownership (TCO).

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 person as metadata. However, simple classifications such as clothing color or gender are insufficient to reliably identify people in complex environments. This necessitates technology where Edge-based AI detects people in real-time and converts them into metadata, capturing not only visible appearance but also unique Biometric Information. The unique identification values generated through this process serve as the vital link that organically connects fragmented video data, forming the foundation for maintaining people identity even within wide-area security systems.

Ultimately, the core competitiveness of next-generation Video Management Systems (VMS) lies in how quickly and accurately they can identify targets using vast amounts of collected metadata. The integration of AI-based feature extraction and Similarity Search technology enables the determination of object identity 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 achieve overwhelming operational efficiency by completing people tracking and path analysis—which previously took hours—in just a few seconds. 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. Core Technology: AI-Based Similarity Search and Vector Embedding

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:

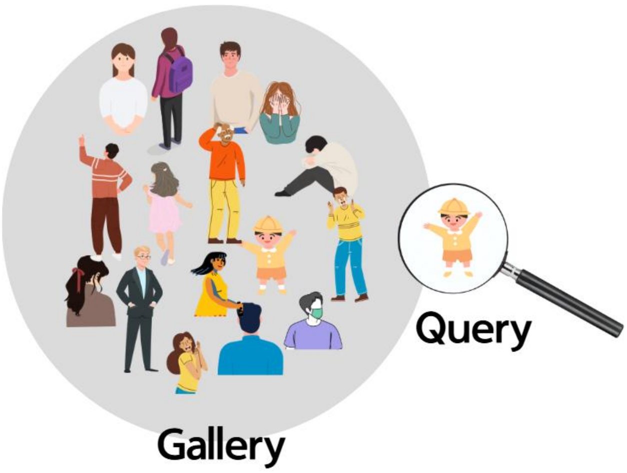


Figure 1. Concept of Gallery and Query

- **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.

Beyond simply locating the same individual, this technology ranks candidates based on the similarity of physical traits—such as body type or facial features. This allows for the rapid reconstruction of a person's movement path across wide-area security environments in the shortest possible time.

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. This process utilizes deep learning models to transform complex image data into meaningful numerical arrays. Edge AI cameras hierarchically analyze key feature points within an image to precisely extract essential visual characteristics, including biometric information, color, texture, and shape. The extracted information is ultimately represented as a Feature Vector—a numerical data format that compresses and organizes only the most critical information.

Figure 2. Example of ‘Feature vector’

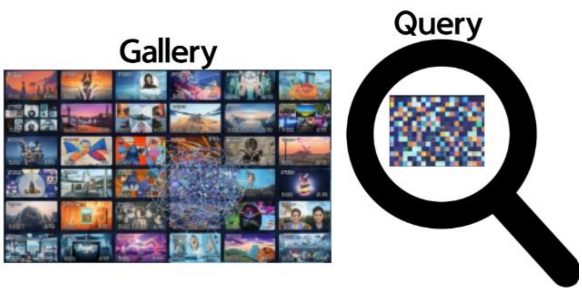
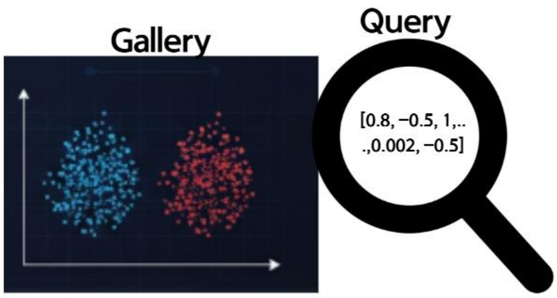
<table>
<tr><th>Image Comparison</th><th>Vector Comparison</th></tr>
<tr><td>
</td><td>
</td></tr>
</table>

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 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 high-speed searches.

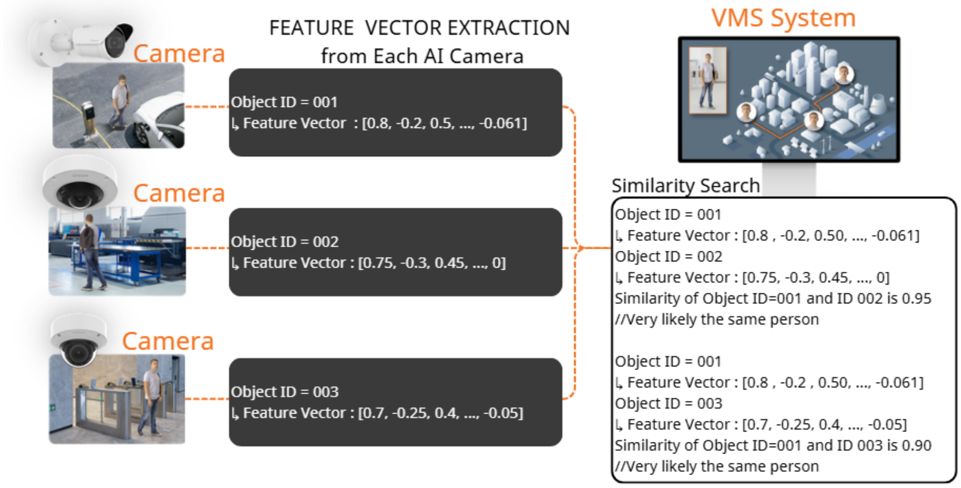


Figure 4. Example of a Feature Vector

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 assigns highly consistent feature values to 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.

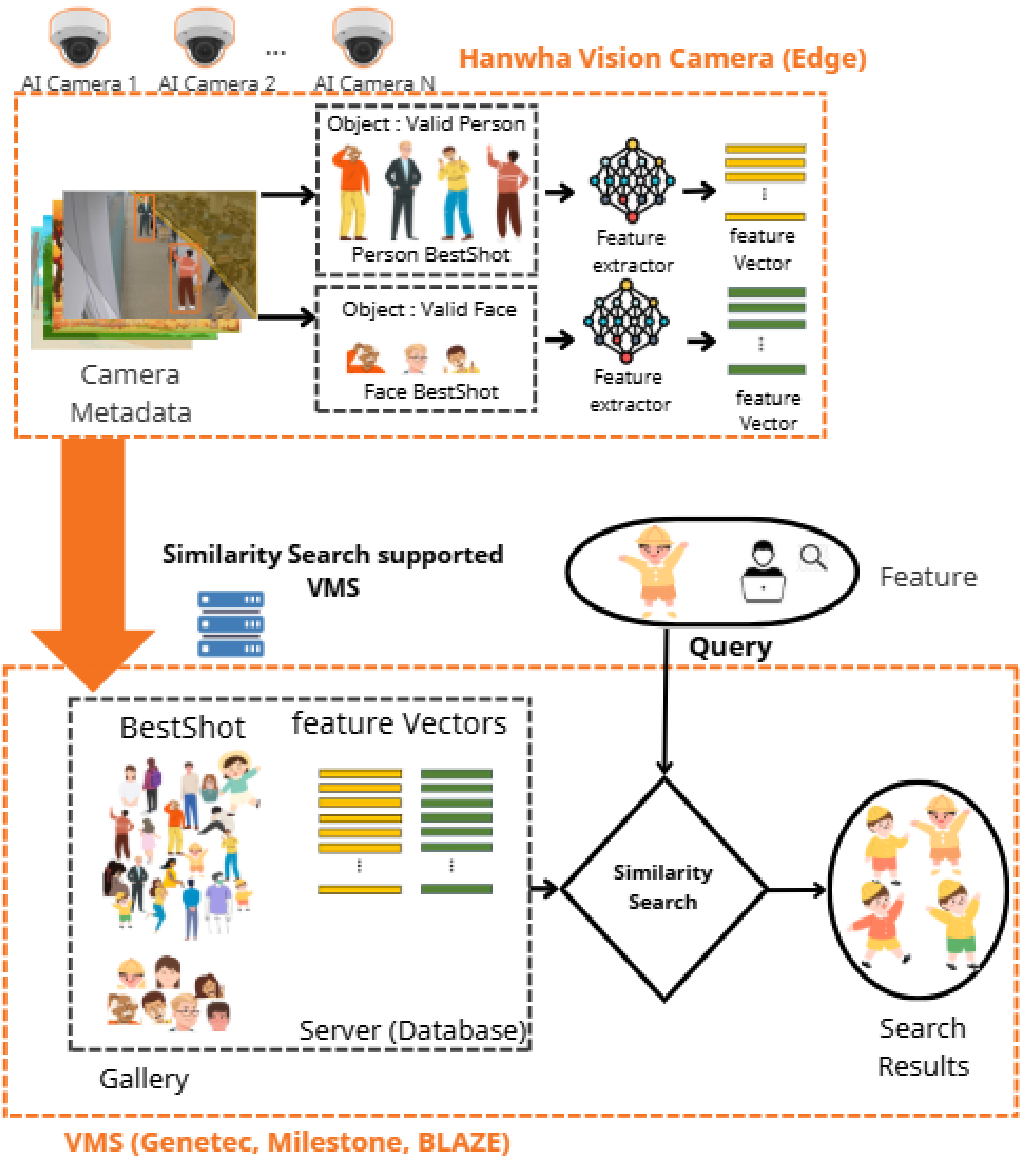


Figure 5. System Architecture of Similarity Search Utilizing Edge Cameras and VMS

Metadata is one of the sub-attributes of the Object ID, as illustrated in the image below. It is located under both the Person and Face categories. The feature vector values can be identified through the metadata shown in Figure 6.

The omitted string between these tags represents the encoded feature vector. Once decoded, it translates to numerical values such as (-0.33276, -0.15478, -0.06964, 0.17028, ...). Ultimately, the similarity search is executed through the mathematical comparison of these precise feature vectors.


```xml
<?xml version="1.0" encoding="UTF-8"?>
<tt:MetadataStream xmlns:tt="http://www.onvif.org/ver10/schema">
  <tt:VideoAnalytics>
    <tt:Frame UtcTime="2026-03-03T13:19:13.659Z" Source="WiseAI">
      <tt:Object ObjectId="6664" Parent="6662">
        <tt:Appearance>
          <tt:Class>
            <tt:Type Likelihood="0.82">Face</tt:Type>
          </tt:Class>
   <tt:CandidateImageRef>/download/WiseAI/ch0/objectid_6664_1772543952759_G.jpg</tt:CandidateImageRef>
          <tt:ReID Samples="257" BytesPerSample="2" Endianness="Little" Version="1.0" Type="Face">
            U7X0sHWsczH0rPGptL ... [omitted] ... AAAgVg==</tt:ReID>
        </tt:Appearance>
      </tt:Object>
      <tt:Object ObjectId="6663" Parent="6662">
        <tt:Appearance>
          <tt:Class>
            <tt:Type Likelihood="0.85">Head</tt:Type>
          </tt:Class>
        </tt:Appearance>
      </tt:Object>
      <tt:Object ObjectId="6662">
        <tt:Appearance>
          <tt:Class>
            <tt:Type Likelihood="0.77">Human</tt:Type>
          </tt:Class>
        </tt:Appearance>
      </tt:Object>
    </tt:Frame>
  </tt:VideoAnalytics>
</tt:MetadataStream>
```


Figure 6. Sample Metadata Including Similarity Vector Values

4.1. 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.

4.2. 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 recognizes it as the same person and displays the results. A lower threshold means that matches will be returned even with a lower degree of vector similarity.

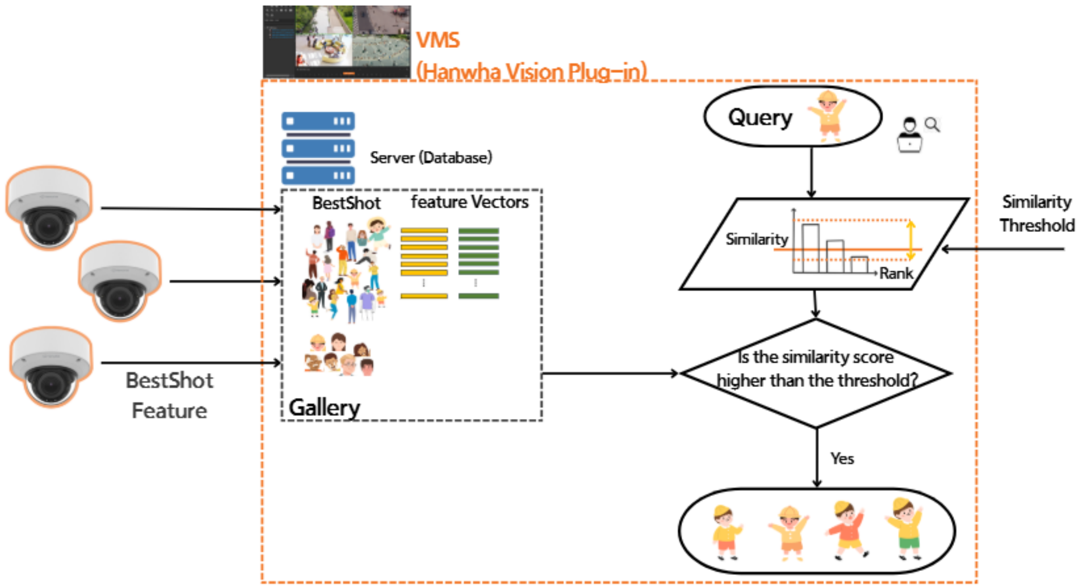


Figure 7. Sensitivity Setting During Similarity Search

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.

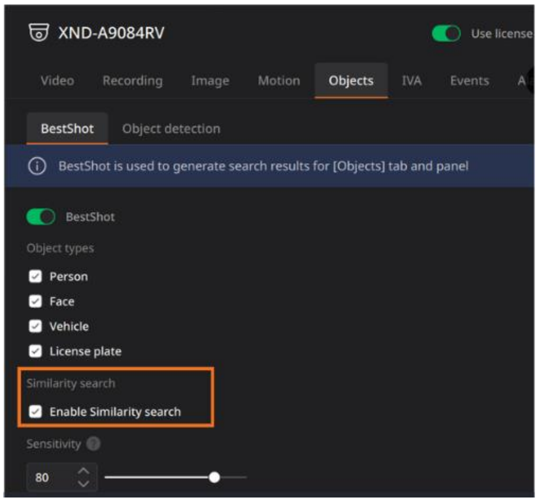


Figure 8. Enabling Similarity Search in the VMS

As previously detailed, similarity search utilizes two distinct types of feature vectors: full-body (Person) and facial (Face). To maximize search accuracy by activating facial feature vectors, users must navigate to the General tab and explicitly consent to the use of biometric information.

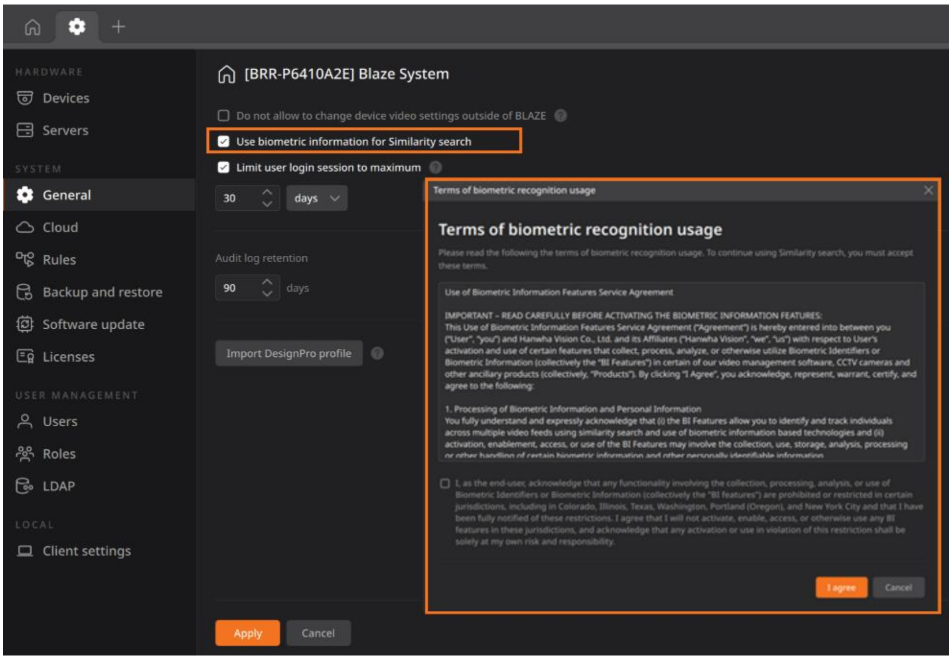


Figure 9. Configuration for Utilizing Facial Feature Vectors

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.

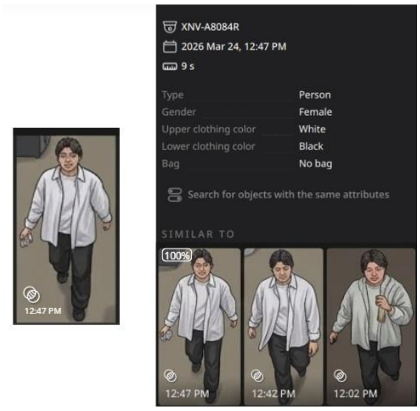


Figure 10. Example of a BestShot Containing a Feature Vector

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' rigorously filters the output, presenting only highly confident, strictly matching results with elevated similarity scores. The similarity results are sorted by confidence, and the time information is displayed alongside them.

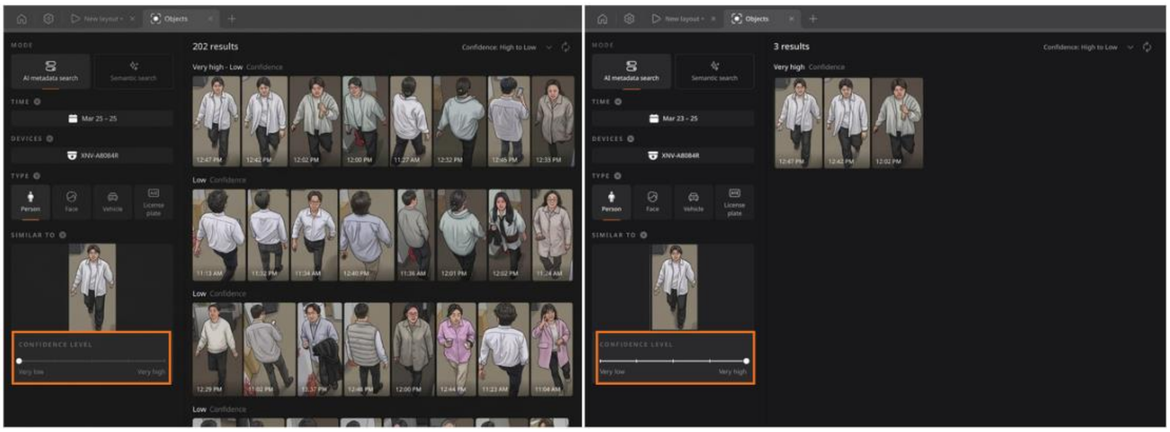


Figure 11. Example of Similarity Sensitivity Adjustment

Ultimately, these search results empower operators to consistently track the identical person across multiple cameras, varying times, and diverse locations. By seamlessly integrating with Hanwha Vision's VMS and cloud solutions, this technology enables highly efficient search and tracking operations, even within large-scale surveillance environments.

4.3. Regulations & System Configurations to Consider Utilizing Face Data

When selecting a specific individual from search results to find similar matches, certain jurisdictions impose restrictions on the use of identifiable information derived from biometric identification systems (Face Recognition) and object attribute classification systems. In particular, adherence to personal data protection laws and video privacy guidelines is paramount. Video footage that allows for facial identification is classified as personal data; thus, explicit consent or a valid legal basis is strictly required for its collection, storage, and utilization.

Under frameworks like the EU GDPR and the EU AI Act, facial recognition is categorized as a high-risk biometric system, placing strict limitations on real-time remote biometric identification. Other countries may also enforce rigorous regulations on public space surveillance due to privacy infringement concerns. Considering these international regulatory landscapes, facial recognition under GDPR demands strict purpose limitation (e.g., security) and data minimization. Furthermore, as the EU AI Act generally prohibits real-time biometric identification in publicly accessible spaces, VMS system administrators must obtain explicit consent to utilize Face Data from the cameras before activating this specific similarity search capability.

VMS solutions must be designed to facilitate data processing consent from users to prepare for potential legal audits. It is important to note that the ultimate responsibility for regulatory compliance lies with the end-user operating the system.

To proactively address stringent global regulatory requirements, such as the EU GDPR and the EU AI Act, Hanwha Vision has successfully implemented and certified its AI Management System (ISO/IEC 42001). The certified scope specifically includes Face Recognition and Re-ID, guaranteeing that biometric information and object attribute data utilized in similarity searches are processed with the highest level of transparency and security in accordance with international management protocols.

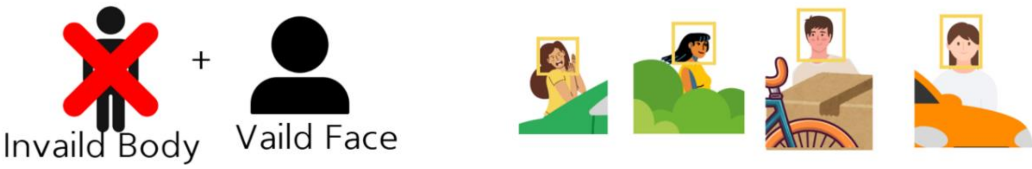
Utilizing Face Data during Hanwha Vision's similarity search significantly enhances matching accuracy. However, Similarity Search remains fully functional even without utilizing Face Data (by relying on Person feature vectors). Nevertheless, users must carefully manage camera settings: if the 'BestShot' feature is turned off, the required metadata will not be transmitted. Similarly, if the 'Similarity Search' option is disabled at the camera level via the VMS, the search functionality becomes entirely unavailable. Note, 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." Optimal capture of both the face and full body is critical for maximizing 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 from Camera, 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.

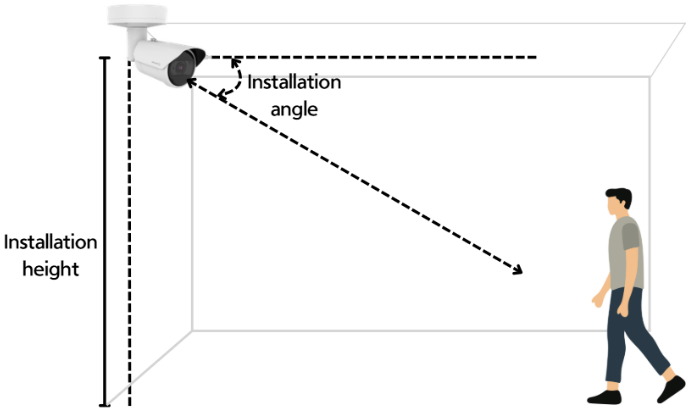


Conversely, if the full-body information is insufficient, searches based on body data will not function properly. This occurs when only the upper body is visible, the subject is seated, or the body is significantly obscured by objects. However, if sufficient frontal facial data is available, the similarity 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, proprietary search technology delivered through the seamless integration of its high-performance AI cameras with its native BLAZE VMS, or via dedicated plug-ins for leading global VMS platforms such as Genetec and Milestone. Furthermore, the ISO/IEC 42001 certification empowers users to establish an intelligent video surveillance environment that is both legally secure and fundamentally trustworthy. By aligning technical innovation with global compliance, Hanwha Vision provides a robust foundation for responsible AI deployment in complex security landscapes.

By utilizing feature vectors extracted directly from the BestShot—a core strength of Hanwha Vision's edge AI cameras—this technology drastically optimizes data volume. This edge-based optimization ensures rapid and highly accurate search and tracking even within vast, large-scale deployments, delivering tangible improvements to security and operational workflows across diverse industries.

Ultimately, this technology overcomes the inherent limitations of conventional metadata or natural language-based search methods currently prevalent in the security sector. By reliably identifying object similarities across multiple cameras, it empowers operators to respond swiftly and effectively during integrated monitoring, presenting a transformative new vision for the future of 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

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