ITUPublications
Recommendations

International Telecommunication Union
Standardization Sector

Recommendation

ITU-T F.743.19 (12/2022)

SERIES F: Non-telephone telecommunication services

Multimedia services

Requirements for intelligent surveillance cameras in intelligent video surveillance systems

ITU-T F-SERIES RECOMMENDATIONS

NON-TELEPHONE TELECOMMUNICATION SERVICES

TELEGRAPH SERVICE	
Operating methods for the international public telegram service	F.1–F.19
The gentex network	F.20–F.29
Message switching	F.30–F.39
The international telemessage service	F.40–F.58
The international telex service	F.59–F.89
Statistics and publications on international telegraph services	F.90–F.99
Scheduled and leased communication services	F.100–F.104
Phototelegraph service	F.105–F.109
MOBILE SERVICE	
Mobile services and multidestination satellite services	F.110–F.159
TELEMATIC SERVICES	
Public facsimile service	F.160–F.199
Teletex service	F.200–F.299
Videotex service	F.300–F.349
General provisions for telematic services	F.350–F.399
MESSAGE HANDLING SERVICES	F.400–F.499
DIRECTORY SERVICES	F.500–F.549
DOCUMENT COMMUNICATION	
Document communication	F.550–F.579
Programming communication interfaces	F.580–F.599
DATA TRANSMISSION SERVICES	F.600–F.699
MULTIMEDIA SERVICES	**F.700–F.799**
ISDN SERVICES	F.800–F.849
UNIVERSAL PERSONAL TELECOMMUNICATION	F.850–F.899
ACCESSIBILITY AND HUMAN FACTORS	F.900–F.999

For further details, please refer to the list of ITU-T Recommendations.

Recommendation ITU-T F.743.19

Requirements for intelligent surveillance cameras in intelligent video surveillance systems

Summary

Recommendation ITU-T F.743.19 specifies the classification of, and scenarios for intelligent analysis functions, in addition to intelligent analysis function and grading requirements for intelligent surveillance cameras (ISCs). The related intelligent analysis functions include video diagnosis, tampering detection, video enhancement, target detection and feature extraction, and object behaviour identification. The basic functions of a camera (see the premises unit specified in Recommendation ITU-T H.626), such as multimedia capturing, multimedia encoding, output alarm signal and parsing pan/tilt/zoom command, lie outside the scope of this Recommendation.

This Recommendation specifies the relevant intelligent analysis function and grading requirements for ISCs used in intelligent video surveillance.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T F.743.19	2022-12-14	16	11.1002/1000/15186

Keywords

Grading, intelligent surveillance camera, requirements.

* To access the Recommendation, type the URL http://handle.itu.int/ in the address field of your web browser, followed by the Recommendation's unique ID. For example, http://handle.itu.int/11.1002/1000/11830-en.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure, e.g., interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents/software copyrights, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the appropriate ITU-T databases available via the ITU-T website at http://www.itu.int/ITU-T/ipr/.

Table of Contents

Recommendation ITU-T F.743.19

Requirements for intelligent surveillance cameras in intelligent video surveillance systems

1 Scope

This Recommendation specifies the classification of, and scenarios for intelligent analysis function, in addition to the intelligent analysis function and grading requirements for intelligent surveillance cameras (ISCs). The basic functions of a camera (see the premises unit (PU) specified in [ITU-T H.626]) such as multimedia capturing, multimedia encoding, output alarm signal and parsing pan/tilt/zoom (PTZ) command, lie outside the scope of this Recommendation.

The scope of this Recommendation includes:

– overview of ISCs for intelligent video surveillance (IVS);

– typical scenarios for ISCs;

– intelligent analysis function and grading requirements for ISCs.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-T F.743.1] Recommendation ITU-T F.743.1 (2015), *Requirements for intelligent visual surveillance*.

[ITU-T G.711] Recommendation ITU-T G.711 (1988), *Pulse code modulation (PCM) of voice frequencies*.

[ITU-T H.222.0] Recommendation ITU-T H.222.0V8 (2021), *Information technology – Generic coding of moving pictures and associated audio information: systems*.

[ITU-T H.264] Recommendation ITU-T H.264V14 (2021), *Advanced video coding for generic audiovisual services*.

[ITU-T H.265] Recommendation ITU-T H.265V8 (2021), *High efficiency video coding*.

[ITU-T H.626] Recommendation ITU-T H.626V2 (2019), *Architecture requirements for video surveillance system*.

[ITU-T H.626.5] Recommendation ITU-T H.626.5V2 (2022), *Architecture for intelligent video surveillance systems*.

[ITU-T H.627] Recommendation ITU-T H.627V2 (2020), *Signalling and protocols for a video surveillance system*.

[ITU-T H.627.2] Recommendation ITU-T H.627.2 (2022), *Requirements and protocols for home surveillance systems*.

[ITU-T T.81] Recommendation ITU-T T.81 (1992), *Information technology – Digital compression and coding of continuous-tone still images – Requirements and guidelines*.

[ITU-T T.800] Recommendation ITU-T T.800 (2019), *Information technology – JPEG 2000 image coding system: Core coding system*.

[ISO/IEC 14496-3] ISO/IEC 14496-3:2019, *Information technology – Coding of audio-visual objects – Part 3: Audio*.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 facial recognition [ITU-T F.743.1]: Detection of the human face in video and extraction of its characteristics. The extracted information and/or image are compared with other images or a database of individuals' images, generating comparison results. Facial recognition can be used in airports, ports, railway stations, coach stations and other public places.

3.1.2 intelligent video diagnosis [ITU-T F.743.1]: Automatic recognition of video images failures, such as snowflakes, scrolling or fuzzy video, partial colour, picture freeze, gain imbalance, pan/tilt/zoom (PTZ) control problems and video signal loss.

3.1.3 premises intelligent video (PIV) [ITU-T H.626.5]: An intelligent identification module in the premises unit (PU). It identifies required information from input video and outputs the result.

3.1.4 region detection [ITU-T F.743.1]: Alarm output that is based on the polygonal regions defined by the user and that is triggered when an object enters or leaves the region.

3.1.5 region stranded [ITU-T F.743.1]: Alarm output that is based on the polygonal regions defined by the user and that is triggered when the retention time of an object exceeds a prescribed threshold.

3.1.6 traffic/pedestrian flow analysis [ITU-T F.743.1]: Motion analysis and feature classification of the video of moving target(s) within the specified single or multiple video monitoring area. The pedestrian flow data can be detected accurately.

3.1.7 tripwire intrusion [ITU-T F.743.1]: The setting off of an alarm when an object crosses a line drawn within the surveillance area.

3.1.8 video surveillance system [ITU-T H.626]: A telecommunication service focusing on video (but including audio and image) application technology, which is used to remotely capture multimedia (such as audio, video, image, alarm signal, etc.), and present them to the end user in a user-friendly manner, based on a managed broadband network with ensured quality, security and reliability.

3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

3.2.1 detection rate: The ratio of the number of correct targets or correct events detected to the number of targets or events that should be detected. The detection rate, r_d, is calculated by

$$r_d = \frac{n_{t+}}{n_{f-} + n_{t+}} \times 100\%$$

where:

n_{t+} is the number of true positive results

n_{f-} is the number of false negative results.

3.2.2 humanoid detection: Activation of an alarm when a human body passes through the monitoring area of a camera.

3.2.3 intelligent video management (IVM): A unit that supports strategies configuration of intelligent application by users and video sources schedule dynamically. The IVM accepts registration, deletion, capability report from IVU and schedule IVU dynamically. It can also store, manage, and schedule those capabilities dynamically. IVM can be invoked by an intelligent application from the centre management unit (CMU).

NOTE – Paraphrased from clause 6 of [ITU-T F.743.1].

3.2.4 intelligent visual surveillance (IVS): A system that can automatically identify specific objects, behaviours or attributes in visual signals. IVS extracts data from visual signals, which are then transmitted or archived so that the visual surveillance system can act accordingly.

NOTE – Paraphrased from clause 1 of [ITU-T F.743.1].

3.2.5 intelligent premises unit (IPU): A unit that adds a premises intelligent video (PIV) module to a premises unit (PU). A PIV is an intelligent recognition module of the PU, which can be achieved and integrated in PU devices. A PIV can recognize required information from input videos, and output recognition results. The recognized information includes that from event triggers and data acquisition.

NOTE – Paraphrased from clause 6 of [ITU-T F.743.1].

3.2.6 object enhancement: An image processing function realizes target object enhancement from the whole video image.

3.2.7 recognition accuracy: The ratio of the number of correctly identified targets or events to the total number of targets or events that should be identified correctly. The recognition accuracy, a_r, is calculated by

$$a_r = \frac{n_{t+}}{n_{ic}} \times 100\%$$

where:

n_{t+} is the number of true positive results

n_{ic} is the number of results that should be identified correctly.

3.2.8 tampering detection: The discovery of a process of malicious alteration of video content, so as to conceal an object, an event or change the meaning conveyed by the imagery in the video.

3.2.9 intelligent surveillance camera (ISC): A kind of intelligent premises unit that can not only capture video, audio, and images but also process the captured images or video and execute particular analysis algorithms, recognize required information and output analysis results including alarms, video structure data and recognition results.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

IPU	Intelligent Premises Unit
ISC	Intelligent Surveillance Camera
IVM	Intelligent Video Management
IVS	Intelligent Video Surveillance
PIV	Premises Intelligent Video
PTZ	Pan/Tilt/Zoom

PU Premises Unit

5 Conventions

In this Recommendation:

– The phrase "**is required to**" indicates a requirement that must be strictly followed and from which no deviation is permitted if conformance to this document is to be claimed.

– The phrase "**is recommended**" indicates a requirement that is recommended but which is not absolutely required. Thus, this requirement need not be present to claim conformance.

– The phrases "**can optionally**" and "**may**" indicate an optional requirement that is permissible, without implying any sense of being recommended. These terms are not intended to imply that the vendor's implementation must provide the option and the feature can be optionally enabled by the network operator or service provider. Rather, it means the vendor may optionally provide the feature and still claim conformance with the specification.

6 Overview of intelligent surveillance cameras in IVS

An ISC is defined in clause 3.

This clause specifies the typical intelligent analysis function classification, intelligent-level classification and each intelligent analysis category-dedicated code.

6.1 Classification of intelligent surveillance cameras

This Recommendation divides ISCs into two categories, whose descriptions follow.

– Consumer-level ISC: An ISC dedicated to home surveillance and private use. The consumer-level ISC is usually used in remote surveillance scenarios in homes, private stores, orchards, courtyards, fish ponds, etc.;

– Public-level ISC: An ISC dedicated to traditional public security scenarios (including public transportation surveillance and public security enhancement), and enterprise smart surveillance scenarios (including in communities, supermarkets, construction sites, kitchens and power infrastructure). Public security ISCs are commonly deployed by governments to monitor public areas to prevent security incidents, while enterprise-used ISCs are commonly deployed by enterprises to improve the public security of their community, and the intelligent operation and management of their workflow.

6.2 Intelligent analysis classification of ISC

The typical intelligent analysis functions for ISCs are divided into several categories that are shown in Figure 6-1. Usually intelligent video diagnosis, tampering detection and video enhancement are basic analysis functions for an ISC to achieve HD video and image in any environment, while the use of intelligent features such as object behaviour analysis, target detection and attribute extraction depends on actual scenarios. For simplification, this Recommendation classifies these intelligent analysis functions as follows.

– Intelligent perception and adaptability capability: This function contains the capability for intelligent video diagnosis, video image enhancement and tampering detection. The intelligent video diagnosis analysis includes that of brightness anomalies, as well as detection of blurring, fogging and colour casts. Video and image enhancement employs process technology to improve quality, e.g., of definition, resolution and contrast, as well as enhancement of the target object in the whole image. Tampering analysis indicates some changes in images that are triggered by some external factors, such as camera coverage by a tree felled by storm winds, disconnection from power supply, theft and physical coverage;

– Vehicle analysis category: This function refers to the capability to automatically detect, identify and analyse traffic target information and traffic flow status of vehicles, including detection of motor vehicles and their features, plus their recognition (e.g., of licence plates, brand, type and colour), traffic statistics, and non-motor vehicle detection and recognition;
– Face or body analysis category: This function refers to the capability to automatically detect and recognize target information, e.g., faces and bodies, including their detection and attribute recognition.
– Behaviour and event analysis category: This function refers to abnormal behaviour, e.g., physical violence, unexpected gatherings and falling, as well as behavioural events occurring in the scene, including monitored region intrusion, when a specific item moves or disappears, a baby cries or an elderly person experiences a fall;

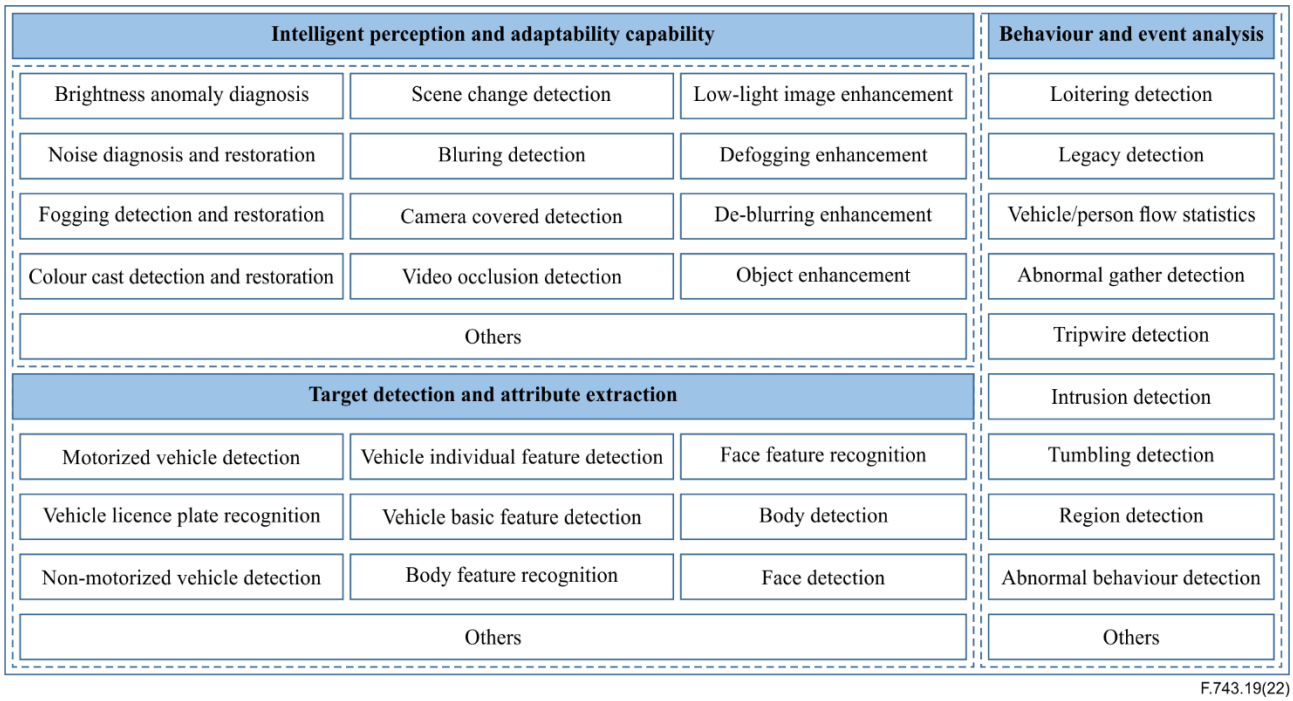

Figure 6-1 – Intelligent analysis function classification of ISC

6.3 Intelligent level classification of ISC

The intelligence level of the camera is graded and assessed according to the performance of the intelligent analysis function. This Recommendation specifies two levels for ISC:

a) general level: level I, which can meet general commercial use of the market;

b) enhanced level: level II, which can meet advanced commercial use of the market.

7 Typical scenarios for intelligent surveillance cameras

This clause defines typical intelligence scenarios and analysis functions for ISC.

7.1 Intelligent perception and adaptability scenarios

7.1.1 Brightness anomaly diagnosis and restoration

Brightness anomaly diagnosis is a basic function for intelligent transportation monitoring; especially at night, vehicle headlights are overexposed in captured images, which makes recognition analysis of vehicle licence plates difficult, due to excessive brightness close to them. This use case is referred to as strong light suppression, which is illustrated in Figure 7-1.

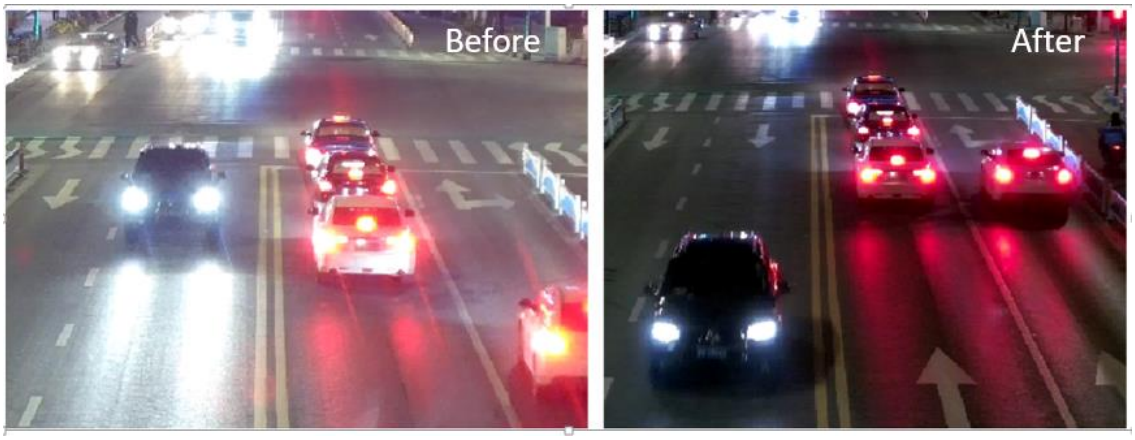

Figure 7-1 – Strong light suppression

A camera with a strong light suppression function can effectively improve the concealment of dark details by the halo around a strong light source, and can make the outline and part of the details of the backlit objects visible without any increase in signal noise.

Another typical scenario related to brightness anomaly diagnosis and restoration is the backlit scene; when the rear of a target object is exposed to sunlight or any light source, there will be parts of the object affected by low brightness, which is illustrated in Figure 7-2.

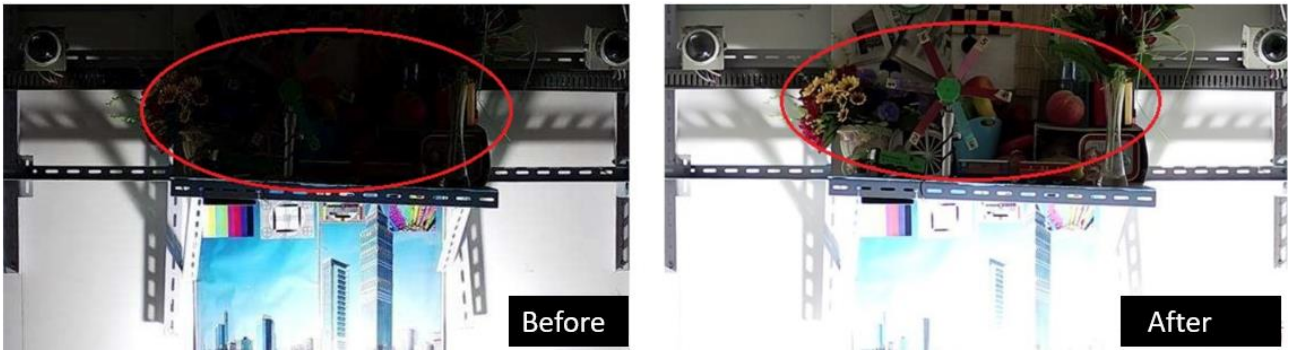

Figure 7-2 – Backlit enhancement

Backlit enhancement can process and enhance the brightness and clarity of images of scenes with dark subjects on which it is difficult to distinguish details, and the resolution does not decrease while the signal noise is not increased.

7.1.2 Blurring detection and restoration

Licence plate and vehicle features are recognized at highway turnpikes and at urban road vehicle checkpoints by taking high-definition vehicle images and analysing them. High vehicle speed affects and blurs image quality, which will cause licence plate recognition to fail. In this situation, blur detection and restoration is needed.

7.1.3 Fogging detection and restoration

When a camera is deployed in a foggy outside environment, the initial video and image will be as shown in Figure 7-3-a, in which the image is low resolution and low definition. After processing by the fogging detection and restoration function (also called the intelligent defogging function) the resolution and definition are improved; in addition, there is no significant increase in signal noise and colour casting. The processed image is shown in Figure 7-3-b.

a **b**

Figure 7-3 – Intelligent defogging function

7.1.4 Object enhancement

Useful information in the image of an object can be enhanced and its visual effect improved for a given application (shown in Figure 7-4). For example, new coding processing can be performed on moving vehicles and human face regions, so that the regional targets that the surveillance camera focuses on are more detailed and clearer, and the overall coding efficiency is not reduced.

Figure 7-4 – Object enhancement

7.1.5 Colour cast detection and correction

Sometimes external factors, e.g., power abnormalities, low temperature and special colour temperature of the environment, will induce colour cast in a video or camera image. A dedicated correction function can detect the colour cast of an image and automatically adjust several colour-related configurations to make the colour of an output image closer to that in reality.

7.1.6 Detection of a camera obstruction

Obstruction mainly affectsthe lens . If a camera lens is obstructed, an alarm is generated to alert the user to perform related operations. Detection of a camera obstruction is used to alert the user if a camera is:

– vandalized with spray paint covering the lens to prevent recording;

– physically covered by a fallen tree or other objects.

7.1.7 Global scene change detection

Global scene change detection is used when a monitoring scene changes, e.g., when someone hits the camera. If the global scene changes, an alarm is generated. This use case can be applied to outdoor scenes that have a number of moving objects, e.g., roads and carparks. Global scene change detection could be used to alert the user if the camera is physically moved or hit by something.

7.2 Vehicle analysis-related scenarios

This clause specifies scenarios related to vehicle detection and recognition.

7.2.1 Vehicle detection and recognition

This scenario is illustrated in clause 5.1.2 of [ITU-T F.743.1].

7.2.2 Vehicle licence plate recognition

Vehicle licence plate recognition is specified in clause 5.3.3 of [ITU-T F.743.1].

7.2.3 Vehicle basic feature recognition

Basic vehicle features, including brand, type, mode, colour, can be detected for recognition.

7.2.4 Vehicle individual feature recognition

Individual features related to a specific vehicle (e.g., detecting whether there are pendants, tissue box or vehicle inspection signs), and any person within it (including a driver or co-pilot not wearing a seat belt or making phone calls) can be detected for recognition.

7.2.5 Vehicle flow statistics

Vehicle flow statistics count traffic flow according to lanes and cycles, including the number of vehicles, vehicle types, and average speeds.

7.2.6 Non-motor vehicle detection

Non-motor vehicles, like bicycles, electric vehicles and tricycles. can be recognized and analysed in non-motor vehicle detection. This kind of non-motor vehicle detection is used to assist traffic analysis and statistics at traffic intersections, as well as to identify electric bicycles and then analyse whether cyclists are wearing helmets.

7.3 Human feature-related scenarios

This claise describes specific scenarios for analysis of the face and body from video or image including theirdetection, feature recognition and recognition.

7.3.1 Face detection

Face detection captures and tracks human faces in areas monitored, and surrounds a target with a highlight rectangle in the image. The scenarios and face recognition are introduced in clause 5.3.2 of [ITU-T F.743.1].

7.3.2 Mask-wearing recognition

Recognition of facial features, e.g., those indicating age, gender or clothing worn, is widely used in transportation security. Especially during a pandemic or in hospital, people may be required to wear masks. It is useful for a camera to integrate a mask detection algorithm, so that mask wearing as an attribute of the human face is integrated with face detection.

7.3.3 Body detection

Body detection is used to distinguish the human body from other target objects, like vehicles or animals, and is widely used to detect perimeter intrusion. It is necessary to exclude certain features, so as to distinguish a human from a dog, for example.

7.3.4 Body feature recognition

Body feature recognition involves analysis of the attributes of the clothing and accessories of personnel detected in the scene monitored. Body feature detection may be used in the following scenarios:

– detection of helmets for people who ride a non-motor vehicle;

– detection of helmets and uniforms worn by workers during security checks on a smart construction site.

7.4 Behaviour analysis scenarios

7.4.1 Tripwire detection

Tripwire detection is used to detect whether a target crosses a line. If so, an alarm is generated. Tripwire detection may be used in the following intrusion prevention scenarios:

– at key locations;

– in restricted areas.

7.4.2 Region detection

For an introduction to region detection, see clause 5.2.1 of [ITU-T F.743.1].

7.4.3 Region stranded detection

For an introduction to region stranded detection, see clause 5.2.1 of [ITU-T F.743.1].

7.4.4 Traffic/pedestrian flow analysis

For an introduction to traffic or pedestrian flow analysis, see clause 5.3.1 of [ITU-T F.743.1].

7.4.5 Abnormal behaviour detection

Abnormal behaviours (e.g., unplanned gatherings of people, running, falls to the ground, climbing and travel in the wrong direction), can be detected. For example, gathering detection determines the density of people in a current scene by counting their images in specified area. When the density of people reaches a threshold, an alarm is generated. Gathering detection can be used to prevent threats to safety in public areas, e.g., bus stations, railway stations or scenic spots, where the maximum flow of people is limited or to monitor the density of people to quickly prevent accidents, including outbreaks of violence, due to overcrowding.

7.4.6 Abnormal sound detection

Abnormal sounds. including babies crying and sudden high-decibel noise, can be detected and tracked. This function is typically used in a home surveillance scenario. When a camera detects abnormal sound, an alarm with a captured image of sound source can be generated and reported to the client of the user. Users can be alerted about what happened in their homes and respond in a timely fashion.

7.4.7 Detection of falls

A fall is a kind of change in posture. Changes in human posture, including falling, slipping and fainting, can be detected. This function is typically used in a home surveillance scenario to look after the elderly or young children. When a camera detects people falling, an alarm with a captured image of the spot is generated and reported to the user. Users can be alerted about what happened and respond in a timely fashion.

7.4.8 Humanoid motion detection

Humanoid motion detection will only raise an alarm when human movement is detected. This type of intelligent analysis is used in warehouses, private houses, banks, etc. Usually, after such detection, the camera can capture an image of the human body and send it to the client. In the meantime, the camera broadcasts an alarm and turns on a double source light-emitting diode source to deter intruders.

7.4.9 High-altitude throwing detection

Objects thrown from the top of buildings can be detected and tracked, and a real-time alarm sent to the client. These objects include but are not limited to rubbish bags, paper balls, cigarette butts, mineral water bottles, knives, porcelain, pots and bowls, as well as bricks. At the same time, high-

altitude throwing detection can effectively exclude environmental factors that affect the results, such as trees, weather change, birds and vibration.

8 Functional architecture of ISC

The functional architecture of an ISC, which is specified in clause 6.2.1 of [ITU-T H.626.5], is shown in Figure 8-1. PIV module is the intelligent analysis function module used to implement specific intelligent analysis, it:

– identifies required information from captured videos and image and outputs analysis results;

– reports the analysis results such as triggered events, alarms and acquired data to the service platform.

For detailed functions of other modules, please see [ITU-T H.626.5].

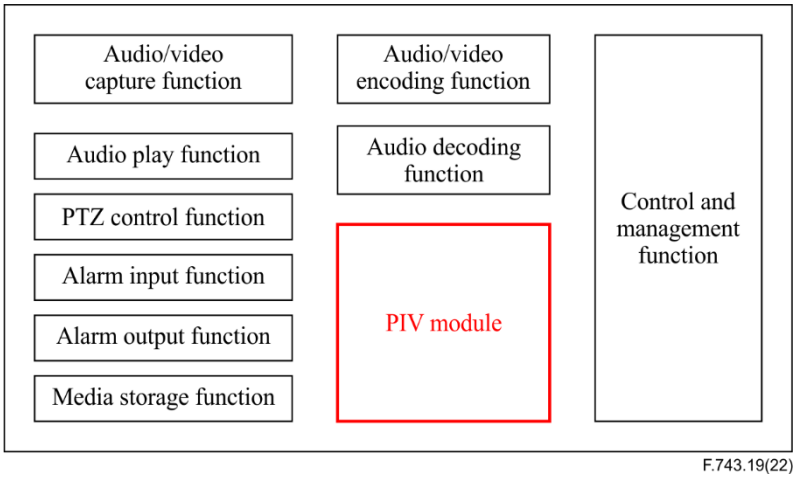

Figure 8-1 – Functional components of an ISC

9 Requirements for consumer-level ISC

9.1 Basic requirements

9.1.1 Common requirements

BR-001: The highest resolution is required to be not lower than 1 920 × 1 080 pixels.

BR-002: A consumer-level ISC is required to support chip computing power not lower than 0.5 T operations/s.

BR-003: A consumer-level ISC is required to support video codec, e.g., as specified in [ITU-T H.264], [ITU-T H.265] or MPEG-4 [ITU-T H.222.0], and support the image format, e.g., as specified in [ITU-T T.81] or [ITU-T T.800].

BR-004: A consumer-level ISC is required to support at least two video streams.

BR-005: A consumer-level ISC is required to support a bi-directional voice intercom function.

BR-006: A consumer-level ISC is recommended to support an infrared night vision function, to ensure the user can obtain the clear video at night.

BR-007: A consumer-level ISC is required to support multiple network access, including wired network connection and wireless connection such as Wi-Fi, 4G and 5G.

BR-008: A consumer-level ISC is required to support both local micro non-volatile flash memory card storage and cloud storage.

BR-009: A consumer-level ISC is required to support two- or three-dimensional digital noise reduction.

9.1.2 Signalling and protocols requirements

PR-001: A consumer-level ISC is required to support the signalling and protocols specified in [ITU-T H.627.2].

9.1.3 Security requirements

SR-001: The authentication and authorization control capability of a consumer-level ISC is required to meet all the following requirements:

a) it is required to support user authorization and authentication;

b) the administrator ID is required to not be easily used fraudulently, and the password is required to be complex and changed regularly;

c) the login failure handling function is required to be enabled, including session ending, number of illegal logins limited and automatic logout;

d) authentication control is required to be performed for important operations of administrator such as restoring configuration and modifying administrator's information when the ISC is remotely managed;

e) the Internet protocol address or media access control address range of the access device is required to be restricted according to the security policy or management needs;

f) the number of users with remote access authority is required to be limited;

g) user rights are required to be separated, and the rights of administrator users and ordinary users are required to be separated.

SR-002: A consumer-level ISC is required to support cryptographic technology or other mechanisms to ensure the integrity of data in the communication process according to business needs, e.g., supporting encrypted transmission.

SR-003: A consumer-level ISC is required to support important data encrypted, including session process, entire message body and important authentication data during the transmission process.

SR-004: A consumer-level ISC is recommended to support video and image privacy mask transmission according to the user configuration.

SR-005: A consumer-level ISC is required to be able to detect and deal with hacker intrusions, including but not limited to tampering with sensitive system files, abnormal super accounts, camera hijacking and process privilege escalation.

SR-006: A consumer-level ISC is required to support querying of operation logs and uploading to the log server or monitoring management platform.

9.2 Intelligent analysis requirements

9.2.1 Humanoid detection

IAR-001: A consumer-level ISC is recommended to support the detection and capture of a humanoid target larger than 80 × 80 pixels in the monitoring image; the humanoid detection response time is required to be no greater than 50 ms indoors and no greater than 150 ms outdoors, and the humanoid accuracy is required to be no less than 95%.

9.2.2 Face detection

IAR-002: A consumer-level ISC is recommended to support the detection of a face target larger than 120 × 120 pixels in the monitoring image, detection of multiple faces at the same time, and outputs the number, size (width, height, and pixels), location and other information about the face.

IAR-003: A consumer-level ISC is recommended to support the capture of no less than 10 effective face targets at the same time. A level-I ISC is required to support a detection rate of no less than 90%. A level-II ISC is required to support a detection rate of no less than 95%.

9.2.3 Motor vehicle detection

IAR-004: A consumer-level ISC is recommended to support the detection and capture of motor vehicle targets larger than 120 × 120 pixels in the monitoring image; the vehicle detection response time should be no greater than 100 ms. A level-I consumer-level ISC is required to support a detection accuracy rate of no less than 95% A level-II consumer-level ISC is required to support a detection accuracy rate of no less than 99%.

9.2.4 Motor vehicle licence plate recognition

IAR-005: A consumer-level ISC is recommended to support the recognition of the number, colour and type of motor vehicle licence plates with a width greater than 100 × 30 pixels in the monitoring area. The accuracy rate of motor vehicle licence plate recognition should be no less than 95% during the day and no less than 90% at night.

9.2.5 Motor vehicle basic feature recognition

IAR-006: A consumer-level ISC is recommended to support the detection and capture of motor vehicle targets larger than 100 × 80 pixels in the monitoring image; the vehicle detection response time should be no greater than 200 ms. A level-I consumer-level ISC is required to support a detection accuracy rate of no less than 90%. A level-II consumer-level ISC is required to support a detection accuracy rate of no less than 95%.

9.2.6 Non-motor vehicle detection

IAR-007: A consumer-level ISC is recommended to support the detection and recognition of non-motor vehicle targets (including electric bicycles, bicycles, scooters, trolleys), larger than 30 × 40 pixels in the monitoring image, the detection accuracy rate should not be less than 90%, and the vehicle detection response time should be no greater than 100 ms.

9.2.7 High altitude throwing detection

IAR-008: A consumer-level ISC is recommended to support the detection and capture of high-altitude throwing targets larger than 5 ×5 pixels in the surveillance image. A level-I consumer-level ISC is required to support a detection accuracy rate of no less than 90%. A level-II consumer-level ISC is required to support a detection accuracy rate of no less than 95%.

10 Requirements for public-level ISC

10.1 Basic requirements

10.1.1 Common requirements

BR-001: A public-level ISC is required to support a basic PU feature (see clause 6 of [ITU-T F.743.1]) including video or audio acquisition, image snapshot, recording or playback, event and alarm, and PTZ control.

BR-002: A public-level ISC is required to support a chip computing power not lower than 1 T operations/s.

BR-003: A public-level ISC is required to support video codec, e.g., as specified in [ITU-T H.264], [ITU-T H.265] or [ITU-T H.222.0], and support an image format, e.g., as specified in [ITU-T T.81] or [ITU-T T.800].

BR-004: A public-level ISC is required to support audio codec, e.g., pulse code modulation [ITU-T G.711] or advanced audio coding [ISO/IEC 14496-3].

BR-005: A public-level ISC is required to support at least two video streams.

10.1.2 Signalling and protocols requirements

PR-001: A public-level ISC is required to support the signalling and protocols specified in [ITU-T H.627], in order to support video and audio transmission, access and control from IVS.

10.1.3 Security requirements

A public-level ISC is required to support the security requirements specified in clause 9.1.3.

10.2 Intelligent analysis requirements

10.2.1 Intelligent perception and adaptability capability requirements

10.2.1.1 Brightness anomaly diagnosis and restoration

IAR-001: A public-level ISC is recommended to support brightness anomaly diagnosis and a restoration function including adjustment when the brightness of an image is too high or too low, enhancement of the whole or part of the brightness of a video or image and strong light suppression.

IAR-002: A public-level ISC is recommended to support automatic adjustment of the high-dynamic range parameter according to the brightness change in the environment and support the backlight compensation function to ensure the object monitored is clear and recognizable.

10.2.1.2 Blurring detection and restoration

IAR-003: A public-level ISC is recommended to support the detection and processing of image blur caused by target movement or an unfocused lens.

IAR-004: A public-level ISC is recommended to have an adaptable de-blurring function, and to improve the resolution and definition of images.

10.2.1.3 Fogging detection and restoration

IAR-005: A public-level ISC is recommended to have an adaptive defogging function that can process video images with low clarity and visibility in foggy weather, and to improve the resolution and definition of video images; obvious colour cast should not be caused after defogging.

10.2.1.4 Object enhancement

IAR-006: A public-level ISC is recommended to enhance the target-related region of an image by improving the resolution and definition of the related region.

IAR-007: A public-level ISC is recommended to support the ability to locally enhance regions related to moving target objects (e.g., human faces and motor vehicle licence plates).

10.2.1.5 Colour cast detection and correction

IAR-008: A public-level ISC is recommended to support processing the colour cast of an image that is caused by the wrong distribution of colour in a specific range, and making the colour of the output image closer to that of reality.

10.2.1.6 Obstructed camera detection

IAR-009: A public-level ISC is recommended to support the detection of obstructed cameras; a level-Ipublic-level ISC is required to support detection of 100% of obstructed cameras; a level-II public-level ISC is required to support detection of more than 50% of obstructed cameras.

10.2.1.7 Scene change detection

IAR-010: A public-level ISC is recommended to support detection of scene changes, e.g., due to the camera being moved or rotated. A level-I public-level ISC is required to support detection of more

than 100% of scenes changed. A level-II public-level ISC is required to support detection of more than 50% of scenes changed.

10.2.2 Vehicle analysis capability requirement

10.2.2.1 Motor vehicle detection

IAR-011: A public-level ISC is recommended to support detection of motor vehicles, and to support the detection and capture of multiple motor vehicle targets at the same time. A level-I public-level ISC is required to support detection and capture of motor vehicle targets larger than 120 × 120 pixels in the monitoring image and achieve a detection rate ≥95% – the detection accuracy rate should not be less than 99%. A level-II public-level ISC is required, in addition to meeting level-I specifications, to achieve a detection rate no less than 95% in the vehicle forward and reverse scenes; furthermore, the accuracy rate should be no less than 99%.

10.2.2.2 Motor vehicle licence plate recognition

IAR-012: A public-level ISC is recommended to support the recognition of the number, colour and type of motor vehicle licence plates with a width greater than 100 × 30 pixels in the monitoring area. In addition, the following recognition accuracy requirements for level-I and level-II ISC should be met:

a) level-I: the accuracy rate of motor vehicle licence plate recognition should be no less than 95% during the day and no less than 90% at night;

b) level-II: on the basis of conforming to level-I specifications, it is required to have the ability to identify whether a licence plate is altered or covered.

10.2.2.3 Motor vehicle basic feature recognition

IAR-013: A public-level ISC is recommended to support the recognition of basic characteristics of motor vehicles and meet the following recognition accuracy requirements for level-I and level-II ISC:

a) level I: an ISC is required to support the analysis and output of attribute information such as type and colour of motor vehicle, the average recognition accuracy of motor vehicle basic characteristics should not be less than 90%;

b) level II: on the basis of level I, the ISC is still required to support the recognition of the brand of motor vehicle, and the average recognition accuracy of the basic characteristics of the motor vehicle should not be less than 90%.

10.2.2.4 Motor vehicle individual feature recognition

IAR-014: A public-level ISC is recommended to support the identification of individual characteristics of motor vehicles, such as whether the main or co-driver is wearing a seat belt, or holding a phone, etc., the average recognition accuracy of individual characteristics of motor vehicles is not less than 80%.

10.2.2.5 Motor vehicle flow statistics

IAR-015: A public-level ISC is recommended to support the counting of motor vehicle traffic targets larger than 120 × 120 pixels and the compilation of statistics of motor vehicle flow according to lane and time conditions, including the number of motor vehicles – the accuracy rate of motor vehicle flow statistics is not less than 90%.

10.2.2.6 Non-motor vehicle detection

IAR-016: A public-level ISC is recommended to support the identification of non-motor vehicles targets with a width greater than 80 pixels in the video image – the detection rate of non-motor vehicles should not be less than 90%, and the accuracy rate should not be less than 90%.

10.2.3 Person analysis capability requirement

10.2.3.1 Face detection

IAR-017: A public-level ISC is recommended to support the detection of face targets and their positions and sizes in scenes monitored, detection of multiple faces at the same time, and outputs the number, size (width, height, and pixels), location and other information about the face. A level-I public-level ISC is required to support the capture of no less than 10 effective face targets at the same time, the detection rate of face detection should be no less than 95%, and the accuracy rate of face detection should be no less than 99%. A level-II public-level ISC is required to support the capture of no less than 50 at the same time. For a valid face target, the rate of face detection should be no less than 95%, and the accuracy rate of face detection should be no less than 99%.

10.2.3.2 Body detection

IAR-018: A public-level ISC is recommended to support the detection of the portrait, as well as its position and size in the monitoring scene, supporting the detection of multiple human targets at the same time, and outputting information such as the number of human targets and the size of each target (width and height pixels). A level-I public-level ISC is required to support the capture of no less than 10 effective portrait targets at the same time. A level-II public-level ISC is required to support the capture of no less than 50 effective portrait targets at the same time.

10.2.3.3 Body feature recognition

IAR-019: A public-level ISC is recommended to support the analysis of and judgement on the attributes of people's clothing, accessories and objects carried that are detected in the monitoring scene, as well as the identification of attributes such as hairstyle, clothing style and colour, body shape, objects carried and irection of travel.

10.2.4 Behaviour analysis capability requirement

10.2.4.1 Tripwire detection

IAR-020: A public-level ISC is recommended to support the detection of tripwires of targets larger than 32 × 32 pixels in the area of interest of the video screen and output event information, and meet the detection rate of not be less than 90% when the false detection rate is not greater than 5%.

10.2.4.2 Region detection

IAR-021: A public-level ISC is recommended to support the detection of the intrusion of targets larger than 32 × 32 pixels in the area of interest of the video screen and output event information; when the false detection rate is not greater than 5%, the detection rate is required to not be less than 90%.

10.2.4.3 Region stranded detection

IAR-022: A public-level ISC is recommended to support the detection of the intrusion of targets larger than 32 × 32 pixels into the area of interest of the video screen that stay longer than the specified time, and output event information; when the false detection rate is not greater than 5%, the detection rate is required to be not less than 90%.

10.2.4.4 Traffic or pedestrian flow analysis

IAR-023: A public-level ISC is recommended to support the detection of the number of people, whose shoulder width is greater than 32 pixels in the target statistical area, alarm and output event information if the threshold value is exceeded; the accuracy of crowd density detection is not less than 90%.

10.2.4.5 Abnormal behaviour detection

IAR-024: A public-level ISC is recommended to support the detection of one or more abnormal behaviours (including running, falling to the ground, climbing, travelling in the wrong direction), of personnel targets larger than 64 × 128 pixels in the area of interest of the video screen and output event information.

10.3 Advanced requirements

10.3.1 Intelligent algorithm sustainability requirements

ADR-001: A public-level ISC is recommended to support on-demand deployment and online upgrade of applications without service interruption.

ADR-002: A public-level ISC is recommended to support the life cycle management of applications, including installation or uninstallation, activation or deactivation, upgrade and licence management.

ADR-003: A public-level ISC is recommended to support licence verification and algorithm model file encryption of an application to protect its intellectual property.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	Tariff and accounting principles and international telecommunication/ICT economic and policy issues
Series E	Overall network operation, telephone service, service operation and human factors
Series F	**Non-telephone telecommunication services**
Series G	Transmission systems and media, digital systems and networks
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Environment and ICTs, climate change, e-waste, energy efficiency; construction, installation and protection of cables and other elements of outside plant
Series M	Telecommunication management, including TMN and network maintenance
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling, and associated measurements and tests
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks, open system communications and security
Series Y	Global information infrastructure, Internet protocol aspects, next-generation networks, Internet of Things and smart cities
Series Z	Languages and general software aspects for telecommunication systems

Published in Switzerland
Geneva, 2023