ITUPublications
Recommendations

International Telecommunication Union
Standardization Sector

Recommendation

# ITU-T F.747.12 (12/2022)

SERIES F: Non-telephone telecommunication services

Multimedia services

# Requirements for artificial intelligence based machine vision system in smart logistics warehouse

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.747.12

## Requirements for artificial intelligence based machine vision system in smart logistics warehouse


### Summary

With the rapid development of industrial automation and logistics technology in accordance with the market demand for high-tech, machine vision technology has begun to enable the automation transformation of logistics warehouse systems. The application of machine vision technology in the field of logistics warehouse has enabled the rapid evolution of goods sorting, goods palletizing and de-palletizing, goods handling and shelf inventory from intensive manual work to intelligence and automation, improving the operational efficiency and management capabilities of logistics warehouse.

Recommendation ITU-T F.747.12 specifies the requirements and framework for artificial intelligence based machine vision system in smart logistics warehouse, and provides use cases. This Recommendation is intended to guide the design and development of machine vision systems in smart logistics warehouse.



### History

| Edition | Recommendation | Approval | Study Group | Unique ID* |
|---|---|---|---|---|
| 1.0 | ITU-T F.747.12 | 2022-12-14 | 16 | 11.1002/1000/15193 |

### Keywords

Artificial intelligence, framework, machine vision system, requirements, smart logistics warehouse.


* 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.747.12

## Requirements for artificial intelligence based machine vision system in smart logistics warehouse

### 1 Scope

This Recommendation specifies requirements and framework for artificial intelligence (AI) based machine vision systems (MVSs) in a smart logistics warehouse (SLW).

The scope of this Recommendation includes:

– overview of AI based MVS in SLW;

– framework of AI based MVS in SLW;

– requirements of AI based MVS in SLW.

Use cases of AI based MVS in SLW are provided in Appendix I.

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

### 3 Definitions

#### 3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

**3.1.1 accuracy** [b-ISO/IEC 2382]: Qualitative measure of the magnitude of error, preferably expressed as a function of the relative error, a high value of this measure corresponding to a small error.

**3.1.2 algorithm** [b-ISO/IEC 11557]: A set of rules for transforming the logical representation of data.

**3.1.3 machine vision** [b-ITU-T F.748.16]: Signal processing to acquire, process and interpret an image or video to support visual analysis for applications, such as automatic inspection, process control and guidance.

#### 3.2 Terms defined in this Recommendation

None.

## 4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

| | |
|---|---|
| 2D | Two Dimensional |
| 3D | Three Dimensional |
| AI | Artificial Intelligence |
| AMR | Automated Mobile Robot |
| DM | Data Matrix |
| MVS | Machine Vision System |
| OCR | Optical Character Recognition |
| QR | Quick Response |
| ROI | Region of Interest |
| SKU | Stock Keeping Unit |
| SLW | Smart Logistics Warehouse |

## 5 Conventions

In this Recommendation:

– The keywords "**is recommended**" indicate a requirement which is recommended but which is not absolutely required. Thus this requirement needs not be present to claim conformance.

## 6 Overview

In smart logistics warehouse (SLW), based on the real-time management of the goods, shelves and logistics warehouse environments, artificial intelligence (AI) enabled visual analysis technology is widely used, which greatly improves the work efficiency of SLW. In the process of goods entering and leaving the SLW, the machine vision system (MVS) (e.g., high-resolution industrial cameras) collects and analyses label information of the goods in real-time, and quickly sorts goods according to the variety, material and weight of the goods. At the same time, it can identify abnormal appearance states such as stains, wrinkles and breakages, and can provide event reminders for abnormal goods, label errors and other issues. Through shelf status recognition, the real-time analysis of the shelf situation is carried out to determine whether the shelf is full and whether the goods are placed in a normal state, to realize the intelligent management of the shelf and improve the utilization rate of the shelf. The automated mobile robot (AMR) is responsible for the handling of goods in SLW. AMR's vision system collects ground conditions and surrounding environmental obstacles information, extracts feature information, matches database information, and completes map creation, positioning and navigation, which improves the efficiency of goods transportation. MVS identifies the stacking state of the goods and calculates the grasping pose, and controls such as the mechanical arm to perform the tasks of automatic grasping, palletizing and de-palletizing of the goods. The surveillance camera conducts real-time monitoring of the potential safety hazards in the SLW environment, the loading situation of the transport vehicle, and the use rate of the platform.

## 7 Framework

Figure 1 shows the framework of AI based MVS in SLW.

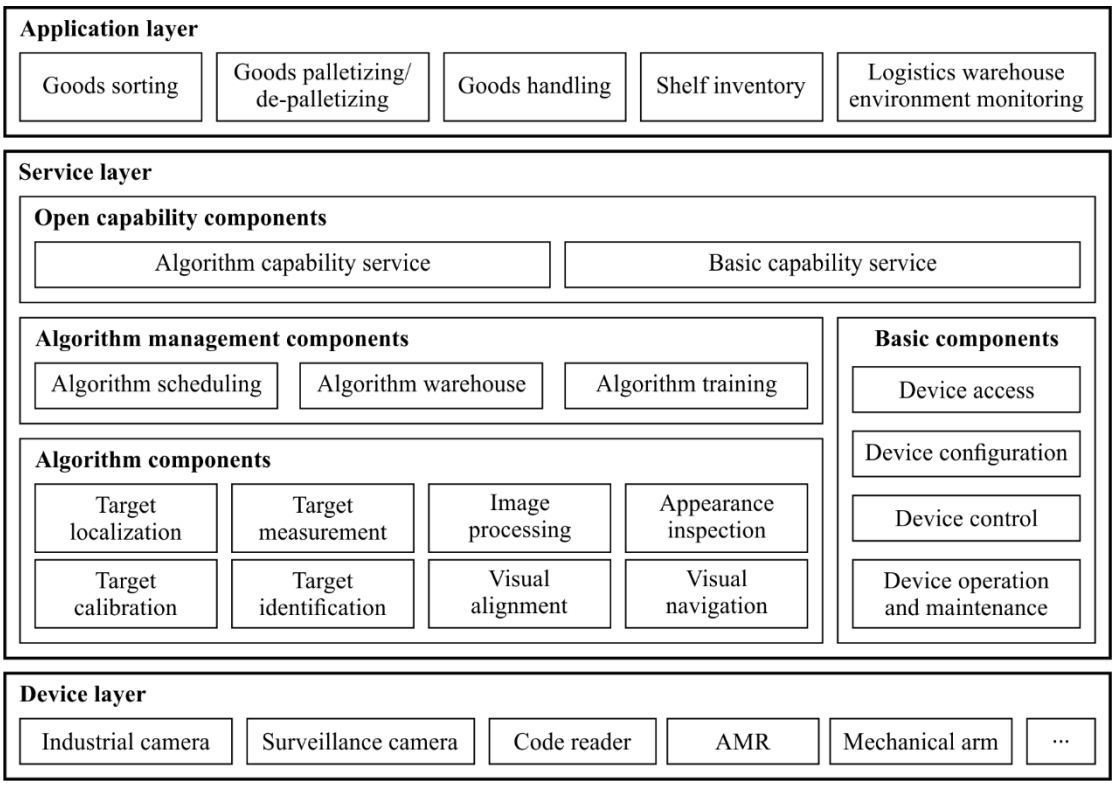


**Figure 1 – Framework of artificial intelligence based machine vision system in a smart logistics warehouse**

The framework of AI based MVS in SLW (hereafter "MVS") consists of three layers: device layer, service layer and application layer.

The device layer represents the devices used for data collection and control in MVS, including industrial cameras, surveillance cameras, code readers, AMRs and mechanical arms, etc.

The service layer consists of four components, known as basic components, algorithm components, algorithm management components and open capability components. Basic components provide basic services for MVS, including device access, device configuration, device control, and device operation and maintenance. The algorithm component provides algorithm services for MVS, including algorithms for target localization, target measurement, image processing, appearance inspection, target calibration, target identification, visual alignment and visual navigation. The algorithm management component provides algorithm-related management services for MVS, including algorithm scheduling, algorithm warehouse management and algorithm training. The open capability component provides interface services for MVS, including algorithm capability services and basic capability services.

The application layer includes service application scenarios such as goods sorting, goods palletizing/de-palletizing, goods handling, shelf inventory, and logistics warehouse environment monitoring.

## 8 Requirements

### 8.1 Device requirements

– MVS is recommended to support data collection, including but not limited to video, image, point cloud, quick response (QR) code and barcode.

– The industrial camera is recommended to support timed exposure and trigger delay functions.

– The industrial camera is recommended to support region of interest (ROI) functions.
– The industrial camera is recommended to support the centre horizontal resolution of output image no less than 85% of image height.
– The code reader is recommended to support barcode format including Code 128, Code 39, etc.
– The code reader is recommended to support QR code format including QR and data matrix (DM), etc.
– The surveillance camera is recommended to support intelligent analysis function, including fire and smoke detection, AMR traffic jam detection, etc.
– AMR is recommended to support QR code navigation and laser navigation.
– AMR is recommended to support obstacle detection and path planning.

## 8.2 Service requirements

### 8.2.1 Requirements of basic components

– MVS is recommended to support device access, including device manual addition, device retrieve addition, and device deletion.
– MVS is recommended to support device access protocols [ITU-T H.627].
– MVS is recommended to support device configuration, including image parameters configuration (frame rate, exposure, gain and gamma), communication parameters configuration and image storage configuration (goods surface original image, goods surface original image cut-out and goods panorama image).
– MVS is recommended to support device control, including data reading, command issuance, device timing, device restart and device reset.
– MVS is recommended to support device operation and maintenance, including the management of device files information, device status, device upgrading and device alarm.

### 8.2.2 Requirements of algorithm components

– MVS is recommended to support target localization algorithms to locate or detect certain features in the image, including feature matching, grayscale matching, graphic localization, rectangle detection, vertex detection, etc.
– MVS is recommended to support target calibration algorithms to realize the conversion from image coordinates to actual coordinates, including checkerboard calibration, N-point calibration, distortion calibration, mapping calibration, etc.
– MVS is recommended to support target measurement algorithms to achieve distance measurement between target object elements in the image, including line-circle measurement, line-line measurement, circle-circle measurement, point-circle measurement, point-line measurement, distance detection, etc.
– MVS is recommended to support target identification algorithms to realize the identification of target objects in the image, including barcode recognition, QR code recognition, optical character recognition (OCR), etc.
– MVS is recommended to support image processing algorithms to realize image pre-processing of the target image, including image filtering, image enhancement, distortion correction, graph scaling and graph stitching, etc.

NOTE – When the image has poor contrast, more burrs, more interference and less obvious features, image pre-processing can be considered.

– MVS is recommended to support visual alignment algorithms to calculate the amount of movement (positional movement and angular movement) required from object alignment to the target object, including single-point alignment, point set alignment, line alignment, etc.
– MVS is recommended to support appearance inspection algorithms to verify and compare target images with standard images to find appearance defects, including detection of printed characters and patterns appearance defects, surface defect detection without texture background, linear edge defect detection, arc edge defect detection, etc.
– MVS is recommended to support visual navigation algorithms, and to complete the autonomous navigation and positioning of the target object by associating the image information collected by the camera with the actual position of the target object.

### 8.2.3 Requirements of algorithm management components

– MVS is recommended to support algorithm scheduling, realizes the scheduling and management of algorithm resources, including algorithm task configuration, algorithm task control (start, execution and cancellation), algorithm time-consuming statistics, and algorithm scheduling result storage.
– MVS is recommended to support algorithm tasks of adding, deleting, starting and stopping.
– MVS is recommended to support algorithm task configuration, including data source devices, data formats, algorithm types, algorithm parameters.
– MVS is recommended to support computing resource management, assign computing resources to algorithm tasks for scheduling, and support viewing computing resource usage situation.
– MVS is recommended to support algorithm warehouse management, including adding, deleting, modifying and querying.
– MVS is recommended to support algorithm training, including sample management, model training, and algorithm deployment.
– MVS is recommended to support sample management, including sample data adding, deleting, modifying and querying.
– MVS is recommended to support sample data annotation, including annotation data adding, deleting, modifying, and querying.
– MVS is recommended to support the training task management, including training task adding, deleting, starting, and stopping, and support querying the progress and results of training tasks.
– MVS is recommended to support model validation and evaluation.
– MVS is recommended to support algorithm deployment.

### 8.2.4 Requirements of open capability components

– MVS is recommended to support algorithm capability services, including algorithm scheduling open capability, algorithm warehouse open capability and algorithm training open capability.
– MVS is recommended to support basic capability services, including device access open capability, device configuration open capability, device control open capability, and device operation and maintenance open capability.

## 8.3 Application requirements

### 8.3.1 Requirements of goods sorting

– MVS is recommended to support the identification of the appearance, size and quantity of goods during the sorting process, and is not recommended to be restricted by the placement direction. The identification speed of the number of parcels is recommended ≥ 5 000 pieces/hour, and the resolution of goods image collection is recommended ≥ 6 million pixels.

– MVS is recommended to support the measurement of the length, width, height of the goods during the sorting process. The volume measurement range is recommended to support 150 mm*150 mm*30 mm to 1 200 mm*1 000 mm*700 mm (regular goods), and the measurement and identification accuracy is recommended less than ±10 mm.

– MVS is recommended to support simultaneous identification of multiple sides of goods in the process of parcel sorting, and it is recommended to support simultaneous identification of six sides of goods.

– MVS is recommended to support the identification of barcodes, QR codes and other label information on the surface of the goods during the sorting process. The code reading rate is recommended ≥ 99%, and the identification speed of a single barcode or QR code is recommended ≤ 100 ms.

– MVS is recommended to support the single identification of multi-goods label information during the sorting process.

– MVS is recommended to support abnormal code identification during the sorting process, including surface dirt, damage, scratches, etc.

– MVS is recommended to support the identification of abnormal status information, including goods accumulation, blockage and overlapping during the sorting process, and identification accuracy rate is recommended ≥ 99%.

– MVS is recommended to support the collection of a complete and clear image of goods during the sorting process, which can clearly identify codes and label information.

### 8.3.2 Requirements of goods palletizing and de-palletizing

– MVS is recommended to support automatic identification of multiple stock keeping units (SKUs).

– MVS is recommended to support the identification of goods abnormal states, including appearance defects, etc.

– MVS is recommended to support the identification of the stacking method and volume size of the goods.

– MVS is recommended to support the calculation of accurate grab location, placement location and trajectory location through intelligent algorithms to control the mechanical arm to quickly grab and stack goods.

### 8.3.3 Requirements of shelf inventory

– MVS is recommended to support the identification of the number of goods on the shelf.

– MVS is recommended to support the identification of abnormal information, including dirty and damaged surfaces of goods.

– MVS is recommended to support the identification of falling goods on the shelf, and is recommended to support the identification of goods that are at risk of falling.

– MVS is recommended to support the identification of vacant and occupied status of shelves.

– MVS is recommended to support the identification of shelf storage space volume.

### 8.3.4 Requirements of goods handling

– MVS is recommended to support the identification of label information on the surface of the AMR transported goods, the identification accuracy is recommended ≥ 99%, and the identification speed is recommended ≤ 500 ms.

– MVS is recommended to support the identification of obstacles and environmental information around the AMR.

– MVS is recommended to support the identification of ground identification mark information.

– MVS is recommended to support shelf label identification, and the identification accuracy is recommended ≥ 99%.

– MVS is recommended to support the automatic planning of handling routes of AMR based on the logistics warehouse map, surrounding obstacles and environmental information.

### 8.3.5 Requirements of logistics warehouse environment monitoring

– MVS is recommended to support the identification of abnormal environmental information, including smoke, fire and water leakage, etc.

– MVS is recommended to support the identification of violent sorting and handling behaviours of workers.

– MVS is recommended to support the identification of AMR traffic jam.

– MVS is recommended to support the identification of falling goods during AMR handling.

– MVS is recommended to support the identification of the accumulation of goods in a specific area.

– MVS is recommended to support the identification of loaded vehicle information, including colour, licence plate, vehicle brand, etc.

– MVS is recommended to support the identification of vehicle loading rate, including empty, half and full, etc.

– MVS is recommended to support the identification of platform usage, including occupied, unoccupied.

# Appendix I

## Use cases of artificial intelligence based machine vision system in a smart logistics warehouse

(This appendix does not form an integral part of this Recommendation.)

Figure I.1 shows a service scenario of smart logistics warehouse.

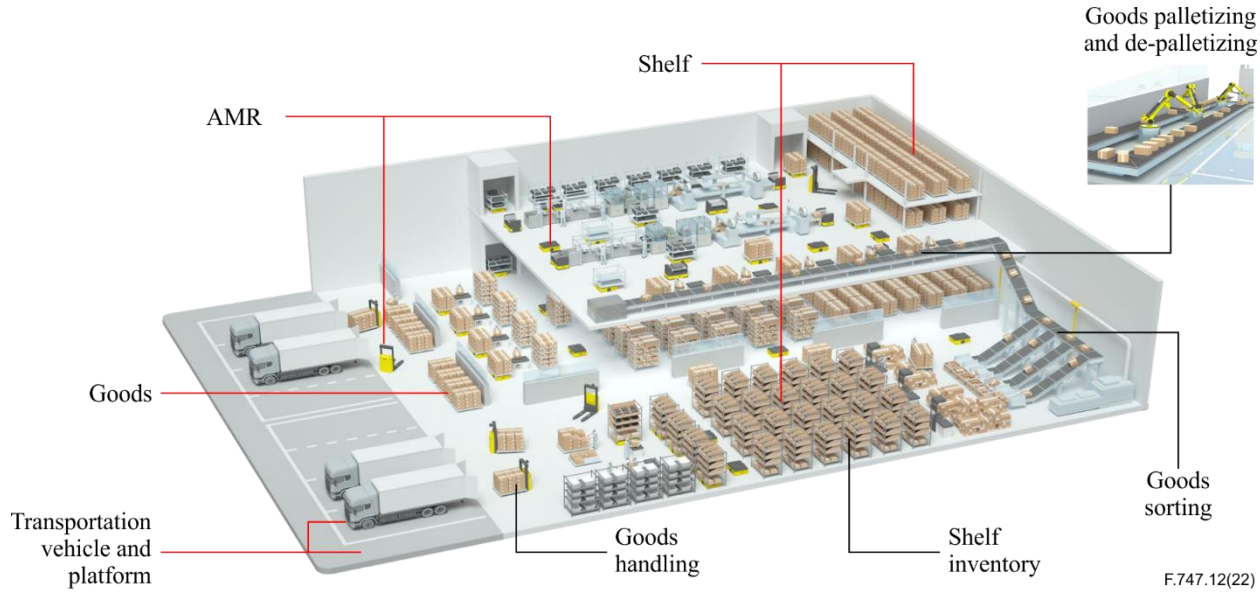


**Figure I.1 – Service scenario of smart logistics warehouse**

### I.1 Goods sorting

In the SLW, the sorting of goods is accomplished by industrial cameras that identify the label, size, and appearance of the goods. Fast identification of labels can collect and classify transportation information such as barcodes and QR codes on goods. Identifying the size of goods can provide a volume basis for subsequent shelf placement and vehicle loading. Identifying the appearance of the goods can determine whether there are abnormalities such as damage, dirt, and scratches on the surface of the goods. At the same time, the abnormal state of the sorting process of the goods on the conveyor belt can be identified, such as the accumulation, blockage and falling of goods.

### I.2 Goods palletizing and de-palletizing

In the SLW, MVS (e.g., three-dimensional (3D) cameras) obtain goods depth information and high-definition two-dimensional (2D) images, generate goods appearance point cloud data, complete goods image positioning and segmentation through visual algorithms, and control mechanical arms to grasp, palletize and de-palletize goods. At the same time, according to the preset accumulation type, MVS identifies and verifies the placement method, and identify the appearance defects of the goods.

### I.3 Shelf inventory

In the SLW, the inventory of goods is completed by industrial cameras that identify the vacancy and occupancy status of the shelf, can count the vacancy rate of the shelf, and provide data support for the goods to be placed on the shelf. MVS identifies the volume of the shelf storage space, which can match the size of the goods to place goods and improve the utilization rate of the shelf. MVS identifies the number of goods on the shelf and carry out an inventory of goods. MVS identifies the

surface of the goods on the shelf, such as damage, dirt and other information, to facilitate the judgment of whether there is an abnormality inside the goods. MVS identifies the goods that fall from the shelf or that are at risk of falling to ensure the safety of goods and people.

### I.4 Goods handling

In the SLW, the handling of goods is completed by AMR. The MVS of the AMR identifies the label information on the surface of the goods and matches the position of the goods on the shelf or transportation vehicle. The AMR machine vision service identifies the surrounding environment and obstacle information, and plans transportation routes and logistics storage maps. AMR identifies ground material information and judges the arrival area.

### I.5 Logistics warehouse environment monitoring

In the SLW, MVS monitors the potential safety hazards of the logistics warehouse environment, including smoke, fire, or water leakage. The MVS monitors the AMR's handling activities, including AMR transportation traffic jams, AMR transportation goods falling, etc. The MVS monitors the accumulation of goods in a specific area, and at the same time monitors the violent sorting and handling behaviour of workers. The MVS collects information on goods transportation vehicles, including vehicle body colour, licence plate, vehicle brand, etc., and collects in-vehicle loading rates to identify empty, half, and full situations.

# Bibliography


[b-ITU-T F.748.16] Recommendation ITU-T F.748.16 (2022), *Requirements for applications and services in smart manufacturing based on machine vision.*

[b-ISO/IEC 2382] ISO/IEC 2382 (2015), *Information technology – Vocabulary*.

[b-ISO/IEC 11557] ISO/IEC 11557 (1992), *Information technology – 3,81 mm wide magnetic tape cartridge for information interchange – Helical scan recording – DDS-DC format using 60 m and 90 m length tapes.*

## 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