

Recommendation

ITU-T Y.4231 (01/2025)

SERIES Y: Global information infrastructure, Internet protocol aspects, next-generation networks, Internet of Things and smart cities

Internet of things and smart cities and communities –
Requirements and use cases

Requirements and capability framework of the Internet of things for vision

ITU-T Y-SERIES RECOMMENDATIONS

Global information infrastructure, Internet protocol aspects, next-generation networks, Internet of Things and smart cities

GLOBAL INFORMATION INFRASTRUCTURE	Y.100-Y.999
INTERNET PROTOCOL ASPECTS	Y.1000-Y.1999
NEXT GENERATION NETWORKS	Y.2000-Y.2999
FUTURE NETWORKS	Y.3000-Y.3499
CLOUD COMPUTING	Y.3500-Y.3599
BIG DATA	Y.3600-Y.3799
QUANTUM KEY DISTRIBUTION NETWORKS	Y.3800-Y.3999
INTERNET OF THINGS AND SMART CITIES AND COMMUNITIES	Y.4000-Y.4999
General	Y.4000-Y.4049
Definitions and terminologies	Y.4050-Y.4099
Requirements and use cases	Y.4100-Y.4249
Infrastructure, connectivity and networks	Y.4250-Y.4399
Frameworks, architectures and protocols	Y.4400-Y.4549
Services, applications, computation and data processing	Y.4550-Y.4699
Management, control and performance	Y.4700-Y.4799
Identification and security	Y.4800-Y.4899
Evaluation and assessment	Y.4900-Y.4999

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

Recommendation ITU-T Y.4231

Requirements and capability framework of the Internet of things for vision

Summary

Recommendation ITU-T Y.4231 provides requirements and capability framework of the Internet of things (IoT) for vision, a typical IoT for vision application workflow as well as some relevant IoT for vision use cases.

With the popularization and development of smart cities and numerous IoT based verticals, the connotation of the Internet of things (IoT) is expanding and is no longer limited to simple sensor data collection and discrimination.

The IoT for vision is one of the main development directions of the IoT which uses digital cameras and other types of sensors to perceive objects of interest (including human beings, vehicles, animals, etc.) for diverse applications based on appropriate analysis.

History *

Edition	Recommendation	Approval	Study Group	Unique ID
1.0	ITU-T Y.4231	2025-01-15	20	11.1002/1000/16026

Keywords

Capability, IoT, requirement, sensors, vision application, visual sensors, workflow.

* To access the Recommendation, type the URL <https://handle.itu.int/> in the address field of your web browser, followed by the Recommendation's unique ID.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, and 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 <https://www.itu.int/ITU-T/ipr/>.

© ITU 2025

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

Table of Contents

		Page
1	Scope	1
2	References.....	1
3	Definitions	2
	3.1 Terms defined elsewhere	2
	3.2 Terms defined in this Recommendation.....	2
4	Abbreviations and acronyms	2
5	Conventions	3
6	Overview of IoT for vision	3
7	Requirements of IoT for vision.....	5
	7.1 Data acquisition	5
	7.2 Data transmission	7
	7.3 Data analysis.....	7
	7.4 Security and privacy	9
	7.5 Management	9
8	Capability framework of IoT for vision.....	10
	8.1 Application layer	10
	8.2 Service support and application support layer.....	10
	8.3 Network layer	11
	8.4 Device layer.....	12
	8.5 Security capabilities.....	13
9	Workflow of IoT for vision applications	13
	Appendix I – Use cases of IoT for vision	15
	I.1 Smart community applications	15
	I.2 Defect detection application in industries	15
	I.3 Livestock applications	16
	I.4 Autonomous driving.....	16
	Bibliography.....	18

Recommendation ITU-T Y.4231

Requirements and capability framework of the Internet of things for vision

1 Scope

This Recommendation specifies the requirements and capability framework of the Internet of things (IoT) for vision, complementing the developments on common requirements and framework of the IoT in terms of vision-related systems. The requirements of the IoT for vision are independent of any specific application.

The scope of this Recommendation includes:

- requirements of IoT for vision;
- capability framework of IoT for vision;
- typical workflow of IoT for vision applications.

Some relevant use cases of the Internet of things for vision are provided in the Appendix.

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.264] Recommendation ITU-T H.264 (V15) (2024), *Advanced video coding for generic audiovisual services*.
- [ITU-T H.265] Recommendation ITU-T H.265 (V10) (2024), *High efficiency video coding*.
- [ITU-T H.266] Recommendation ITU-T H.266 (V3) (2023), *Versatile video coding*.
- [ITU-T M.3030] Recommendation ITU-T M.3030 (2002), *Telecommunications Markup Language (tML) framework*.
- [ITU-T P.1202.2] Recommendation ITU-T P.1202.2 (2013), *Parametric non-intrusive bitstream assessment of video media streaming quality – Higher resolution application area*.
- [ITU-T T.805] Recommendation ITU-T T.805 (2012) | ISO/IEC 15444-6:2013, *Information technology – JPEG 2000 image coding system: Compound image file format*.
- [ITU-T Y.4101] Recommendation ITU-T Y.4101/Y.2067 (2017), *Common requirements and capabilities of a gateway for Internet of things applications*.
- [ITU-T Y.4113] Recommendation ITU-T Y.4113 (2016), *Requirements of the network for the Internet of things*.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 access network [b-ITU-T Q.1742.11]: A network that connects access technologies (such as a radio access network) to the core network.

3.1.2 core network [b-ITU-T Y.101]: A portion of the delivery system composed of networks, systems equipment and infrastructures, connecting the service providers to the access network.

3.1.3 device [b-ITU-T Y.4000]: With regard to the Internet of things, this is a piece of equipment with the mandatory capabilities of communication and the optional capabilities of sensing, actuation, data capture, data storage and data processing.

3.1.4 gateway [ITU-T Y.4101]: A unit in the Internet of things which interconnects the devices with the communication networks. It performs the necessary translation between the protocols used in the communication networks and those used by devices.

3.1.5 Internet of things (IoT) [b-ITU-T Y.4000]: A global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies.

NOTE 1 – Through the exploitation of identification, data capture, processing and communication capabilities, the IoT makes full use of things to offer services to all kinds of applications, whilst ensuring that the security and privacy requirements are fulfilled.

NOTE 2 – From a broader perspective, the IoT can be perceived as a vision with technological and societal implications.

3.2 Terms defined in this Recommendation

None.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

AI	Artificial Intelligence
AVC	Advanced Video Coding
BPP	Bit Per Pixel
BPS	Bit Per Second
CNN	Convolutional Neural Network
CSV	Comma Separated Values
DASH	Dynamic Adaptive Streaming over HTTP
GPU	Graphics Processing Unit
HAS	Hypertext Transfer Protocol Based Adaptive Streaming
HDR	High Dynamic Range
HEVC	High Efficiency Video Coding
IMT	International Mobile Telecommunication
IoT	Internet of Things
JSON	JavaScript Object Notation

MJPEG	Motion Joint Photographic Expert Group
ML	Machine Learning
PaaS	Platform as a Service
PII	Personally Identifiable Information
QoS	Quality of Service
RFID	Radio Frequency Identification
VVC	Versatile Video Coding
XML	Extensible Markup Language

5 Conventions

In this Recommendation:

The keywords "is required to" indicate a requirement which must be strictly followed and from which no deviation is permitted if conformance to this document is to be claimed.

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.

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/service provider. Rather, it means the vendor may optionally provide the feature and still claim conformance with the specification.

6 Overview of IoT for vision

An IoT for vision system is characterized by the use of digital cameras as the primary sensors complemented by other type of sensors in order to perceive objects of interest for diverse applications. It combines IoT, network, and artificial intelligence (AI) technologies to incorporate visual information and related services into the IoT and to support advanced visual applications. So that visual information can be timely and accurately acquired, transmitted and analysed. The IoT for vision system can receive the user instructions and achieve control of sensors and actuators. Since their conception, the usage of digital cameras in IoT deployment has continued to grow and today it is practically pervasive in almost all areas of life.

IoT for vision plays an important role in smart cities, traffic crossroads, and the densely populated areas in the city centre of the main cities from all over the world have large numbers of cameras to capture visual information. However, the visual data are usually not well utilized, as a large amount of the visual data is stored in terminal devices, edge platform, and cloud platform for backup purposes. With AI technical development, the most important objects including pedestrians and vehicles can be detected for different tasks, thereby playing an important role in urban management and early warning. IoT for vision can assist in pedestrian and vehicle management and provide relevant information for possible traffic jams and subsequent measures.

The following IoT technical components are identified to support IoT for vision systems, as shown in Figure 6-1:

- Device [b-ITU-T Y.4000]: Visual sensors and other type of sensors are used for data acquisition. Visual sensors are the devices collecting visual raw data. Visual sensors include, but are not limited to, 2D cameras, 3D cameras, and point cloud cameras [b-Zahner].

Visual sensors may be complemented by other sensors which can provide other information in support of the data analysis.

NOTE 1 – Other sensors include, but are not limited to, global navigation satellite system (GNSS) sensors, radio frequency identification (RFID) sensors, infrared sensors, and pressure sensors. Due to