ITUPublications
Recommendations

International Telecommunication Union
Standardization Sector

Recommendation

ITU-T F.740.7 (09/2023)

SERIES F: Non-telephone telecommunication services

Multimedia services

Reference architecture and requirements for a mobile terminal computational photography system

ITU-T F-SERIES RECOMMENDATIONS

Non-telephone telecommunication services

TELEGRAPH SERVICE	F.1-F.109
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	F.110-F.159
Mobile services and multidestination satellite services	F.110-F.159
TELEMATIC SERVICES	F.160-F.399
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	F.550-F.599
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.740.7

Reference architecture and requirements for a mobile terminal computational photography system

Summary

Recommendation ITU-T F.740.7 specifies the reference architecture for a mobile terminal computational photography system, which includes three functional blocks, namely the optical imaging block, image processing block and application block. This Recommendation also describes requirements for the mobile terminal computational photography system.

This Recommendation can facilitate cooperation between manufacturers, and enhance developing efficiency, so as to provide end users with better photography experiences. The aim is to define the reference architecture and requirements for the mobile terminal photography system.

History *

Edition	Recommendation	Approval	Study Group	Unique ID
1.0	ITU-T F.740.7	2023-09-13	16	11.1002/1000/15617

Keywords

Computational photography, mobile terminal, MTCPS, reference architecture, requirements.

* 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, 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.740.7

Reference architecture and requirements for a mobile terminal computational photography system

1 Scope

This Recommendation specifies the reference architecture for a mobile terminal computational photography system, which includes an optical imaging block, image processing block and application block, as well as a description of each functional block. This Recommendation also describes requirements for the mobile terminal computational photography system.

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.

None.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

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

3.1.2 application [b-ITU-T Y.101]: A structured set of capabilities, which provide value-added functionality supported by one or more services.

3.1.3 mobile terminal [b-ITU-T X.1121]: An entity that has wireless network access function and connects a mobile network for data communication with application servers or other mobile terminals.

3.1.4 service [b-ITU-T Y.101]: A structure set of capabilities intended to support applications.

3.2 Terms defined in this Recommendation

This Recommendation defines the following term:

3.2.1 mobile terminal computational photography system (MTCPS): A system with optical devices and computing units, which realizes computational photography relevant applications through the cooperation of hardware and software on a mobile terminal.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

3D 3-Dimensional

API Application Programming Interface

AR Augmented Reality

HDR High Dynamic Range

ISP Image Signal Processor

MTCPS Mobile Terminal Computational Photography System

OIS Optical Image Stabilization

SDK Software Development Kit

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.

6 Overview

With the rapid development of imaging and computing technology on mobile terminals such as smart phones, digital photography has become ubiquitous in the field of digital culture, which also promotes the progress of computational photography. The mobile terminal computational photography system (MTCPS) has transformed digital culture by making photography accessible, influencing visual communication through social media, and democratizing creativity while introducing new visual aesthetics. First, the accessibility and ubiquity of MTCPS has enabled widespread participation in photography, contributing to a culture of image or video creation and sharing. Second, MTCPS could integrate with social media platforms, and transform visual communication within digital culture, by allowing users to process photography contents instantly, and to convey thoughts and emotions. Finally, MTCPS has democratized creativity, enabling individuals to express themselves artistically without specialized equipment or technical expertise. The capabilities of MTCPS have also introduced new visual aesthetics, shaping the visual language and artistic practices of digital culture.

In the process of optical imaging on mobile terminals, the components used for imaging are limited by size and power consumption and the hardware for receiving optical signals and controlling the hardware is quite different from the traditional photography equipment. Under some complicated circumstances, problems may arise when converting the optical signals into digital images via the hardware imaging system, such as severe image noise in low-light environments, narrow field of view, frustrating effects under backlight conditions, and low quality of distant imaging. These problems may become a hindrance for day-to-day photography of common mobile users and for some specific industry image applications.

Mobile terminal computational photography system (MTCPS) refers to a system that creates image applications through imaging control, image computing and processing, and that combines artificial intelligence (AI) and mobile Internet technology to reconstruct scenes, light and objects relying on limited optical devices and computing units on the mobile terminal. MTCPS could use different calculation function units on a mobile terminal hardware for digital image signal sensing, computer vision analysis, semantic segmentation, image understanding, and synthetic transformation, which break the limitations of the traditional imaging model to increase and extend the capability of data acquisition in the traditional camera.

This Recommendation specifies a reference architecture and details of each functional block, including the optical imaging block, the image software block and the application block. Through

the cooperation of the hardware and software, the MTCPS can process various image data types and satisfy the applications in typical scenarios such as photography, education, entertainment, robotics, and augmented reality.

In order to improve the development efficiency and quality, and facilitate industrial collaboration, the basic requirements for MTCPS, which include availability, reliability, credibility and scalability, are also described in this Recommendation.

7 Reference architecture for MTCPS

A reference architecture of MTCPS is shown as Figure 1, which includes an optical imaging block, an image processing block, and an application block.

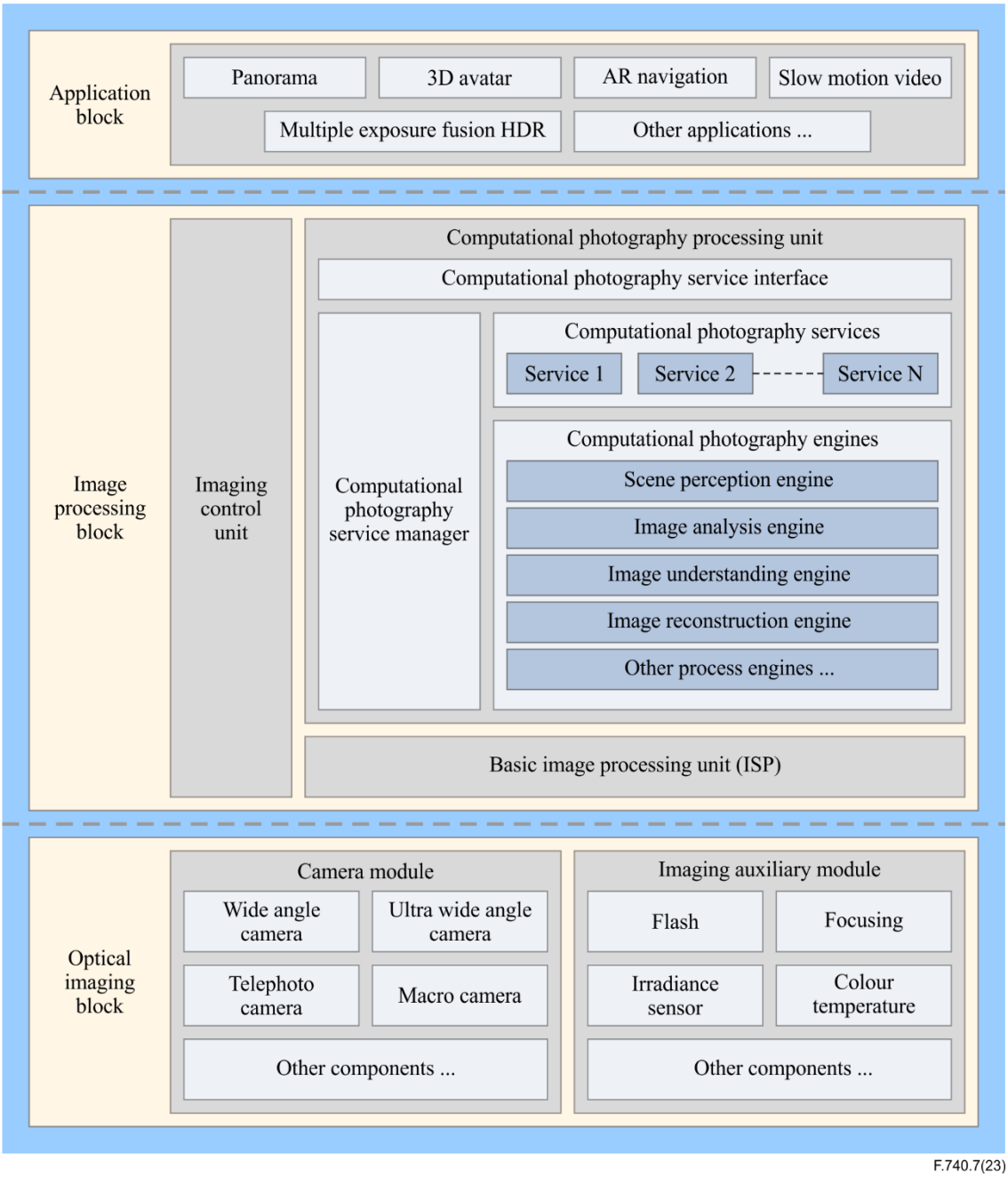

Figure 1 – Reference architecture for MTCPS

When photography applications within the application block are activated by a user, the relevant imaging control instructions will be launched, and the camera module in the optical imaging block will generate and output image data stream, which will be processed by the basic image and computational photography processing unit in the image processing block. Finally, the processed image data will be transmitted back to the application block.

The imaging control instructions include not only the control instructions issued to the optical imaging block (via the imaging control unit) to trigger optical imaging, but also the image processing instructions issued to the basic image processing unit and computational photography processing unit.

The image processing indicates a series of processing works that the image data undergoes, after being generated from the optical sensor, and finally fed back to the applications.

7.1 Optical imaging block

A complete optical imaging block includes the following parts:

– **Camera module.** This module is a necessity for an imaging block, which usually includes but is not confined to sensors, lens, focus actuators, and optical image stabilization (OIS) devices, constituting a camera. It is worth noting that the camera module may contain one or more cameras. With the support of the computational photography processing unit, multi-camera collaboration becomes an important feature that distinguishes MTCPS from other digital camera systems.

– **Imaging auxiliary module.** This module includes components related to the following aspects: flash, focusing, colour temperature and irradiance metering, etc.

7.2 Image processing block

The image processing block includes the following three parts: an imaging control unit, a basic image processing unit and a computational photography processing unit.

– **Imaging control unit.** While translating an optical signal to a digital image signal, the imaging control unit will control the camera hardware according to the real-time statistical information of the image. This unit includes the process of focusing, exposure and flash by controlling lens, sensor and flashlight.

– **Basic image processing unit.** The basic image processing unit mainly realizes the common and basic image processing for digital image signal output from the hardware imaging system. This unit usually includes format conversion, dimension transformation (such as resizing or cropping), bad point correction, noise reduction, highlight suppression, backlight compensation, colour enhancement, lens shadow correction, and content encoding, etc.

– **Computational photography processing unit.** The computational photography processing unit mainly carries out complex image processing towards the single or multiple ways of the digital image signal. The computational photography processing unit is the core of the MTCPS functional module. MTCPS not only supports multi-lens, multi-sensor or multi-spectral cameras to generate better image quality, but also includes several image processing engines to implement the more advanced computer vision algorithms and deep learning algorithms. These technologies obtain more accurate image information and provide the adaptive function combination for different capture and processing demands. For example, MTCPS could provide specialized effects like 3D avatar, AI moon mode or multi-camera images fusion.

  MTCPS ensures flexibility and scalability in development, through the structured interface and internal modules in the computational photography processing unit, providing

convenience for development on the one hand, and ease of management on the other. The computational photography processing unit comprises the following entities:

- **Computational photography service interface.** This interface is the entrance of control instructions and image data between the computational photography processing unit and the peripheral units. It also provides an interface for development (via SDK or API) and management, and hides complicated details that users may not need to know.
- **Computational photography service manager.** The computational photography service manager forwards control instructions to one or more related computational photography engines, organizing and scheduling the computational photography algorithms distributed within each engine to run dynamically in the form of threads or processes.
- **Computational photography engines.** There are mainly four computational photography engines, namely the scene perception engine, the image analysis engine, the image understanding engine and the image reconstruction engine. Each computational photography engine contains a variety of related computational photography algorithms, providing corresponding image processing abilities, and to some extent, extending the capabilities of digital cameras in photography.
- **Scene perception engine** identifies and extracts useful scene information such as atmospheric situation, irradiation conditions, colour temperature, depth range, camera or subject motion status, etc., for subsequent processing by other engines.
- **Image analysis engine** analyses image and relevant information to perform various computer vision tasks, for instance, object detection, recognition and tracking, image segmentation, etc.
- **Image understanding engine** is responsible for integrating the low-level image information and high-level semantic information from other engines to perform photographic scene understanding and decision-making, for instance on, events happening in the scene, behaviour of subject, focus setting of the tracking object, best shutter timing, etc.
- **Image reconstruction engine** reconstructs the scene, light and the subject being photographed, by integrating the information and decision from other engines, presenting a higher quality or more favourable image to the end user.

7.3 Application block

The application block, abstracted as the one part of the reference architecture with which users interact directly, plays an important role in MTCPS. First, the application block contains a variety of applications, such as dual-camera bokeh, panorama, slow motion video, 3D avatar, AI moonshot, multiple camera fusion, and so on, to meet demands in various application scenarios. In addition, the application block also provides application management, content display, resource storage, network interconnection, social sharing and other common capabilities of mobile terminals.

8 Requirements for MTCPS

There are mainly five types of requirements for MTCPS, which are related to general requirements (including use, development and manufacture), service, security, scalability, and reliability, as described below.

8.1 General requirements

8.1.1 Application user requirements

USR-001: MTCPS is recommended to support users to view the information of the deployed applications, including name, version, status, license and other information of the application.

USR-002: MTCPS is recommended to support users to manage applications, including installation, activation, deactivation, upgrade and uninstallation of the application in the system.

USR-003: MTCPS is recommended to support users to manage image contents, including storing, removing, transmitting and other operations.

8.1.2 Developer requirements

DEV-001: MTCPS is recommended to provide a software development kit (SDK) or application programming interface (API), as well as an auditing and deploying mechanism for developing image application and image processing software.

DEV-002: MTCPS is recommended to provide mechanisms to restrict developers from making changes to infrastructure of the hardware and software system, to prevent corruption or functional failure.

8.1.3 Manufacturer requirements

MAN-001: MTCPS is recommended to provide mechanisms to allow manufacturers making changes to infrastructure of the hardware and software system, to support operations including addition, deletion, modification, replacement, or providing of customized functional modules.

8.2 Service requirements

8.2.1 Requirements for computational photography imaging control services

SRV-001: MTCPS is recommended to provide computational photography requests parsing service for image applications, to translate the abstract photography requests to specific imaging control commands.

SRV-002: MTCPS is recommended to provide imaging control services. The service controls the cameras and related auxiliary modules by receiving and executing imaging control commands.

8.2.2 Requirements for computational photography image processing services

SRV-003: MTCPS is recommended to provide basic image processing services, such as automatic focusing, automatic exposure, automatic white balance, basal colour correction, noise reduction, etc.

SRV-004: MTCPS is recommended to provide computational photography related image processing services, such as scene depth detection, object detection and recognition, semantic classification, semantic segmentation, multi-frame noise reduction, multi-frame fusion, etc.

SRV-005: MTCPS is recommended to adopt image data transmission mechanisms to ensure efficient and reliable transmissions within and among the optical imaging block, image processing block and application block.

8.2.3 Requirements for algorithm services in computational photography engines

SRV-006: MTCPS is recommended to provide application services for computational photography algorithm modules, including but not limited to query, selection, and combination, etc.

SRV-007: MTCPS is recommended to provide management services for computational photography algorithm modules, including but not limited to loading, scheduling, unloading, and communication, etc.

SRV-008: MTCPS is recommended to provide update services for computational photography algorithm modules.

8.3 Security requirements

8.3.1 System security requirements

SEC-001: MTCPS is required to provide security assurance during the system operation to avoid health or safety hazards caused by hardware or software failures.

SEC-002: MTCPS is recommended to provide security mechanisms to prevent malicious software from affecting the system operation.

8.3.2 Content security requirements

SEC-003: MTCPS is recommended to provide security mechanisms to protect end-user's privacy from unauthorized access.

SEC-004: MTCPS is recommended to provide security mechanisms to ensure the integrity and copyrights of the image application software, image processing software, and end-users' private content from data corruption or unauthorized use.

8.4 Scalability requirements

SCA-001: MTCPS is required to support the scalability of hardware devices, image processing software and image application software.

8.5 Reliability requirements

REL-001: MTCPS is recommended to pass the system operation stability test to ensure the normal operation of the system under heavy load for a period of time.

REL-002: MTCPS is recommended to support system suspension, sleeping or entering low power consumption mode for a period of time, and could be quickly switched to normal mode when necessary.

Appendix I

Typical application scenarios of MTCPS

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

Typical application scenarios of MTCPS are described in this appendix.

I.1 Mobile terminal photography scenarios

People can use MTCPS-based terminals (such as smart phones) to obtain higher-quality photography effects. It is a common situation to take photos in dark and backlit environments. With the denoising, brightening, and deblurring functions provided by MTCPS, users can get clearer, brighter and wider dynamic ranges photos, even under less ideal lighting conditions.

People can also use MTCPS-based terminals to achieve various photography effects, such as colour style filters, lighting effects, portrait beautification, panoramic stitching, etc., thereby enhancing the photography experience.

I.2 Education application scenarios

Students can use MTCPS-based terminals to learn unfamiliar vocabularies independently. After taking pictures of the given text materials, words or sentences, from which information will be extracted via the text detection and recognition functions. Together with the voice synthesis and pronunciation functions (maybe from 3rd party applications), students can learn new vocabularies easier.

Teachers can use MTCPS-based terminals to achieve rapid marking and correcting of students' assignments or examination submissions. After taking pictures of the submission (e.g., examination paper), and using the detection and recognition functions to locate and identify the contents submitted by students, inspection could be completed automatically, thereby reducing the teachers' burden.

Terminals with MTCPS could also be used to perform non-contact photographing and scanning of a specific object, and then perform the three-dimensional reconstruction of the object to build a three-dimensional model that can be further displayed and printed for teaching purposes.

I.3 Entertainment application scenarios

People can use MTCPS-based terminals to take pictures, and generate interesting photographic effects, such as cartooning and stylizing scenes or characters. The user could also add, delete and modify the background and foreground objects in the scene. After that, the user may share the edited content to others through the network transmission function.

People can also use MTCPS-based terminals for posture and motion detection to achieve more interesting interactions. Through the detection of the key points, the skeleton information could be extracted to estimate the posture and movement of a certain person, which can be used to control other software programs or obtain feedback (visually or acoustically) to realize more interesting human-computer or interpersonal interactions.

I.4 Telecommuting scenarios

More and more people have to work from home because of the pandemic. MTCPS-based terminals could provide background blur and virtual background functions to protect people's home privacy. Furthermore, relying on the portrait segmentation capability provided by MTCPS-based terminals, the video conference or remote meeting can integrate the speaker and the presentation content

within the same display window to perform a more professional presentation and create a better meeting experience.

I.5 Robot application scenarios

MTCPS can also be built into robots to realize scene perception and navigation. Firstly, features and landmarks of a certain target could be detected by a monocular or multi-view vision system. Secondly, taking advantage of above information, a scene depth calculation and 3D reconstruction system can easily identify obstacles and rebuild surrounding environment. Finally, after a series of complicated calculations using the information of obstacles, robot positioning and navigation can become possible.

I.6 Augmented reality scenarios

MTCPS-based terminals can also be used in augmented reality (AR) related applications. Unlike traditional two-dimensional image capturing techniques, MTCPS can reconstruct the 3D information and superimpose virtual information of certain objects, providing an immersive experience when capturing on mobile terminals.

Bibliography

[b-ITU-T X.1121] Recommendation ITU-T X.1121 (2004), *Framework of security technologies for mobile end-to-end data communications*.

[b-ITU-T Y.101] Recommendation ITU-T Y.101 (2000), *Global Information Infrastructure terminology: Terms and definitions*.

[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