ITUPublications
Recommendations

International Telecommunication Union
Standardization Sector

Recommendation

ITU-T F.746.17 (12/2022)

SERIES F: Non-telephone telecommunication services

Multimedia services

Requirements for media processing services

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.746.17

Requirements for media processing services

Summary

Recommendation ITU-T F.746.17 identifies the functional requirements for the media processing services. In particular, the scope of this Recommendation includes functional requirements and application scenarios. Media processing services utilize a set of techniques including cloud computing, computing resource virtualization, and job queue processing to dynamically control and manage computing resources which improves scalability, flexibility and availability. This Recommendation specifies the functional requirements of general requirements, service provision requirements, service management requirements, security considerations, etc.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T F.746.17	2022-12-14	16	11.1002/1000/15191

Keywords

Artificial intelligence (AI), audio processing, content delivery network (CDN), media processing service (MPS), video services.

* 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.746.17

Requirements for media processing services

1 Scope

This Recommendation identifies the requirements for media processing services. In particular, the scope of this Recommendation includes functional requirements and application scenarios.

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 artificial intelligence (AI) [b-ITU-T F.749.13]: An interdisciplinary field, usually regarded as a branch of computer science, dealing with models and systems for the performance of functions generally associated with human intelligence, such as reasoning and learning.

3.1.2 cloud computing [b-ITU-T Y.3500]: Paradigm for enabling network access to a scalable and elastic pool of shareable physical or virtual resources with self-service provisioning and administration on-demand.

3.1.3 content delivery network (CDN) [b-ITU-T F.750]: A network optimized for delivering digital content.

3.1.4 platform as a service (PaaS) [b-ITU-T Y.3500]: Cloud service category (see clause 3.2.10) in which the cloud capabilities type (see clause 3.2.4 in [b-ITU-T Y.3500]) provided to the cloud service customer (see clause 3.2.11 in [b-ITU-T Y.3500]) is a platform capabilities type (see clause 3.2.31 in [b-ITU-T Y.3500]).

3.2 Terms defined in this Recommendation

This Recommendation defines the following term:

3.2.1 media processing service (MPS): A cloud service for transcoding multimedia data, which can provide a cost-effective, easy-to-use, elastic, and highly scalable method to convert audio and video into formats suitable for playing on PCs, TVs and mobile interfaces, and can also conduct deep learning based technology on massive data to analyse the content including text, voice and other scenarios of the media in multiple modes, and to achieve intelligent review, content understanding, intelligent editing, etc.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

3GP	3rd Generation Partnership Project (3GPP) file format
AI	Artificial Intelligence
AR	Augmented Reality
ASIC	Application Specific Integrated Circuit
AVI	Audio Video Interleaved
CDN	Content Delivery Network
CPU	Central Processing Unit
DAG	Direct Acyclic Graph
DRM	Digital Rights Management
FLV	Flash Video
FPGA	Field Programmable Gate Array
GPU	Graphics Processing Unit
HD	High-Definition
IO	Input/Output
M3U8	Moving Picture Experts Group Audio Layer 3 Uniform Resource Locator-UTF8
MP4	Moving Picture Experts Group 4
MPG/MPEG	Moving Picture Experts Group
MPS	Media Processing Service
PaaS	Platform as a Service
PGC	Professionally Generated Content
PIP	Picture in Picture
SD	Standard-Definition
UGC	User Generated Content
URL	Uniform Resource Locator
UTF-8	Unicode Transformation Format
VR	Virtual Reality

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.

– The keywords "can optionally" and "may" indicate an optional requirement which 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/service provider. Rather, it means the vendor may optionally provide the feature and still claim conformance with the specification.

6 Introduction

With the development of Internet technology, the network traffic brought by multimedia content, especially audio and video services has had an explosive growth. Meanwhile, video services have been widely used in entertainment, finance, education, medical and other fields. 4K/8K and virtual reality (VR) / augmented reality (AR) video services will be popular. The rise of live broadcasting and short video services has driven the development of professionally generated content (PGC) and user generated content (UGC). With these massive media contents, how to provide superior service to ensure that media content can be distributed quickly, reliably, and safely is the most significant issue to be solved currently.

Traditionally, content delivery network (CDN) is used to deliver content from the head-end of content sources to the terminal devices. CDN only transmits media content without any processing. With these requirements of new video services, media contents are required to be processed before transmission through the CDN. For example, media contents are required to be transcoded adaptively according to the quality of the network. UGC contents should not contain specific features. Some image and text contents are required to be detected, located, and identified to provide personalized service for consumers.

Regarding the huge amount of video data and the variety of media processing, the way of processing in the local computing system cannot meet the current computing resource requirements of video services. At present, computing technology is widely applied. Media processing services are based on the concept of cloud computing technology. Media processing services provide a cost-effective, easy-to-use, elastic, on-demand, and highly scalable method to process audio and video in various video services.

In a typical media processing service flow from end to end, video contents are captured by the content provider. After the video is uploaded to the cloud server, media processing services begin to work. In order to adapt to different network environments and terminals, the video will be usually transcoded in the cloud. Some personalized value-added services are provided, which include data statistics service, content review service, real-time watermarking service and so on. After media processing, the video contents are distributed and speeded up to the terminal devices by the CDN. Video contents are decoded and played on the terminal devices.

Figure 6-1 shows the reference model for media processing services (MPSs).

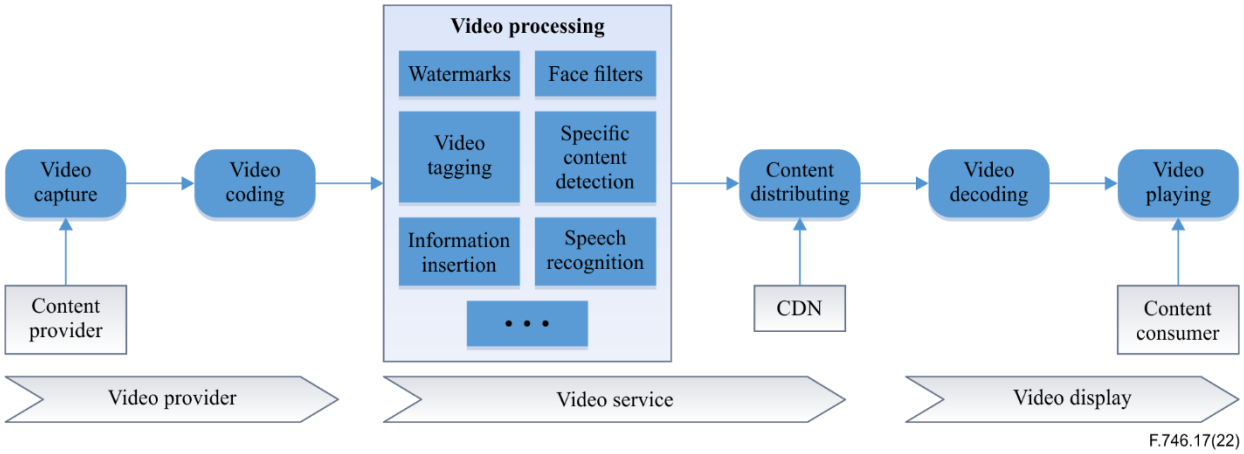

Figure 6-1 – Reference model for media processing services

7 Overview of media processing services

Media processing services utilize a set of techniques including cloud computing, computing resource virtualization, and job queue processing to dynamically control and manage various computing resources, which improves scalability, flexibility and availability.

Figure 7-1 illustrates the typical framework of media processing services.

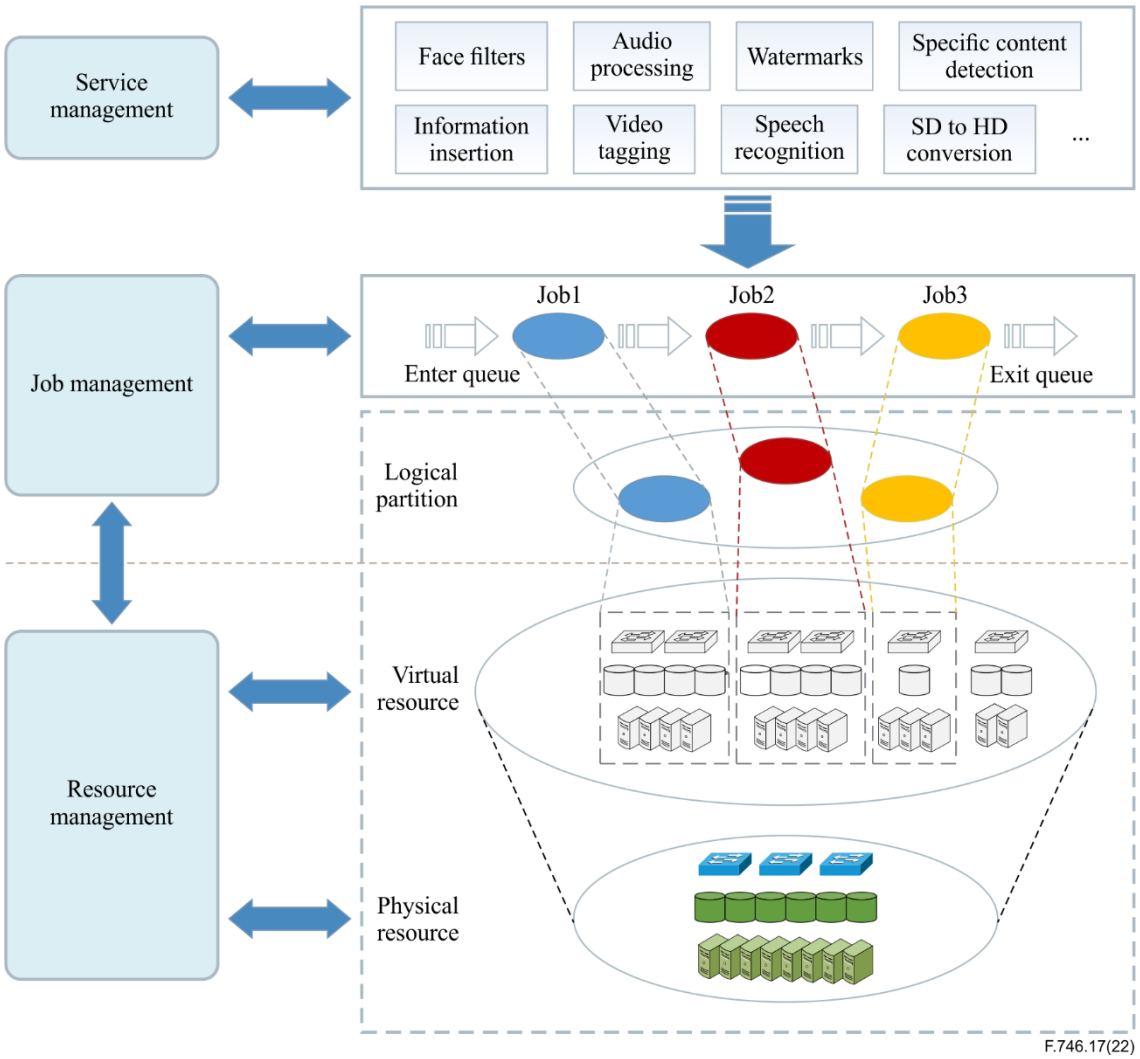

Figure 7-1 – Typical framework of media processing services

A typical media processing service generally consists of many physical resource nodes which are located in a resource pool supported by the cloud computing technology. The nodes are deployed as physical resources such as servers and storage which include many types of computing units, such as the central processing unit (CPU), graphics processing unit (GPU), field programmable gate array (FPGA), and the application specific integrated circuit (ASIC). These computing units are the main hardware to execute various media processing services. The physical resources are abstracted to the virtual resources such as virtual machines and virtual storage, which form a virtual resources pool utilizing virtualization technology. Computing units are virtualized into many virtual computing units. In order to provide service to multiple vendors, virtual resources and virtual computing units are aggregated and isolated according to the vendors' service requirements. Every logical partition is allocated to a specific job. Resource management capability is responsible to manage physical resources and virtualized resources.

A job is a media processing job according to the vendor's requirements. A job consists of three main features which are input resources, media processing service type and parameters, and output video or data. Input resources are original media such as videos, files, or streams that vendors want to process. Media processing service types are mentioned in scenarios while parameters are set by the vendors according to their requirements. Output videos and data are products of media processing, which are various types or formats according to the requirements.

Job management capability is responsible for managing the life span of a job. It is also responsible for job scheduling management, selecting computing resources, and logical partition to execute jobs based on workload and resource loads. Job scheduling can dispatch jobs to different queues for execution according to information such as media processing service type, job priority, submission time, etc.

Service management capability is responsible for managing all kinds of media processing services. Service management provides a portal for vendors to select their services and set up parameters. Service management submits this information to the job management, job management then configures the related logical partitions and physical nodes, and executes the jobs by scheduling.

The logical partition is managed by job management which controls and configures the logical partition in terms of the vendors' requirements. When the requirements of a logical partition are changing, job management requires resource management to perform functions such as adding or diminishing the virtual resources to the logical partition.

8 Requirements

This clause addresses the requirements for media processing services.

8.1 General requirements

Media processing service is a platform as a service (PaaS) service. It provides various types of services by concatenating or assembling various media related processing function modules. These related modules involve the mux layer, codec-stream layer, and transport protocol layer. It includes various processing conversions such as video image processing, audio processing, etc. All of these functions can be combined. Since each module requires different computing resources, network input/output (IO), storage, etc., thus different hardware environments required for each module need to be scheduled together for parallel or serial processing to meet the various functional requirements of users. Due to the inconsistent processing time of each function module, it may cause them to be influenced by each other. This could lead the entire processing link to be too long in time. Since some types of media processing services need more processing capabilities to be consumed than that of the common services, they can also meet the same efficiency requirements of the media processing. The system is recommended or required to provide the following capabilities:

GR-01: The system is required to support the orchestration capability of the function modules based on the direct acyclic graph (DAG). By freely combining various processors, the functions could get integrity and variety.

GR-02: The system is recommended to support the setting of the efficiency requirements of a certain function. For example, a 30-minute video can be processed within 5 minutes or 30 minutes.

GR-03: The system is required to support heterogeneous capabilities to coordinate different types of devices to complete the same job, and to achieve a balance between performance and cost. Algorithms of media processing may be based on different computing units and hardware (such as CPU, GPU, FPGA, ASIC, etc.) It is necessary to complete specific processing sub-job by the corresponding computing units and hardware, and finally serially or parallel coordinate these processing results to complete a job.

In general, the system is a distributed system, which is required to support streaming and parallel processing.

8.2 Service provision requirements

SPR-01: The system is required to provide load-balance capability, which supports the concurrent processing of multiple jobs and avoid excessive load on a single hardware node, which affects the processing efficiency.

SPR-02: The system is recommended to support elastic expansion and contraction of computing resources on time or on demand, thereby improving processing efficiency and reducing costs.

SPR-03: The system is recommended to support grey releasing or upgrading among different versions. The version iteration cycle of cloud services is not fixed, and there are jobs running during release. In order to avoid the risk of potential defects introduced by the released version, which may cause the system to be fully affected, the grey releasing or upgrading can be performed in different arrangements such as groups, intervals, etc., and ways such as splitting dimensions on a user's importance/priorities, job function module, regions of service supporting/infrastructure, etc.

SPR-04: The system is recommended to support the online job duplication ability for verification of functions. In order to ensure the stability of the version release and reduce the introduction of potential defects, A/B testing can be performed for online jobs and the test results can be compared. Some jobs may be suitable for running A/B testing both online, for an original job and offline for a duplicated job.

SPR-05: The system is recommended to support cloud-native deployment capabilities with the container. It is convenient for operation, maintenance, management and control.

SPR-06: The system is required to provide basic capabilities such as logging, end-to-end link inspection, capability usage / occupancy ratio monitoring and alarm, cost and billing, etc.

8.3 Availability and scalability requirements

ASR-01: The system is required to provide unitized (in logical dimension) and rationalized (in physical dimension) deployment capabilities, and can provide high availability at the region-level (divided following service supporting / infrastructure, etc.) When an input file (such as a video segment) comes from region A with lower availability than region B where an MPS system is located, the availability of service after the MPS is processed may not be lower than that in region A. When the input file (such as a video segment) comes from region A with higher availability than region B where the MPS system is located, the availability will be the same as that in region B.

ASR-02: The system is required to support extendibility and can dynamically expand or contract resources to quickly respond to service requirements when service scale quickly increases or decreases.

ASR-03: The system is required to provide the alarm capability with pre-set values for alarm conditions or thresholds.

ASR-04: The system is recommended to provide the back-off and degradation mechanism to ensure the stability of the basic functions for the services as far as possible under abnormal conditions.

ASR-05: The system is recommended to support automatic or manual intervention for failed jobs, then get compensation via quick rerunning. It should prioritize the requirement to restore jobs from users as normal under a flexible strategy.

8.4 Service management requirements

The system is required to provide capabilities such as operable, maintainable, management, control, configurable, traceable and troubleshooting.

SMR-01: Operability, maintainability, manageability and controllability: It is required to include batch submission of jobs, job queue management (with job processing priority) and job cancellation, etc.

SMR-02: Configurability: It is required to be changed only through configuration without modifying the hardware / software.

SMR-03: Trackability: It is required to include the status of the overall capability usage / occupancy ratio and the status of specific jobs.

SMR-04: Troubleshooting capability: According to the end-to-end link log, it is required to support the problem that could be located at the system-level and job-level.

9 Security considerations

The system is required to provide security for accessing data. The address of the input video or output video processed by the system should be the address of the stored object bucket, rather than the uniform resource locator (URL) which is accessible on the public network, to ensure the security of the reading / writing data of the system.

SC-01: The system is required to provide security for the storage service of data. The object storage service selected by the system should have permission configuration management and security control.

SC-02: The system is required to provide security for the operation of the system. As a public cloud service, the system should have a unified account management system consistent with other cloud products and services so that it can build mechanisms such as account permission management, service subscribing or unsubscribing, and anti-attack.

SC-03: The system is required to provide compliance for the user's service data. The system must not store the user's video content, and the logs related to the user's service data shall be accessed only within the specific area/region scope and must not be accessed across borders or areas/regions. The related log data should be cleared periodically. The alarm information related to the logs should also be desensitized.

SC-04: The system is required to provide security for the operation and maintenance of the system itself. The core management and control application service of the system is the operation and maintenance system of the internal platform, which is separate from the public cloud management and control system. The internal operation and maintenance system should be operated under multi-role with different permission levels.

SC-05: The system is recommended to provide encryption of the user's data. The system needs to provide an encryption (i.e., digital rights management (DRM)) function for video content, which can be used as content protection for distribution and playback.

Appendix I

Application scenarios

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

The typical service application scenarios for media processing services are the following.

I.1 Face filtering

Face filtering service offers various filters and methods such as beautifying the facial features for anchors in the Internet broadcast services.

I.2 Audio processing

Audio processing service offers audio mixing, in-ear monitoring, and other audio processing technologies for anchors in the Internet broadcast services.

I.3 Watermarks

Watermarks service helps to add an identifying image or text as watermarks in the output video to safeguard video copyright for content providers.

I.4 Specific content detection

Specific content detection service offers intelligent identification of specific content to reduce the risk of violations for content providers and users.

I.5 Information insertion

Information insertion services can insert supplemental information in live streaming, which is suitable for live streaming quizzes, presentations, and other scenarios.

I.6 Video tagging

Video tagging service offers the ability to provide video classification and multi-dimension content tags, which is suitable for personalized recommendation and searching services for content providers.

I.7 Speech recognition

Speech recognition service offers multiple speech recognition solutions to transfer audio information into texts in various scenarios, such as meetings, courts and interviews.

I.8 SD to HD conversion

Standard-definition (SD) to high-definition (HD) conversion service is able to convert standard-definition videos to high-definition videos by removing the compression noise of transcoding technology.

I.9 Video smart cover

Video smart cover service offers an automatic capture of wonderful pictures such as the cover of the video content by artificial intelligence (AI) analysis, which can be applied to the automatic cover selection of new and old videos and improve the click rate and user experience of the video.

I.10 Video format converting

In order to be compatible with different demands for video storage, transferring and playing, a video format converting service is needed. Video format includes a 3rd generation partnership project

(3GPP) file format (3GP), audio video interleaved (AVI), flash video (FLV), moving picture experts group 4 (MP4), moving picture experts group audio layer 3 uniform resource locator – unicode transformation format (UTF-8) (M3U8), moving picture experts group (MPG/MPEG), etc.

I.11 Resolution enhancement

The resolution of old historical videos is low. It is necessary to convert the low resolution of the original video into high resolution, which is called resolution enhancement.

I.12 Video clipping and stitching

Video clipping and stitching service offers multiple videos to be edited, spliced, cut, and processed to make a new video programme.

I.13 Picture in picture (PIP)

PIP service offers the service to embed one video as part of the frame into another video.

I.14 Encrypting video

Encrypting video service offers an encrypted content to avoid specific downloading and disseminating. Services such as online education, paid video, and other paid contents use this service.

Bibliography

[b-ITU-T F.749.13] Recommendation ITU-T F.749.13 (2021), *Framework and requirements for civilian unmanned aerial vehicle flight control using artificial intelligence*.

[b-ITU-T F.750] Recommendation ITU-T F.750 (2005), *Metadata framework*.

[b-ITU-T Y.3500] Recommendation ITU-T Y.3500 (2014), *Information technology – Cloud computing – Overview and vocabulary*.

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