International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

H.626

(11/2019)

SERIES H: AUDIOVISUAL AND MULTIMEDIA SYSTEMS

Broadband, triple-play and advanced multimedia services – Advanced multimedia services and applications

Architectural requirements for a video surveillance system

Recommendation ITU-T H.626

ITU-T

ITU International Telecommunication Union

ITU-T H-SERIES RECOMMENDATIONS

AUDIOVISUAL AND MULTIMEDIA SYSTEMS

CHARACTERISTICS OF VISUAL TELEPHONE SYSTEMS	H.100–H.199
INFRASTRUCTURE OF AUDIOVISUAL SERVICES	
General	H.200–H.219
Transmission multiplexing and synchronization	H.220–H.229
Systems aspects	H.230–H.239
Communication procedures	H.240–H.259
Coding of moving video	H.260–H.279
Related systems aspects	H.280–H.299
Systems and terminal equipment for audiovisual services	H.300–H.349
Directory services architecture for audiovisual and multimedia services	H.350–H.359
Quality of service architecture for audiovisual and multimedia services	H.360–H.369
Telepresence, immersive environments, virtual and extended reality	H.420–H.439
Supplementary services for multimedia	H.450–H.499
MOBILITY AND COLLABORATION PROCEDURES	
Overview of Mobility and Collaboration, definitions, protocols and procedures	H.500–H.509
Mobility for H-Series multimedia systems and services	H.510–H.519
Mobile multimedia collaboration applications and services	H.520–H.529
Security for mobile multimedia systems and services	H.530–H.539
Security for mobile multimedia collaboration applications and services	H.540–H.549
VEHICULAR GATEWAYS AND INTELLIGENT TRANSPORTATION SYSTEMS (ITS)	
Architecture for vehicular gateways	H.550–H.559
Vehicular gateway interfaces	H.560–H.569
BROADBAND, TRIPLE-PLAY AND ADVANCED MULTIMEDIA SERVICES	
Broadband multimedia services over VDSL	H.610–H.619
Advanced multimedia services and applications	**H.620–H.629**
Ubiquitous sensor network applications and Internet of Things	H.640–H.649
IPTV MULTIMEDIA SERVICES AND APPLICATIONS FOR IPTV	
General aspects	H.700–H.719
IPTV terminal devices	H.720–H.729
IPTV middleware	H.730–H.739
IPTV application event handling	H.740–H.749
IPTV metadata	H.750–H.759
IPTV multimedia application frameworks	H.760–H.769
IPTV service discovery up to consumption	H.770–H.779
Digital Signage	H.780–H.789
E-HEALTH MULTIMEDIA SYSTEMS, SERVICES AND APPLICATIONS	
Personal health systems	H.810–H.819
Interoperability compliance testing of personal health systems (HRN, PAN, LAN, TAN and WAN)	H.820–H.859
Multimedia e-health data exchange services	H.860–H.869
Safe listening	H.870–H.879

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

Recommendation ITU-T H.626

Architectural requirements for a video surveillance system

Summary

The purpose of Recommendation ITU-T H.626 is to define a video surveillance system based on IP networks. The video surveillance system provides the display and storage functions of the multimedia (such as video, audio and image) captured by multiple remote cameras over an IP network for multiple users. It also provides other functions, such as remote control, alarm and linkage action, recording and playback.

This Recommendation defines the model, architecture, entities and reference points of the video surveillance system. This Recommendation also describes the hierarchy model for deployment and interworking between the video surveillance system and other multimedia systems.

This revision updates the title of this Recommendation, deletes requirements and service flows, and revises the architecture and reference points.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T H.626	2011-05-14	16	11.1002/1000/11306
2.0	ITU-T H.626 (V2)	2019-11-29	16	11.1002/1000/14109

Keywords

Alarming, camera, IP network, PTZ control, recording and playback, video surveillance.

* 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, 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 TSB patent database at http://www.itu.int/ITU-T/ipr/.

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	3
5	Conventions	5
6	Functional architecture	5
6.1	Functional model	5
6.2	Functional architecture	5
6.3	System interworking models	11
7	Reference points	13
7.1	Description of related protocols	13
7.2	Reference point Pc: CU-PU	15
7.3	Reference point Mp/Mp': MS-PU	15
7.4	Reference point Cm: CCS-MS	15
7.5	Reference point Mc/Mc': MS-CU	15
7.6	Reference point Ps: CCS-PU	15
7.7	Reference point Cs: CCS-CU	15
7.8	Reference point Ds/Ds': MDU-MSU	16
8	Interworking with other multimedia systems	16
8.1	Interworking with videoconferencing	16
8.2	Interworking with IPTV	17
Bibliography		18

Recommendation ITU-T H.626

Architectural requirements for a video surveillance system

1 Scope

The video surveillance system is a telecommunication system focusing on video (and audio) application technology, which is used to remotely capture multimedia (such as audio, video, image and various alarm signals), and present them to end users in a friendly manner (including accessibility aspects), based on a managed broadband network with ensured quality, security and reliability.

The objective of this Recommendation is to define the architectural of the video surveillance system, which meets the needs of video surveillance requirements defined in [ITU-T F.743]. This Recommendation defines the model, functional architecture, entities and reference points of video surveillance 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.

[ITU-T F.743] Recommendation ITU-T F.743 (2019), *Requirements and service description for video surveillance.*

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

[ITU-T G.723.1] Recommendation ITU-T G.723.1 (2006), *Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s.*

[ITU-T G.729] Recommendation ITU-T G.729 (2012), *Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction (CS-ACELP).*

[ITU-T H.264] Recommendation ITU-T H.264 (2019), *Advanced video coding for generic audiovisual services.*

[ITU-T H.265] Recommendation ITU-T H.265 (2019), *High efficiency video coding.*

[ITU-T H.626.3] Recommendation ITU-T H.626.3 (2018), *Architecture for visual surveillance system interworking.*

[IETF RFC 4571] IETF RFC 4571 (2006), *Framing Real-time Transport Protocol (RTP) and RTP Control Protocol (RTCP) Packets over Connection-Oriented Transport.*

[IETF RFC 6086] IETF RFC 6086 (2011), *Session Initiation Protocol (SIP) INFO Method and Package Framework.*

[IETF RFC 6665] IETF RFC 6665 (2012), *SIP-Specific Event Notification.*

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

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

3.1.2 centre control server [ITU-T H.626.3]: A device located at the centre part of the video surveillance system. It is used for centralized system management, service operation management, and access control.

NOTE – The centre control server consists of centre management unit and service control unit.

3.1.3 customer [b-ITU-T M.60]: An entity which receives services offered by a service provider based on a contractual relationship. It may include the role of a network user.

NOTE – In this Recommendation, 'customer' has the same meaning as 'end user'.

3.1.4 electronic program guide (EPG) [b-ITU-T J.200]: A program table that is presented electronically.

3.1.5 functional architecture [b-ITU-T Y.2012]: A set of functional entities, and the reference points between them, used to describe the structure of an NGN. These functional entities are separated by reference points, and thus, they define the distribution of functions.

NOTE 1 – The functional entities can be used to describe a set of reference configurations. These reference configurations identify which reference points are visible at the boundaries of equipment implementation and between administrative domains.

NOTE 2 – The definition is not only applicable to NGNs, but also to other IP packet switch based networks.

3.1.6 functional entity [b-ITU-T Y.2012]: An entity that comprises an indivisible set of specific functions. Functional entities are logical concepts, while groupings of functional entities are used to describe practical, physical implementations.

3.1.7 interface [b-ITU-T Y.101]: A shared boundary between two functional units.

NOTE – An interface is defined by various characteristics pertaining to the functions, physical interconnections, signal exchanges, and other characteristics as appropriate.

3.1.8 media [b-ITU-T Y.101]: Plural of medium.

3.1.9 medium [b-ITU-T Y.101]: a) Specific physical support for transmission or storage of information. b) Type of presentation of information (i.e., video, audio, text, etc.).

3.1.10 reference point [b-ITU-T Y.2012]: A conceptual point at the conjunction of two non-overlapping functional entities that can be used to identify the type of information passing between these functional entities.

NOTE – A reference point may correspond to one or more physical interfaces between pieces of equipment.

3.1.11 role [b-ITU-T Y.110]: The role is a business activity which fits in a value chain. The role is constrained by the smallest scale of business activity which could exist independently in the industry and so a marketplace will exist for every relationship between roles.

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

3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

3.2.1 centre management unit: A device located at the main part of a video surveillance system. It is used for centralized system management, service operation management, etc. The main functions include management, authorization, accounting, charging, location and presence.

3.2.2 customer unit: A device located at the customer part of a video surveillance system and used to present multimedia information (such as audio, video, image, alarm signal, etc.) to the end user.

3.2.3 media distribution unit: A device located at the centre part of a video surveillance system. It is used to transport media from the premises unit (PU) to the customer unit (CU). The main functions include media receiving, media processing, media routing, media transmission, media forwarding and media replication, etc.

3.2.4 media server: A device located at the centre part of a video surveillance system. It is used to forward real-time media stream as well as store, retrieve and replay historical media stream. The media server receives the media stream from premises unit or other media server and forwards the media stream to other customer unit or media server. It consists of media distribution unit and media storage unit.

3.2.5 media storage unit: A device located at the centre part of a video surveillance system. It is used to retrieve, store media and provide a streaming media serving capability. The main functions include media storage, media serving, media indexing and media downloading, etc.

3.2.6 premises unit: A device located at the remote part of a video surveillance system and used to capture multimedia information (such as audio, video, image, alarm signal, etc.) from a surveilled object.

3.2.7 service control unit: A device located at the main part of a video surveillance system. It is used to access service control and signal call control between the PU and the CU. The main functions include access registration, access authentication, identification, authorization, call control, location, presence, and target media serving function selection.

3.2.8 service platform: A series of devices and subsystems located at the centre part of a video surveillance system. It is used to integrate all of the capabilities and provide video surveillance services to customers. The main functions include service-control function, media switching, distribution, storage, and control and management.

3.2.9 surveillant: The individual performing surveillance.

3.2.10 surveilled object: The target (such as site, human, and related environment) on which surveillance is performed.

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

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

AAA Authentication, Authorization and Accounting

AAC Advanced Audio Coding

AVC Advanced Video Coding

CCS Central Control Server

CDN Content Delivery Network

CDR Charging Data Record

CMU	Centre Management Unit
CPG	Control Protocol Gateway
CU	Customer Unit
ECCS	Extended Control Command Set
EMCS	Extended Media Command Set
EPG	Electronic Programme Guide
GUI	Graphical User Interface
HEVC	High Efficiency Video Coding
HVSS	Heterogeneous Video Surveillance System
ID	Identification
IP	Internet Protocol
IPC	IP Camera
IPTV	Internet Protocol Television
MCU	Multimedia Control Unit
MDU	Media Distribution Unit
MG	Media Gateway
MOC	Monitor Centre
MS	Media Server
MSU	Media Storage Unit
NGN	Next Generation Network
NVR	Network Video Recorder
PC	Personal Computer
PTZ	Pan/Tilt/Zoom
PU	Premises Unit
RTP	Real-time Transport Protocol
RTCP	RTP Control Protocol
SCU	Service Control Unit
SDP	Session Description Protocol
SIP	Session Initiation Protocol
SRG	Signal Routing Gateway
STB	Set-Top Box
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
VS	Video Surveillance
VSS	Video Surveillance System

5 Conventions

None.

6 Functional architecture

This clause describes the basic functional model, functional architecture, functional entities and system interworking which pertain to video surveillance (VS).

6.1 Functional model

The basic functional model for VS consists of customer units (CUs), Pus and a service platform. The interface Pi is the interface between Pus and the service platform, and the interface Ci is the interface between CUs and the service platform. This is illustrated in Figure 6-1.

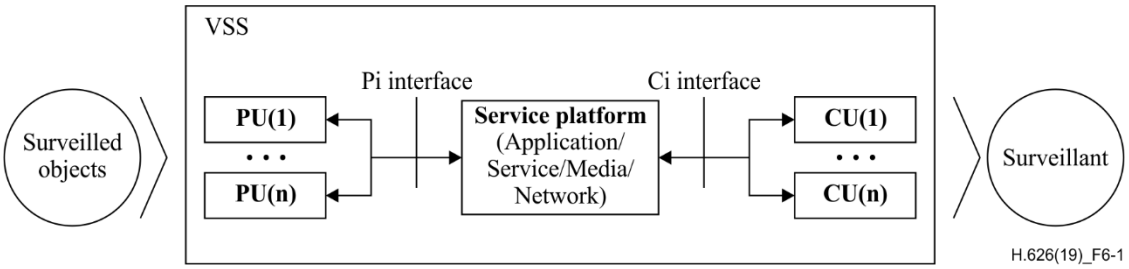

Figure 6-1 – Functional model of video surveillance system

6.2 Functional architecture

This clause provides the functional architecture of the video surveillance system (VSS).

6.2.1 Overview of entities

The VSS functional architecture could be decomposed into eight units:

– Centre management unit (CMU)
– Service-control unit (SCU)
– Media storage unit (MSU)
– Media distribution unit (MDU)
– Premises unit (PU)
– Customer unit (CU)
– Centre control server (CCS)
– Media server (MS)

The units shown in Figure 6-2 are described in clauses 6.2.2 to 6.2.9 and the interfaces in clause 7.

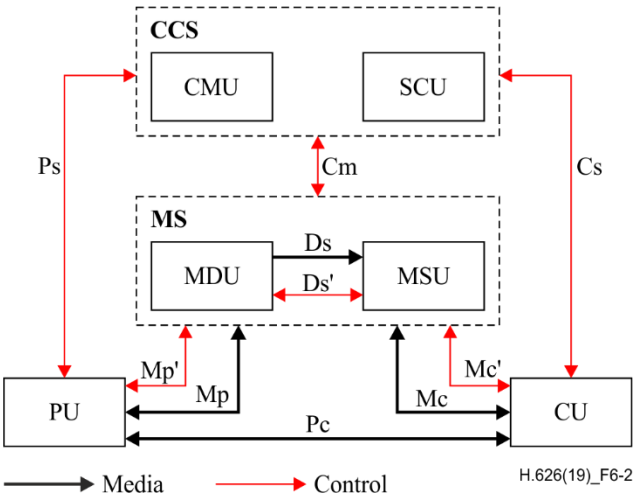

Figure 6-2 – Functional architecture of VSS

6.2.2 CMU: centre management unit

The centre management unit (CMU) is used for centralized system management, service operation management, etc.

Accordingly, the main functions of the CMU include management, authorization, accounting and charging, location and presence. The detailed functional features are:

– Management: managing content, service, network, PU, etc.

– Authorization: authorizing access right.

– Accounting and charging: providing charging data record (CDR) and charging policy, etc.

6.2.3 SCU: service-control unit

The service-control unit (SCU) is used to access service control and signal call control between the PU and the CU.

Accordingly, SCU main functions include access registration, access authentication, identification, authorization, call control, location, presence, and target media serving function selection. The detailed functional features are:

– Access registration: receiving the access registration request from the CU and the PU, and registering them with the registration management server.

– Access authentication: receiving the access authentication request from the CU and the PU, and as an authentication server function, forward (proxy) authentication signals between the authentication, authorization and accounting (AAA) servers.

– Identification: identifying the CU and the PU, and comparing identities with subscriber data and terminal data to determine matching results.

– Authorization: authorize CU and PU access rights.

– Call control: mediate the call control session with the PU and the CU, establish and maintain media transport path information (media routing table, i.e., MRT). The related session methods could include setup, tear down, update, modify and keep live, in order to access, route and relay.

– Location: Locating the PU, if required, providing the placement information of the PU and the CU.

– Presence: presenting the status of the CU and the PU, such as on-line, offline, active, inactive, etc.

– Clock synchronization: providing unified time criteria and clock synchronization control.

– Media serving function selection: selecting one media serving function unit (i.e., media distribution unit and related media storage unit) when one SCU controls multiple media serving function units.

6.2.4 CCS: centre control server

The CCS is used for centralized system management, service operation management, and access control. It consists of the CMU and the SCU.

6.2.5 MDU: media distribution unit

The media distribution unit (MDU) is used to transport media from the PU to the CU.

Accordingly, the main functions of the MDU include media receiving, media processing, media routing, media transmission, media forwarding, media replication, etc. The detailed functional features are:

– Media receiving: receiving media (stream and files) from the PU.

– Media processing: muxing/demuxing, decoding/encoding and transcoding media that is received by the MDU, as required.

– Media routing: seeking media routing information between the PU and the CU in the media routing table in order to decide on the media transport path and destination routing port.

– Media transmission: transmitting media to a CU or multiple CUs on request, according to media routing information.

NOTE – Transmission from one PU to multiple CUs requires conversion from a unicast media channel to a multicast media channel).

– Media forwarding: forwarding media from the current MDU (Domain A) to another MDU (Domain B), in order to access the remote CU (Domain B) under the control of the SCU.

– Media replication: replicating media for point-to-multipoint transmission, if required.

NOTE – The alternative approach is to convert the unicast media channel to a multicast media channel. For this situation, the media replication is not needed.

The MDU internal process is illustrated in Figure 6-3 and the VS end-to-end media transport process is illustrated in Figure 6-4.

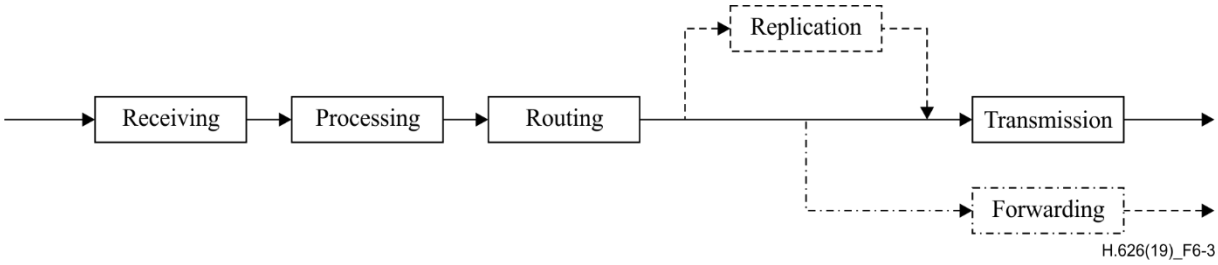

Figure 6-3 – MDU internal process

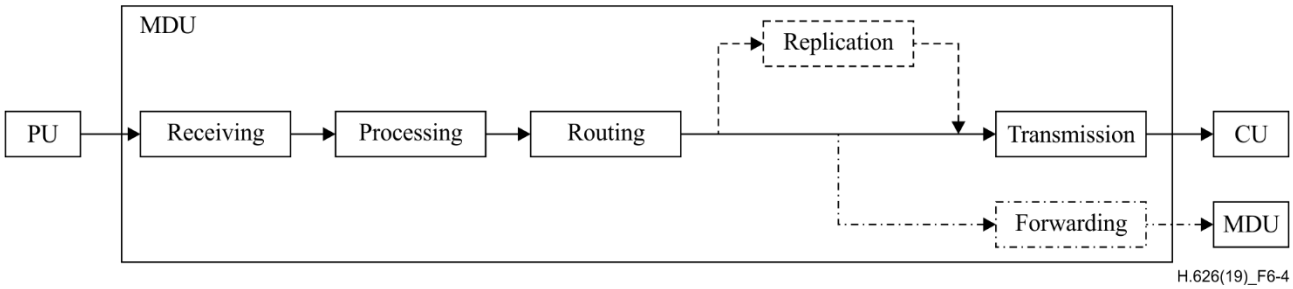

Figure 6-4 – VS E2E media-transport process

The four media transport modes among PU, CU and MDU are:

– Type A: Conversation mode (point-to-point), see Figure 6-5.

– Type B: Distribution mode (point-to-multipoint), see Figure 6-6.

– Type C: Collecting mode (multipoint-to-point), see Figure 6-7.

– Type D: Conferencing mode (multipoint-to-multipoint), see Figure 6-8.

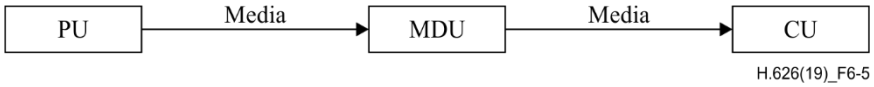

Figure 6-5 – Type A media transport mode (point-to-point mode)

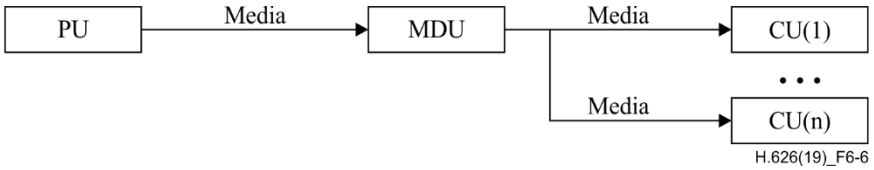

Figure 6-6 – Type B media transport mode (point-to-multipoint mode)

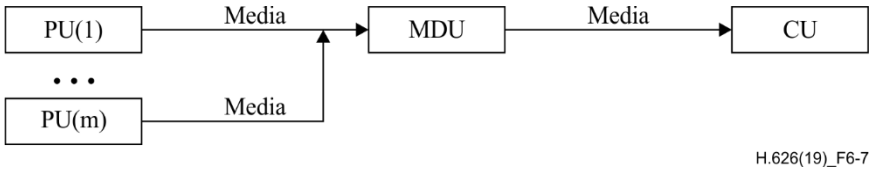

Figure 6-7 – Type C media transport mode (multipoint-to-point mode)

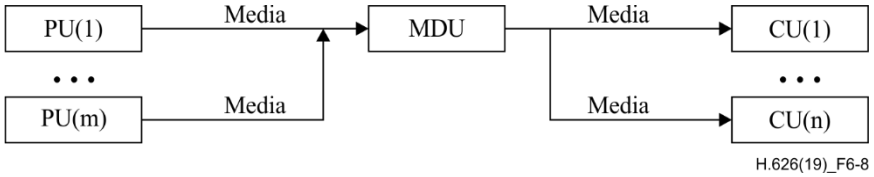

Figure 6-8 – Type D media transport mode (multipoint-to-multipoint mode)

6.2.6 MSU: media storage unit

The media storage unit (MSU) is used to retrieve, store media and provide a streaming media serving capability.

Accordingly, MSU main functions include media storage, media serving, media indexing, media downloading, etc. The detailed functional features are:

– Media retrieval: retrieving audio and/or video media streams and image files from PU and/or MDU.

– Media storage: storing recorded audio, video and image files from PU and/or MDU.

NOTE 1 – The images from the PU can be stored directly in the MSU, without the control of the MDU.

NOTE 2 – Alternatively, images can be stored by the independent image file storage server (IFSS), which is out of the scope of the MSU. For this situation, IFSS could not be controlled by the MDU.

– Media serving: serving audio and video to one CU or multiple CUs with media streaming on request of playback service.

– Media indexing: indexing audio, video and images files when stored for further media searching on request.

The MSU internal process is illustrated in Figure 6-9.

The VS end-to-end media storage and serving process is illustrated in Figure 6-10.

The media transport mode for PU and playback service among MDU, MSU and CU is illustrated in Figure 6-11a and Figure 6-11b.

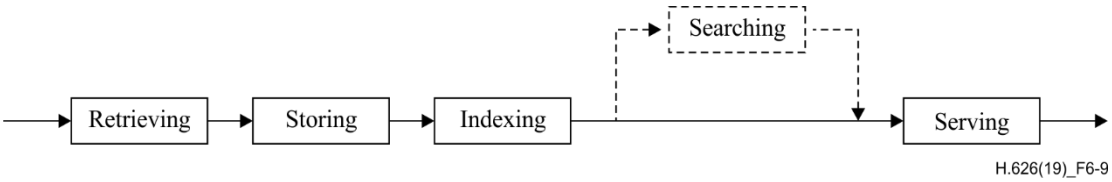

Figure 6-9 – MSU internal process

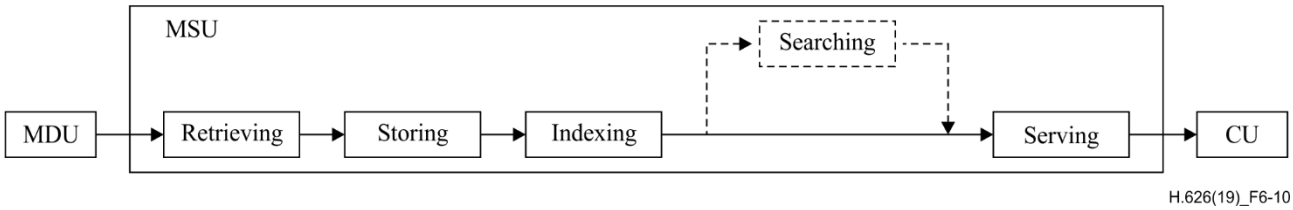

Figure 6-10 – VS end-to-end media storage and serving process

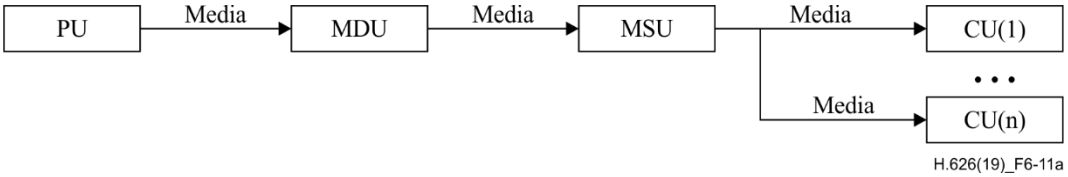

Figure 6-11a – Media transport mode – PU playback service

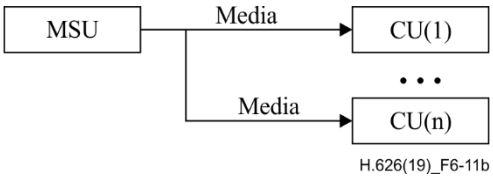

Figure 6-11b – Media transport mode – platform playback service

6.2.7 MS: media server

The MS is used to forward real-time media stream as well as store, retrieve and replay historical media stream. The media server receives the media stream from premises unit or other media server and forwards the media stream to other client unit or media server. It consists of the media distribution unit and the media storage unit.

6.2.8 PU: premises unit

The PU refers to the premise subsystem within the VSS. It implements the following functions:

– Captures multimedia information (such as audio, video, image, alarm signals, etc.) from the surveilled object;

– Encodes multimedia (audio, video, image);

– Outputs alarm signal to external linkage device;

– Parses pan/tilt/zoom (PTZ) command codes and transmit them to the devices in order to control remote direction and zoom;

– Provides network transport, and transmits bidirectional media stream to other entities and alarm signals.

The media storage function is optional for the PU.

Audio decoding and play function is optional for the PU.

Figure 6-12 shows the function composition of the PU.

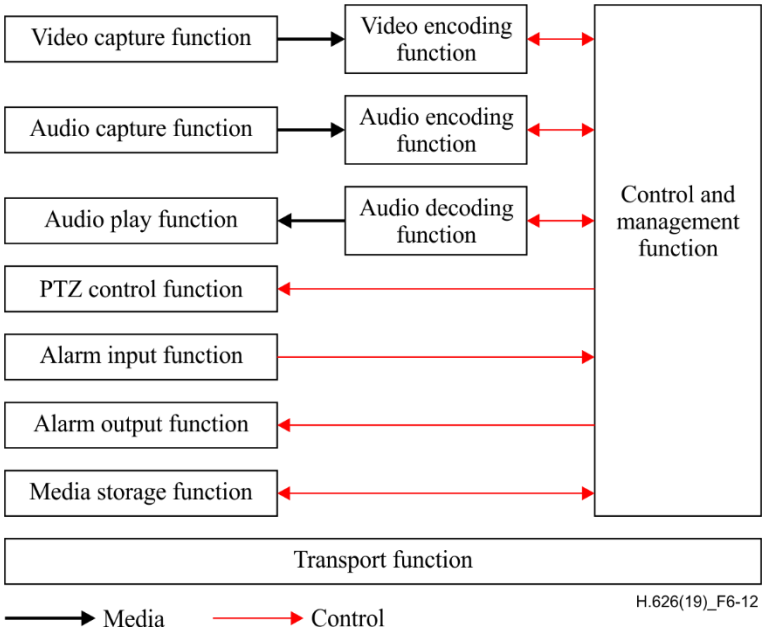

Figure 6-12 – Function composition of the PU

The devices of the PU include video capture devices such as various cameras and audio capture devices like microphones, pan and tilt devices, alarm input and output devices, and audio/video/image encoding devices.

The PU main reference internal process is as follows.

After encoding by an encoding device of the video stream captured by a camera, of a voice stream captured by a microphone, and of alarm information collected by an alarm device, the PU sends the compressed information or media stream to the service platform or the customer unit that requests this information for real-time surveillance. Meanwhile, the PU extracts the control signal sent from the service platform, parsing the command code and transmitting the PTZ control command to pan and tilt the device to implement PTZ control.

6.2.9 CU: customer unit

The CU is the customer unit within the VSS. It is used to present multimedia information (such as audio, video, image, and alarm signal) to the end user. It implements the following functions:

– Multimedia decoding function;

– Audio play and video/image display function;

– Console interface for the end user to operate the VSS.

Figure 6-13 shows the CU functional composition.

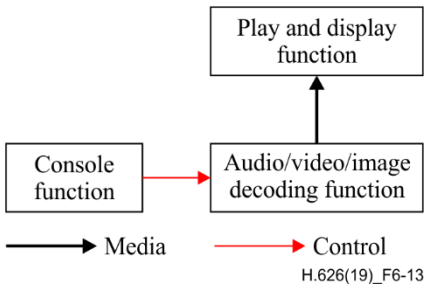

Figure 6-13 – CU functional composition

The CU main reference process is as follows.

After decoding and transcoding the media stream transported from the service platform or the PU by the decoding function entity, the media is output to the play and display function entity for presentation to the end user. The play and display function entity can implement some advanced functions, such as the multi-picture surveillance, multi-screen display, multi-screen alternation (e.g., auto-patrol), caption overlay, and designated-area shielding.

The console function provides the graphical user interface (GUI) to operate and control the VSS for end use. Through the console function, the authorized end user can, e.g., select and display remote video (audio is optional) of one or more surveilled objects, receive alarm signals and trigger linkage actions, set and control the display model, configure the auto-record or alarm conditions, start or stop recoding, and search and play the recorded file.

The CU could be based on personal computer (PC), monitor centre (MOC), mobile phone, embedded video decoder, embedded web page client, Internet protocol television (IPTV) set top box (STB), etc.

6.3 System interworking models

This clause introduces the interworking models for VSS, including interworking model for homogeneous video surveillance systems, interworking model for heterogeneous video surveillance systems, and deployment models for video surveillance interworking.

6.3.1 Interworking for homogeneous video surveillance systems

Figure 6-14 illustrates the interworking model for homogeneous video surveillance systems. As shown in Figure 6-14, VSS A and VSS B are two homogeneous video surveillance systems, and this means that the two systems have the same architecture, transmission protocol, control signals, and data formats. One video surveillance system is composed of CCS, MS, PU, CU as well as other functional entities. For the interworking of homogeneous video surveillance systems, a new system functional entity, which is signal routing gateway (SRG), is used to route the control signals between the two systems, and the media can be received and forwarded through the MSs of the two systems.

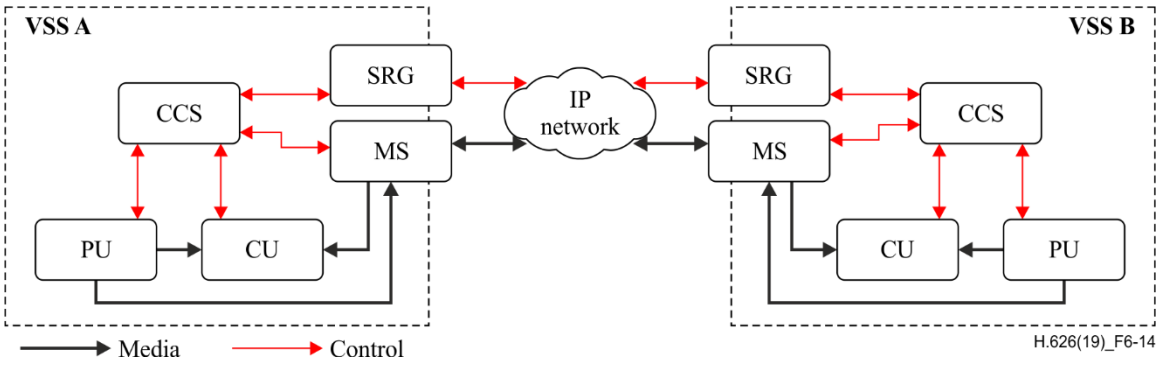

Figure 6-14 – Interworking model for homogeneous video surveillance systems

6.3.2 Interworking for heterogeneous video surveillance systems

Figure 6-15 illustrates the interworking model for heterogeneous video surveillance systems. As shown in Figure 6-15, the heterogeneous video surveillance system (HVSS) has a different architecture, transmission protocol, control signals and data formats from the standardized video surveillance system defined in this Recommendation. The VSS can communicate with the HVSS through the control protocol gateway (CPG) and media gateway (MG). The CPG is used for the switching network transmission protocols, control signals and device addresses between the VSS and the HVSS. The MG is used for switching media transmission protocols and media data coding formats between the VSS and the HVSS.

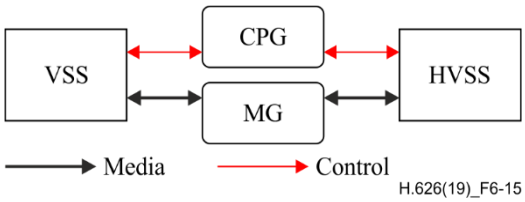

Figure 6-15 – Interworking model for heterogeneous video surveillance systems

6.3.3 Topology models for video surveillance interworking

There are two topology models for interworking among different video surveillance systems, including the peer-to-peer topology model and the hierarchical topology model.

In the peer-to-peer topology model, as shown in Figure 6-16, two adjacent VSSs are in the peer-to-peer relationship. When the surveillance resources need to be shared between two systems, the initiator SRG in one system initiates registration to the target SRG in another system actively. The communication between two systems can only be performed when the initiator SRG has been authenticated by the target SRG. In this flat interworking topology structure model, the signal and media streams are directly forwarded between two systems through SRG and MS respectively.

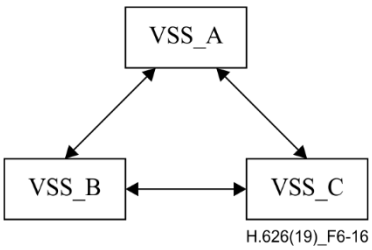

Figure 6-16 – Peer-to-peer topology model

In the hierarchical topology model, as shown in Figure 6-17, two adjacent VSSs are in the superior-subordinate relationship. The subordinate SRG in one system initiates registration to the

superior SRG in another system actively, and communication between the adjacent systems can only be performed when the subordinate SRG has been authenticated by the superior SRG. In this multilevel interworking topology structure model, the signal and media streams are forwarded layer by layer through CPG and MG respectively.

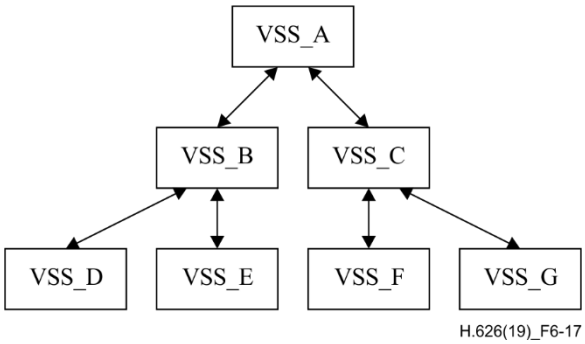

Figure 6-17 – Hierarchical topology model

NOTE – [ITU-T H.626.3] describes the architecture and topology models for video surveillance systems interworking in more detail.

7 Reference points

This clause defines reference points.

7.1 Description of related protocols

7.1.1 General description

The information transmission, switch and control of information such as command, video, audio, etc., within the VSS shall conform to the communication protocols specified in clause 7.1.1. The structure of communication is shown in Figure 7-1.

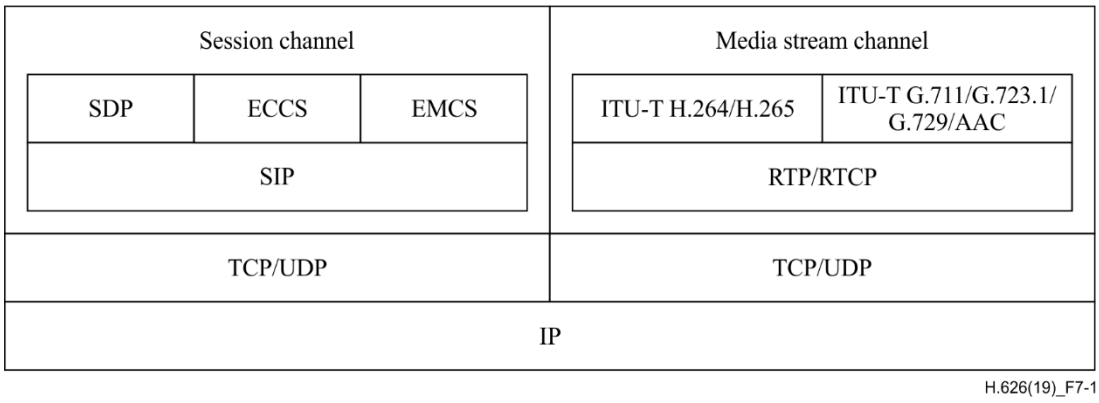

Figure 7-1 – Structure of communication

The VSS shall establish two transmission channels, the session channel and the media stream channel, during video and audio transmission and control. The session channel is used for establishing sessions and transmission system control signals among devices. The media stream channel is used for transmitting video and audio stream, and the compressed video and audio stream is transmitted through real-time transport protocol/RTP control protocol (RTP/RTCP). The network layer of the VSS shall support IP, and the transport layer shall support transmission control protocol (TCP) and user datagram protocol (UDP).

7.1.2 Session protocol requirements

The following are session protocol requirements (SPRs):

SPR-1: Session initiation protocol

The session control of applications as secure registration, live video streaming, historical video and audio playback, etc., adopt the session initiation protocol (SIP) [b-IETF RFC 3261] method, such as REGISTER and INVITE. The historical video and audio playback control is implemented with the INFO method adopt INFO method and package framework [IETF RFC 6086]. The session control of such applications as PU control, information query and alarm event notification and distribution are implemented with the MESSAGE method adopting SIP extensions for instant messaging. SUBSCRIBE and NOTIFY methods adopt SIP specific event notification [IETF RFC 6665].

SPR-2: Session description protocol

The session negotiation and media negotiation between the relevant devices of the network system in the session establishment process adopt session description protocol (SDP), mainly including the session, media and time related information description. The session negotiation and the media negotiation information shall be carried and transported through SIP messages.

SPR-3: Extended control command set

The control commands of the PU control, alarm messages and device directory information of VSS adopt the extended control command set (ECCS). ECCS is carried and transported within the message body of SIP MESSAGE.

SPR-4: Extended media command set

The historical video and audio playback control commands adopt the extended media command set (EMCS) to control play, fast forward, pause, stop, dragging the mouse and play, etc., of video and audio stream. EMCS is carried and transported within the message body of SIP INFO.

7.1.3 Media protocol requirements

The following are media protocol requirements (MPRs):

MPR-1: Media transmission protocol

In VSS, the transmission of media stream shall support RTP, and the PU shall be able to control bit rate peak. RTP load shall apply one of the following two formats: video and audio based on PS packaging or video and audio elementary stream. The transmission of the media stream based on the UDP shall adopt the RTP. The transmission of the media stream based on the TCP shall comply with [IETF RFC 4571], providing timestamp information in real-time transmission and synchronization of various streams. It adopts RTCP to provide a reliable guarantee for data package transport in order as well as flow control and congestion control.

MPR-2: Media compression

The video compression adopts [ITU-T H.264] (AVC), [ITU-T H.265] (HEVC) Recommendation, etc. The audio compression adopts [ITU-T G.711], [ITU-T G.723.1], [ITU-T G.729] and advanced audio coding (AAC) Recommendation.

MPR-3: Media storage and packaging format

The storage and packaging of the video and audio media stream in the VSS shall be in PS format.

7.1.4 VSS ID

All devices (servers, cameras, NVRs, etc.) and user accounts in the VSS should have a unique ID within the network. At the same time, the zone code and device type should be reflected in the ID. The unified ID is 24-digit, which is reflected in Figure 7-2 and adopts the following rules.

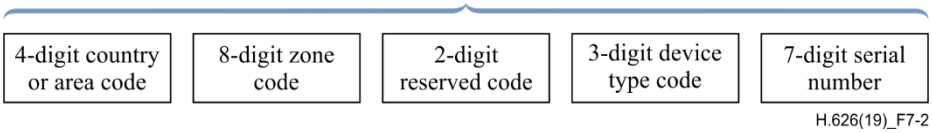

Figure 7-2 – 24-digit digital character scheme

The ID consists of the following five parts:

1) 4-digit country or area code: Representative of the country or region in which the VSS is deployed. It is recommended to use the international postal code.
2) 8-digit zone code: It is recommended to use postal code or customized zone code.
3) 2-digit reserved code: It is reserved for industry classification or other usage.
4) 3-digit device type code: It is used to distinguish different devices, such as IPC, DVR, NVR, CCS, etc.
5) 7-digit serial number: It generally follows numerical order.

7.2 Reference point Pc: CU-PU

Reference point Pc lies between the CU and the PU.

It is used to deliver interaction signals directly between the CU and the PU, as well as media streams from PU to CU.

7.3 Reference point Mp/Mp': MS-PU

Reference point Mp/Mp' lies between the MS and the PU.

It is used by the MS to distribute and store media from the PU using file and streaming methods.

7.4 Reference point Cm: CCS-MS

Reference point Cm lies between the CCS and the MS.

It is used by the CCS to control the MS for media distribution, media selection and media location, media storage and media serving, and related media operations. The MDU is used for media distribution, media selection and media location, while MSU is used for media storage and media serving, and related media operations.

7.5 Reference point Mc/Mc': MS-CU

Reference point Mc/Mc' lies between the MS and the CU.

It is used by the MS to distribute, store and deliver media to, or from the CU, using file and streaming methods.

7.6 Reference point Ps: CCS-PU

Reference point Ps lies between CCS and PU.

It is used by the CCS to control PU registration, access authentication, authorization, and monitoring of the status of the PU (such as online/offline, active/inactive status), and to control PU, such as PTZ control, remote reboot, configure parameters and upgrade the software, etc.

7.7 Reference point Cs: CCS-CU

Reference point Cs lies between the CCS and the CU.

It is used by the CCS to control CU registration, access authentication, authorization, monitoring of the CU status (such as online/offline, active/inactive status), and receiving the PTZ control operation commands from the CU.

7.8 Reference point Ds/Ds': MDU-MSU

Reference point Ds/Ds' lies between the MDU and the MSU.

It is used by the CCS to control media scheduling. The MSU captures media streams from the MDU and store them in the storage medium.

8 Interworking with other multimedia systems

8.1 Interworking with videoconferencing

The interworking and convergence between VS and videoconferencing systems is based on the consideration that surveilled media stream is introduced into videoconferencing system through the media stream server in the service platform, CUs or PUs of the VSS.

The following are instances of interworking between VS and videoconferencing systems:

Instance 1: A surveilled video stream is introduced into the videoconferencing system through the media stream server in the service platform of the VSS.

– A streaming server is added between the service platform of the VSS and the videoconferencing system. The streaming server stores the media resource providing for video conferencing from the VSS, and publishes the link addresses of the media resource to an electronic programme guide (EPG) server.

– A videoconferencing terminal accesses the EPG server to select and obtain the media address.

– The videoconferencing terminal initiates sending of the media stream request to the streaming server with the media address, and obtains the media stream from the streaming server. It then transfers it to the multimedia control unit (MCU) or other terminals in the video conferencing system.

Instance 2: A surveilled video stream is introduced into the videoconferencing system through the CU in the VSS.

– The CU in the VSS creates connections with the videoconferencing terminal, including signalling and media communication.

– The CU decodes the media stream and transfers it from the VSS to the videoconferencing terminal. This media stream is served as one of the input media sources for the videoconferencing terminal.

– The videoconferencing terminal broadcast this media stream within the videoconferencing system.

Instance 3: A surveilled video stream is directly introduced into the videoconferencing system through the PU in the VSS.

– The PU in the VSS creates connections with the MCU in the videoconferencing system.

– The MCU regards this PU as a virtual videoconferencing terminal and convenes it into the video conferencing.

– The MCU handles the media stream from the PU like from other videoconferencing terminals.

8.2 Interworking with IPTV

The interworking and convergence between VS and IPTV systems is based on the following considerations:

– The VSS can provide the VS service to the IPTV end user via interworking with the IPTV system.

– The IPTV end user can access the VSS via IPTV terminal devices (e.g., set-top box), which can act as the end-user functions.

– The IPTV end user can access the VSS via IPTV EPG, which can act as the VS service entry.

– The audio, video and alarm signal information from the VSS can be inserted into the closest related IPTV content delivery network (CDN) system, and be delivered to IPTV terminal devices (e.g., set-top box).

– The IPTV end user can play audio and video and get alert signals information from the VSS, and can also control the PTZ functions of the video camera with authorization.

The following describes an instance of interworking between VS and IPTV:

– PUs, the IPTV system and user terminals connect to the same bearer network used for multimedia services.

– PUs capture video and audio information from the surveilled objects, transcode and encode into the media stream.

– PUs send the media stream to media servers in the IPTV system. The media servers process, store and manage the media stream data.

– User terminals connect to the PUs through the IPTV system, view the real-time video and audio information captured by the Pus, and also take remote control of the PUs.

– User terminals can also view the recording media data from PUs stored in media servers in the IPTV system.

Bibliography

[b-ITU-T H.222.0] Recommendation ITU-T H.222.0 (2018) | ISO/IEC 13818-1:2019, *Information technology-Generic coding of moving pictures and associated audio information: Systems.*

[b-ITU-T H-Sup.1] ITU-T H-series Recommendations – Supplement 1 (1999), *Application profile – Sign language and lip-reading real-time conversation using low bit rate video communication.*

[b-ITU-T J.200] Recommendation ITU-T J.200 (2010), *Worldwide common core – Application environment for digital interactive television services.*

[b-ITU-T M.60] Recommendation ITU-T M.60 (1993), *Maintenance terminology and definitions.*

[b-ITU-T M.3320] Recommendation ITU-T M.3320 (1997), *Management requirements framework for the TMN X-Interface.*

[b-ITU-T Q.1290] Recommendation ITU-T Q.1290 (1998), *Glossary of terms used in the definition of intelligent networks.*

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

[b-ITU-T Y.110] Recommendation ITU-T Y.110 (1998), *Global Information Infrastructure principles and framework architecture*.

[b-ITU-T Y.2012] Recommendation ITU-T Y.2012 (2010), *Functional requirements and architecture of next generation networks*.

[b-IETF RFC 768] IETF RFC 768 (1980), *User datagram protocol.*

[b-IETF RFC 793] IETF RFC 793 (1981), *Transmission control protocol.*

[b-IETF RFC 3428] IETF RFC 3428 (2002), *Session Initiation Protocol (SIP) Extension for Instant Messaging.*

[b-IETF RFC 3550] IETF RFC 3550 (2003), *RTP: A Transport Protocol for Real-Time Applications.*

[b-IETF RFC 4566] IETF RFC 4566 (2006), *SDP: Session Description Protocol.* [b-IETF RFC 3261]IETF RFC 3261 (2002), *SIP: Session Initiation Protocol.*

[b-ISO/IEC 14496-3] ISO/IEC 14496-3:2009, *Information technology-Coding of audio-video objects-Part 3:Audio.*

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

Printed in Switzerland
Geneva, 2020