

International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

H.704

(08/2020)

SERIES H: AUDIOVISUAL AND MULTIMEDIA SYSTEMS
IPTV multimedia services and applications for IPTV –
General aspects

Enhanced user interface framework for IPTV terminal device – Gesture control interface

Recommendation ITU-T H.704



International
Telecommunication
Union

Documents
Document Sharing Hub

ITU-T

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.704

Enhanced user interface framework for IPTV terminal device – Gesture control interface

Summary

Recommendation ITU-T H.704 defines the general requirements, functional elements and interfaces supporting enhanced capability of user interaction by gesture recognition and controlling over Internet protocol TV (IPTV) terminal devices, based on the enhanced user interface (UI) framework defined in Recommendation ITU-T H.703. Those functional elements are described in the gesture controlling enabler and gesture recognition enabler defined in this Recommendation. Moreover, the procedures of interaction between the gesture recognition device and the gesture-controlled device are defined with the recommended information used in the interaction.

This Recommendation enables the gesture controlling feature in the enhanced IPTV user interface defined in Recommendation ITU-T H.703. With those features, users can control the operation of IPTV applications in an IPTV terminal device in a convenient, natural and comfortable way.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T H.704	2020-08-13	16	11.1002/1000/14343

Keywords

Enhanced user interface, gesture control, gesture control enabler, gesture interface, gesture recognition, IPTV.

* 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/>.

© ITU 2020

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

Table of Contents

		Page
1	Scope	1
2	References.....	1
3	Definitions	1
	3.1 Terms defined elsewhere	1
	3.2 Terms defined in this Recommendation.....	1
4	Abbreviations and acronyms	2
5	Conventions	3
6	Introduction	3
	6.1 The concept enhanced user interface functions for gesture control	3
	6.2 Functional block overview	5
7	Requirements	6
	7.1 Gesture control functional block requirements	6
	7.2 Gesture controlling enabler requirements	6
	7.3 Gesture recognition enabler requirements.....	6
8	Gesture UI functionality and interfaces.....	7
	8.1 Gesture control functional block	7
	8.2 Interfaces between functional components	11
9	High-level procedure flows	12
	9.1 Procedure of gesture control enabled device discovery	13
	9.2 Procedure of gesture event subscription.....	14
	9.3 Procedure of user gesture event recognition	15
	Appendix I – Examples of descriptions related to the gesture-enabled service.....	17
	I.1 Example of an XML description for a gesture controlling device – STB.....	17
	I.2 Example of an XML description for a gesture-enabled service within a gesture controlling device.....	17
	Appendix II – Four scenarios where an IPTV TD can connect to a gesture capturing sensor	19
	II.1 IPTV TD with embedded gesture capturing sensor	19
	II.2 IPTV TD with external gesture capturing sensor	19
	II.3 IPTV TD with remote gesture capturing sensor.....	19
	II.4 IPTV TD with remote gesture capturing sensor combined with recognition device	19
	Appendix III – Summary of relevant standardization activity.....	20
	Bibliography.....	21

Recommendation ITU-T H.704

Enhanced user interface framework for IPTV terminal device – Gesture control interface

1 Scope

This Recommendation describes the gesture event features, general requirements and the functionalities of the framework to support enhanced user interface for an Internet protocol TV (IPTV) terminal device. This Recommendation focuses on gesture control in a high-level approach.

The purpose of the enhanced user interface (UI) framework is to define the functional elements supporting enhanced capability of user interaction over an IPTV terminal device regarding a gesture control interface.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-T H.703] Recommendation ITU-T H.703 (2016), *Enhanced user interface framework for IPTV terminal devices*.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following term defined elsewhere:

3.1.1 video sensor [b-ITU-T H.742.0]: A device that has image processing ability, such as image compression, image manipulation with the video data captured by a camera. A video sensor can generate video sensor information alone, by processing video data when it has enough computational power. Otherwise, some functions of the video sensor are processed through external systems and the video sensor generates intermediate video sensor information.

3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

3.2.1 event packet: Gesture description information, or control commands related to a gesture, or even to a gesture image.

3.2.2 gesture: A shape and/or movement generated by a user's hand, arm, or other body parts. A gesture can be predefined.

3.2.3 raw gesture data: Non-processed gesture data directly input from a motion detect device.

3.2.4 gesture description: Information of the analysed raw gesture data.

3.2.5 gesture command: The operation instruction related to a recognizable gesture. One gesture description may be associated with different commands. For example, "raise arm" may refer to "power on" or "volume up".

3.2.6 gesture list: A textual description of a gesture, the identification of the gesture or of the gesture images.

3.2.7 gesture recognition: Analysis of raw gesture data and its translation as an understandable gesture description.

3.2.8 gesture event subscription: Action performed on a gesture recognition device that allows the list of gestures or gesture commands that are received by the target gesture-controlled IPTV terminal device (TD) to be recognized and accepted.

3.2.9 gesture-recognition device: A device which is equipped with a gesture capturing sensor so that it can recognise a user's gesture and send event packet(s) to a gesture-controlled device according to the gesture-controlled device's subscription. A gesture-recognition device has capabilities such as networking, computing, storage, etc.

3.2.10 gesture-controlled device: A device which is intended to be controlled by gesture control commands.

3.2.11 gesture capturing sensor: A device that has gesture data processing ability, such as gesture data/image compression, gesture data/image manipulation with raw data captured by a somatosensory/video sensor.

3.2.12 gesture capturing sensor information: Extracted and derived information from raw gesture data, for example, shapes, outlines, directions, duration, or movement.

3.2.13 gesture-controlled device discovery: Function enabling a gesture-recognition device to discover and identify one or several gesture-controlled devices using computing identification information such as IP address and device ID, and visual features such as QR code or optical signal detected by a visual computation mechanism..

3.2.14 visual feature: A label which can be recognised with a visual computing mechanism. For example, a device's appearance picture, bar code, quick response (QR) code or optical signal, etc.

3.2.15 recognizable gesture: A gesture which is able to be identified and processed by a gesture recognition device or software, depending on the recognition algorithm, filter conditions and other predefined recognition principles.

4 Abbreviations and acronyms

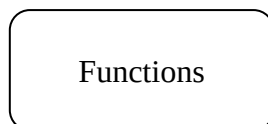
This Recommendation uses the following abbreviations and acronyms:

API	Application Programming Interface
EPG	Electronic Programme Guide
EUIF	Enhanced User Interface Functions
DNS	Domain Name System
IPTV	Internet Protocol TV
mDNS	Multicast DNS
QR	Quick Response
RFID	Radio Frequency Identification
SDK	Software Development Kit
STB	Set-Top-Box
TD	Terminal Device
UI	User Interface
UPnP	Universal Plug and Play

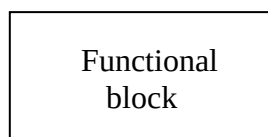
5 Conventions

The following conventions are used 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 Recommendation is to be claimed.
- The keywords "is prohibited from" indicate a requirement which must be strictly followed and from which no deviation is permitted, if conformance to this Recommendation is to be claimed.
- The keywords "is recommended" indicate a requirement which is recommended but which is not absolutely required. Thus, this requirement need not be present to claim conformance.
- The keywords "is not recommended" indicate a requirement which is not recommended but which is not specifically prohibited. Thus, conformance with this Recommendation can still be claimed even if this requirement is present.
- The keywords "can optionally" indicate an optional requirement which is permissible, without implying any sense of being recommended. This term is 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 this Recommendation.
- The keyword "functions" is defined as a collection of functionalities. It is represented by the following symbol in the context of IPTV architecture:



- The keyword "functional block" is defined as a group of functionalities that have not been further subdivided at the level of detail described in this Recommendation. It is represented by the following symbol in the context of IPTV architecture:



6 Introduction

6.1 The concept enhanced user interface functions for gesture control

A gesture control interface used in an IPTV service helps an end-user to control and interact with the IPTV terminal device without the use of a conventional remote controller. Users can easily change the IPTV application status, e.g., channel change, volume up and down, according to their gestures or special body motions. Currently, many terminal devices have been embedded with a dedicated gesture/motion detect device, such as a gyroscope, G-sensor, etc. Such equipment could help users create a gesture with a control intention. However, in this case the gesture detection, gesture recognition and gesture-controlled functions are merged in one IPTV terminal device (TD). The IPTV TD is only interacting with one user. In the current market, the technology of gesture