International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

J.604

(01/2014)

SERIES J: CABLE NETWORKS AND TRANSMISSION OF TELEVISION, SOUND PROGRAMME AND OTHER MULTIMEDIA SIGNALS

Transport of Large Screen Digital Imagery

Requirements for a scalable video transmission system over cable networks

Recommendation ITU-T J.604

International Telecommunication Union

Recommendation ITU-T J.604

Requirements for a scalable video transmission system over cable networks

Summary

Recommendation ITU-T J.604 describes requirements for a scalable video transmission system over cable networks, which is one of the transmission methods for providing additional high resolution video services.

High resolution video formats such as 4K are expected to be one of the ultra-realistic video formats that will be applied to future video services over cable network infrastructures. In the beginning of 4K video (ITU-R BT.1769/ITU-R BT.2020) services, both high definition television (HDTV) and 4K video services will be simultaneously provided.

Considering the cost of constructing the infrastructure, the current and new services will be provided over the same cable network infrastructure. Furthermore, in the case where the 4K video service will be additionally provided over the existing cable network infrastructure, the signal receiver (e.g., the set top box (STB) device) should be replaced with one that supports the 4K service. However, it is difficult to change all STB devices in the existing service at the same time, and therefore the 4K service will coexist with the current HDTV service.

Distribution systems using scalable video coding technology are one of the effective methods for providing simultaneous services. Within a general scalable video coding method, the smallest resolution video is coded as the low layer, and then the higher resolution video is coded as the higher layer by encoding the difference information from the lower resolution video.

From the viewpoint of compatibility to the current cable services which employs radio frequency (RF)-based broadcasting, one solution is that the existing HDTV service over RF is also used as the low layer, and then the difference information between 4K video and HDTV may be transmitted as the higher layer over the IP interface (data over cable service interface specification (DOCSIS) ITU-T J.222.0).

This Recommendation includes requirements regarding general issues, scalability of video service, and system performance.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T J.604	2014-01-13	9	11.1002/1000/12105-en

* 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

Recommendation ITU-T J.604

Requirements for a scalable video transmission system over cable networks

1 Scope

This Recommendation specifies requirements with regard to a scalable video transmission system over cable networks, which is one of the approaches for transmitting an additional video service (e.g., 4K) over cable networks on top of the existing video service (e.g., high definition television (HDTV)). Furthermore, this system supports scalable video coding bitstreams, which consist of a multi-layer bitstream structure, and that complies with video coding technology standards. In order to enable the provision of simultaneous video services over cable networks, this system transmits individual layer bitstreams over radio frequency (RF) and IP (data over cable service interface specification (DOCSIS) [ITU-T J.222.0]).

This Recommendation is most suitable for the case where a bit rate reduction is the highest priority because the available transmission capacity of the IP (DOCSIS) interface is insufficient for transmitting several higher quality video programs (e.g., 4K), or the remaining RF bandwidth is not sufficient to convey such programs. Another applicable case is that the gain of scalable coding is sufficiently larger than independent coding, or more specifically, scalable transmission is capable of reducing the bit amount of the higher layer (quality) stream against a simultaneously coded stream. Here, an assumption is that the higher quality video and the legacy quality video have identical angles of view, and it helps to achieve good coding gain in scalable coding.

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 J.222.0] Recommendation ITU-T J.222.0 (2007), *Third-generation transmission systems for interactive cable television services – IP cable modems: Overview*.

3 Definitions

3.1 Terms defined elsewhere

None.

3.2 Terms defined in this Recommendation

None.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

AVC Advanced Video Coding

DOCSIS Data Over Cable Service Interface Specification

HD High Definition

HDTV High Definition Television

HEVC High Efficiency Video Coding

QAM Quadrature Amplitude Modulation

RF Radio Frequency

STB Set Top Box

SVC Scalable Video Coding

5 Conventions

In this Recommendation, the following conventions are used.

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 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 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 "**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 the specification.

In the body of this Recommendation and its annexes, the words *shall*, *shall not*, *should*, and *may* sometimes appear, in which case they are to be interpreted, respectively, as *is required to*, *is prohibited from*, *is recommended* and *can optionally*. The appearance of such phrases or keywords in an appendix or in material explicitly marked as *informative* are to be interpreted as having no normative intent.

6 Total configuration

The configuration of the scalable video transmission system in this Recommendation is shown in Figure 1. The major processing of the system includes: acquisition, generation, encoding and decoding.

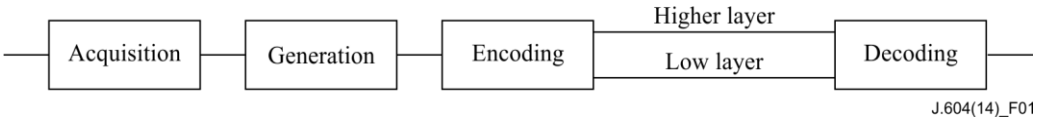

Figure 1 – Configuration of the scalable video transmission system

6.1 Acquisition

A high quality video signal (e.g., high resolution, high frame rate or high signal-to-noise ratio (SNR)) is acquired by either a camera or another video source signal.

6.2 Generation

For input to the scalable video transmission system, several types of videos are generated from the original video signal. In the case of a simultaneous service, the low quality video may be generated from the high quality video by using a down-converting filter.

6.3 Encoding

Multi-format videos are efficiently encoded based on scalable video coding technology.

Initially, the low quality video is encoded as the low layer of scalable video coding, which is compatible with single layer video coding. Before encoding the high quality video on the higher layer, the low layer signal is locally decoded and up-converted to the high quality video. The locally decoded signal contributes to efficient estimation for the higher layer coding. Therefore, the difference signal information between the high quality video and the low quality video is encoded on the higher layer.

Regarding the transmission process of scalable video transmission systems over cable networks, each layer bitstream is independently transmitted over a different interface. The low layer and the higher layer are transmitted over RF and IP (DOCSIS [ITU-T J.222.0]) interfaces, respectively.

6.4 Decoding

A multi-format video stream is decoded based on scalable video decoding technology. The bitstreams of the low layer and the higher layer are sequentially decoded for presenting original video since the higher layer decoding process requires the decoded and up-converted low layer signal.

7 Requirements

The following requirements from the standpoint of transmission service operators are considered.

7.1 General

(1) Simultaneous service

Video service based on this transmission system is required to be compatible with the current cable service. In the case that the current service (e.g., high definition television (HDTV)) is based on quadrature amplitude modulation (QAM), an additional service (e.g., 4K) using the QAM and/or IP interface is required not to affect the current service.

(2) Transmission

Each layer bitstream of scalable video is recommended to be independently transmitted over a different interface.

(3) Decoding process

In order to decode a scalable video stream, the decoding process for an additional service (e.g., 4K) is required to synchronize each layer of bitstream that is individually transmitted over QAM and IP interfaces.

7.2 Scalability

The transmission system is required to support either one or both types of the following scalabilities:

1. Resolution scalability

   The low layer bitstream contains the encoded video sequence with the minimum resolution, and the higher layer bitstream(s) contains the encoded video sequence(s) with higher resolution(s).

2. Temporal scalability

   The low layer bitstream contains the encoded video sequence with the lowest frame rate, and the higher layer bitstream(s) contains the encoded video sequence(s) with the higher frame rate(s).

7.3 Functionality

(1) Bitstream conformance

The transmitted bitstream is recommended to be compliant with international standards of transport and video coding.

(2) Error correction

Error correction information is recommended to be included in the transport stream for each layer bitstream. In the case where the transport stream is partially damaged, the transmission error is recommended to be corrected on the decoder side.

7.4 Performance

(1) Processing time

The processing time for decoding the scalable video stream is required to be faster than real-time displaying.

(2) Latency

The arrival latency to the decoder side between bitstream layers, which includes transmission latency between QAM and IP interfaces, is required to be limited to a range of time which enables the decoding process to synchronize between bitstream layers.

Appendix I

Example of a scalable video transmission system based on the existing standard

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

I.1 Example configuration of a scalable video transmission system

This appendix describes an example configuration of a scalable video transmission system based on ITU-T H.264 scalable video coding (SVC) and ITU-T H.222.0 in the case of resolution scalability of 4K and high definition (HD).

I.1.1 Generation

For input to a scalable video transmission system, several resolution videos are generated from an original 4K video signal. In the case of a simultaneous service of 4K and HDTV, the HD resolution video is generated from the 4K video by using a sub-sampling filter.

I.1.2 Encoding

Multi-resolution videos are encoded based on the [b-ITU-T H.264] SVC standard. Initially, HDTV is encoded as an SVC base layer, which has compatibility with single layer video coding by ITU-T H.264 advanced video coding (AVC) standard. Before encoding the 4K video on the enhancement layer, the coded base layer signal is locally decoded and up-sampled to 4K resolution. The locally decoded signal contributes to the efficient estimation for the enhancement layer coding. Therefore, the difference signal information between the 4K video and HDTV is encoded on the enhancement layer.

Regarding the transmission process of a scalable video transmission system over cable networks, each layer bitstream is independently transmitted over different interfaces based on [b-ITU-T H.222.0]. Considering simultaneous services of upcoming 4K and existing HDTV over cable networks, the base layer (HDTV) and the enhancement layer (4K) should be transmitted over RF and IP (DOCSIS) interfaces, respectively.

I.1.3 Decoding

A multi-resolution video stream is decoded based on ITU-T H.264 SVC. The bitstreams of the base layer and the enhancement layer are sequentially decoded since the enhancement layer decoding process requires the decoded and up-sampled base layer signal.

Appendix II

Example of a scalable video transmission system using several video coding standards

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

II.1 Example configuration of a scalable video transmission system

This appendix describes an example configuration of a scalable video transmission system using two video coding standards: ITU-T H.264 AVC and ITU-T H.265 high efficiency video coding (HEVC)) in the case of resolution scalability of HD and 4K.

II.1.1 Generation

For input to a scalable video transmission system, several resolution videos are generated from an original 4K video signal. In the case of simultaneous service of 4K and HDTV, the HD resolution video is generated from the 4K video by using a sub-sampling filter.

II.1.2 Encoding

Multi-resolution videos are encoded based on both ITU-T H.264 AVC and ITU-T H.265 HEVC. Initially, HDTV is encoded based on ITU-T H.264 AVC as the low layer. Before encoding the 4K video on the higher layer, the coded low layer signal is locally decoded and up-sampled to 4K resolution. The locally decoded signal contributes to efficient estimation for the higher layer coding. Therefore, the difference signal information between the 4K video and the HDTV is encoded by using ITU-T H.265 HEVC as the higher layer.

An offset value is added to the difference signal before ITU-T H.265 HEVC coding because a negative number will be included as a pixel value in the original difference signal. Then, the corrected signal is encoded as a video signal expressed as 9 bit depth.

Regarding the transmission process of a scalable video transmission system over cable networks, each layer bitstream is independently transmitted over a different interface based on [b-ITU-T H.222.0]. Considering simultaneous services of upcoming 4K and existing HDTV over cable networks, the low layer (HDTV) and the higher layer (4K) should be transmitted over RF and IP (DOCSIS) interfaces, respectively.

II.1.3 Decoding

Multi-resolution video streams are decoded based on [b-ITU-T H.264] AVC and [b-ITU-T H.265] HEVC. The bitstreams of the low layer and higher layer are sequentially decoded since the higher layer decoding process requires the decoded and up-sampled low layer signal.

Bibliography

[b-ITU-T H.222.0] Recommendation ITU-T H.222.0 | ISO/IEC 13818-1 (2012), *Information technology – Generic coding of moving pictures and associated audio information: Systems*.

[b-ITU-T H.264] Recommendation ITU-T H.264 | ISO/IEC 14496-10 (2012), *Advanced video coding for generic audiovisual services*.

[b-ITU-T H.265] Recommendation ITU-T H.264 (2013), *High efficiency video coding*.

[b-ITU-R BT.1769] Recommendation ITU-R BT.1769 (2008), *Parameter values for an expanded hierarchy of LSDI image formats for production and international programme exchange*.

[b-ITU-R BT.2020] Recommendation ITU-R BT.2020 (2012), *Parameter values for ultra-high definition television systems for production and international programme exchange*.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	General tariff principles
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	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	Terminals and subjective and objective assessment methods
Series Q	Switching and signalling
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 and next-generation networks
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2014