TTC標準
Standard

JJ-40.30

Audiovisual Communication System on SIP Network

Version 1.1

Established on October 26, 2012

THE TELECOMMUNICATION TECHNOLOGY COMMITTEE

TTC Telecommunication Technology Committee

Contents

<Reference information>

1. Relationship with International Recommendations

This standard has no particular relation to international recommendations.

2. History of Revision

Revision	Date	Description
Version 1.1	October 26, 2012	Initial publication

3. Industrial Property Rights

Information regarding submission of "The Policy for the Handling of Industrial Property Rights" associated with this standard is available on TTC's Web page.

4. Contact

Media Coding Working Group

1. Introduction

This standard was created to help enable the successful interconnection of real-time, conversational audiovisual communication systems on a Session Initiation Protocol (SIP) network.

2. Scope

This standard addresses basic systems that have the following communications functions:

- Bidirectional symmetrical communication
- Point-to-point communication
- Multimedia communication that simultaneously uses audio communication and video communication functions

The communications functions, e.g., multipoint communications using a multipoint control unit (MCU), which are beyond the above functions, will be studied as part of the next stage of this standard.

The main items that have been studied for this standard are the recommended values for the parameters related to media signals (audio, video, and data signals) and the recommended method of capability negotiation on SDP. For details on capability negotiation on SDP, refer to TR-1020 [TR-1020]. This standard is intended to define the interconnection specifications for audiovisual communication systems that are connected to a SIP network. This version of the standard applies only to those SIP networks that are compliant with the SIP/SDP interface between users and networks that are defined by TTC Standard JT-Q3402 [Q3402].

3. References

This standard references the following documents:

[G711] "Pulse Code Modulation (PCM) of Voice Frequencies," TTC Standard JT-G711, Ver. 4.0, The Telecommunication Technology Committee, April 2001

[G711.1] "Wideband Embedded Extension for G.711 Pulse Code Modulation," TTC Standard JT-G711.1, Ver. 1.0, The Telecommunication Technology Committee, February 2009

[G722] "7 kHz audio-coding within 64 kbit/s," TTC Standard JT-G722 Ver. 3.0, The Telecommunication Technology Committee, May 2008

[H224] "A real time control protocol for simplex applications using the H.221 LSD/HSD/MLP channels," ITU-T Recommendation H.224, ITU-T, January 2005

[H263] "Video Coding for Low Bitrate Communication," TTC Standard JT-H263, Ver. 3.3, The Telecommunication Technology Committee, November 2009

[H264] "Advanced Video Coding for Generic Audiovisual Services," TTC Standard JT-H264, Ver. 5, The Telecommunication Technology Committee, November 2010

[H281] "A far end camera control protocol for videoconferences using H.224," ITU-T Recommendation H.281, November 1994

[H323] "Packet-based multimedia communication systems," TTC Standard JT-H323, Ver. 6.0, The Telecommunication Technology Committee, May 2008

[MPEG-4 Audio] "Information technology - Coding of audio-visual objects - Part 3: Audio," ISO/IEC 14496-3, Ver. 3, ISO, December 2005

[MPEG-4 Video] "Information technology - Coding of audio-visual objects - Part 2: Visual," ISO/IEC 14496-2, Ver. 3, ISO, June 2004

[Q3402] "NGN UNI Signalling Profile (Protocol Set 1)," TTC Standard JT-Q3402, Ver. 1.0, The Telecommunication Technology Committee, May 2009

[BT601] "Studio encoding parameters of digital television for standard 4:3 and wide screen 16:9 aspect ratios," ITU-R Recommendation BT.601-6, ITU-R, January 2007

[BT709] "Parameter values for the HDTV standards for production and international programme exchange," ITU-R Recommendation BT.709-5, ITU-R, April 2002

[RFC3016] "RTP Payload Format for MPEG-4 Audio/Visual Streams," TTC Standard JF-IETF-RFC3016, The Telecommunication Technology Committee, May 2009

[RFC3551] "RTP Profile for Audio and Video Conferences with Minimal Control," TTC Standard JF-IETF-STD65, Ver. 1, The Telecommunication Technology Committee, June 2005

[RFC3984] "RTP Payload Format for H.264 Video," TTC Standard JF-IETF-RFC3984, Ver. 1.0, The Telecommunication Technology Committee, May 2009

[RFC4573] "MIME Type Registration for RTP Payload Format for H.224," RFC 4573, IETF, July 2006

[RFC4585] "Extended RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/AVPF)," TTC Standard JF-IETF-RFC4585, Ver. 1.0, The Telecommunication Technology Committee, March 2008

[RFC4629] "RTP Payload Format for ITU-T Rec. H.263 Video," TTC Standard JF-IETF-RFC4629, Ver. 1.0, The Telecommunication Technology Committee, May 2010

[RFC5104] "Codec Control Messages in the RTP Audio-Visual Profile with Feedback (AVPF)," TTC Standard JF-IETF-RFC5104, Ver. 1.0, The Telecommunication Technology Committee, May 2008

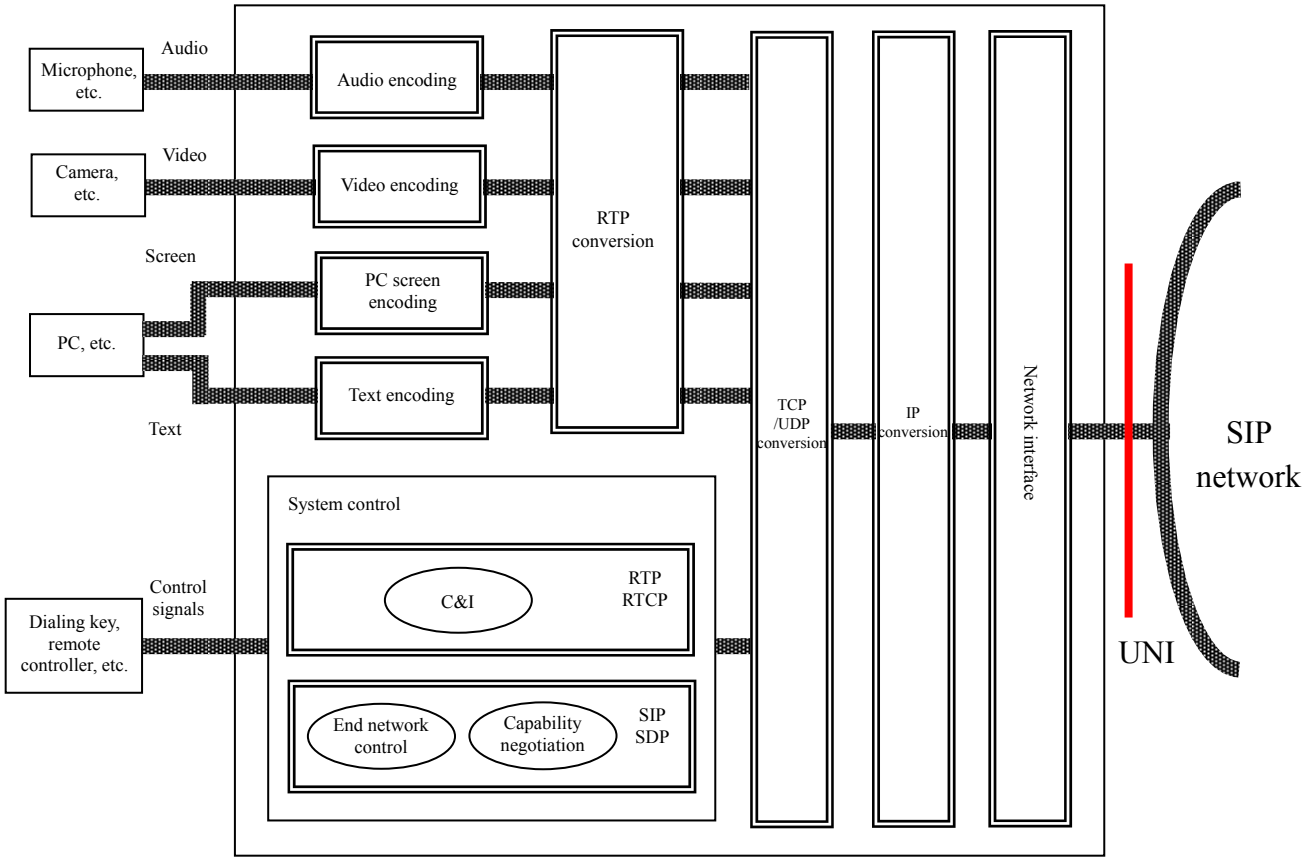
[RFC5391] "RTP Payload Format for ITU-T Recommendation G.711.1," TTC Standard JF-IETF-RFC5391, Ver. 1.0, The Telecommunication Technology Committee, May 2009

[TR-1014] "General Overview of NGN Architecture," TTC Technical Report TR-1014, Ver. 1.0, The Telecommunication Technology Committee, June 2006

[TR-1020] "Technical Report on SDP Media Negotiation over NGN," TTC Technical Report TR-1020, Ver. 1.0, The Telecommunication Technology Committee, May 2009

4. System Model

Figure 4-1 is a schematic of a typical sending terminal with a focus on the data flow in order to explain sending in the audiovisual communication system. The capability of each component and whether each component is applicable vary depending on the system profile.



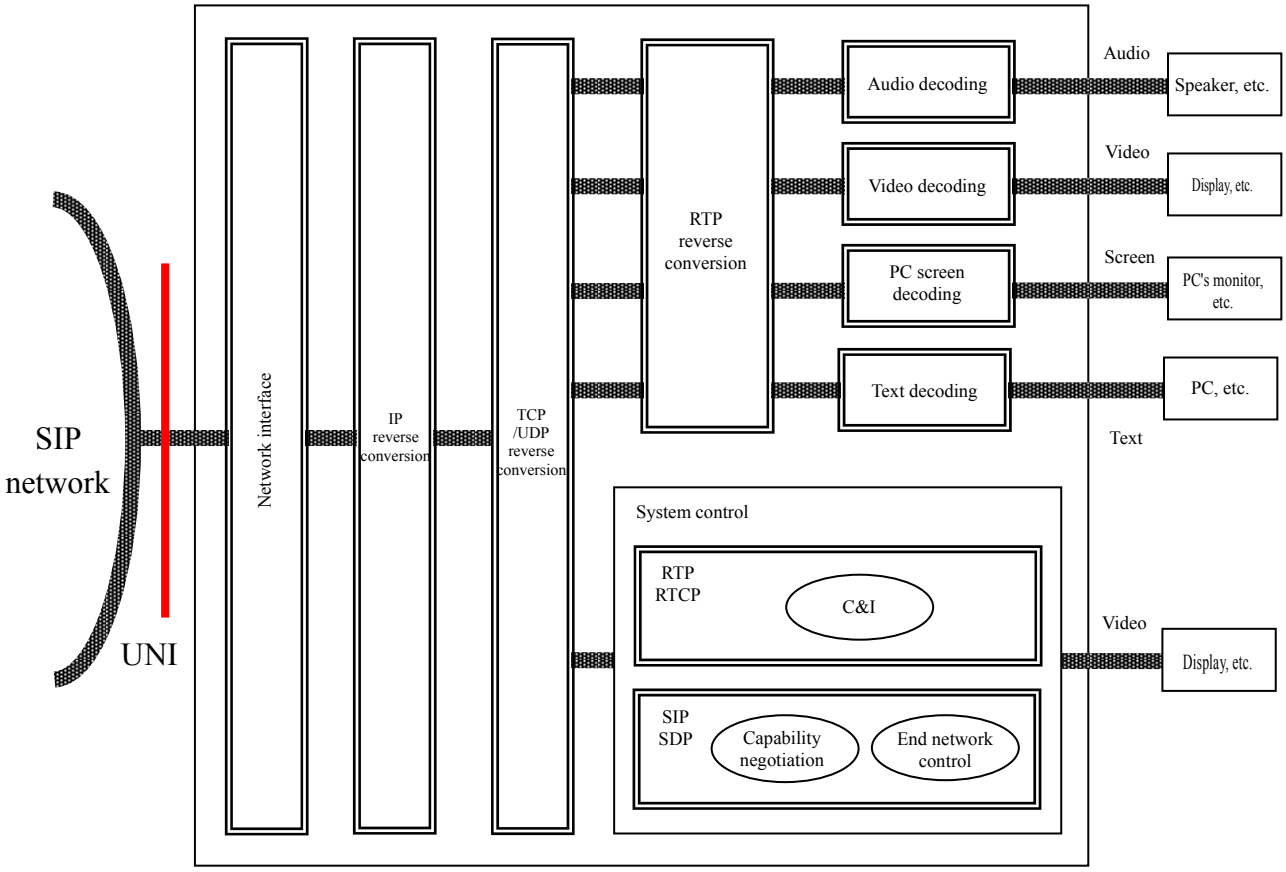
C&I: Control and indication

UNI: User-to-network interface

Note: C&I signal transfer on a protocol other than RTP or RTCP requires further study.

Figure 4-1/JJ-40.30: Sending terminal model with a focus on data flow

Figure 4-2 is a schematic of a typical receiving terminal with a focus on the data flow in order to explain receiving in the audiovisual communication system. The capability of each component and whether each component is applicable vary depending on the system profile.



C&I: Control and indication

UNI: User-to-network interface

Note: The C&I signal transfer on a protocol other than RTP or RTCP requires further study.

Figure 4-2/JJ-40.30: Receiving terminal model with a focus on data flow

5. System Profile

The capability to be supported by a terminal varies greatly according to the terminal's application. To enable the interconnection of different terminals, this standard defines the minimum capabilities the terminals must support for given applications as the system profiles listed in Table 5-1. Each terminal intended to support the recommended specifications of this standard must implement at least one of the system profiles.

Table 5-1/JJ-40.30: System profiles defined by this standard

System profile	Feature	Outline
AVSIP-1	MPEG-4 QCIF video	Defines the capability of videotelephony at a low bit rate. This system profile enables bidirectional audio and video communication. It is assumed that the terminal is used for a mobile application.
AVSIP-1.5	MPEG-4 QVGA video	Defines the capability of videotelephony and videoconferencing with QVGA or the equivalent quality. This system profile enables bidirectional audio and video communication. The terminal must be interconnected with an AVSIP-1 terminal.
AVSIP-2a	MPEG-4 VGA video	Defines the capability of videotelephony and videoconferencing with SD or the equivalent quality. This system profile enables bidirectional audio and video communication. The terminal must be interconnected with an AVSIP-1 or AVSIP-1.5 terminal.
AVSIP-2b	H.264 SD video	Defines the capability of videotelephony and videoconferencing with SD quality. This system profile enables the control of remote cameras, the transmission of PC screens, as well as bidirectional audio and video communication.
AVSIP-3	H.264 720p video	Defines the capability of videoconferencing with HD quality. This system profile enables the control of remote cameras, the transmission of PC screens, as well as bidirectional audio and video communication. It is recommended that the terminal be interconnected with an AVSIP-2b terminal.
AVSIP-4	H.264 1080i video	Defines the capability of videoconferencing with full-HD quality. This system profile enables the control of remote cameras, the transmission of PC screens, as well as bidirectional audio and video communication.

(Note) In this standard, "SD quality" refers to the picture quality provided when the resolution of the video source format is 704 by 480 pixels; "HD quality" refers to the picture quality provided when the resolution of the video source format is 1,280 by 720 pixels; and "full-HD quality" refers to the picture quality provided when the resolution of the video source format is 1,920 by 1,080 pixels.

6. Required Characteristics of Media and Encoding

Each terminal must have the media data processing capability required for communication, and perform encoding and decoding using the specified encoding and decoding methods.

6.1 Audio Media

6.1.1 Required Characteristics

Table 6-1 lists the audio-band processing capabilities defined by this standard. Other capabilities may be implemented as options.

Table 6-1/JJ-40.30: Audio bandwidth defined by this standard

Audio bandwidth	Usage example
3.4 kHz	Videotelephony with standard quality
7 kHz	Videotelephony and videoconferencing with SD quality
21 kHz	Videoconferencing with HD or full-HD quality

Table 6-2 lists the audio channel processing capabilities defined by this standard. Other capabilities may be implemented as options.

Table 6-2/JJ-40.30: Audio channels defined by this standard

Audio channel	Usage example
1 channel	Monaural audio
2 channels	Stereo audio

6.1.2 Audio Codecs

Table 6-3 lists the audio codec capabilities defined by this standard. Other capabilities may be implemented as options.

Table 6-3/JJ-40.30: Audio codec capabilities defined by this standard

Audio codec	Applicable standard	Bandwidth	Bit rate	Usage example
G.711μ-law	JT-G711 [G711]	3.4 kHz	64 kbit/s	Videotelephony with standard quality
G.722	JT-G722 [G722]	7 kHz	64 kbit/s	Videotelephony with SD quality
G.711.1	JT-G711.1 [G711.1]	7 kHz	96 kbit/s	Videotelephony and videoconferencing with SD quality
MPEG-4 AAC-LC	ISO/IEC 14496-3 [MPEG-4 Audio]	21 kHz	96 kbit/s, 192 kbit/s	Videoconferencing with HD or full-HD quality
MPEG-4 AAC-LD	ISO/IEC 14496-3 [MPEG-4 Audio]	21 kHz	128 kbit/s, 192 kbit/s	Videoconferencing with HD or full-HD quality

G.711.1 audio codec defined by this standard supports the encoding and decoding capabilities listed in Table 6-4.

Table 6-4/JJ-40.30: G.711.1 audio codec defined by this standard

Item	Content	Specification on SDP
Sampling rate	16 kHz	Specify "16000" for "clock rate" in the "a=rtpmap" line.
Packetization period	20 ms	Specify "a=ptime:20".
Mode	4	Specify "mode-set=4" in the "a=fmtp" line.

MPEG-4 AAC-LC audio codec defined by this standard supports the encoding and decoding capabilities listed in Table 6-5.

Table 6-5/JJ-40.30: MPEG-4 AAC-LC audio codec defined by this standard

Item	Content	Specification on SDP
Sampling rate	48 kHz	Specify "3" (48 kHz) for "samplingFrequencyIndex" in the "config" parameter of the "a=fmtp" line.
RTP time stamp rate	90 kHz	Specify "90000" for "clock rate" in the "a=rtpmap" line.
Packetization period	21.33 ms	Specify "a=ptime:20". (Note 1)
Profile	AAC Profile	
Level	Level2 (0x29)	Specify "profile-level-id=41" in the "a=fmtp" line.
Codec type	AAC-LC	Specify "object=2" in the "a=fmtp" line.
No. of channels	Monaural or stereo	Specify a value for "channelConfiguration" in the "config" parameter of the "a=fmtp" line. ("1" [monaural] or "2" [stereo])
Bit rate (Codec bandwidth)	96 kbit/s (monaural), 192 kbit/s (stereo)	Specify "bitrate=96000" (monaural) or "bitrate=192000" (stereo).
Bit rate (Layer 3 bandwidth)	192 kbit/s (monaural), 384 kbit/s (stereo) (Note 2)	Specify "b=AS:192" (monaural) or "b=AS:384" (stereo). (Note 2)

(Note 1) On SDP, negotiation is performed by setting "a=ptime:20". When, however, communication uses MPEG-4 AAC-LC (21 kHz), the system operates with a packetization period of 21.33 ms.

(Note 2) Reference value: The bit rate for the layer 3 bandwidth must be determined considering the burst characteristics and shaper tuning.

MPEG-4 AAC-LD audio codec defined by this standard supports the encoding and decoding capabilities listed in Table 6-6.

Table 6-6/JJ-40.30: MPEG-4 AAC-LD audio codec defined by this standard

Item	Content	Specification on SDP
Sampling rate	48 kHz	Specify "3" (48 kHz) for "samplingFrequencyIndex" in the "config" parameter of the "a=fmtp" line.
RTP time stamp rate	90 kHz	Specify "90000" for "clock rate" in the "a=rtpmap" line.
Packetization period	10.67 ms	Specify "a=ptime:20". (Note 1)
Profile	Low Delay Audio Profile	
Level	Level4 (0x19)	Specify "profile-level-id=25" in the "a=fmtp" line.
Codec type	AAC-LD	Specify "object=23" in the "a=fmtp" line.
No. of channels	Stereo	Specify "2" (stereo) for "channelConfiguration" in the "config" parameter of the "a=fmtp" line.
Bit rate (Codec bandwidth)	128 kbit/s (64 kbit/s × 2 ch[stereo]), 192 kbit/s (96 kbit/s × 2 ch [stereo])	Specify "bitrate=128000" (128 kbit/s) or "bitrate=192000" (192 kbit/s).
Bit rate (Layer 3 bandwidth)	256 kbit/s, 384 kbit/s (Note 2)	Specify "b=AS:256" or "b=AS:384". (Note 2)

(Note 1) On SDP, negotiation is performed by setting "a=ptime:20". When, however, communication uses MPEG-4 AAC-LD (21 kHz), the system operates with a packetization period of 10.67 ms.

(Note 2) Reference value: The bit rate for the layer 3 bandwidth must be determined considering the burst characteristics and shaper tuning.

6.1.3 Audio Media Specifications of System Profiles

Table 6-7 lists the specifications of the audio characteristics required for each terminal and the codec that must be implemented on each terminal.

Table 6-7/JJ-40.30: Audio media specifications to be implemented on each terminal

Item		Audio media specifications					
		Profile AVSIP-1	Profile AVSIP-1.5	Profile AVSIP-2a	Profile AVSIP-2b	Profile AVSIP-3	Profile AVSIP-4
Audio bandwidth	3.4 kHz	Mandatory	Mandatory	Mandatory	Mandatory	Mandatory	Mandatory
	7 kHz	Optional	Optional	Recommended	Recommended	Recommended	Recommended
	21 kHz	Optional	Optional	Optional	Optional	Optional	Optional
No. of channels	1 (monaural)	Mandatory	Mandatory	Mandatory	Mandatory	Mandatory	Mandatory
	2 (stereo)	Optional	Optional	Optional	Optional	Optional	Optional
Codec	G.711μ-law	Mandatory	Mandatory	Mandatory	Mandatory	Mandatory	Mandatory
	G.722	Optional	Optional	Recommended (Note)	Recommended (Note)	Recommended (Note)	Recommended (Note)
	G.711.1	Optional	Optional	Recommended (Note)	Recommended (Note)	Recommended (Note)	Recommended (Note)
	MPEG-4 AAC-LC	Optional	Optional	Optional	Optional	Optional	Optional
	MPEG-4 AAC-LD	Optional	Optional	Optional	Optional	Optional	Optional

(Note)This standard recommends the implementation of either one or both of G.722 and G.711.1 audio codecs.

6.2 Video Media

This standard defines camera and PC screen images as video media. For the camera images, side-by-side 3D images are defined as an optional function in addition to 2D images.

6.2.1 Required Characteristics

Table 6-8 lists the image format processing capabilities (with a focus on the encoder input point) defined by this standard. Other capabilities may be implemented as options.

Table 6-8/JJ-40.30: Video source formats defined by this standard

Video source	Video source format				Usage example
	Resolution and scan method	Frame rate (per second)	Scan method	Aspect ratio of screen	
Camera image (Note 4)	QCIF: 176 x 144, p	30 (Note 2)	Progressive	4:3	Videotelephony with mobile quality
	QVGA: 320 x 240, p	30 (Note 2)	Progressive	4:3	Videotelephony with QVGA quality
	VGA: 640 x 480, p	30 (Note 2)	Progressive	4:3	Videotelephony with SD or equivalent quality
	SD: 704 x 480 (Note 1), p	30 (Note 2)	Progressive	4:3	Videotelephony with SD quality
	720p: 1280 x 720, p	60 (Note 3)	Progressive	16:9	Videoconferencing with HD quality
	1080i: 1920 x 1080, i	30 (Note 2)	Interlaced	16:9	Videoconferencing with full-HD quality
PC screen image	VGA: 640 x 480, p	60 or more	Progressive	4:3	PC screen
	SVGA: 800 x 600, p	60 or more	Progressive	4:3	PC screen
	XGA: 1024 x 768, p	60 or more	Progressive	4:3	PC screen

(Note 1) The video source format includes 720 pixels per line, with encoding applied to 704 of the 720 pixels. This format is called "4SIF" ("SIF" stands for "source input format").

(Note 2) The frame rate of 30 fps must be replaced with 29.97 fps (30000/1001 = 29.97) when the system conforms to the television standard.

(Note 3) The frame rate of 60 fps must be replaced with 59.94 fps (60000/1001 = 59.94) when the system conforms to the television standard.

(Note 4) For a 3D image, the frame images that are combined in a side-by-side format are defined as the video source format.

The color matrix defined by ITU-R BT.601 [BT601] is used for SD video, and the color matrix defined by ITU-R BT.709 [BT709] is used for 720p and 1080i videos.

Table 6-9 lists the video channel processing capabilities defined by this standard. Other capabilities may be implemented as options.

Table 6-9/JJ-40.30: Video channels defined by this standard

Video channel	Video source	Usage example
1 channel	Camera image only	Videotelephony and videoconferencing
2 channels	One channel for camera image One channel for PC screen image	Videoconferencing

6.2.2 Video Codecs

Table 6-10 lists the video codec capabilities defined by this standard.

Table 6-10/JJ-40.30: Video codecs defined by this standard

Video codec	Applicable standard	Usage
MPEG-4 video	ISO/IEC 14496-2 [MPEG-4 Video]	Camera image for videotelephony
H.263	JT-H263 [H263]	PC screen
H.264	JT-H264 [H264]	Camera image and PC screen for videoconferencing

MPEG-4 video codec defined by this standard supports the encoding and decoding capabilities listed in Table 6-11. Note that the frame rate of 30 fps must be replaced with 29.97 fps (30000/1001 = 29.97) when the system conforms to the television standard.

Table 6-11/JJ-40.30: MPEG-4 video codec defined by this standard

Item	Content	
	QCIF video	
	Specification	Specification on SDP
Profile	Simple Profile	Specify "profile-level-id=8" on the "a=fmtp" line.
Level	Level 0	
Maximum frame rate	15 fps	This parameter specifies the operating frame rate. The use of 15 fps is recommended. When using the recommended value, omit the "a=framerate" line or specify the following: a=framerate:15
Maximum bit rate (Layer 3 bandwidth)	48 kbit/s	This parameter specifies the operating bit rate. When using 48 kbit/s, specify the following: b=AS:48

Item	Content	
	QVGA video	
	Specification	Specification on SDP
Profile	Simple Profile	Specify "profile-level-id=3" on the "a=fmtp" line.
Level	Level 3	
Maximum frame rate	15 fps	This parameter specifies the operating frame rate. The use of 15 fps is recommended. When using the recommended value, omit the "a=framerate" line or specify the following: a=framerate:15
Maximum bit rate (Layer 3 bandwidth)	384 kbit/s	This parameter specifies the operating bit rate. When using 384 kbit/s, specify the following: b=AS:384

Item	Content	
	VGA video	
	Specification	Specification on SDP
Profile	Simple Profile	Specify "profile-level-id=4" on the "a=fmtp" line.
Level	Level 4a	
Maximum frame rate	30 fps	This parameter specifies the operating frame rate. The use of 30 fps is recommended. When using the recommended value, omit the "a=framerate" line or specify the following: a=framerate:30
Maximum bit rate (Layer 3 bandwidth)	2 Mbit/s	This parameter specifies the operating bit rate. When using 2 Mbit/s, specify the following: b=AS:2000

H.263 video codec defined by this standard supports the encoding and decoding capabilities listed in Table 6-12.

Table 6-12/JJ-40.30: H.263 video codec defined by this standard

Item	Content	
	Specification	Specification on SDP
Version	H.263-1998	Designate "H263-1998" for "encoding name" on the "a=rtpmap" line.
Profile	Baseline Profile	
Resolution	VGA, SVGA, XGA	Designate the following on the "a=fmtp" line: VGA: CUSTOM=640,480,1;PAR=1,1;QCIF=1 SVGA: CUSTOM=800,600,1;PAR=1,1;QCIF=1 XGA: CUSTOM=1024,768,1;PAR=1,1;QCIF=1
Maximum frame rate	30 fps	This parameter specifies the operating frame rate. The use of 30 fps is recommended. When using the recommended value, omit the "a=framerate" line or designate the following: a=framerate:30
Maximum bit rate (Layer 3 bandwidth)	2 Mbit/s	This parameter specifies the operating bit rate. When using 2 Mbit/s, specify the following: b=AS:2000

H.264 video codec defined by this standard supports the encoding and decoding capabilities listed in Table 6-13.

Table 6-13/JJ-40.30: H.264 video codec defined by this standard

Item	Content	
	SD video	
	Specification	Specification on SDP
Profile	Baseline Profile	Specify "profile-level-id=42c01e" or "profile-level-id=42e01e" on the "a=fmtp" line. (Note 1)
Level	3.0	
Maximum frame rate	30 fps	This parameter specifies the operating frame rate. The use of 30 fps is recommended. When using the recommended value, omit the "a=framerate" line or specify the following: a=framerate:30
Maximum bit rate (Layer 3 bandwidth)	2 Mbit/s	This parameter specifies the operating bit rate. When using 2 Mbit/s, specify the following: b=AS:2000

Item	Content	
	720p video	
	Specification	Specification on SDP
Profile	Baseline Profile	Specify "profile-level-id=42c01f" or "profile-level-id=42e01f" on the "a=fmtp" line. (Note 2)
Level	3.1	
Maximum frame rate	30 fps	This parameter specifies the operating frame rate. The use of 30 or 15 fps is recommended. When using 30 fps, omit the "a=framerate" line or specify the following: a=framerate:30 When using 15 fps, specify the following: a=framerate:15
Maximum bit rate (Layer 3 bandwidth)	5 Mbit/s	This parameter specifies the operating bit rate. When using 5 Mbit/s, specify the following: b=AS:5000

Item	Content	
	1080i video	
	Specification	Specification on SDP
Profile	High Profile	Specify "profile-level-id=640028" on the "a=fmtp" line. (Note 3)
Level	4.0	
Maximum frame rate	30 fps (60 field/s)	This parameter specifies the operating frame rate. The use of 30 fps (60 field/s) is recommended. When using the recommended value, omit the "a=framerate" line or specify the following: a=framerate:30
Maximum bit rate (Layer 3 bandwidth)	10 Mbit/s	This parameter specifies the operating bit rate. When using 10 Mbit/s, specify the following: b=AS:10000

Item	Content	
	PC video	
	Specification	Specification on SDP
Profile	Baseline Profile	Specify the following value on the "a=fmtp" line: VGA: profile-level-id=42c01e or profile-level-id=42e01e SVGA or XGA: profile-level-id=42c01f or profile-level-id=42e01f
Level	VGA: 3.0 SVGA or XGA: 3.1	
Maximum frame rate	30 fps	This parameter specifies the operating frame rate. The use of 30 fps is recommended. When using the recommended value, omit the "a=framerate" line or specify the following: a=framerate:30
Maximum bit rate (Layer 3 bandwidth)	2 Mbit/s	This parameter specifies the operating bit rate. When using 2 Mbit/s, specify the following: b=AS:2000

(Note 1) When profile-level-id is "42c01e" or "42e01e", profile_idc is "66" (Baseline Profile), constraint_set0_flag is "1" (compliant with the constraint on Baseline Profile), constraint_set1_flag is "1" (compliant with the constraint on Main Profile), and level_idc is "30" (level 3.0).

(Note 2) When profile-level-id is "42c01f" or "42e01f", profile_idc is "66" (Baseline Profile), constraint_set0_flag is "1" (compliant with the constraint on Baseline Profile), constraint_set1_flag is "1" (compliant with the constraint on Main Profile), and level_idc is "31" (level 3.1).

(Note 3) When profile-level-id is "640028", profile_idc is "100" (High Profile) and level_idc is "40" (level 4.0).

H.264 video codec defined by this standard must be able to decode the stream that satisfies the conditions listed in Table 6-14 and Table 6-15.

Table 6-14/JJ-40.30: H.264 video codec bit stream defined by this standard

H.264 syntax element		SD video	720p video	1080i video
seq_parameter_set_rbsp()	constraint_set1_flag	1 (Note 1)	1 (Note 1)	0
	num_reorder_frames	0	0	0
	max_dec_frame_buffering	1	1	1 or 2 (Note 2)
	entropy_coding_mode_flag	0 (CAVLC)	0 (CAVLC)	1 (CABAC)
	chroma_format_idc	-	-	1
pic_parameter_set_rbsp(), slice_header()	num_ref_idx_l0_active_minus1	0	0	0 or 1 (Note 3)
slice_header()	slice_type	0, 2, 5, or 7	0, 2, 5, or 7	0, 2, 5, or 7

(Note 1) This setting implies that the bit stream can be decoded by Main Profile. Note that this setting therefore involves constraints on the maximum number of slices per frame and the error resilience tools that can be used.

(Note 2) The value of this element must be "1" when frame_mbs_only_flag is "1", or "2" otherwise.

(Note 3) The value of this element must be "1" when field_pic_flag is "1" and such two fields as top and bottom fields are to be referenced. The value of this element must be "0" in other cases. Note that, when field_pic_flag is "1" and num_ref_idx_l0_active_minus1 is "1", the two most recently decoded fields are referenced, or, when field_pic_flag is "1" and num_ref_idx_l0_active_minus1 is "0", the most recently decoded field is referenced.

Table 6-15/JJ-40.30: nal_unit_type of the H.264 video bit stream defined by this standard

H.264 nal_unit_type	Meaning	Handling		
		Interpretation required	Ignorable	Use inhibited
0	Unspecified			√
1	Coded slice of a non-IDR picture	√		
2	Coded slice data partition A			√
3	Coded slice data partition B			√
4	Coded slice data partition C			√
5	Coded slice of an IDR picture	√		
6	Supplemental enhancement information		√ (Note)	
7	Sequence parameter set	√		
8	Picture parameter set	√		
9	Access unit delimiter		√	
10	End of sequence		√	
11	End of stream		√	
12	Filler data		√	
13	Sequence parameter set extension			√
14 to 18	Reserved			√
19	Coded slice of an auxiliary coded picture without partitioning			√
20 to 23	Reserved			√
24 to 31	Unspecified			√

(Note) The distinction between a 2D image and a side-by-side 3D image is indicated by the frame packing arrangement SEI ("SEI" stands for "supplemental enhancement information"). Therefore, interpreting the frame packing arrangement SEI enables the sending/receiving equipment that receives side-by-side 3D image to not only manually switch between 2D and 3D displaying according to its display capability but also automatically switch between 2D and 3D displaying.

The color matrix of the video to be encoded and decoded with H.264 video codec conforms to the characteristic requirements described in Section 6.1.1. Table 6-16 lists the parameters that must be set to specify this color matrix with H.264 syntax.

Table 6-16/JJ-40.30: Color matrix of the H.264 video bit stream defined by this standard

H.264 syntax element		SD video	720p video 1080i video
vui_parameters()	colour_primaries	2 or 6	1 or 2
	transfer_characteristics	2 or 6	1 or 2
	matrix_cofficients	2 or 6	1 or 2

(Note) If a parameter is not encoded, its value will be "2".

When an SVGA video (800 by 600 pixels) is encoded with H.264 video codec, the number of encoded lines and the number of display lines differ by eight lines. Therefore, the parameters listed in Table 6-17 must be set in a sequence parameter set to specify the range of the encoded area to be displayed.

Table 6-17/JJ-40.30: SVGA display parameter specifications of H.264 video codec

H.264 syntax name	Value	Description
• • •		
pic_width_in_mbs_minus1	49	"Number of horizontal macro blocks of decoded image" – 1
pic_height_in_map_units_minus1	37	"Vertical mapping unit of decoded image" – 1 The mapping unit is 16 pixels for 4:2:0 progressive.
• • •		
frame_cropping_flag	1	For SVGA video, the value of this parameter must be "1".
if(frame_cropping_flag) {		
frame_crop_left_offset	0	Pixels are displayed from the left edge of the decoded image.
frame_crop_right_offset	0	Pixels are displayed up to the right edge of the decoded image.
frame_crop_top_offset	0	Pixels are displayed from the top edge of the decoded image.
frame_crop_bottom_offset	4	The 8 pixels (4 × 2 pixels) from the bottom edge of the decoded image are not displayed.
}		

It is recommended that the IDs of the sequence parameter set (SPS) and picture parameter set (PPS) to be used for H.264 video encoding should be varied according to the resolution.

When encoding a side-by-side 3D image with H.264 video codec, the frame packing arrangement SEI can be assigned in units of sequence to enable the receiving equipment to automatically distinguish a 3D image from a 2D image. Table 6-18 lists the recommended values for the frame packing arrangement SEI. The receiving equipment that complies with the 3D function (an option of this standard) must have the capability to interpret the frame packing arrangement SEI shown in Table 6-18.

The system defined by this standard interprets frames that are combined in a side-by-side format as video source frames. Therefore, the aspect ratio value specified in the sequence parameter set must be the same as that of the video source format of the 2D image (for example, a square pixel aspect ratio for 720p video).

Table 6-18/JJ-40.30: Frame packing arrangement SEI specifications of H.264 video codec

H.264 syntax name	Value	Description
frame_packing_arrangement_id	0	
frame_packing_arrangement_cancel_flag	0	0: Subsequent fields are valid.
frame_packing_arrangement_type	3	Side-by-side 3D image format
quincunx_sampling_flag	0	quincunx_sampling is not applied.
content_interpretation_type	1	1: The left half of the frame is the left-eye image.
spatial_flipping_flag	0	0: Neither the left half nor the right half of frame is inverted.
frame0_flipped_flag	0	Always 0
field_views_flag	0	Always 0
current_frame_is_frame0_flag	0	Always 0
frame0_self_contained_flag	0	The left-eye image might reference the right-eye image.
frame1_self_contained_flag	0	The right-eye image might reference the left-eye image.
frame0_grid_position_x	0	Horizontal position of the top-left pixel of the left-eye image
frame0_grid_position_y	0	Vertical position of the top-left pixel of the left-eye image
frame1_grid_position_x	0	Horizontal position of the top-left pixel of the right-eye image
frame1_grid_position_y	0	Vertical position of the top-left pixel of the right-eye image
frame_packing_arrangement_reserved_byte	0	Always 0
frame_packing_arrangement_repetition_period	1	This SEI is valid during the sequence (until the next IDR).
frame_packing_arrangement_extension_flag	0	Always 0

6.2.3 Video Media Specifications of System Profiles

Table 6-19 lists the video channel processing capabilities that must be implemented by each terminal.

Table 6-19/JJ-40.30: Video media specifications to be implemented by each terminal

Item		Video media specification					
		Profile AVSIP-1	Profile AVSIP-1.5	Profile AVSIP-2a	Profile AVSIP-2b	Profile AVSIP-3	Profile AVSIP-4
Camera image format	QCIF: 176x144	Mandatory	Mandatory	Mandatory	Optional	Optional	Optional
	QVGA: 320x240	Unsupported	Mandatory	Mandatory	Optional	Optional	Optional
	VGA: 640x480	Unsupported	Unsupported	Mandatory	Optional	Optional	Optional
	SD: 704x480	Unsupported	Unsupported	Unsupported	Mandatory	Recommended	Optional
	720p: 1280x720	Unsupported	Unsupported	Unsupported	Unsupported	Mandatory	Optional
	1080i: 1920x1080	Unsupported	Unsupported	Unsupported	Unsupported	Optional	Mandatory
PC screen image format	VGA: 640x480	Optional	Optional	Optional	Conditionally mandatory	Conditionally mandatory	Conditionally mandatory
	SVGA: 800x600	Optional	Optional	Optional	Conditionally mandatory	Conditionally mandatory	Conditionally mandatory
	XGA: 1024x768	Optional	Optional	Optional	Conditionally mandatory	Conditionally mandatory	Conditionally mandatory
No. of channels	1 (camera image)	Mandatory	Mandatory	Mandatory	Mandatory	Mandatory	Mandatory
	2 (camera image + PC screen image)	Optional	Optional	Optional	Optional	Optional	Optional
Camera image encoding (2D)	MPEG-4 Video	Mandatory	Mandatory	Mandatory	Optional	Optional	Optional
	H.264	Optional	Optional	Optional	Mandatory	Mandatory	Mandatory
Camera image encoding (3D)	MPEG-4 Video	Unsupported	Unsupported	Unsupported	Unsupported	Unsupported	Unsupported
	H.264	Optional	Optional	Optional	Optional	Optional	Optional
PC screen image encoding	H.263	Optional	Optional	Optional	Conditionally mandatory (Note 2)	Conditionally mandatory (Note 2)	Conditionally mandatory (Note 2)
	H.264	Optional	Optional	Optional	Conditionally mandatory (Note 2)	Conditionally mandatory (Note 2)	Conditionally mandatory (Note 2)

(Note 1) "Conditionally mandatory" indicates that the item is mandatory when PC screen image display is implemented as an optional capability.

(Note 2) This standard recommends the implementation of either or both of H.263 and H.264 video codecs.

7. End-to-End Control

This standard is intended for systems that use only one video codec capability. Systems that use multiple video codec capabilities will be studied in the future.

7.1 Capability Negotiation

Capability negotiation must comply with TR-1020 [TR-1020]. This standard defines the negotiation procedure with an SDP offer/answer model and an additional fallback function (to retry the call upon receiving an error response). Using the fallback function enables the offering and answering terminals to mutually recognize available media information before sending and receiving a media stream and can prevent incorrect charging (the user is charged for service upon the connection of the line, but does not actually receive video).

The terminal that assumes fallback first sends an SDP offer with the maximum capability of the local terminal. Upon receiving a 488 response, the local terminal judges the "warn-code" in the Warning header of the response, and resends an SDP offer (falls back) with a profile recommended by this standard. Figure 7-1 shows an example of the flow of fallback. For the definition of "warn-code" shown in the example, refer to TR-1020 [TR-1020].

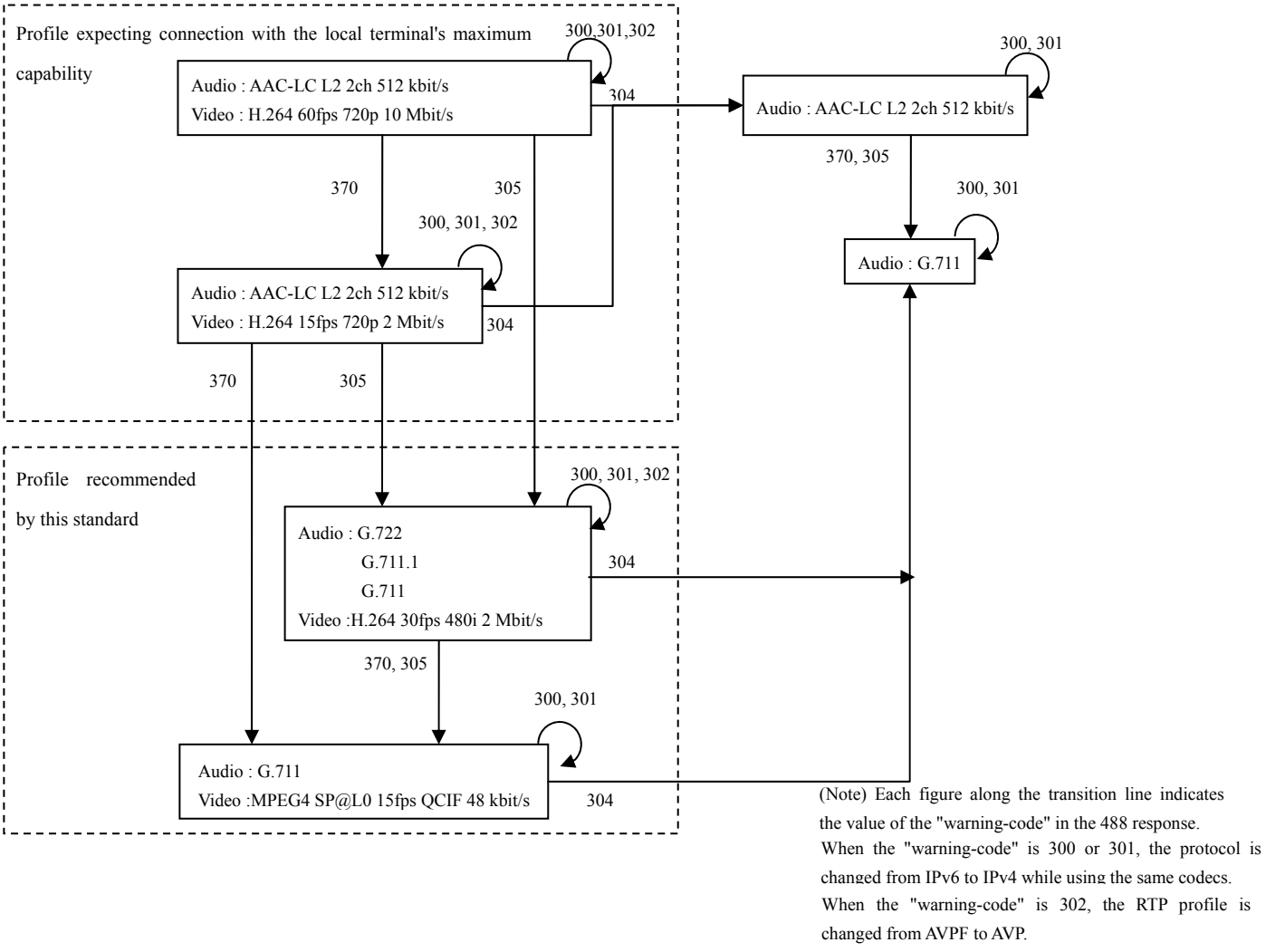


Figure 7-1/JJ-40.30: Example of the flow of fallback

7.2 SDP Description

SDP descriptions must comply with TR-1020 [TR-1020]. This standard defines the SDP descriptions specific to audiovisual communication systems.

Usually, the codec parameters to be used are described in-band. A problem, however, may occur if the in-band codec parameters cannot be interpreted and video or other data is not displayed, even after the call is established. To avoid such a problem, this standard stipulates that the available codec tool set (profile) or the like (which is especially important for determining whether the video codecs to be used can decode data) must be specified in the SDP description, and that, before the call is established, both terminals must determine whether the codecs of both terminals can decode data. Note that the maximum frame rate and maximum bit rate must be specified in the "a=framerate" and "b=AS" parameters, respectively.

Note, however, that the bit rate and frame rate of video data may be negotiated through an exchange of offer and answer according to the following rule provided they are described in the same AVSIP profile:

- Bit rate

  In the SDP offer, the offering terminal must describe the maximum bit rate the offering terminal can use for sending and receiving. In the SDP answer, the answering terminal must describe a bit rate that does not exceed the bit rate described in the SDP offer. Both the offering and answering terminals must send and receive video streams at the bit rate described in the SDP answer.

For details on the contents of the SDP description and the terminal operations at offering and answering, see Annex A.

7.3 Communication Mode Setting

This item will be studied in the future.

7.4 C&I

This item will be studied further in the future.

7.4.1 Remote Camera Control

Remote camera control in a terminal that is compliant with this standard is an optional capability.

To use the remote camera control capability, the requirements of RFC 4573 [RFC4573] must be observed.

The RFC 4573 [RFC4573] document defines media types for H.224, specifies those remote camera control procedures for H.281 and H.224 that are described in Annex Q [H323] of TTC Standard JT-H323, and describes the details of the SDP message contents. Based on these definitions and specifications, camera control information can be sent as RTP packets in a network.

7.4.2 Other Control

This item will be studied in the future.

7.5 Control by RTP or RTCP

This section describes the RTCP feedback messages that are used as codec control messages for feedback from the receiving terminal to the sending terminal and the error recovery based on the feedback. For the specifications of the RTP audiovisual profile (AVPF), refer to JF-IETF-RFC4585 [RFC4585] and JF-IETF-RFC5104 [RFC5104].

This standard recommends the use of AVPF with H.264 video codec. This standard defines the applicable ranges of the feedback messages to be implemented by the terminal among those described in the above documents. Table 7-1 and Table 7-2 show the applicable ranges of the feedback messages.

Table 7-1/JJ-40.30: Applicable ranges of feedback messages (described in JF-IETF-RFC4585)

Feedback message	Applicable range
Generic ACK	Optional
Picture Loss Indication (PLI)	Recommended
Slice Loss Indication (SLI)	Optional
Reference Picture Selection Indication (RPSI)	Optional
Application Layer Feedback	Optional

How the terminal behaves upon the reception of the PLI message may depend on the implementation. This standard does not define the behavior.

Table 7-2/JJ-40.30: Applicable ranges of feedback messages (described in JF-IETF-RFC5104)

Feedback message	Applicable range
Temporary Maximum Media Stream Bit Rate Request (TMMBR)	Optional
Temporary Maximum Media Stream Bit Rate Notification (TMMBN)	Optional
Full Intra Request (FIR)	Mandatory for both sending and receiving
Temporal-Spatial Trade-off Request (TSTR)	Optional
Temporal-Spatial Trade-off Notification (TSTN)	Optional
H.271 Video Back Channel Message (VBCM)	Optional

When a terminal has the capability to send and receive the feedback messages listed in Table 7-1 and Table 7-2, the terminal must express the capability to the distant terminal with the "a=rtcp-fb" line when sending an SDP offer or answer. The sending of feedback messages that the distant terminal is incapable of handling is not recommended. If the "a=rtcp-fb" line is not described in the SDP offer or answer from the distant terminal, the local terminal must determine that the distant terminal does not have the capability to handle any feedback messages. If the capability to handle FIR is not expressed in the "a=rtcp-fb" line, the local terminal must not return an answer with the RTP/AVPF profile but instead return a 488 response to fall back to the RTP/AVP profile.

When the terminal that has the capability to handle FIR in the "a=rtcp-fb" line receives an FIR message, the terminal must respond by sending an I frame (called "IDR picture" in H.264 video codec, "I-VOP" in MPEG-4 Video codec, or "I-picture" in H.263 video codec).

8. End Network Control

The terminal must perform call connection using SIP and SDP in accordance with JT-Q3402 [Q3402].

9. Multiplexing/demultiplexing of Multimedia

To send and receive a bit stream of media (video, audio or data) signal encoded by a specified coding method on an IP network, the terminal that is compliant with this standard must convert the bit stream into IP packets through the application of a prespecified method. After conversion into IP packets, the sending terminal can send the bit stream to the receiving terminal via the IP network.

This chapter describes the rules for converting a bit stream into IP packets and also converting them back. This chapter also describes the matters to be considered to ensure compliance with this standard.

9.1 RTP Conversion

To handle encoded data on an IP network, a terminal that is compliant with this standard must convert the media data into the RTP payload format. Details of the RTP payload format vary according to the coding method, and must conform to the corresponding technical standard shown in Table 9-1.

Table 9-1/JJ-40.30: RTP payload formats

Media	Codec	Applicable standard
Video	H.263 (1998 version)	RFC 4629 [RFC4629]
	H.264	RFC 3984 [RFC3984]
	MPEG-4 Video	RFC 3016 [RFC3016]
Audio	G.711	RFC 3551 [RFC3551]
	G.722	
	G.711.1	RFC 5391 [RFC5391]
	MPEG-4 AAC	RFC 3016 [RFC3016]

9.1.1 H.263 Video Codec

To send and receive the H.263 video bit stream that is compliant with H.263 (1998 version) video codec in RTP format, the RTP payload format defined by RFC 4629 [RFC4629] must be used.

9.1.2 H.264 Video Codec

To send and receive the H.264 video bit stream that is compliant with H.264 video codec in RTP format, the RTP payload format defined by RFC 3984 [RFC3984] must be used.

RFC 3984 [RFC3984] defines seven types of payload structure because the structure of the RTP payload format varies with the method used to aggregate the NAL units of the H.264 video bit stream in the RTP payload. This standard recommends the use of either one or both of the single NAL unit packet and FU-A packet in non-interleaved mode (packetization-mode=1) among the seven types of payload structure, and excludes the use of other packets, including the STAP-A packet. In particular, the FU-A packet should be used for the High Profile using CABAC.

When an IDR frame is inserted, the PPS/SPS information to be referenced by the IDR frame must be attached to the top of every IDR frame before the frame is sent. When RTP/AVPF is not used, the periodic insertion of an IDR frame or intra-macroblock refresh process is recommended to refresh the image. Note that there may be some encoders that do not periodically insert an IDR frame. If the connectivity of a decoder with such an encoder is considered, the decoder is also required to perform, upon the occurrence of a decoding error, a process that does not assume the reception of an IDR frame. For example, the decoder should not perform a process in which presentation of a decoded image is suspended until an IDR frame is decoded.

This standard recommends that SPS information should be specified as the "sprop-parameter-sets" value in the

"a=fmtp" line of the SDP description when the resolutions that can be received within the range of the level specified in the SDP description are constrained. Then, the RTP packets carrying the H.264 bit stream to be sent by the sending terminal must contain the bit stream that corresponds to the codec parameter value. Although this standard recommends the use of the SPS information sending procedure using SDP, the same SPS information as the "sprop-parameter-sets" value must be included in the RTP packets in order to improve interoperability.

9.1.3 MPEG-4 Video Codec

To send and receive an MPEG-4 Video bit stream that is compliant with MPEG-4 Video codec in the RTP format, the RTP payload format defined by RFC 3016 [RFC3016] must be used.

This standard requires a DCI value to be specified as the "config" value in the "a=fmtp" line of the SDP description. Then, the RTP packets carrying the MPEG-4 Video bit stream to be sent by the sending terminal must contain the bit stream that corresponds to the DCI value. Although this standard defines the DCI value sending procedure using SDP as mandatory, the same DCI value as the "config" value must be included in the RTP packets in order to improve interoperability.

It is recommended that resync_marker be used as the error resilient tool, while the use of data_partioned or reversible_vlc should be avoided as the error resilient tool.

When RTP/AVPF is not used, it is recommended that an I-VOP frame be inserted periodically, with the config parameters attached to the top of the I-VOP frame.

At RTP packetization, one of the three fragmentation methods, such as (a) DCI information only, (b) DCI information and video information, and (d) video information only, must be used among those defined by RFC 3016 [RFC3016].

The maximum size of the video packet may be specified by the administration.

9.1.4 G.711 Audio or G.722 Audio Codec

To send and receive an audio bit stream compliant with G.711 or G.722 audio codec in the RTP format, the RTP payload format defined by RFC 3551 [RFC3551] must be used.

9.1.5 G.711.1 Audio Codec

To send and receive an audio bit stream compliant with G.711.1 Audio codec in the RTP format, the RTP payload format defined by RFC 5391 [RFC5391] must be used.

When using G.711.1 Audio codec, a G.711μ-law audio stream must be used as the base stream.

9.1.6 MPEG-4 Audio Codec

To send and receive an MPEG-4 Audio bit stream compliant with MPEG-4 Audio codec in the RTP format, the RTP payload format defined by RFC 3016 [RFC3016] must be used.

The MPEG-4 Audio bit stream must be in LATM format.

9.1.7 Media Packet Sending Based on the Results of Bandwidth Negotiation

Each terminal performs encoding according to the results of negotiation based on the offer/answer model, and then sends the generated media packets.

The bit rate value specified in the "b=" line of the SDP answer is applied as the bit rate for sending media packets. Note that the information specified in the "b=" line of the SDP description includes the header information of Layer 3 and Layer 4, RTP header, RTP payload header, and the bit rate of the elementary stream.

Note, however, that the "b=" line of the SDP description may be omitted when the audio codec being used has a specific bit rate as in the case of G.711 codec. In such cases, media packets are sent at the bit rate specific to the audio codec.

10. Definition of UNI

Each terminal is connected to a network via a UNI defined by TR-1014 [TR-1014], and connects a call using SIP and SDP and media sending and receiving using RTP in accordance with JT-Q3402 [Q3402].

Regarding media sending and receiving, special attention must be paid to the QoS functions defined by JT-Q3402 [Q3402]. For the sending traffic conditions allowed by the network for the bandwidth negotiated on SDP, refer to Annex g of JT-Q3402 [Q3402].

Annex A SDP Description and Negotiation Using SDP

(This annex forms part of the specifications.)

A.1 Overview

Negotiation using SIP/SDP as described in this standard complies with JT-Q3402 [Q3402] and TR-1020 [TR-1020]. This annex addresses the SDP descriptions specific to individual audio and video codecs and the procedure for checking the codec parameters in the case of a negotiation using SIP/SDP performed by an audiovisual communication system in an SIP network environment.

A.2 SDP Descriptions and Negotiation Procedures

This section and its subsections specify the SDP descriptions and negotiation procedures related to video and audio codecs. This section and its subsections describe the details of the lines and parameters of SDP descriptions, signal conditions, and the terminal operations for SDP offering and answering, regarding the codecs that are used for the AVSIP profiles defined in Chapter 6.

In the "Signal condition" column of the tables shown below, the condition of each line or option is represented by a condition code, e.g., "m" (mandatory) or "o" (optional), from a dynamic viewpoint as to whether the line or option appears in an interface signal on the UNI. For example, when the condition code for a line or option is "m" (mandatory), the line or option must always be described in the message that is sent or received by the terminal. Table A-1 lists the definitions of the condition codes.

Table A-1/JJ-40.30: Definitions of condition codes

Condition code	Definition
m	The relevant parameter is mandatory. A mandatory parameter of an SDP offer must always exist in the SDP offer and be understandable by the terminal receiving the offer in accordance with the corresponding RFC. Similarly, a mandatory parameter of an SDP answer must always exist in the SDP answer and be understandable by the terminal processing the answer in accordance with the corresponding RFC.
m*	The relevant parameter should exist in each SDP description. The terminal or network receiving the SDP description must be prepared for the case in which the header field corresponding to the relevant parameter does not exist.
o	The relevant parameter is optional. An optional parameter may exist in an offer or answer. When an optional parameter exists in an offer or answer, the parameter must be understood in accordance with the corresponding RFC by the receiving terminal and the operation corresponding to the parameter must be performed.
-	The relevant parameter is not applied. It must not exist in any offer or answer.
c	Application of the relevant parameter is conditional, and depends on the context of the SDP process.

A.2.1 Video Codecs

A.2.1.1 MPEG-4 Video codec

Table A-2 lists the codec parameters to be specified in the SDP media description when MPEG-4 Video [MPEG-4 Video] codec is used, as well as the terminal operations performed when sending and receiving SDP.

Table A-2/JJ-40.30: Codec parameters described on SDP (MPEG-4 Video codec)

Codec	SDP description		Signal condition		Terminal operation at sending/receiving SDP	
	Line	Parameter	Send	Receive	Offering	Answering
MPEG-4 Video	b=AS		m	m	• Setting is mandatory. When using an AVSIP profile, specify a value according to Table 6-11.	• Return a 200 response only when sending and receiving can be done with the offered value for the bandwidth. When communication cannot be done, return a 488 response. • In the SDP answer to return the 200 response, set the same value as the offered value.
	a=rtpmap	payload type	m	m	• Specify a value from 96 to 127 according to RFC 3016.	• Return the same "payload type" value as the offered value.
		encoding name	m	m	• Specify "MP4V-ES" according to RFC 3016.	• Specify "MP4V-ES" according to RFC 3016.
		clock rate	m	m	• Specify "90000" according to RFC 3016.	• Specify "90000" according to RFC 3016.
		encoding parameters	-	o	• Omit this parameter according to RFC 4566 and RFC 3016.	• Return a 488 response when the parameter is set in the offer. • Do not set the value in the answer.
	a=fmtp	profile-level-id	m	o	• Setting is mandatory. When using an AVSIP profile, specify a value according to Table 6-11.	• When the "profile-level-id" parameter is not set, interpret the parameter as meaning that "1" is specified according to RFC 3016. • Return a 200 response only when sending and receiving can be done with the offered profile and level. When communication cannot be done, return a 488 response (warn-code is "305"). • In the SDP answer to return the 200 response, set the same value as the offered value.

		config	m	o	• Specify the values of the decoder configuration information (DCI) allowed in the profile and level specified in the "profile-level-id" parameter within the scope of ISO/IEC 14496-2. • When an AVSIP profile is used, the values of the screen size parameters (video_object_layer_width [horizontal number of pixels] and video_object_layer_height [vertical number of pixels]) specified as DCI values must satisfy the specifications listed in Table 6-8 and Table 6-11.	• When the "config" parameter is not set, return a 200 response if the answerer can support the maximum capability in the profile and level specified by the "profile-level-id" parameter (Note 1). When communication cannot be done, return a 488 response. • When the "config" parameter is set, return a 200 response only if sending and receiving as well as displaying for the user can be done with the screen size specified as the DCI value (video_object_layer_width [horizontal number of pixels] and video_object_layer_height [vertical number of pixels]) (Note 1). If communication cannot be done, return a 488 response (warn-code is "305"). When Simple Profile is used, the answerer must accept the offered values of other "config" parameters and be able to perform decoding with those values, provided the offered values are within the respective ranges defined by ISO/IEC 14496-2.
		Other parameters in the "a=fmtp" line	-	o	• Omit the parameters. (They are not defined by RFC 3016 [RFC3016].)	• Ignore the parameters.
	a=framerate		o	o	• The parameter may be set. When using an AVSIP profile, specify a value according to Table 6-11.	• When the "a=framerate" line is not specified, interpret the parameter as meaning that the maximum frame rate for the relevant AVSIP profile is offered. • Return a 200 response only when communication can be done with a frame rate that does not exceed the offered rate. When communication cannot be done, return a 488 response. • In the SDP answer to return the 200 response, set a value that does not exceed the offered value. (Note 2)

(Note 1) The offerer performs communication by using the "config" parameter values described in the received answer; the answerer performs communication by using the "config" parameter values described in the received offer.

(Note 2) Both the offerer and answerer perform communication by using the values in the "a=framerate" line described in the answer.

A.2.1.2 H.263 video codec

Table A-3 lists the codec parameters to be specified in the SDP media description when H.263 [H263] video codec is used, as well as the terminal operations performed when sending and receiving SDP.

Table A-3/JJ-40.30: Codec parameters described on SDP (H.263 video codec)

Codec	SDP description		Signal condition		Terminal operation at sending/receiving SDP	
	Line	Parameter	Send	Receive	Offering	Answering
H.263	b=AS		m	m	• Setting is mandatory. When using an AVSIP profile, specify a value according to Table 6-12.	• Return a 200 response only when sending and receiving can be done with the offered value for the bandwidth. When communication cannot be done, return a 488 response. • In the SDP answer to return the 200 response, set the same value as the offered value.
	a=rtpmap	payload type	m	m	• Specify a value from 96 to 127 according to RFC 4629.	• Return the same "payload type" value as the offered value.
		encoding name	m	m	• Specify "H263-1998" according to RFC 4629.	• Specify "H263-1998" according to RFC 4629.
		clock rate	m	m	• Specify "90000" according to RFC 4629.	• Specify "90000" according to RFC 4629.
		encoding parameters	-	o	• Omit this parameter according to RFC 4566 and RFC 4629.	• Return a 488 response when the parameter is set in the offer. • Do not set the value in the answer.
	a=fmtp	QCIF	m	o	• Setting is mandatory. (Note 1)	• Reference this parameter only for communication using QCIF. Return a 200 response when sending and receiving as well as displaying for the user can be done with the offered value. Otherwise, return a 488 response. • When using an AVSIP profile, ignore this parameter. (Note 1)
		CUSTOM	o	o	• When using an AVSIP profile, specify a value according to "PC screen image format" in Table 6-19.	• Return a 200 response when sending and receiving as well as displaying for the user can be done with the specified "CUSTOM" parameter value. Otherwise, return a 488 response. • In the SDP answer to return the 200 response, set the same screen resolution as the offered "CUSTOM" parameter value, and a frame rate that does not exceed the offered frame rate. (Note 2)

		PAR	o	o	• When using an AVSIP profile, specify a value according to Table 6-12.	• When the "PAR" parameter is not specified, interpret the parameter as meaning that the "12:11" is offered according to RFC 4629. • Return a 200 response only when sending and receiving as well as displaying for the user can be done with the offered aspect ratio of pixels. Otherwise, return a 488 response. • In the SDP answer to return the 200 response, set the same "PAR" value as the offered value.
		Other parameters described in RFC 4629	o	o	• These parameters may be set according to their definitions by RFC 4629.	• Interpret the parameters according to specifications of RFC 4629, and return a 200 response only when communication (both sending and receiving) can be done. When communication cannot be done, return a 488 response.
	a=framerate		o	o	• This parameter may be set. When setting, specify a value that does not conflict with the "a=fmtp" line parameters, e.g., CUSTOM, specifying a frame rate. • When using an AVSIP profile, specify a value according to Table 6-12.	• If the "a=framerate" line is specified, return a 200 response only when communication can be done with a frame rate that does not exceed the offered frame rate. When communication cannot be done, return a 488 response. In the SDP answer to return the 200 response, set a value that does not exceed the offered value (Note 2). The value to be set must not conflict with the "a=fmtp" line parameters, e.g., CUSTOM, specifying a frame rate. • If the "a=framerate" line is not specified, do not specify the "a=framerate" line in the answer but instead specify only the "a=fmtp" line parameters, e.g., CUSTOM, that specify the frame rate.

(Note 1) Although the "QCIF" parameter is a mandatory parameter based on the definition by RFC 4629, the parameter need not be referenced by the system that uses an AVSIP profile according to the specification defined in Table 6-12.

(Note 2) Both the offerer and answerer perform communication by using the values in the "a=framerate" line described in the answer.

A.2.1.3 H.264 video codec

Table A-4 lists the codec parameters to be specified in the SDP media description when H.264 [H264] video codec is used, as well as the terminal operations performed when sending and receiving SDP.

Table A-4/JJ-40.30: Codec parameters described on SDP (H.264 video codec)

Codec	SDP description		Signal condition		Terminal operation at sending/receiving SDP	
	Line	Parameter	Send	Receive	Offering	Answering
H.264	b=AS		m	m	• Setting is mandatory. When using an AVSIP profile, specify a value according to Table 6-13.	• Return a 200 response only when sending and receiving can be done with the offered value for the bandwidth. When communication cannot be done, return a 488 response. • In the SDP answer to return the 200 response, set the same value as the offered value.
	a=rtpmap	payload type	m	m	• Specify a value from 96 to 127 according to RFC 3984.	• Return the same "payload type" value as the offered value.
		encoding name	m	m	• Specify "H264" according to RFC 3984.	• Specify "H264" according to RFC 3984.
		clock rate	m	m	• Specify "90000" according to RFC 3984.	• Specify "90000" according to RFC 3984.
		encoding parameters	-	o	• Omit this parameter according to RFC 4566 and RFC 3984.	• Return a 488 response when the parameter is set in the offer. • Do not set the value in the answer.
	a=fmtp	profile-level-id	m	o	• Setting is mandatory. When using an AVSIP profile, specify a value according to Table 6-13.	• When the "profile-level-id" parameter is not set, interpret the parameter as meaning that Baseline Profile Level 1 is offered according to RFC 3984. • Return a 200 response only when sending and receiving can be done with the offered profile and level. When communication cannot be done, return a 488 response (warn-code is "305"). • In the SDP answer to return the 200 response, set the same values as the offered values. Note, however, that the value of "constraint_set2_flag" may be different.
		packetization-mode	c1	o	• Specify a value when the value is not the default ("0": single NAL mode). When using a fragmentation unit, specify its use explicitly. (Note 1)	• When the "packetization-mode" parameter is not set, interpret the parameter as meaning that "0" (single NAL mode) is specified according to RFC 3984. • Return a 200 response only when sending and receiving can be done with the offered "packetization-mode" value. When communication cannot be done, return a 488 response (warn-code is "305"). • In the SDP answer to return the 200 response, set the same value as the offered value.

		sprop-parameter-sets	o	o	• This parameter may be set according to the definition by RFC 3984. • When an AVSIP profile is used, the setting of this parameter is recommended. When the resolutions available for sending and receiving are constrained, describe the resolutions available for sending and receiving as SPS information. Note that resolution information need not be described in the "sprop-parameter-sets" parameter if the resolutions available for sending include that resolution which is mandatory for AVSIP and the resolutions available for receiving are not constrained.	• When the "sprop-parameter-sets" parameter is not specified, perform communication with a resolution that can be set in the level specified by profile-level-id. • When using an AVSIP profile, describe the resolutions available for sending and receiving in the "sprop-parameter-sets" parameter if the offered resolutions include the resolutions available for sending. If all the offered resolutions are available for receiving, the resolutions need not be described in the parameter. In either case, return a 200 response. If the SPS information excludes the offered resolutions available for sending and receiving, return a 488 response.
		max-mbps	c2	o	• These parameters may be set according to their definitions by RFC 3984.	• Interpret the parameters according to the specifications of RFC 3984, and return a 200 response only when communication (both sending and receiving) can be done. When communication cannot be done, return a 488 response. • When the values of the "max-mbps," "max-fs," "max-br," "max-cpb," and "max-dpb" parameters are offered, set the same values in the answer to return the 200 response.
		max-fs	c2	o		
		max-cpb	c2	o		
		max-dpb	c2	o		
		max-br	c2	o		
		redundant-pic-cap	c2	o		
		parameter-add	o	o	• These parameters may be set according to their definitions by RFC 3984. • When using an AVSIP profile, omit these parameters. (Note 2)	
		Other parameters described in RFC 3984	c3	o		
	a=framerate		o	o	• The parameter may be set. When using an AVSIP profile, specify a value according to Table 6-13.	• When the "a=framerate" line is not specified, interpret the parameter as meaning that the maximum frame rate for the relevant AVSIP profile is offered. • Return a 200 response only when communication can be done with a frame rate that does not exceed the offered frame rate (Note 3). When communication cannot be done, return a 488 response. • In the SDP answer to return the 200 response, set a value that does not exceed the offered value.

c1:	Setting of this parameter is not mandatory in the single NAL mode. Setting of this parameter is mandatory when using a packetization mode other than the single NAL mode.
c2:	According to Annex A of H.264, the setting of this parameter is mandatory when using a value not less than the value corresponding to the profile and level specified by profile-level-id.
c3:	Setting of this parameter is mandatory when its setting is required by the specification of RFC 3984 according to the setting of the codec to be used.
(Note 1)	The single NAL mode can be applied even in the case of broadband communication if the slice size of H.264 video is controlled. The use of the single NAL mode, however, will lower the encoding efficiency because the overhead due to the slice header increases and , the intra prediction over the slice size and the motion prediction for the motion vector information become unusable. Therefore, a fragmentation unit should be used when the slice size cannot be controlled within the MTU size or when the encoding efficiency must be considered. A fragmentation unit should also be used when CABAC is used because CABAC requires the consideration of problems in its implementation such as the lower efficiency of compressing shorter slices and the relatively larger computing load at the top of slices.
(Note 2)	When using an AVSIP profile, other parameters need not be set because their defaults defined by RFC 3984 will be used or communication will be performed within the range of the capabilities of the specified profile and level.

A.2.2 Audio Codecs

A.2.2.1 G.711μ-law audio codec

Table A-5 lists the codec parameters to be specified in the SDP media description when G.711μ-law [G711] audio codec is used, as well as the terminal operations performed when sending and receiving SDP.

Table A-5/JJ-40.30: Codec parameters described on SDP (G.711μ-law audio codec)

Codec	SDP description		Signal condition		Terminal operation at sending/receiving SDP	
	Line	Parameter	Send	Receive	Offering	Answering
G.711μ-law	a=rtpmap	payload type	m	m	• Specify "0" according to RFC 3551.	• Specify "0" according to RFC 3551.
		encoding name	m	m	• Specify "PCMU" according to RFC 3551.	• Specify "PCMU" according to RFC 3551.
		clock rate	m	m	• Specify "8000" according to RFC 3551.	• Specify "8000" according to RFC 3551.
		encoding parameters	-	o	• Omit this parameter.	• Return a 488 response when the parameter is set in the offer. • Do not set the value in the answer.
	a=ptime		m	o	• Specify "20" according to TR-1020.	• When the "ptime" line is not set, interpret the parameter as meaning that "20" is specified according to JT-Q3402. • Return a 200 response only when sending and receiving can be done with the offered packetization period. When communication cannot be done, return a 488 response. • In the SDP answer to return the 200 response, set the same value as the offered value.

A.2.2.2 G.722 audio codec

Table A-6 lists the codec parameters to be specified in the SDP media description when G.722 [G722] audio codec is used, as well as the terminal operations performed at sending and receiving SDP.

Table A-6/JJ-40.30: Codec parameters described on SDP (G.722 audio codec)

Codec	SDP description		Signal condition		Terminal operation at sending/receiving SDP	
	Line	Parameter	Send	Receive	Offering	Answering
G.722	a=rtpmap	payload type	m	m	• Specify "9" according to RFC 3551.	• Specify "9" according to RFC 3551.
		encoding name	m	m	• Specify "G722" according to RFC 3551.	• Specify "G722" according to RFC 3551.
		clock rate	m	m	• Specify "8000" according to RFC 3551.	• Specify "8000" according to RFC 3551.
		encoding parameters	-	o	• Omit this parameter.	• Return a 488 response when the parameter is set in the offer. • Do not set the value in the answer.
	a=ptime		m	o	• Specify "20" according to TR-1020.	• When the "ptime" line is not set, interpret the parameter as meaning that "20" is specified in a similar manner to the specification for G.711μ-law. • Return a 200 response only when sending and receiving can be done with the offered packetization period. When communication cannot be done, return a 488 response. • In the SDP answer to return the 200 response, set the same value as the offered value.

A.2.2.3 G.711.1 audio codec

Table A-7 lists the codec parameters to be specified in the SDP media description when G.711.1 [G711.1] audio codec is used, as well as the terminal operations performed at sending and receiving SDP.

Table A-7/JJ-40.30: Codec parameters described on SDP (G.711.1 audio codec)

Codec	SDP description		Signal condition		Terminal operation at sending/receiving SDP	
	Line	Parameter	Send	Receive	Offering	Answering
G.711.1	a=rtpmap	payload type	m	m	• Specify a value from 96 to 127 according to RFC 5391.	• Return the same "payload type" value as the offered value.
		encoding name	m	m	• Specify "PCMU-WB" according to RFC 5391.	• Specify "PCMU-WB" according to RFC 5391.
		clock rate	m	m	• Setting is mandatory. When using an AVSIP profile, specify a value according to Table 6-4.	• Return a 200 response only when sending and receiving can be done with the offered value. When communication cannot be done, return a 488 response. • In the SDP answer to return the 200 response, set the same value as the offered value.
		encoding parameters	-	o	• Omit this parameter.	• Return a 488 response when the parameter is set in the offer. • Do not set the value in the answer.
	a=fmtp	mode-set	m	o	• Setting is mandatory. When using an AVSIP profile, specify a value according to Table 6-4.	• Return a 200 response only when sending and receiving can be done with the offered mode. When communication cannot be done, return a 488 response. • In the SDP answer to return the 200 response, set the same value as the offered value.
		Other parameters in "a=fmtp" line	-	o	• Omit the parameters. (They are not defined by RFC 5391.)	• Ignore these parameters.
	a=ptime		m	o	• Setting is mandatory. When using an AVSIP profile, specify a value according to Table 6-4.	• When the "ptime" line is not set, interpret the parameter as meaning that "20" is specified in a similar manner to the specification for G.711μ-law. • Return a 200 response only when sending and receiving can be done with the offered packetization period. When communication cannot be done, return a 488 response. • In the SDP answer to return the 200 response, set the same value as the offered value.

A.2.2.4 MPEG-4 Audio codec

Table A-8 lists the codec parameters to be specified in the SDP media description when MPEG-4 Audio [MPEG-4 Audio] codec is used, as well as the terminal operations performed at sending and receiving SDP.

Table A-8/JJ-40.30: Codec parameters described on SDP (MPEG-4 Audio codec)

Codec	SDP description		Signal condition		Terminal operation at sending/receiving SDP	
	Line	Parameter	Send	Receive	Offering	Answering
MPEG-4 Audio	b=AS		m	m	• Setting is mandatory. When using an AVSIP profile, specify a value according to Table 6-5.	• Return a 200 response only when sending and receiving can be done with the offered value for the bandwidth. When communication cannot be done, return a 488 response. • In the SDP answer to return the 200 response, set the same value as the offered value.
	a=rtpmap	payload type	m	m	• Specify a value from 96 to 127 according to RFC 3016.	• Return the same "payload type" value as the offered value.
		encoding name	m	m	• Specify "MP4A-LATM" according to RFC 3016.	• Specify "MP4A-LATM" according to RFC 3016.
		clock rate	m	m	• Setting is mandatory. When using an AVSIP profile, specify a value according to Table 6-5.	• Return a 200 response only when sending and receiving can be done with the offered value. When communication cannot be done, return a 488 response.
		encoding parameters	-	o	• Omit this parameter.	• Return a 488 response when the parameter is set in the offer. • Do not set the value in the answer.
	a=fmtp	profile-level-id	m	o	• Setting is mandatory. When using an AVSIP profile, specify a value according to Table 6-5.	• When the "profile-level-id" parameter is not set, interpret the parameter as meaning that "30" (Natural Audio Profile/Level 1) is specified according to RFC 3016. • Return a 200 response only when sending and receiving can be done with the offered value. When communication cannot be done, return a 488 response.
		object	m	o		• Return a 200 response only when sending and receiving can be done with the offered value. When communication cannot be done, return a 488 response.
		config	m	o		
		bitrate	m	o		
		cpresent	c1	o	• The parameter may be set according to RFC 3016. When using an AVSIP profile, specify "1".	• When the "cpresent" parameter is not set, interpret the parameter as meaning that "1" is specified according to RFC 3016. • In the SDP answer to return the 200 response, set the same value as the offered value.
		Other parameters in the "a=fmtp" line	-	o	• Omit the parameters. (They are not defined by RFC 3016.)	• Ignore these parameters.

	a=ptime	m	o	• Setting is mandatory. When using an AVSIP profile, specify a value according to Table 6-5.	• When the "ptime" line is not set, interpret the parameter as meaning that "20" is specified in a similar manner to the specification for G.711μ-law. • Return a 200 response only when sending and receiving can be done with the offered packetization period. When communication cannot be done, return a 488 response. • In the SDP answer to return the 200 response, set the same value as the offered value.
c1: "0" must be set in this parameter when the config information is not described in the RTP payload.					

Annex B Method of Resolution Negotiation for H.264 Video Codec

(This annex forms part of the specifications.)

B.1 Overview

In the media negotiation for H.264 video codec defined by this standard, the SPS information representing resolutions is described in the "sprop-parameter-sets" parameter contained in the "a=fmtp" line of an SDP description, and the parameter is checked by the offerer and answerer. The operations of the offerer and answerer are described in Table A-4. This annex particularly notes the detailed method of describing the "sprop-parameter-sets" parameter.

B.2 Method of Describing the Resolution

The offerer can represent the resolutions available for sending and receiving by SPS information and describe the SPS information as values of the "sprop-parameter-sets" parameter. Note, however, that, when the H.264 video codec profile and level to be used correspond to the characteristics of the AVSIP profile defined in this standard, the values of the "sprop-parameter-sets" parameter must include the mandatory resolution defined for the AVSIP profile.

The answerer can represent the resolutions available for sending and receiving by SPS information and describe the SPS information as values of the "sprop-parameter-sets" parameter. Note, however, that, when the H.264 video codec profile and level to be used correspond to the characteristics of the AVSIP profile defined in this standard, the values of the "sprop-parameter-sets" parameter must include the mandatory resolution defined for the AVSIP profile.

If SPS information is not described in the "sprop-parameter-sets" parameter, a resolution in the range specified for the H.264 video codec level to be used is interpreted as being available for receiving, and in this case, if the H.264 profile and level described in the SDP description corresponds to the characteristics of the AVSIP profile defined in this standard, the mandatory resolution of the AVSIP profile is interpreted as being available for sending.

The SPS information that is described in the "sprop-parameter-sets" parameter must not conflict with the SPS description in the H.264 video stream to be sent.

B.3 Resolution of the Stream To Be Sent

When a terminal receives an SDP description that includes SPS information, the terminal selects one of the resolutions described in the SPS information, and sends video with the selected resolution.

When a terminal receives an SDP description that excludes SPS information, the terminal can send video with a resolution specified for the relevant H.264 video codec level.

When a terminal receives an SDP description that excludes SPS information or includes multiple items of SPS information, the terminal interprets the description as meaning that the resolution of the video to be sent can be switched dynamically during communication. If a terminal cannot receive a stream that involves switching of the resolution, the receiving terminal must describe only one item of SPS information specifying a resolution available for receiving in the SDP description.

B.4 Examples of Offer/Answer Negotiation

This section gives examples of the Offer/Answer Negotiation. The operation examples assume that the same H.264 profile and level are used for each resolution.

B.4.1 When the Offerer and Answerer Share a Resolution Available for Sending and Receiving

B.4.1.1 Example 1

The offerer can send and receive only the 1080i video of H.264 High Profile Level 4.0. The answerer can send and receive both the 1080i and 720p videos of H.264 High Profile Level 4.0. The following shows an example in which communication is started with H.264 High Profile Level 4.0 and the resolution of 1080i as a result of the Offer/Answer Negotiation:

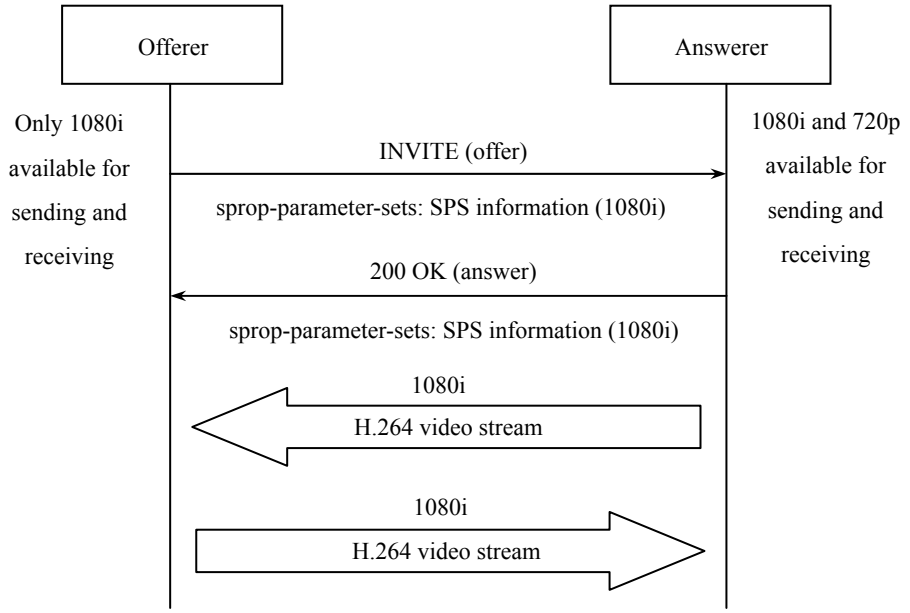


Figure B-1/JJ-40.30: When the offerer and answerer share a resolution available for sending and receiving (example 1)

[Offer]

An example of the setting of the "a=fmtp" line in the offer is shown. Here, the settings in the offer are H.264 High Profile Level 4.0, non-interleaved mode (packetization-mode=1), and 1080i as the video source format.

```
m=video 49170 RTP/AVP 105
a=rtpmap:105 H264/90000
a=fmtp:105
profile-level-id=640028;packetization-mode=1;sprop-parameter-sets=Z2QAKKwspAHgER9o
```

[Answer]

An example of the setting of the "a=fmtp" line in the answer is shown. Here, the settings in the answer are H.264 High Profile Level 4.0, non-interleaved mode (packetization-mode=1), and 1080i as the video source format.

```
m=video 49170 RTP/AVP 105
a=rtpmap:105 H264/90000
a=fmtp:105
profile-level-id=640028;packetization-mode=1;sprop-parameter-sets=Z2QAKKwspAHgER9o
```

B.4.1.2 Example 2

The offerer can send and receive both the 1080i and 720p videos of H.264 High Profile Level 4.0. The answerer can send and receive only the 1080i video of H.264 High Profile Level 4.0. The following shows an example in which communication is started with H.264 High Profile Level 4.0 and the resolution of 1080i as a result of the Offer/Answer Negotiation:

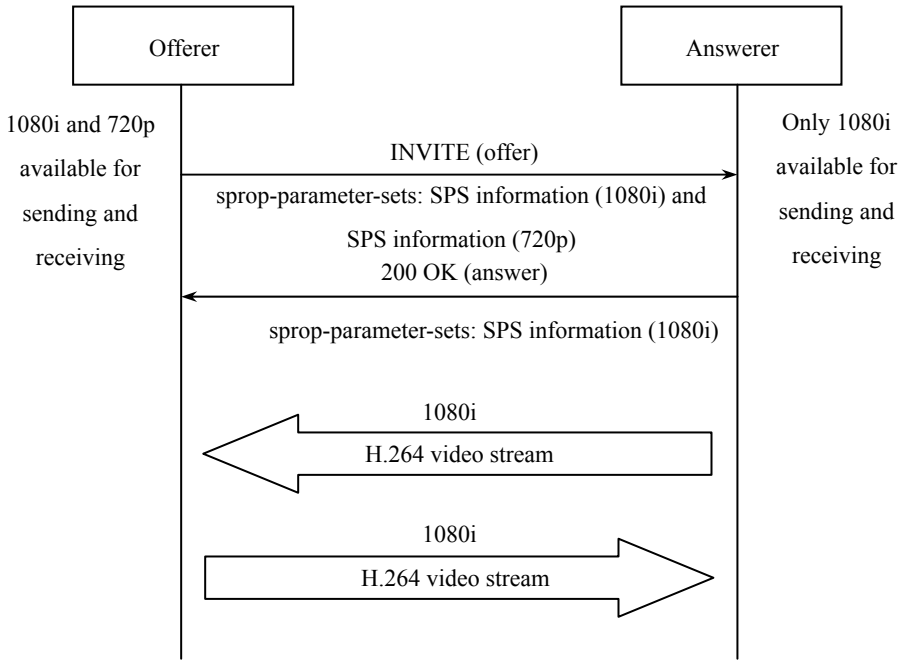


Figure B-2/JJ-40.30: When the offerer and answerer share a resolution available for sending and receiving (example 2)

[Offer]

An example of the setting of the "a=fmtp" line in the offer is shown. Here, the settings in the offer are H.264 High Profile Level 4.0, non-interleaved mode (packetization-mode=1), and 1080i or 720p as the video source format.

```
m=video 49170 RTP/AVP 105
a=rtpmap:105 H264/90000
a=fmtp:105
profile-level-id=640028;packetization-mode=1;sprop-parameter-sets=Z2QAKKwspAHgER9o,
Z2QAKKwspAFAFuQ=
```

[Answer]

An example of the setting of the "a=fmtp" line in the answer is shown. Here, the settings in the answer are H.264 High Profile Level 4.0, non-interleaved mode (packetization-mode=1), and 1080i as the video source format.

```
m=video 49170 RTP/AVP 105
a=rtpmap:105 H264/90000
a=fmtp:105
profile-level-id=640028;packetization-mode=1;sprop-parameter-sets=Z2QAKKwspAHgER9o
```

B.4.1.3 Example 3

The offerer can send and receive both the 1080i and 720p videos of H.264 High Profile Level 4.0. The answerer can send and receive both the 720p and XGA videos of H.264 High Profile Level 4.0. The following shows an example in which communication is started with H.264 High Profile Level 4.0 and the resolution of 720p as a result of the Offer/Answer Negotiation:

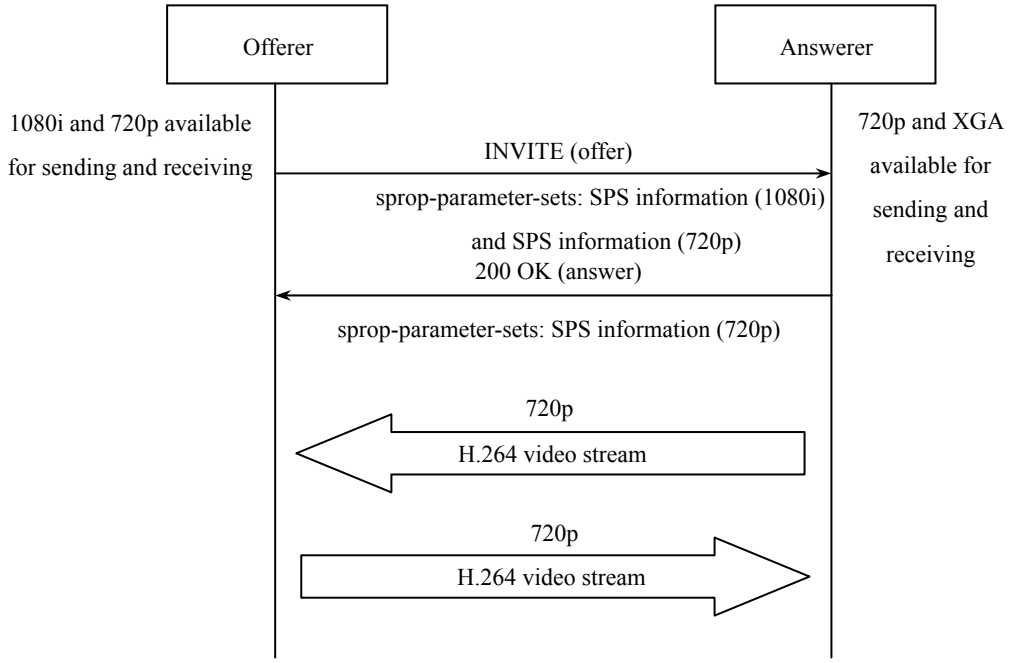


Figure B-3/JJ-40.30: When the offerer and answerer share a resolution available for sending and receiving (example 3)

[Offer]

An example of the setting of the "a=fmtp" line in the offer is shown. Here, the settings in the offer are H.264 High Profile Level 4.0, non-interleaved mode (packetization-mode=1), and 1080i or 720p as the video source format.

```
m=video 49170 RTP/AVP 105
a=rtpmap:105 H264/90000
a=fmtp:105
profile-level-id=640028;packetization-mode=1;sprop-parameter-sets=Z2QAKKwspAHgER9o,
Z2QAKKwspAFAFuQ=
```

[Answer]

An example of the setting of the "a=fmtp" line in the answer is shown. Here, the settings in the answer are H.264 High Profile Level 4.0, non-interleaved mode (packetization-mode=1), and 720p as the video source format.

```
m=video 49170 RTP/AVP 105
a=rtpmap:105 H264/90000
a=fmtp:105
profile-level-id=640028;packetization-mode=1;sprop-parameter-sets=Z2QAKKwspAFAFuQ=
```

B.4.2 When the Offerer Can Send Video of the Mandatory Resolution and Receive Video of Any Resolution

B.4.2.1 Example 1

The offerer can send video of at least the resolution (1080i) that is mandatory for AVSIP-4 among those specified by H.264 High Profile Level 4.0, and receive video of any resolution. The answerer can send and receive both the 1080i and XGA videos of H.264 High Profile Level 4.0. The following shows an example in which communication is started with H.264 High Profile Level 4.0 and the resolution of 1080i or XGA as a result of the Offer/Answer Negotiation. Note that, in this example, the resolution can be switched between 1080i and XGA at any time during communication.

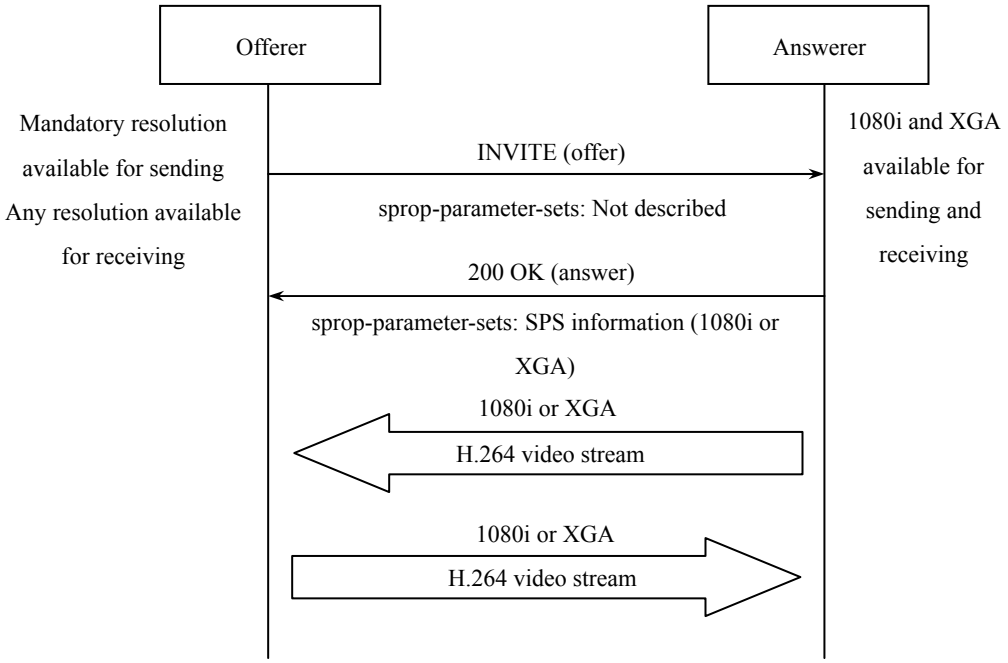


Figure B-4/JJ-40.30: When the offerer can send video of the mandatory resolution and receive video of any resolution (example 1)

[Offer]

An example of the setting of the "a=fmtp" line in the offer is shown. Here, the settings in the offer are H.264 High Profile Level 4.0 and non-interleaved mode (packetization-mode=1).

```
m=video 49170 RTP/AVP 105
a=rtpmap:105 H264/90000
a=fmtp:105 profile-level-id=640028;packetization-mode=1
```

[Answer]

An example of the setting of the "a=fmtp" line in the answer is shown. Here, the settings in the answer are H.264 High Profile Level 4.0, non-interleaved mode (packetization-mode=1), and 1080i or XGA as the video source format.

```
m=video 49170 RTP/AVP 105
a=rtpmap:105 H264/90000
a=fmtp:105
profile-level-id=640028;packetization-mode=1;sprop-parameter-sets=Z2QAKKwspAHgER9o,
Z2QAKKwspAEAGGQ=
```

B.4.2.2 Example 2

The offerer can at least send video of the resolution (1080i) that is mandatory for AVSIP-4 among those specified by H.264 High Profile Level 4.0 and receive video of any resolution. The answerer can send both the 1080i and XGA videos of H.264 High Profile Level 4.0 and receive video of any resolution. The following shows an example in which the answerer sends a 1080i or XGA stream to the offerer and the offerer sends a stream with the resolution available to the offerer as a result of the Offer/Answer Negotiation. Note that, in this example, "sprop-parameter-sets" is not described in the SDP description because both the offerer and answerer can receive the video of any resolution.

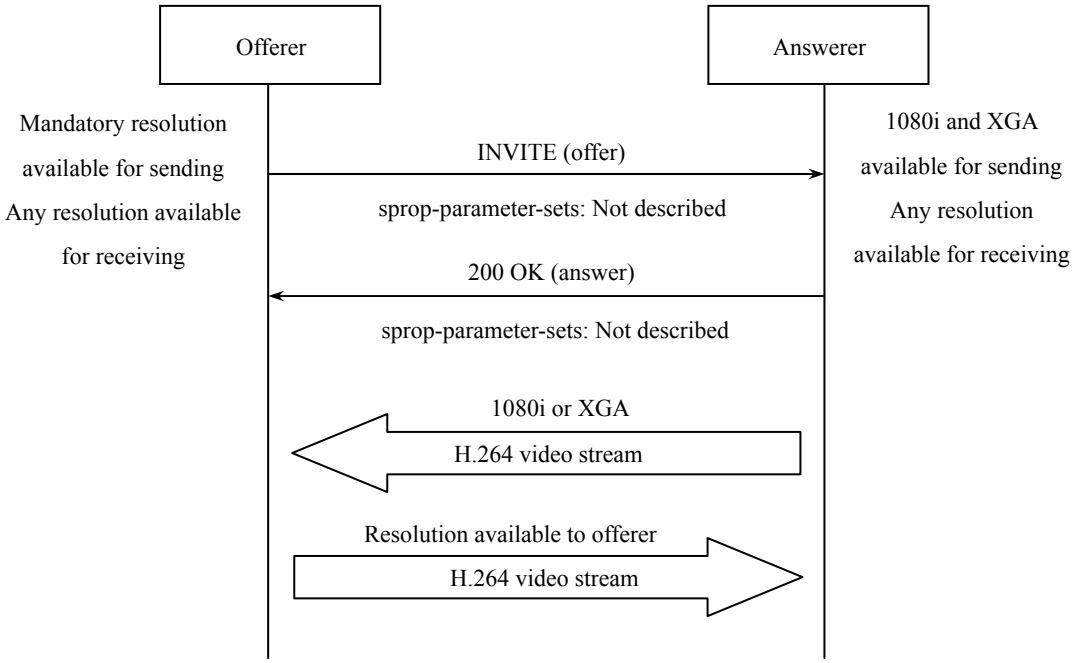


Figure B-5/JJ-40.30: When the offerer can send video of the mandatory resolution and receive video of any resolution (example 2)

[Offer]

An example of the setting of the "a=fmtp" line in the offer is shown. Here, the settings in the offer are H.264 High Profile Level 4.0 and non-interleaved mode (packetization-mode=1).

```
m=video 49170 RTP/AVP 105
a=rtpmap:105 H264/90000
a=fmtp:105 profile-level-id=640028;packetization-mode=1
```

[Answer]

An example of the setting of the "a=fmtp" line in the answer is shown. Here, the settings in the answer are H.264 High Profile Level 4.0 and non-interleaved mode (packetization-mode=1).

```
m=video 49170 RTP/AVP 105
a=rtpmap:105 H264/90000
a=fmtp:105 profile-level-id=640028;packetization-mode=1
```

B.4.3 When the Offerer and Answerer Can Send Video of the Mandatory Resolution and Receive Video of Any Resolution

The offerer can at least send video of the resolution (1080i) that is mandatory for AVSIP-4 among those specified by H.264 High Profile Level 4.0 and receive the video of any resolution. The answerer can at least send video of the resolution (1080i) that is mandatory for AVSIP-4 among those specified by H.264 High Profile Level 4.0 and receive video of any resolution. The following shows an example in which streams are sent with a resolution available to both the offerer and answerer as a result of the Offer/Answer Negotiation. Note that, in this example, "sprop-parameter-sets" is not described in the SDP description because both the offerer and answerer can receive video of any resolution.

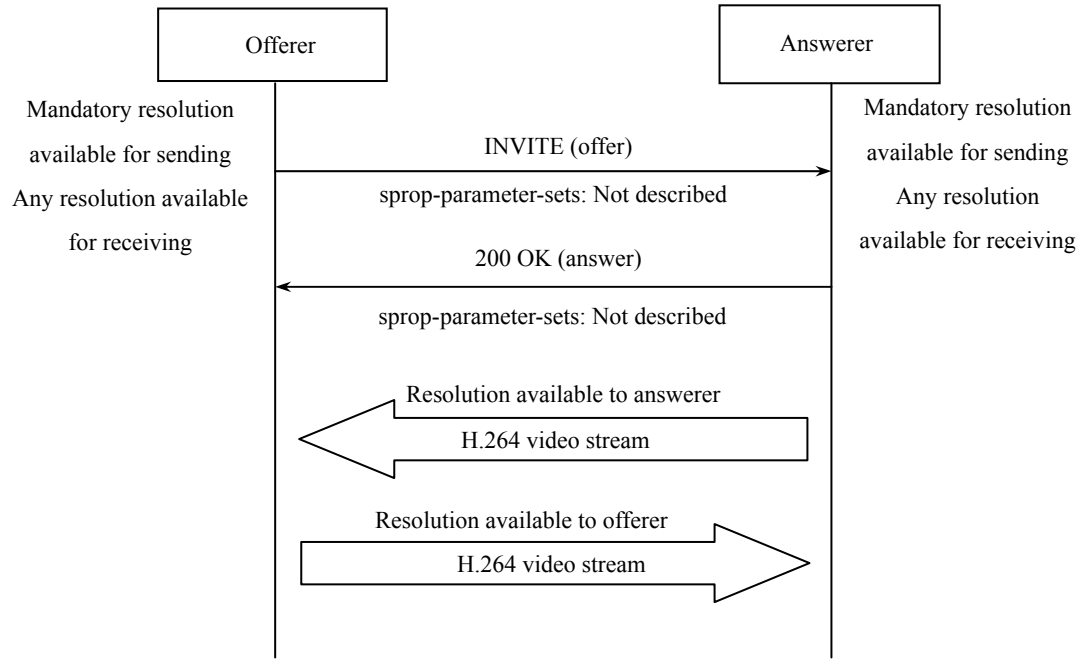


Figure B-6/JJ-40.30: When the offerer and answerer can send video of the mandatory resolution and receive video of any resolution

[Offer]

An example of the setting of the "a=fmtp" line in the offer is shown. Here, the settings in the offer are H.264 High Profile Level 4.0 and non-interleaved mode (packetization-mode=1).

```
m=video 49170 RTP/AVP 105
a=rtpmap:105 H264/90000
a=fmtp:105 profile-level-id=640028;packetization-mode=1
```

[Answer]

An example of the setting of the "a=fmtp" line in the answer is shown. Here, the settings in the answer are H.264 High Profile Level 4.0 and non-interleaved mode (packetization-mode=1).

```
m=video 49170 RTP/AVP 105
a=rtpmap:105 H264/90000
a=fmtp:105 profile-level-id=640028;packetization-mode=1
```

B.4.4 When the Offerer and Answerer Do Not Share Any Resolution Available for Sending and Receiving

The offerer can send and receive only 1080i video of H.264 High Profile Level 4.0. The answerer can send and receive only 720p video of H.264 High Profile Level 4.0. The following shows an example in which the answerer returns an error response and communication is not started because the offerer and answerer do not have common capabilities. Note that, after receiving the error response, the offerer may originate a call again with a different AVSIP profile.

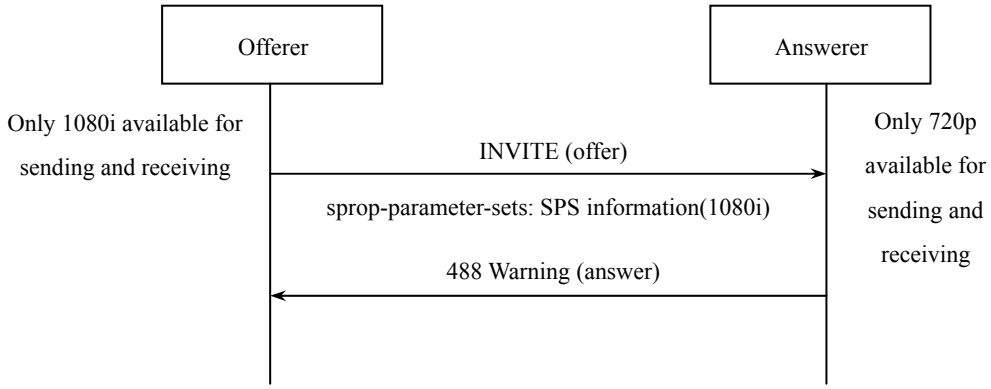


Figure B-7/JJ-40.30: When offerer and answerer do not share any resolution available for sending and receiving

[Offer]

An example of the setting of the "a=fmtp" line in the offer is shown. Here, the settings in the offer are H.264 High Profile Level 4.0, non-interleaved mode (packetization-mode=1), and 1080i as the video source format.

```
m=video 49170 RTP/AVP 105
a=rtpmap:105 H264/90000
a=fmtp:105
profile-level-id=640028;packetization-mode=1;sprop-parameter-sets=Z2QAKKwspAHgER9o
```

Appendix I　　Method for Calculating the Value To Be Specified in "b=" Line for Using Video

(This appendix does not form part of the specifications. It is for reference only.)

I.I　Introduction

This appendix describes an example of how to calculate the value that is specified in the "b=" line of an SDP description for sending video on RTP. Specifically, this appendix describes, using an example, how to calculate the value to be specified in the "b=" line from the bit rate of the video stream that is output from the encoder because the value in the "b=" line of an SDP description must include the overheads incurred by Layers 3 and 4.

I.II　Preconditions

This section describes the preconditions for the calculation.

Note that the following preconditions are intended only for the study of the calculation formula and do not constrain the implementation of this standard.

- The encoding mode is constant bit rate (CBR).
- The CSRC and extended header of RTP are not used.
- The packetization size (Note 1) is constant.
- With MPEG-4 Video codec, the video stream divided into units of packetization size is converted into a single IP packet.
- With H.264 video codec, a single NAL unit packet is used when the packetization size allows the stream to be converted into a single IP packet. When the packetization size does not allow the stream to be converted into a single IP packet, the stream is converted into multiple IP packets by using a fragmentation unit. Here, it is assumed that FU-A is used as the fragmentation unit.

(Note 1) The packetization size refers to the size of the video packet for MPEG-4 Video codec or the size of the slice for H.264 video codec. As the packetization size is smaller, the number of packets is larger. Accordingly, the header overhead resulting from the increase in the number of IP packets will increase the required bandwidth.

I.III Parameters

The table below lists the parameters related to the calculation.

Table I-1/JJ-40.30: Parameters for calculating the value to be specified in the "b=" line

Parameter	Variable name	Remarks
Bit rate of the video stream	x [bit/s]	
Frame rate of the video stream	y [fps]	
Packetization size	z [bit]	
IP header length	a [bit]	IPv4 (without options): 20 bytes (160 bits) IPv6 (without options): 40 bytes (320 bits)
UDP header length	b [bit]	8 bytes (64 bits)
RTP header length	c [bit]	Header without CSRC and extended header: 12 bytes (96 bits)
RTP payload header length	d [bit]	MPEG-4 Video or H.264 (single NAL unit packet): 0 bits H.264 (FU-A): 2 bytes (16 bits)
RTP payload length	e [bit]	Length of the payload data that can be set per RTP packet. The RTP payload length must be of a value that does not cause the total data length after RTP packetization to exceed the MTU size. For MPEG-4 Video or H.264 (single NAL unit packet), the RTP payload length is equal to the packetization size.
Average No. of bits per frame	x / y [bit]	
Average No. of packets per frame	((x / y) / z) * (z / e) [packet]	The digits after the decimal point are rounded out.

I.IV Calculation Formula

Bit rate to be specified in the "b=" line = Bit rate of the video stream

+ ((IP header length + UDP header length + RTP header length + RTP payload header length)

* number of packets/second)

= Bit rate of the video stream

+ ((IP header length + UDP header length + RTP header length + RTP payload header length)

* (average No. of packets per frame * frame rate of the video stream))

= x + (a + b + c +d) * ((x / y) / z * (z / e)) * y

= x + (a + b + c +d) * x/e

The digits after the decimal point in the calculation result are rounded out.

I.V Note

- The calculation formula described above is designed without consideration for the burst characteristics of the video (for example, the I frame includes a large number of coded bits). Therefore, the result of calculation using the above calculation formula requires the addition of, or multiplication by, a certain constant to adjust to the characteristics of the encoder.

I.VI Reference Values

The table below lists the values to be referenced, when the above calculation formula is used, as the values that are to be specified in the "b=" line when using individual system profiles.

Here, the transfer protocol is assumed to be IPv4.

Table I-2/JJ-40.30: Relationships between system profiles and "b=" line values

System profile	Parameter	Value in "b=" line
AVSIP-1 (Mobile)	MPEG-4 Video Simple Profile Level 0 QCIF / 15fps / 43kbit/s Packetization size = 180 bytes	48
AVSIP-1.5 (MPEG-4 QVGA)	MPEG-4 Video Simple Profile Level 3 QVGA / 15fps / 365kbit/s Packetization size = 1200 bytes	384
AVSIP-2a (MPEG-4 SD)	MPEG-4 Video Simple Profile Level 4a VGA / 30fps / 1932kbit/s Packetization size = 1200 bytes	2000
AVSIP-2b (H.264 SD)	H.264 Baseline Profile 3.0 4SIF / 30fps / 1932kbit/s Packetization size = 1200 bytes Not using fragmented packets	2000
AVSIP-3 (H.264 720p)	H.264 Baseline Profile 3.1 720p / 30fps / 4834kbit/s Packetization size = 1200 bytes Not using fragmented packets	5000
AVSIP-4 (H.264 1080i)	H.264 High Profile 4.0 1080i / 30fps(60field/s) / 9657kbit/s Packetization size = 1200 bytes Using fragmented packets	10000

Appendix II Examples of Using 3D Image Communication

(This appendix does not form part of specifications. It is for reference only.)

II.I Introduction

This appendix gives some examples of using the communication system for sending 3D images and describes the points to be noted in operation.

II.II 3D Image Format

As described in Section 6.2, the 3D image format to be applied is a side-by-side format. Details of the format are described in the frame packing arrangement SEI specifications of H.264 in Section 6.2.1, "Required Characteristics." Figure II-1 shows its frame structure, which contains the right-eye image in the right half and the left-eye image in the left half.

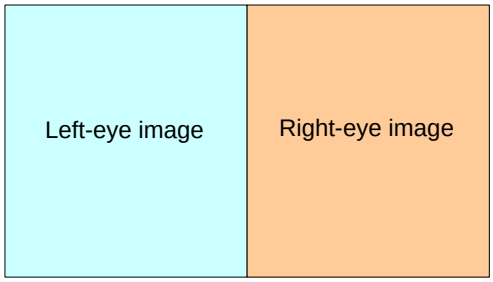


Figure II-1/JJ-40.30: Frame structure of a side-by-side image

II.III Display of 3D Images on a Display Unit

II.III.1 Automatic 3D/2D Image Display by a Display-Unintegrated 3D Image Receiving Terminal on a Display Unit Supporting 3D Images

When a display-unintegrated 3D image receiving terminal outputs a 3D image to a display unit (even when the display unit supports 3D images), the display unit will determine that the image is a 2D image unless the terminal notifies the display unit that the image is a side-by-side 3D image, and thus displays the image frames as shown in Figure II-1. When the 3D image receiving terminal connects to the display unit in compliance with the HDMI standard that supports 3D format, the terminal can notify the display unit that the image is a side-by-side 3D image, and, therefore, enables automatic switching between 3D image display and 2D image display.

Even if the display unit does not support the 3D-compatible HDMI standard, the receiving user can operate the display unit to switch between 2D image display and side-by-side 3D image display.

II.III.2 Display of 3D Images as 2D Images on a Display Unit Supporting 3D Images

Display units supporting 3D image display are classified into two types: glasses-type and naked-eye type. The naked-eye 3D display units can always display a 3D image. The glasses-type 3D display unit may sometimes be required to display a 3D image as a 2D image on the display unit, for example, when the user does not wear the 3D glasses.

To satisfy this requirement, the 3D image receiving terminal outputs, to the display unit, the left-eye image of a side-by-side image as a 2D image that is enlarged to the full size of the frame, as shown in Figure II-2 as operated by the receiving user.

On the other hand, in the case of 3D display by a display-unintegrated 3D image receiving terminal, the terminal outputs a side-by-side 3D image, as is, to the 3D display unit. Then, the display unit can enlarge the left-eye image of

the side-by-side image to the full size of the frame as shown in Figure II-2 and display it as a 2D image according to the user's operation on the display unit (not on the receiving terminal).

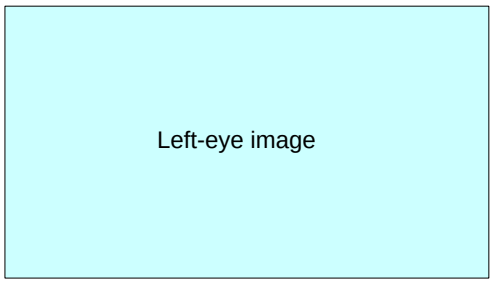


Figure II-2/JJ-40.30: Typical frame structure in the case of the 2D display of a side-by-side image

II.III.3 3D Image Display by a Display-Unintegrated 3D Image Receiving Terminal on a Display Unit That Does Not Support 3D Images

When a display-unintegrated 3D image receiving terminal connects to a display unit that does not support 3D images, the 3D image receiving terminal must have a function to output, to the display unit, the left-eye image of a side-by-side image as a 2D image, enlarged to the full size of the frame. Note that, if the 3D image receiving terminal does not have the said function, the output image will be displayed as shown in Figure II-1 on the display unit.

II.IV Examples of Sending and Receiving a 3D Image

When a 3D image is to be sent or received, the sending terminal and/or receiving terminal might not support 3D images. Also, SIP does not provide a function for negotiating whether each terminal supports 3D images. This section provides some notes on sending and receiving operations with some use cases.

II.IV.1 When Both the Sending and Receiving Terminals Support 3D Images

When a 2D image is sent, the 3D image receiving terminal determines that it is receiving a 2D image because the H.264 video stream for the 2D image excludes the frame packing arrangement SEI. Therefore, the 3D image receiving terminal can automatically receive and display the 2D image correctly.

When a 3D image is sent, the receiving terminal that does not support 3D images can recognize that the sent image is a side-by-side image shown in Figure II-1 by interpreting the frame packing arrangement SEI in the H.264 video stream. Therefore, the 3D image receiving terminal can automatically receive and display the 3D image correctly. Note that, when a display-unintegrated 3D image receiving terminal connects to a display unit in compliance with the HDMI standard that supports 3D format, the terminal can automatically determine whether the connected display unit supports 3D images, and display the sent image appropriately using the method described in Section II.III.1 or II.III.2.

II.IV.2 When the Sending Terminal Supports 3D Images But the Receiving Terminal Does Not Support 3D Images

When a 2D image is sent, the receiving terminal can receive and display the 2D image correctly.

On the other hand, when a 3D image is sent, the receiving terminal not supporting 3D images cannot interpret the frame packing arrangement SEI in the H.264 video stream, and, therefore, displays the side-by-side frame image shown in Figure II-1, as is. In such a case, the user of the receiving terminal should give the user of the sending terminal a verbal instruction, via the voice channel or by some other method, to switch the sending image from 3D to 2D. Then, the receiving terminal not supporting 3D images will be able to display the sent image as a 2D image correctly.

II.IV.3 When the Sending Terminal Does Not Support 3D Images But the Receiving Terminal Does Support 3D Images

The sending terminal always sends a 2D image. The 3D image receiving terminal determines that it is receiving a 2D image because the H.264 video stream for the 2D image excludes the frame packing arrangement SEI. Therefore, a 3D image receiving terminal can automatically receive and display a 2D image correctly.

II.IV.4 When Neither the Sending Nor Receiving Terminals Support 3D Images

The sending terminal always sends a 2D image. Therefore, the receiving terminal that does not support 3D images can automatically receive and display 2D images correctly.