

TTC標準
Standard

JJ-40.30

SIP 網における
オーディオビジュアル通信システム

Audiovisual communication system
on SIP network

第 2.0 版

2013 年 5 月 23 日制定

一般社団法人
情報通信技術委員会

THE TELECOMMUNICATION TECHNOLOGY COMMITTEE



本書は、一般社団法人情報通信技術委員会が著作権を保有しています。
内容の一部又は全部を一般社団法人情報通信技術委員会の許諾を得ることなく複製、転載、
改変、転用及びネットワーク上での送信、配布を行うことを禁止します。

目次

<参考>	5
1 はじめに	6
2 検討対象	6
3 参考文献	6
4 システムモデル	9
5 システムプロファイル	11
6 メディアの所要特性と符号化	12
6.1 音声メディア	12
6.1.1 所要特性	12
6.1.2 符号化	12
6.1.3 システムプロファイルの音声メディア仕様	15
6.2 映像メディア	16
6.2.1 所要特性	16
6.2.2 符号化	17
6.2.3 システムプロファイルの映像メディア仕様	26
7 エンド・エンド制御	27
7.1 能力交換	27
7.1.1 オーディオビジュアル通信における SDP 表現	27
7.1.2 フォールバックを軽減するためのコーデックの併記	27
7.2 通信モード設定	30
7.3 C&I	30
7.3.1 遠隔カメラ制御	30
7.3.2 その他の制御	30
7.4 RTP/RTCP による制御	31
8 マルチメディア多重/分離	32
8.1 RTP 変換	32
8.1.1 H.263 ビデオ	32
8.1.2 H.264 ビデオ	32
8.1.3 MPEG-4 ビデオ	33
8.1.4 G.711 オーディオ、G.722 オーディオ	34
8.1.5 G.711.1 オーディオ	34
8.1.6 MPEG-4 オーディオ	34
8.1.7 帯域のネゴシエーション結果に基づくメディアパケット送出	34
付属資料 A NGN を利用する際の留意事項	35
A.1 NGN を利用する際の関連仕様	35
A.2 誤課金防止のためのフォールバックを利用した能力交換方法	35
付属資料 B SDP の記述及びネゴシエーションの詳細	36
B.1 概要	36
B.2 SDP の記述内容とネゴシエーション手順	36
B.2.1 映像コーデック	37
B.2.2 音声コーデック	46
付録 I 映像利用時の b=行に指定する値の算出方法	51

I.I	はじめに.....	51
I.II	前提条件.....	51
I.III	パラメータ.....	52
I.IV	計算式.....	52
I.V	留意点.....	53
I.VI	参考値.....	53
付録 II	3D 映像通信の使用例.....	54
II.I	はじめに.....	54
II.II	3D 映像フォーマット.....	54
II.III	3D 映像のディスプレイ表示.....	54
II.III.1	一体型でない 3D 映像受信端末の 3D 映像対応ディスプレイでの 3D 映像/2D 映像自動表示.....	54
II.III.2	3D 映像対応ディスプレイでの 3D 映像の 2D 表示.....	54
II.III.3	一体型でない 3D 映像受信端末の 3D 映像非対応ディスプレイへの 3D 映像表示.....	55
II.IV	3D 映像の送受信例.....	55
II.IV.1	送信端末と受信端末が 3D 映像対応の場合.....	55
II.IV.2	送信端末が 3D 映像対応で受信端末が 3D 映像非対応の場合.....	55
II.IV.3	送信端末が 3D 非対応で受信端末が 3D 映像対応の場合.....	55
II.IV.4	送信端末と受信端末が 3D 映像非対応の場合.....	56

<参考>

1. 国際勧告等の関連

本標準に関する国際勧告はない。

2. 改版の履歴

版数	制定日	改版内容
第 1.0 版	2012 年 5 月 24 日	初版
第 1.1 版	2012 年 10 月 26 日	誤記訂正
第 2.0 版	2013 年 5 月 23 日	AVSIP 既定の見直し、非 NGN 環境の考慮

3. 工業所有権

本標準に関わる「工業所有権の実施の権利に係る確認書」の提出状況は、TTC ホームページで公開されている。

4. 標準作成部門

メディア符号化専門委員会

1 はじめに

従来 TV 会議システムをはじめとするオーディオビジュアル通信は閉域網で H.323 を用いて行われてきたが、NGN などのキャリア網が整備され、SIP を用いた通信が利用されるようになってきている。また、社内網やインターネット網など NGN 以外でも SIP を用いた通信も行われ始めている。

そこで、本標準は、SIP 網におけるリアルタイム会話型オーディオビジュアル通信システムの相互接続に資するため作成した。このような通信システムの代表的な利用例としては TV 電話と TV 会議があるが、両者の歴史的経緯によってサポートすべき機能が異なっており、本標準ではその両者を明確に分け AVSIP と呼ばれるプロファイルによって各用途に適した通信機能とその能力交換方法を規定している。

さらに、本標準の末尾に NGN を利用した際の留意事項を参考資料として記載している。

2 検討対象

本標準では、下記特質の基本的なシステムを検討対象とした。

- ・ 双方向対称型通信
- ・ ポイント・ポイント通信
- ・ 音声通信と映像通信の機能を同時に使うマルチメディア通信

これらを超える通信機能、例えば MCU を用いたマルチポイント通信機能の検討は次のステップとする。

主な検討項目は、メディア信号（音声、映像、データ）に関するパラメータの推奨値ならびに SDP による能力交換の推奨方法である。また、本標準は RFC3261[RFC3261]に規定された SIP 及び、RFC4566[RFC4566]に規定された一般的な SIP/SDP 網に接続されるオーディオビジュアル通信システムの相互接続のための仕様を規定することを目的とする。

3 参考文献

本標準で参照する文献を以下に示す。

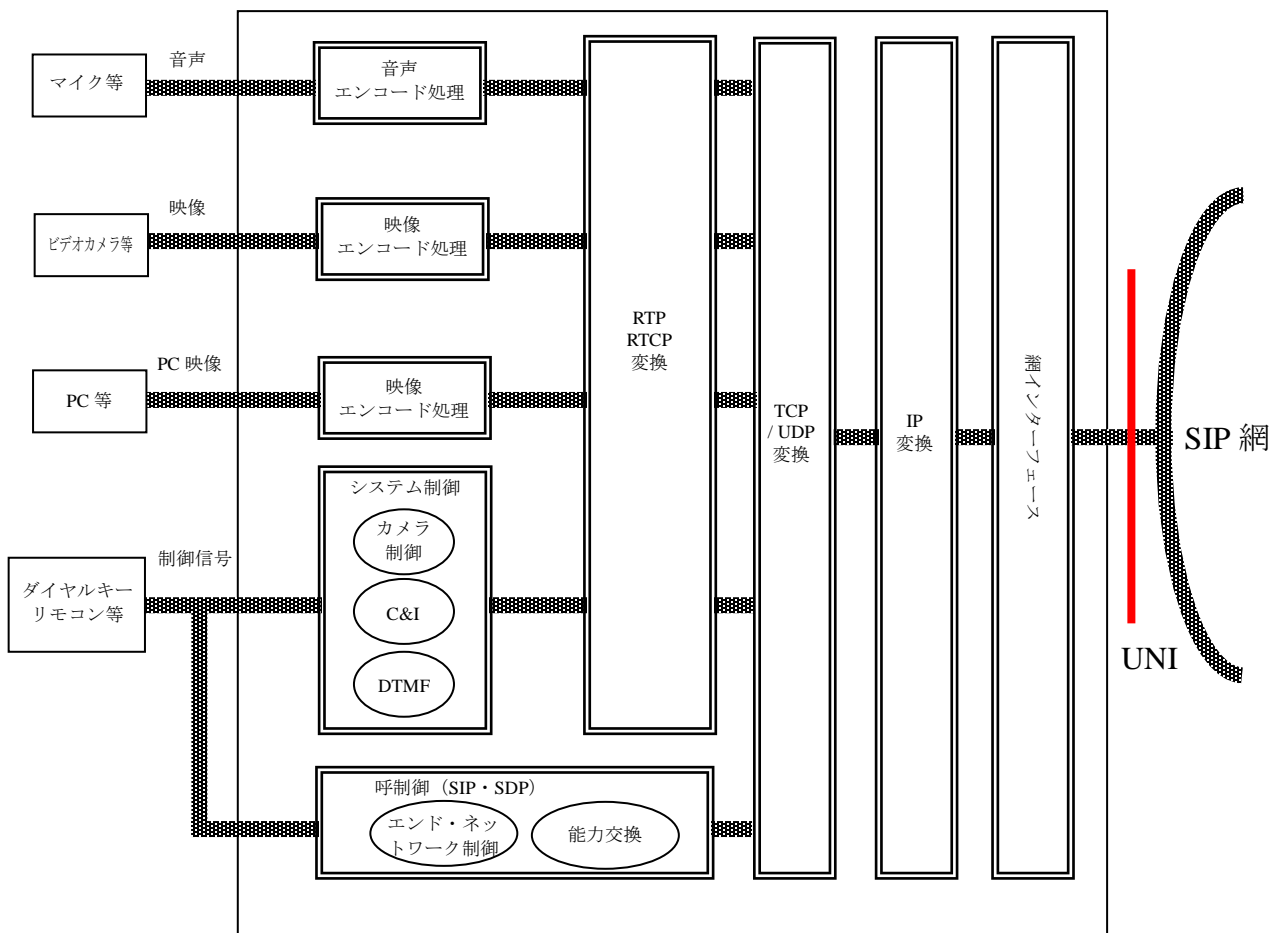
- [G711] “音声周波数帯域信号の PCM 符号化方式”, TTC 標準 JT-G711 第 4.0 版, 情報通信技術委員会 (The Telecommunication Technologies Committee), 2001 年 4 月
- [G711.1] “G.711.1 パルス符号変調に対する広帯域エンベデッド拡張”, TTC 標準 JT-G711.1 第 1.0 版, 情報通信技術委員会 (The Telecommunication Technologies Committee), 2009 年 2 月
- [G722] “64kbit/s 以下の 7kHz オーディオ符号化方式”, TTC 標準 JT-G722 第 3.0 版, 情報通信技術委員会 (The Telecommunication Technologies Committee), 2008 年 5 月
- [H224] “A real time control protocol for simplex applications using the H.221 LSD/HSD/MLP channels”, ITU-T 勧告 H.224, ITU-T, 2005 年 1 月
- [H263] “低ビットレート通信用ビデオ符号化方式”, TTC 標準 JT-H263 第 3.3 版, 情報通信技術委員会 (The Telecommunication Technologies Committee), 2009 年 11 月
- [H264] “オーディオビジュアルサービス全般のための高度ビデオ符号化方式”, TTC 標準 JT-H264 第 7 版, 情報通信技術委員会 (The Telecommunication Technologies Committee), 2012 年 11 月
- [H281] “A far end camera control protocol for videoconferences using H.224”, ITU-T 勧告 H.281, 1994 年 11 月

- [H323] "パケットに基づくマルチメディア通信システム", TTC 標準 JT-H323 第 6.0 版, 情報通信技術委員会 (The Telecommunication Technologies Committee), 2008 年 5 月
- [MPEG-4 Audio] "Information technology- Coding of audio-visual objects- Part 3: Audio", ISO/IEC 14496-3 第 3 版, ISO, 2005 年 12 月
- [MPEG-4 Video] "Information technology- Coding of audio-visual objects- Part 2: Visual", ISO/IEC 14496-2 第 3 版, ISO, 2004 年 6 月
- [Q3402] "NGN UNI シグナリングプロファイル プロトコルセット 1 (NGN UNI Signalling Profile (Protocol Set 1))", TTC 標準 JT-Q3402 第 1.0 版, 情報通信技術委員会 (The Telecommunication Technologies Committee), 2009 年 5 月
- [BT601] "Studio encoding parameters of digital television for standard 4:3 and wide screen 16:9 aspect ratios", ITU-R 勧告 BT.601-6, ITU-R, 2007 年 1 月
- [BT709] "Parameter values for the HDTV standards for production and international programme exchange", ITU-R 勧告 BT.709-5, ITU-R, 2002 年 4 月
- [RFC3016] "MPEG-4 Audio/Visual ストリームの RTP ペイロード形式 (RTP Payload Format for MPEG-4 Audio/Visual Streams)", TTC 標準 JF-IETF-RFC3016, 情報通信技術委員会 (The Telecommunication Technologies Committee), 2009 年 5 月
- [RFC3261] "SIP: セッション開始プロトコル", TTC 標準 JF-IETF-RFC3261, 情報通信技術委員会 (The Telecommunication Technologies Committee), 2005 年 6 月
- [RFC3264] "セッション記述プロトコル (SDP) を使ったオファー/アンサーモデル", TTC 標準 JF-IETF-RFC3264, 情報通信技術委員会 (The Telecommunication Technologies Committee), 2005 年 6 月
- [RFC3551] "最小限の制御による音声とビデオ会議のための RTP プロファイル (RTP Profile for Audio and Video Conferences with Minimal Control)", TTC 標準 JF-IETF-STD65 第 1 版, 情報通信技術委員会 (The Telecommunication Technologies Committee), 2005 年 6 月
- [RFC4566] "SDP: セッション記述プロトコル", TTC 標準 JF-IETF-RFC4566, 情報通信技術委員会 (The Telecommunication Technologies Committee), 2007 年 3 月
- [RFC4573] "MIME Type Registration for RTP Payload Format for H.224", RFC 4573, IETF, 2006 年 7 月
- [RFC4582] "The Binary Floor Control Protocol (BFCP)", RFC 4582, IETF, 2006 年 11 月
- [RFC4585] "RTCP をベースとしたフィードバックのための拡張 RTP プロファイル (RTP/AVPF) (Extended RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/AVPF))", TTC 標準 JF-IETF-RFC4585 第 1.0 版, 情報通信技術委員会 (The Telecommunication Technologies Committee), 2008 年 3 月
- [RFC4629] "ITU-T 勧告 H.263 の RTP ペイロード形式", TTC 標準 JF-IETF-RFC4629 第 1.0 版, 情報通信技術委員会 (The Telecommunication Technologies Committee), 2010 年 5 月
- [RFC4733] "DTMF デジタル、電話トーン、電話信号のための RTP ペイロード", TTC 標準 JF-IETF-RFC4733, 情報通信技術委員会 (The Telecommunication Technologies Committee), 2009 年 5 月
- [RFC5104] "フィードバックを伴う RTP AV プロファイル (AVPF) のコーデック制御メッセージ (Codec Control Messages in the RTP Audio-Visual Profile with Feedback (AVPF))", TTC 標準 JF-IETF-RFC5104 第 1.0 版, 情報通信技術委員会 (The Telecommunication Technologies Committee), 2008 年 5 月
- [RFC5391] "G.711.1 の RTP ペイロード形式 (RTP Payload Format for ITU-T Recommendation G.711.1)

- ”, TTC 標準 JF-IETF-RFC5391 第 1.0 版, 情報通信技術委員会 (The Telecommunication Technologies Committee) , 2009 年 5 月
- [RFC6184] ” RTP Payload Format for H.264 Video” , RFC6184, IETF, 2011 年 5 月
- [TR-1014] ”NGN アーキテクチャの概要(General overview of NGN architecture)”, TTC 技術レポート TR-1014 第 1.0 版, 情報通信技術委員会 (The Telecommunication Technologies Committee), 2006 年 6 月
- [TR-1020] ”NGN における SDP メディア能力交換に関するインタフェース技術レポート”, TTC 技術レポート TR-1020 第 1.0 版, 情報通信技術委員会 (The Telecommunication Technologies Committee), 2009 年 5 月

4 システムモデル

オーディオビジュアル通信システムにおける送信処理について、データの流れから見た典型的な送信端末モデルを図 4-1 に示す。システムプロファイルによっては、各部の能力や適用/非適用が変わってくる。



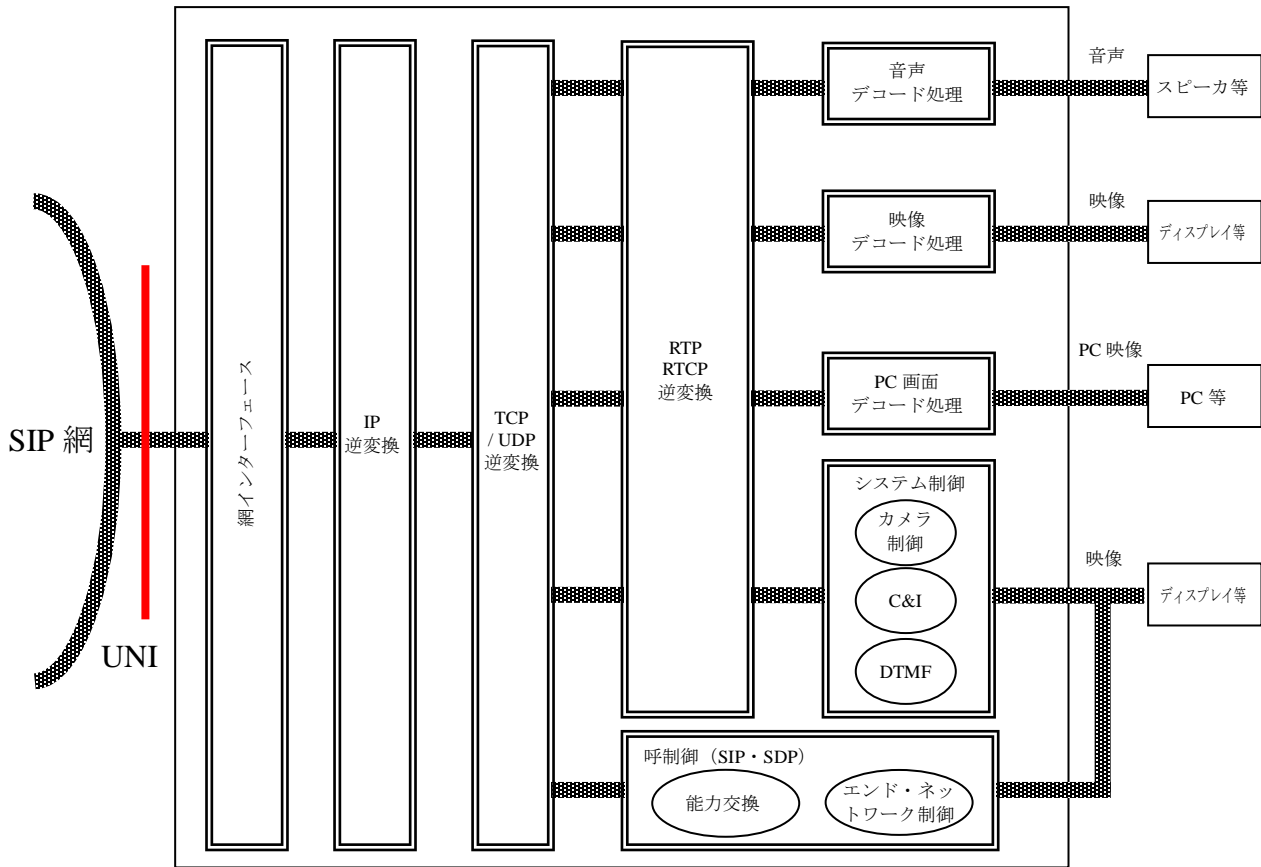
C&I : Control and Indication、制御および表示信号

UNI : User-to-Network Interface、ユーザ・網インタフェース

注 : RTP/RTCP 以外の C&I 転送は更に検討を要する。

図 4-1/JJ-40.30 データの流れから見た送信端末モデル

オーディオビジュアル通信システムにおける受信処理について、データの流れから見た典型的な受信端末モデルを図 4-2 に示す。システムプロファイルによっては、各部の能力や適用／非適用が変わってくる。



C&I : Control and Indication、制御および表示信号

UNI : User-to-Network Interface、ユーザ・網インタフェース

注 : RTP/RTCP 以外の C&I 転送は更に検討を要する。

図 4-2/JJ-40.30 データの流れから見た受信端末モデル

5 システムプロファイル

端末は、適用アプリケーションによってサポートする能力が大きく異なる。そこで異なる端末間の相互接続に資するため、アプリケーションに応じた実装すべき最低限の能力を表 5-1 に示すシステムプロファイルという区分で定義することにする。本標準の推奨仕様をサポートする端末は、少なくとも 1 つ以上のシステムプロファイルを実装しなければならない。また、AVSIP-4b-i を除き、サポートしたシステムプロファイル以下のレベルサポートをすること。

表 5-1/JJ-40.30 本標準で定義するシステムプロファイル

システムプロファイル	主な用途	特徴	概要
AVSIP-1a	TV 電話	MPEG-4 QCIF 映像	低ビットレートでテレビ電話ができる能力を定義する。 音声と映像の双方向通信が可能である。モバイル用途を想定する。
AVSIP-1.5a	TV 電話	MPEG-4 QVGA 映像	QVGA 品質相当でテレビ電話ができる能力を定義する。 音声と映像の双方向通信が可能である。
AVSIP-2a	TV 電話	MPEG-4 VGA 映像	SD 品質相当でテレビ電話ができる能力を定義する。 音声と映像の双方向通信が可能である。。
AVSIP-2b	TV 会議	H.264 SD 映像	SD 品質でテレビ会議ができる能力を定義する。 音声と映像の双方向通信の他に、遠隔カメラ制御や PC 画面伝送が可能である。
AVSIP-3a	TV 電話	H.264 720p 映像	HD 品質でテレビ電話ができる能力を定義する。 音声と映像の双方向通信が可能である。
AVSIP-3b	TV 会議	H.264 720p 映像	HD 品質でテレビ会議ができる能力を定義する。 音声と映像の双方向通信の他に、遠隔カメラ制御や PC 画面伝送が可能である。
AVSIP-4b-i	TV 会議	H.264 1080i 映像	フル HD 品質(1080i)でテレビ会議ができる能力を定義する。 音声と映像の双方向通信の他に、遠隔カメラ制御や PC 画面伝送が可能である。 なお、このプロファイルでは、他のプロファイルとの下位互換は行わなくてもよい。
AVSIP-4b-p	TV 会議	H.264 1080p 映像	フル HD 品質(1080p)でテレビ会議ができる能力を定義する。 音声と映像の双方向通信の他に、遠隔カメラ制御や PC 画面伝送が可能である。

(注) 本標準では、映像ソースフォーマットの解像度が 704x480 である場合を SD 品質と、1280x720 である場合を HD 品質と、1920x1080 である場合をフル HD 品質と表記する。

6 メディアの所要特性と符号化

各端末は、通信時に必要なメディアデータの処理能力を実装し、指定された符号化方式でエンコード・デコードを行う。

6.1 音声メディア

6.1.1 所要特性

本標準で定義する音声帯域の処理能力は表 6-1 に示す通りである。これ以外の能力は、オプションとして実装してもよい。

表 6-1/JJ-40.30 本標準で定義する音声帯域

音声帯域	適用用途例
3.4 kHz	標準品質のテレビ電話
7 kHz	SD 品質のテレビ電話、テレビ会議
21 kHz	HD 品質/フル HD 品質のテレビ会議

また、本標準で定義する音声チャンネルの処理能力は表 6-2 に示す通りである。これ以外の能力はオプションとして実装してもよい。

表 6-2/JJ-40.30 本標準で定義する音声チャンネル

音声チャンネル	適用用途例
1 チャンネル	モノラル音声
2 チャンネル	ステレオ音声

6.1.2 符号化

本標準で定義する音声符号化は表 6-3 に示す通りである。これ以外の能力はオプションとして実装してもよい。

表 6-3/JJ-40.30 本標準で定義する音声符号化

音声符号化	使用標準	帯域	ビットレート	適用用途例
G.711 μ -law	JT-G.711 [G.711]	3.4 kHz	64 kbit/s	標準品質のテレビ電話
G.722	JT-G.722 [G.722]	7 kHz	64 kbit/s	テレビ会議
G.711.1	JT-G.711.1 [G.711.1]	7 kHz	96 kbit/s	テレビ電話
MPEG-4 AAC-LC	ISO/IEC 14496-3 [MPEG-4 Audio]	21 kHz	96 kbit/s, 192 kbit/s	HD 品質/フル HD 品質のテレビ会議
MPEG-4 AAC-LD	ISO/IEC 14496-3 [MPEG-4 Audio]	21 kHz	128 kbit/s, 192 kbit/s	HD 品質/フル HD 品質のテレビ会議

本標準で定義する G711.1 は、表 6-4 に示されるエンコード・デコード能力をサポートする。

表 6-4/JJ-40.30 本標準で定義する G711.1

項目	内容	SDP 上の記述
サンプリングレート	16 kHz	a=rtpmap 行の clock rate に 16000 を指定
パケット化周期	20 ms	a=ptime:20 を指定
モード	4	a=fmtp 行に mode-set=4 を指定

本標準で定義する MPEG-4 AAC-LC は、表 6-5 に示されるエンコード・デコード能力をサポートする。

表 6-5/JJ-40.30 本標準で定義する MPEG-4 AAC-LC

項目	内容	SDP 上の記述
サンプリングレート	48 kHz	a=fmtp 行の config 内 samplingFrequencyIndex に 3(48kHz)を指定
RTP タイムスタンプレート	90 kHz	a=rtpmap 行の clock rate に 90000 を指定
パケット化周期	21.33 ms	a=ptime:20 を指定 (注 1)
プロファイル	AAC Profile	
レベル	Level2 (0x29)	a=fmtp 行に profile-level-id=41 を指定
コーデック種別	AAC-LC	a=fmtp 行に object=2 を指定
チャンネル数	モノラル, ステレオ	a=fmtp 行の config 内 channelConfiguration で指定 (1: モノラル, 2: ステレオ)
ビットレート (コーデック帯域)	96 kbit/s(モノラル), 192 kbit/s(ステレオ)	bitrate=96000 (モノラル), bitrate=192000 (ステレオ)を指定
ビットレート (レイヤ 3 帯域)	192 kbit/s(モノラル), 384 kbit/s(ステレオ) (注 2)	b=AS:192 (モノラル), b=AS:384 (ステレオ) (注 2)

(注 1) SDP 上は a=ptime:20 としてネゴシエーションするが、MPEG-4 AAC-LC (21kHz) で通信を行う場合は、21.33ms のパケット化周期で動作する。

(注 2) 参考値：レイヤ 3 帯域は、バースト性とシェイパのチューニングを考慮した値とすること。

本標準で定義する MPEG-4 AAC-LD は、表 6-6 に示されるエンコード・デコード能力をサポートする。

表 6-6/JJ-40.30 本標準で定義する MPEG-4 AAC-LD

項目	内容	SDP 上の記述
サンプリングレート	48 kHz	a=fmtp 行の config 内 samplingFrequencyIndex に 3(48kHz)を指定
RTP タイムスタンプレート	90 kHz	a=rtpmap 行の clock rate に 90000 を指定
パケット化周期	10.67 ms	a=ptime:20 を指定 (注 1)
プロファイル	Low Delay Audio Profile	
レベル	Level4 (0x19)	a=fmtp 行に profile-level-id=25 を指定
コーデック種別	AAC-LD	a=fmtp 行に object=23 を指定
チャンネル数	ステレオ	a=fmtp 行の config 内 channelConfiguration で指定 (2:ステレオ)
ビットレート (コーデック帯域)	128 kbit/s(64kbit/s×2ch ステレオ), 192 kbit/s(96kbit/s×2ch ステレオ)	bitrate=128000 (128kbit/s), bitrate=192000 (192kbit/s)を指定
ビットレート (レイヤ 3 帯域)	256 kbit/s, 384 kbit/s (注 2)	b=AS:256, b=AS:384 (注 2)

(注 1) SDP 上は a=ptime:20 としてネゴシエーションするが、MPEG-4 AAC-LD (21kHz) で通信を行う場合は、10.67ms のパケット化周期で動作する。

(注 2) 参考値：レイヤ 3 帯域は、バースト性とシェイパーのチューニングを考慮した値とすること。

6.1.3 システムプロファイルの音声メディア仕様

各端末の所要音声特性と実装すべき符号化は、表 6-7 に従わなければならない。

表 6-7/JJ-40.30 各端末が実装する音声メディアの仕様

項目		音声メディアの仕様						
		プロファイル AVSIP-1a	プロファイル AVSIP-1.5a	プロファイル AVSIP-2a	プロファイル AVSIP-3a	プロファイル AVSIP-2b	プロファイル AVSIP-3b	プロファイル AVSIP-4b-i AVSIP-4b-p
音声帯域	3.4 kHz	必須	必須	必須	必須	オプション	オプション	オプション
	7 kHz	オプション	オプション	推奨	推奨	必須	必須	必須
	21 kHz	オプション	オプション	オプション	オプション	オプション	オプション	オプション
チャンネル数	1 (モノラル)	必須	必須	必須	必須	必須	必須	必須
	2 (ステレオ)	オプション	オプション	オプション	オプション	オプション	オプション	オプション
符号化	G711 μ -law	必須	必須	必須	必須	推奨	推奨	推奨
	G722	オプション	オプション	推奨(注)	推奨(注)	必須	必須	必須
	G711.1	オプション	オプション	推奨(注)	推奨(注)	オプション	オプション	オプション
	MPEG-4 AAC-LC	オプション	オプション	オプション	オプション	オプション	オプション	オプション
	MPEG-4 AAC-LD	オプション	オプション	オプション	オプション	オプション	オプション	オプション

(注) 本標準は、G722 と G711.1 のいずれか、あるいは両方の実装を推奨する。

6.2 映像メディア

本標準では、カメラ映像と PC 画面映像を映像メディアとして規定する。カメラ映像では 2D 映像に加え、オプション機能として Side-by-Side の 3D 映像も規定する。

6.2.1 所要特性

本標準で定義する映像フォーマット（エンコーダ入力点に着目）の処理能力は表 6-8 に示す通りである。これ以外の能力はオプションとして実装してもよい。

表 6-8/JJ-40.30 本標準で定義する映像ソースフォーマット

映像ソース	映像ソースフォーマット				適用用途例
	解像度と走査形式	フレームレート (秒)	スキャン形式	画面アスペクト比	
カメラ映像(注 4)	QCIF : 176x144, p	30(注 2)	プログレッシブ	4:3	モバイル品質のテレビ電話
	QVGA : 320x240, p	30(注 2)	プログレッシブ	4:3	QVGA 品質のテレビ電話
	VGA : 640x480, p	30(注 2)	プログレッシブ	4:3	SD 品質相当のテレビ電話
	SD : 704x480(注 1), p	30(注 2)	プログレッシブ	4:3	SD 品質のテレビ会議
	720p : 1280x720, p	60(注 3)	プログレッシブ	16:9	HD 品質のテレビ電話/会議
	1080i : 1920x1080, i	30(注 2)	インターレース	16:9	フル HD 品質のテレビ会議
	1080p : 1920x1080, p	30(注 2)	プログレッシブ	16:9	フル HD 品質のテレビ会議
PC 画面映像	VGA : 640x480, p	≥60	プログレッシブ	4:3	PC 画面
	SVGA : 800x600, p	≥60	プログレッシブ	4:3	PC 画面
	XGA : 1024x768, p	≥60	プログレッシブ	4:3	PC 画面

(注 1) 映像ソースフォーマットは 1 ラインあたり 720 画素で、符号化はそのうちの 704 画素を対象とする。このフォーマットは 4SIF (SIF: Source Input Format) と呼ばれる。

(注 2) フレームレート 30 fps は、テレビジョン標準に従う場合は 30000/1001=29.97 fps を意味する。

(注 3) フレームレート 60 fps は、テレビジョン標準に従う場合は 60000/1001=59.94 fps を意味する。

(注 4) 3D 映像では、Side-by-Side で合成されたフレーム映像を映像ソースフォーマットとする。

カラーマトリクスは 720p・1080i・1080p 映像では ITU-R BT.709 [BT709]を使用する。SD 映像では ITU-R BT.709[BT709]もしくは ITU-R BT.601 [BT601]を使用する。

6.2.2 符号化

本標準で定義する映像符号化は表 6-9 に示す通りである。

表 6-9/JJ-40.30 本標準で推奨する映像符号化

映像符号化	使用標準	用途
MPEG-4 ビデオ	ISO/IEC 14496-2 [MPEG-4 Video]	テレビ電話のカメラ映像
H.263	JT-H263 [H263]	PC 画面
H.264	JT-H264 [H264]	テレビ会議のカメラ映像や PC 画面

本標準で定義する MPEG-4 ビデオは、表 6-10 に示すエンコード・デコード能力をサポートする。なお、フレームレート 30 fps は、テレビジョン標準に従う場合は $30000/1001=29.97$ fps を意味する。

表 6-10/JJ-40.30 本標準で定義する MPEG-4 ビデオ

項目	内容	
	QCIF 映像	
	仕様	SDP 上の記述
プロファイル	Simple Profile	a=fmtp 行に profile-level-id=8 を指定する
レベル	Level 0	
最大フレームレート	15 fps	動作フレームレートを示す。15 fps を利用することを推奨し、次のように記載すべきである。ただし、15 fps の場合は省略可能である。 a=framerate:15
最大ビットレート (レイヤ 3 帯域)	48 kbit/s	動作ビットレートを示す。48 kbit/s の場合次のように指定 b=AS:48

項目	内容	
	QVGA 映像	
	仕様	SDP 上の記述
プロファイル	Simple Profile	a=fmtp 行に profile-level-id=3 を指定する
レベル	Level 3	
最大フレームレート	15 fps	動作フレームレートを示す。15 fps を利用することを推奨し、次のように記載すべきである。ただし、15 fps の場合は省略可能である。 a=framerate:15
最大ビットレート (レイヤ 3 帯域)	384 kbit/s	動作ビットレートを示す。384 kbit/s の場合次のように指定 b=AS:384

項目	内容	
	VGA 映像	
	仕様	SDP 上の記述
プロファイル	Simple Profile	a=fmtp 行に profile-level-id=4 を指定する
レベル	Level 4a	
最大フレームレート	30 fps	動作フレームレートを示す。30 fps を利用することを推奨し、次のように記載すべきである。ただし、30 fps の場合は省略可能である。 a=framerate:30
最大ビットレート (レイヤ 3 帯域)	2 Mbit/s	動作ビットレートを示す。2 Mbit/s の場合次のように指定 b=AS:2000

本標準で定義する H.263 は、表 6-11 に示すエンコード・デコード能力をサポートする。

表 6-11/JJ-40.30 本標準で定義する H.263

項目	内容	
	仕様	SDP 上の記述
版数	H.263-1998	a=rtpmap 行の encoding name で H263-1998 を指定
プロファイル	Baseline profile	
解像度	VGA、SVGA、XGA	a=fmtp 行に次のように指定 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
最大フレームレート	30 fps	動作フレームレートを示す。30 fps を利用することを推奨し、次のように記載すべきである。ただし、30 fps の場合は省略可能である。 a=framerate:30
最大ビットレート (レイヤ 3 帯域)	2 Mbit/s	動作ビットレートを示す。2 Mbit/s の場合次のように指定 b=AS:2000

本標準で定義する H.264 は、表 6-12 に示すエンコード・デコード能力をサポートする。

表 6-12/JJ-40.30 本標準で定義する H.264

項目	内容	
	SD 映像	
	仕様	SDP 上の記述
プロファイル	Constrained Baseline profile	a=fmtp 行に profile-level-id=42c01e または 42e01e を推奨 (注 1)(注 4)
レベル	3.0	
最大フレームレート	30 fps	動作フレームレートを示す。30 fps を利用することを推奨し、 次のように記載すべきである。ただし、30 fps の場合は省略可 能である。 a=framerate:30
最大ビットレート (レイヤ 3 帯域)	2 Mbit/s	動作ビットレートを示す。2 Mbit/s の場合次のように指定 b=AS:2000

項目	内容	
	720p 映像	
	仕様	SDP 上の記述
プロファイル	Constrained Baseline profile	a=fmtp 行に profile-level-id=42c01f または 42e01f を推奨 (注 2)(注 4)
レベル	3.1	
最大フレームレート	30 fps (AVSIP-3b) 15fps (AVSIP-3a)	動作フレームレートを示す。30 fps または 15 fps を利用するこ とを推奨し、30 fps を利用する場合は次のように記載すべきで ある。ただし、30 fps の場合は省略可能である。 a=framerate:30 15 fps を利用する場合は次のように記載すべきである。 a=framerate:15
最大ビットレート (レイヤ 3 帯域)	5 Mbit/s	動作ビットレートを示す。5 Mbit/s の場合次のように指定 b=AS:5000

項目	内容	
	1080i 映像	
	仕様	SDP 上の記述
プロファイル	High profile	a=fmtp 行に profile-level-id=640028 を指定(注 3)
レベル	4.0	
最大フレームレート	30 fps (60 field/s)	動作フレームレートを示す。30 fps (60 field/s) を利用することを推奨し、次のように記載すべきである。ただし、30 fps (60 field/s) の場合は省略可能である。 a=framerate:30
最大ビットレート (レイヤ 3 帯域)	10 Mbit/s	動作ビットレートを示す。10 Mbit/s の場合次のように指定 b=AS:10000

項目	内容	
	1080p 映像	
	仕様	SDP 上の記述
プロファイル	Constrained High profile	a=fmtp 行に profile-level-id=640c28 を指定(注 3)
レベル	4.0	
最大フレームレート	30 fps	動作フレームレートを示す。30 fps を利用することを推奨し、次のように記載すべきである。ただし、30 fps の場合は省略可能である。 a=framerate:30
最大ビットレート (レイヤ 3 帯域)	10 Mbit/s	動作ビットレートを示す。10 Mbit/s の場合次のように指定 b=AS:10000

項目	内容	
	PC 映像	
	仕様	SDP 上の記述
プロファイル	Constrained Baseline profile	a=fmtp 行に次のように指定 VGA : profile-level-id=42c01e または 42e01e(注 1)(注 4) SVGA、XGA : profile-level-id=42c01f または 42e01f (注 2)(注 4)
レベル	VGA : 3.0 SVGA、XGA : 3.1	
最大フレームレート	30 fps	動作フレームレートを示す。30 fps を利用することを推奨し、推奨値を利用する場合は、次のように記載すべきである。ただし、30fps の場合は省略可能である。 a=framerate:30
最大ビットレート (レイヤ 3 帯域)	2 Mbit/s	動作ビットレートを示す。2 Mbit/s の場合次のように指定 b=AS:2000

- (注 1) 42c01e/42e01e は、profile_idc が 66 (ベースライン・プロファイル)、constraint_set0_flag が 1 (ベースライン・プロファイルの制約に適合)、constraint_set1_flag が 1 (メイン・プロファイルの制約に適合)、level_idc が 30 (レベル 3.0)。
- (注 2) 42c01f/42e01f は、profile_idc が 66 (ベースライン・プロファイル)、constraint_set0_flag が 1 (ベースライン・プロファイルの制約に適合)、constraint_set1_flag が 1 (メイン・プロファイルの制約に適合)、level_idc が 31 (レベル 3.1)。
- (注 3) 640028・640c28 は、profile_idc が 100 (ハイ・プロファイル)、level_idc が 40 (レベル 4.0)。
- (注 4) オファー受信時、2 バイト目 (profile-iop) が e0,c0 でない場合も 2 進数表記で x1xx0000 (x は 0 または 1) であればデコード可能と判断できる。(16 進数で f0,d0,70,60,50,40)

本標準で定義する H.264 は、表 6-13 と表 6-14 に示す条件を満たすストリームをデコードできることを必須とする。

表 6-13/JJ-40.30 本標準で定義する H.264 ビットストリーム

H.264 シンタックス要素		SD 映像	720p 映像	1080i 映像 1080p 映像
seq_parameter_set_rbsp()	constraint_set1_flag	1 (注 1)	1 (注 1)	0
	num_reorder_frames	0	0	0
	max_dec_frame_buffering	1	1	1,2 (注 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	0,1 (注 3)
slice_header()	slice_type	0,2,5,7	0,2,5,7	0,2,5,7

- (注 1) メイン・プロファイルでデコード可能であることを示す。これに伴い、フレームあたりの最大スライス数に制約があること、また使用可能なエラー耐性ツールに制約があることに注意。
- (注 2) frame_mbs_only_flag = 1 の場合は 1、それ以外は 2
- (注 3) field_pic_flag = 1 でトップ・フィールドとボトム・フィールドの 2 つのフィールドを参照する場合は 1、それ以外は 0。なお、field_pic_flag = 1 で num_ref_idx_l0_active_minus1 = 1 の場合は、直前にデコードした 2 フィールド、field_pic_flag = 1 で num_ref_idx_l0_active_minus1 = 0 の場合は、直前にデコードした 1 フィールドを参照する。

表 6-14/JJ-40.30 本標準で定義する H.264 ビットストリームの nal_unit_type

H.264 nal_unit_type	意味	扱い		
		解釈 必須	無視 可能	使用 禁止
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		○ (注)	
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..18	Reserved			○
19	Coded slice of an auxiliary coded picture without partitioning			○
20..23	Reserved			○
24..31	Unspecified			○

(注) 2D 映像であるか Side-by-Side の 3D 映像であるかの識別は、Frame packing arrangement SEI (Supplemental Enhancement Information) で通知される。従って、Frame packing arrangement SEI を解釈することで、Side-by-Side の 3D 映像を受信した送受信機器が、ディスプレイ能力に合わせた 2D/3D 表示の手動切替に加え、2D/3D 表示の自動切替の仕組みの実現が可能になる。

H.264 で符号化する映像のカラーマトリクスは、6.2.1 節で示される所要特性に従う。このカラーマトリクスをシンタックスで明示するには表 6-15 で示されるパラメータを設定する。

表 6-15/JJ-40.30 本標準で定義する H.264 ビットストリームのカラーマトリクス

H.264 シンタックス要素		SD 映像	720p 映像 1080i 映像 1080p 映像
vui_parameters()	colour primaries	1,2,6	1,2
	transfer characteristics	1,2,6	1,2
	matrix coefficients	1,2,6	1,2

(注) パラメータが符号化されていない場合は 2 となる。

H.264 で SVGA (800x600 画素) 映像を符号化する場合、符号化ライン数と表示ライン数に 8 ラインの違いがあるため、シーケンス・パラメータ・セットに、表 6-16 に示すパラメータを設定して、符号化した領域の表示範囲を示す。

表 6-16/JJ-40.30 H.264 の SVGA 表示パラメータ規定

H.264 シンタックス名	値	説明
...		
pic_width_in_mbs_minus1	49	デコード画像の水平マクロブロック数-1
pic_height_in_map_units_minus1	37	デコード画像の垂直マップ単位-1 4:2:0 プログレッシブでは、マップ単位=16 画素
...		
frame_cropping_flag	1	SVGA では 1 とする
if(frame_cropping_flag) {		
frame_crop_left_offset	0	デコード画像の左端から画素を表示する
frame_crop_right_offset	0	デコード画像の右端まで画素を表示する
frame_crop_top_offset	0	デコード画像の上端から画素を表示する
frame_crop_bottom_offset	4	デコード画像の下から 4×2=8 画素を表示しない
}		

H.264 で符号化する場合、使用する SPS (Sequence Parameter Set) と PPS (Picture Parameter Set) の id は、異なる解像度では異なる値に設定することを推奨する。

H.264 で Side-by-Side の 3D 映像を符号化する場合、シーケンス単位で Frame packing arrangement SEI を付与し、受信機で 2D 映像と 3D 映像の自動識別を行うことができる。表 6-17 に Frame packing arrangement SEI の推奨値を示す。本標準のオプションである 3D 機能に対応する受信機は、表 6-17 の Frame packing arrangement SEI の解釈を必須とする。

また、本標準のシステムでは Side-by-Side で合成されたフレームを映像ソースのフレームと解釈するため、シーケンス・パラメータ・セットに記載するアスペクト比は、2D 映像の映像ソースフォーマットと同じ値 (例えば、720p ではスクエアピクセル) を記述する。

表 6-17/JJ-40.30 H.264 の Frame packing arrangement SEI 規定

H.264 シンタックス名	値	説明
frame_packing_arrangement_id	0	
frame_packing_arrangement_cancel_flag	0	0 : 以下に続くフィールドが有効
frame_packing_arrangement_type	3	Side-by-Side 3D 映像フォーマット
quincunx_sampling_flag	0	quincunx_sampling は運用しない
content_interpretation_type	1	1 : 画像左半分が左眼用画像
spatial_flipping_flag	0	0 : 画像の左右半分いずれも反転しない
frame0_flipped_flag	0	0 固定
field_views_flag	0	0 固定
current_frame_is_frame0_flag	0	0 固定
frame0_self_contained_flag	0	左眼用画像が右眼用画像を参照する可能性あり
frame1_self_contained_flag	0	右眼用画像が左眼用画像を参照する可能性あり
frame0_grid_position_x	0	左眼用画像の左上画素の水平位置
frame0_grid_position_y	0	左眼用画像の左上画素の垂直位置
frame1_grid_position_x	0	右眼用画像の左上画素の水平位置
frame1_grid_position_y	0	右眼用画像の左上画素の垂直位置
frame_packing_arrangement_reserved_byte	0	0 固定
frame_packing_arrangement_repetition_period	1	本 SEI はシーケンス中 (次の IDR まで) 有効
frame_packing_arrangement_extension_flag	0	0 固定

6.2.3 システムプロファイルの映像メディア仕様

各端末が実装する映像チャンネルの処理能力は、表 6-18 に従わなければならない。

また、PC 画面映像に関しては、映像符号化として H.264、H.263 を利用することとし、TV 電話 (AVSIP-XXa) ではオプション、TV 会議 (AVSIP-XXb) では少なくとも一つのコーデック／画像サイズに対応することを推奨する。3D のカメラに関しては映像符号化方式として H.264 を用いることとし、すべてのプロファイルにおいてオプション扱いとする。

表 6-18/JJ-40.30 各端末が実装する映像メディアの仕様

TV 電話 AVSIP-XXa<カメラ映像>

項目		映像メディアの仕様			
		プロファイル AVSIP-1a	プロファイル AVSIP-1.5a	プロファイル AVSIP-2a	プロファイル AVSIP-3a
カメラ映像 フォーマット	QCIF : 176x144	MPEG4	MPEG4	MPEG4	MPEG4(※o)
	QVGA : 320x240	非対応	MPEG4	MPEG4	MPEG4(※o)
	VGA : 640x480	非対応	非対応	MPEG4	MPEG4(※o)
	SD : 704x480	非対応	非対応	非対応	非対応
	720p : 1280x720	非対応	非対応	非対応	H.264

(※o)該当部分の対応に関してはオプション扱い。

TV 会議 AVSIP-XXb <カメラ映像>

項目		映像メディアの仕様			
		プロファイル AVSIP-2b	プロファイル AVSIP-3b	プロファイル AVSIP-4b-i	プロファイル AVSIP-4b-p
カメラ映像 フォーマット	SD : 704x480	H.264	H.264	H.264(※o)	H.264
	720p : 1280x720	非対応	H.264	H.264(※o)	H.264
	1080i : 1920x1080	非対応	非対応	H.264	H.264(※o)
	1080p : 1920x1080	非対応	非対応	H.264(※o)	H.264

(※o) 該当部分の対応に関してはオプション扱い

7 エンド・エンド制御

本標準では、映像ストリームが 1 個だけの場合を対象とし、複数の場合は今後の検討とする。

7.1 能力交換

能力交換は、SDP のオファー・アンサーモデル RFC4566[RFC4566]に従う。本章では、オーディオビジュアルに固有の留意点と、フォールバック（エラーレスポンス受信時の再発信）軽減、相互接続性向上のための能力交換の詳細を規定する。

7.1.1 オーディオビジュアル通信における SDP 表現

通常、使用するコーデックパラメータをインチャネルで記述するが、呼が成立後にインチャネルでのコーデックパラメータを解釈できずに映像等が表示されない問題が発生する場合がある。そこで、本標準では、映像コーデックがデコード可能であるかどうかを判定する上で特に重要である、使用可能な符号化ツールセット（プロファイル）等を SDP で記述することとし、呼が成立する前に互いのコーデックがデコード可能かを判断するよう規定する。なお、最大フレームレートと最大ビットレートは、`a=framerate`、`b=AS` で記述する。

ただし、映像のビットレートとフレームレートについては、以下のルールに従い、1 回のオファー・アンサーで能力交換を行ってもよい。

- ビットレート

オファーの SDP にはオファー側端末が送受信可能な最大ビットレートを記載し、アンサーの SDP にはアンサー側端末がオファーに記載された値以下のビットレートを記載する。オファー側端末、アンサー側端末ともに、アンサーに記載されたビットレートで映像ストリームの送受信を行う。

SDP の記述内容及びオファー・アンサー時の端末動作に関する詳細は、付属資料 B に示す。

7.1.2 フォールバックを軽減するためのコーデックの併記

本仕様では、オファー時に AVSIP で定められていないものも含めメディア・コーデックを複数 SDP に記載し発信する方式を推奨する。その際、アンサー時には以下の方針で応答を返すことを推奨する

[推奨する動作仕様]

- 対応可能なメディアが一つでも存在すれば、エラーレスポンスを返さず応答をする。
- オファーに対応可能なメディアのコーデックが複数記述されている場合は、メディアを一つのみ選択して応答する。ただし、音声メディアの能力として `telephone-event`[RFC4733]に限り他のメディアと重複して応答してもよい。
- 未対応のメディアに関しては `port0` を指定する。

以下に能力交換の SDP のオファー／アンサーの例を示す。

7.1.2.1 複数メディアの同時併記オファーおよび、アンサー例

【発信の SDP】

```
v=0
o=- 0 0 IN IP4 10.35.197.4
s=-
c=IN IP4 10.35.197.4
t=0 0
m=audio 5004 RTP/AVP 100 9 0 101
a=rtpmap:100 MP4A-LATM/90000
a=fmtp:100 profile-level-id=24;object=23;bitrate=64000
a=rtpmap:9 G722/8000
a=rtpmap:0 PCMU/8000
a=rtpmap:101 telephone-event/8000
a=fmtp:101 0-15
a=sendrecv
m=video 49170 RTP/AVP 121 105 96
b=AS:10000
a=rtpmap:121 H264/90000
a=fmtp:121 profile-level-id=640028;packetization-mode=1
a=rtpmap:105 H264/90000
a=fmtp:105 profile-level-id=42e01f;packetization-mode=0
a=rtpmap:96 MP4V-ES/90000
a=fmtp:96 profile-level-id=4; config=000001B004000001B5090000001010000012100C48D8800
F514043C1463
a=framerate:30
a=sendrecv
```

ここで、音声は MPEG4,G722、PCMU および DTMF 送受信用の telephone-event、映像は H264 の(最大画像サイズ)1080i と 720p、および MPEG4 の VGA の能力が併記されている。また、映像の b 行において複数コーデックの中で考えられる最大のビットレートサイズを提示している。

【応答 SDP】

```
v=0
o=- 0 0 IN IP4 10.35.195.7
s=-
c=IN IP4 10.35.195.7
t=0 0
m=audio 5004 RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=rtpmap:101 telephone-event/8000
a=fmtp:101 0-15
m=video 49170 RTP/AVP 96
b=AS:2000
a=rtpmap:96 MP4V-ES/90000
a=fmtp:96 profile-level-id=4; config=000001B004000001B509000001010000012100C48D8800
F514043C1463
a=framerate:30
```

応答例は以下ようになる。音声に関しては **telephone-event** 及び、コーデックとして **PCMU** を 1 つ選択するとともに、映像に関しては **MPEG4** ビデオを選択している。

7.1.2.2 port0 応答によるアンサー例

アンサー側の TV 会議装置が遠隔カメラ制御未対応で、オファー側の遠隔カメラ制御に該当する **m=application** を port0 で応答している例を以下に示す。

【発信 SDP】

```
v=0
o=- 0 0 IN IP4 10.35.197.4
s=-
c=IN IP4 10.35.197.4
t=0 0
m=audio 5004 RTP/AVP 9 0
a=rtpmap:9 G722/8000
a=rtpmap:0 PCMU/8000
a=sendrecv
m=video 49170 RTP/AVP 105
b=AS:2000
a=fmtp:105 profile-level-id=42e01f;packetization-mode=0
a=framerate:30
a=sendrecv
m=application 5362 RTP/AVP 100
b=AS:7
a=rtpmap:100 H224/4800
a=sendrecv
```

その際のポート 0 対応は以下のように行う。

【応答 SDP】

```
v=0
o=- 0 0 IN IP4 10.35.197.4
s=-
c=IN IP4 10.35.197.4
t=0 0
m=audio 5004 RTP/AVP 0
a=rtpmap:0 PCMU/8000
m=video 49170 RTP/AVP 105
b=AS:2000
a=fmtp:105 profile-level-id=42e01f;packetization-mode=0
a=framerate:30
m=application 0 RTP/AVP 100
```

7.1.2.3 デュアルビデオ制御のための BFCP 発信への対応

会議映像とともに PC 映像を提示するデュアルビデオを制御するため BFCP[RFC4582]に対応する端末が存在する。そのような端末は、付属資料 A 記載の a=fmtp 行以外に、メディアを識別するための a=label 行 [RFC4574]をオファーに付与する。端末が BFCP に未対応の場合には、そのオファーを a=label 行を無視し応答を返す方がフォールバック回数軽減の観点からは望ましい。

7.2 通信モード設定

今後の検討とする。

7.3 C&I

今後の検討とする。

7.3.1 遠隔カメラ制御

本標準に準拠する端末の遠隔カメラ制御はオプションである。

遠隔カメラ制御を利用する場合、RFC 4573 [RFC4573]に従わなければならない。

RFC 4573 [RFC4573]は、H.224 のメディアタイプについて定めた文書であり、TTC 標準 JT-H323 付属資料 Q [H323]に記載されている H.281/H.224 の遠隔カメラ制御手順について示され、SDP メッセージの記述内容について詳述している。これによって、カメラ制御情報は RTP パケットとしてネットワークに流すことができる。

7.3.2 その他の制御

今後の検討とする。

7.4 RTP/RTCP による制御

受信端末から送信端末へのフィードバックおよびフィードバックをベースとしたエラー回復のために、コーデック制御メッセージとして使用する RTCP について示す。オーディオ・ビジュアル・プロファイル (AVPF) に対する規定については、原則、JF-IETF-RFC4585 [RFC4585] と JF-IETF-RFC5104 [RFC5104] を参照する。

本標準では、H.264 に関して AVPF の使用を推奨する。上記文書で複数存在するフィードバックメッセージのうち、端末が実装するメッセージの適用範囲を規定する。表 7-1、表 7-2 にフィードバックメッセージの適用範囲を示す。

表 7-1/JJ-40.30 フィードバックメッセージの適用範囲 (JF-IETF-RFC4585 記載)

フィードバックメッセージ	適用範囲
Generic ACK	オプション
Picture Loss Indication (PLI)	推奨
Slice Loss Indication (SLI)	オプション
Reference Picture Selection Indication (RPSI)	オプション
Application Layer Feedback	オプション

ここで PLI を受信したときの端末の振る舞いとして、FIR と同様に I フレーム (H.264 では IDR ピクチャ、MPEG-4 ビデオでは I-VOP、H.263 では I ピクチャと呼ばれる) を送信することを推奨する。

表 7-2/JJ-40.30 フィードバックメッセージの適用範囲 (JF-IETF-RFC5104 記載)

フィードバックメッセージ	適用範囲
Temporary Maximum Media Stream Bit Rate Request (TMMBR)	オプション
Temporary Maximum Media Stream Bit Rate Notification (TMMBN)	オプション
Full Intra Request (FIR)	送受信とも必須
Temporal-Spatial Trade-off Request (TSTR)	オプション
Temporal-Spatial Trade-off Notification (TSTN)	オプション
H.271 Video Back Channel Message (VBCM)	オプション

端末は表 7-1 及び表 7-2 に示されるフィードバックメッセージの送受信能力を有する場合、SDP のオファー及びアンサー時に、a=rtcp-fb 行を用いて能力を表明する。相手端末が能力を有していない種別のフィードバックメッセージを送信することは推奨されない。なお、a=rtcp-fb 行が記載されていない場合は、相手端末がフィードバックメッセージの能力を有していないと解釈する。a=rtcp-fb 行で FIR の能力が表明されていない場合は、RTP/AVPF でのアンサーを行わず、488 応答を返すことで、RTP/AVP へのフォールバックを促すものとする。

a=rtcp-fb 行で FIR の能力を表明した端末は FIR を受信した場合に、I フレーム (H.264 では IDR ピクチャ、MPEG-4 ビデオでは I-VOP、H.263 では I ピクチャと呼ばれる) を送信して応答しなければならない。

8 マルチメディア多重／分離

本標準に準拠した端末は、映像信号や音声信号を所定の符号化方式でエンコードしたビットストリームを IP ネットワーク上で送受信するためには、あらかじめ指定された方式で IP パケットに変換する。IP パケットに変換することで、送信端末から受信端末へ IP ネットワーク経由でビットストリームを送ることが可能になる。

本章では、ビットストリームを IP パケットに変換・逆変換する規定について述べる。また、本標準に準拠する上で考慮すべき事項について述べる。

8.1 RTP 変換

本標準に準拠した端末は、符号化データを IP ネットワーク上で取り扱うために、RTP ペイロードフォーマットに変換しなければならない。RTP ペイロードフォーマットの内容は符号化方式によって異なり、表 8-1 で示される所定の技術標準に従う必要がある。

表 8-1/JJ-40.30 RTP ペイロードフォーマット

メディア	符号化方式	使用標準
映像	H.263 (1998 版)	RFC 4629 [RFC4629]
	H.264	RFC 6184 [RFC6184]
	MPEG-4 ビデオ	RFC 3016 [RFC3016]
音声	G.711	RFC 3551 [RFC3551]
	G.722	
	G.711.1	RFC 5391 [RFC5391]
	MPEG-4 AAC	RFC 3016 [RFC3016]

8.1.1 H.263 ビデオ

H.263 (1998 版) に準拠した H.263 ビットストリームを RTP 形式で送受信するため、RFC 4629 [RFC4629] に規定される RTP ペイロードフォーマットを用いる。

8.1.2 H.264 ビデオ

H.264 に準拠した H.264 ビットストリームを RTP 形式で送受信するため、RFC 6184 [RFC6184] に規定される RTP ペイロードフォーマットを用いる。

RFC 6184 [RFC 6184] では、H.264 ビットストリームの NAL ユニットのどのように RTP ペイロードに集約させるかによってフォーマット構造が異なり、7 種類のペイロード構造を定義している。

そのうち、本標準ではパケットタイプに関してはシングル NAL, FU-A, STAP-A の 3 種を対象とし、非インターリーブモードを利用することとする。

H264 はビットストリームを送信する単位を NALU と呼んでいるが、シングル NAL ユニットの 1 つのパケットに 1 つの NALU を格納するパケットタイプである。これに対して、FU-A パケットは 1 つのパケットに複数の NALU を格納し送信するパケットタイプで通信データ量の多い TV 電話に利用されており、特に CABAC を使用するハイ・プロファイルの場合はスライスの NALU 長が大きいため FU-A の使用が適切である。

また、STAP-A は複数のパケットに分割し 1 つの NALU を送信するパケットタイプであり、高解像度の映像通信を低ビットレート帯域で行うために有効である。

本標準ではこれらのパケットタイプのサポートを SDP の packetization-mode によって能力交換を行う。packetization-mode が 0 の場合はシングル NAL のみを、1 の場合はシングル NAL、FU-A、STAP-A すべてを、

付与しない場合はシングル NAL のサポートを行えなくてはならない。さらに、既存の端末との相互接続のために AVSIP-3a, AVSIP-3b, AVSIP-4b-i, AVSIP-4b-p に対応する際は FU-A の受信能力サポートを推奨する。

表 8-2/JJ-40.30 H.264 パケットサイズ

packetization-mode (パケットタイプ)	プロファイル AVSIP-3a	プロファイル AVSIP-2b	プロファイル AVSIP-3b	プロファイル AVSIP-4b-i	プロファイル AVSIP-4b-p
0,付与なし (シングル NAL)	必須	必須	必須	オプション	オプション
1 (FU-A,STAP-A シングル NAL)	オプション	推奨	推奨	必須	必須

IDR フレームを挿入する場合には全ての IDR フレームの前に必ずその IDR で参照する PPS/SPS 情報を付加して送信することを必須とし、RTP/AVPF を利用しない場合は、画像のリフレッシュのために、定期的な IDR フレームあるいはイントラ・マクロブロック・リフレッシュを挿入することを推奨する。なお、IDR フレームを定期的には挿入しないエンコーダが存在する可能性も否定できない。そういったエンコーダとの接続性を考慮するなら、デコーダはデコードエラー発生時に、IDR フレームを受信することを前提としない処理も必要である。例えば、IDR フレームをデコードするまで復号画像の描画を停止するような処理は行わないことが挙げられる。

8.1.3 MPEG-4 ビデオ

MPEG-4 ビデオに準拠した MPEG-4 ビデオビットストリームを RTP 形式で送受信するため、RFC 3016 [RFC3016]に規定される RTP ペイロードフォーマットを用いる。

本標準では、DCI 値を SDP の a=fmtp 行にある config 値に指定しなければならない。ここで、送信側が送出する MPEG-4 ビデオビットストリームを載せた RTP パケットには、DCI 値に相当するビットストリームを含むべきである。本標準では SDP を使った DCI 値送出手順を必須としているものの、相互接続性の向上を考慮して、RTP パケットにも config 値と同一の DCI 値を含めるべきである。

誤り耐性ツールは resync_marker の利用を推奨し、data_partitioned、reversible_vlc は利用しないことを推奨する。

また、RTP/AVPF を利用しない場合は、定期的な I-VOP を挿入することを推奨し、I-VOP の前に config パラメータを付加することを必須とする。

RTP パケット化する際には、RFC 3016 [RFC3016] に規定されている分割方式のうち、a) DCI 情報のみ、b)DCI 情報+ビデオ情報、d) ビデオ情報のみ、のいずれかを使用する。

VideoPacket の最大サイズは、通信事業者に指定される場合がある。

8.1.4 G.711 オーディオ、G.722 オーディオ

G.711 や G.722 に準拠したオーディオビットストリームを RTP 形式で送受信するため、RFC 3551 [RFC3551] に規定される RTP ペイロードフォーマットを用いる。

8.1.5 G.711.1 オーディオ

G.711.1 オーディオに準拠したオーディオビットストリームを RTP 形式で送受信するため、RFC 5391 [RFC5391] に規定される RTP ペイロードフォーマットを用いる。

G.711.1 オーディオを利用する場合は、ベースとなるストリームとして G.711 μ -law を使用する。

8.1.6 MPEG-4 オーディオ

MPEG-4 オーディオに準拠した MPEG-4 オーディオビットストリームを RTP 形式で送受信するため、RFC 3016 [RFC3016] に規定される RTP ペイロードフォーマットを用いる。

MPEG-4 オーディオビットストリームのフォーマット構造は、LATM フォーマットでなければならない。

8.1.7 帯域のネゴシエーション結果に基づくメディアパケット送出

各端末は、オファー・アンサーモデルによるネゴシエーションの結果に従いエンコードを行い、生成されたメディアパケットを送出する。

メディアパケットを送出する際のビットレートは、アンサーに指定された b 行の値を利用する。この際、SDP の b 行に指定するビットレートには、レイヤ 3、レイヤ 4 のヘッダ情報、RTP ヘッダ、RTP ペイロードヘッダ、エレメンタリストリームのビットレートが含まれることに留意する。

ただし、G.711 のように音声コーデックが固有のビットレートを持つ場合、SDP の b 行は省略される場合がある。その場合、音声コーデック固有のビットレートに従い、メディアパケットを送信する。

付属資料A NGN を利用する際の留意事項

(本付属資料は仕様の一部である。)

A.1 NGN を利用する際の関連仕様

NGN を利用する際には、網に接続する UNI の規定として TR-1014 [TR-1014]、SIP/SDP を用いた呼接続と RTP などメディアの送受信の規定として JT-Q3402 [Q3402]に従う。

メディアの送受信に関しては、特に JT-Q3402 [Q3402]が有する QoS 機能に関して留意すべきである。SDP によりネゴシエーションした帯域に対して網が許容する送出トラフィック条件については、JT-Q3402 [Q3402] の付属資料 g を参照すること。

A.2 誤課金防止のためのフォールバックを利用した能力交換方法

NGN を利用する際の能力交換例として、SDP のオファー・アンサーモデルにフォールバック（エラーレスポンス受信時の再発信機能）を利用した能力交換方法を記述する。フォールバックを利用することにより、メディアストリームを送受信する以前に利用可能なメディア情報が相互で認識可能となり、誤課金（回線接続が行われたにも関わらず映像の受信ができない状態で課金がされてしまう）を防ぐことができる。

フォールバックを想定した端末は、最初に自端末の最大能力での接続を期待した SDP でオファーを試み、488 レスポンスを受信した場合、Warning ヘッダの warn-code を判定し、本標準で推奨するプロファイルの SDP で再オファー（フォールバック）することとする。図 A-1 にフォールバックフローの一例を示す。ここに現れる warn-code の定義は TR-1020 [TR-1020]参照。

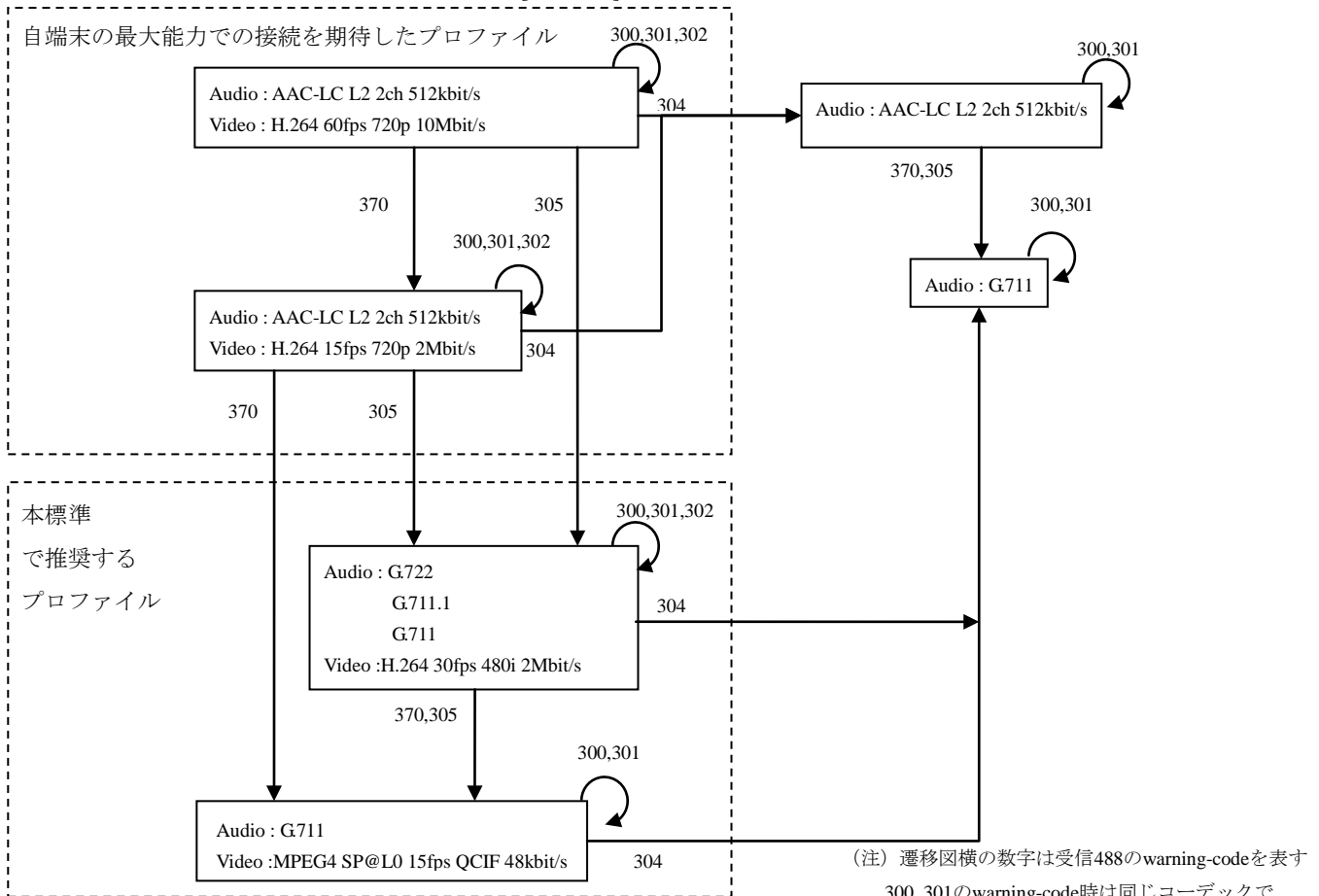


図 A-1/JJ-40.30 フォールバックフローの一例

付属資料B SDP の記述及びネゴシエーションの詳細

(本付属資料は仕様の一部である。)

B.1 概要

本標準で示す SIP/SDP を用いたネゴシエーションに関しては RFC4566[RFC4566]に準拠するが、本付属資料は、SIP 網環境下のオーディオビジュアル通信システムが SIP/SDP を用いてネゴシエーションを行う場合における、オーディオ及びビデオのコーデックに固有の SDP の記述内容とパラメータのチェック手順について特記する。

B.2 SDP の記述内容とネゴシエーション手順

本節及び従属節では、映像コーデック及び音声コーデックに関して、SDP の記述内容とネゴシエーション手順に関して規定する。第 6 章に規定した AVSIP プロファイルで利用するコーデックに関して、具体的な SDP 行及びパラメータと、信号条件、SDP オファー時及び SDP アンサー時の端末動作について記載している。

信号条件の欄では、UNI におけるインタフェース上の信号に現れるか、現れないかという観点（ダイナミックビュー）で m (mandatory) や o (optional) といった条件コードの表現を行っている。例えば、m (mandatory) であれば、当該の行またはオプションは、端末が送信または受信するメッセージに必ず記述されなければならない。条件コードの定義を付表 B-1 に示す。

付表 B-1/JJ-40.30 条件コードの定義

条件コード	定義
m	当該のパラメータは、必須 (mandatory) である。オファーの SDP 中の必須のパラメータは存在していなくてはならず、また、オファーを受ける側で各 RFC に従い理解され得なくてはならない。同じくアンサーの SDP 中の必須のパラメータは存在していなくてはならず、また、アンサーを処理する側で各 RFC に従い理解され得なくてはならない。
o	当該のパラメータは選択的 (optional) である。選択的とは、当該のパラメータは、オファーやアンサーに存在しても良い。また当該のパラメータがオファーやアンサー内に存在した場合には、RFC に従い受信側で理解され、対応する動作が行われなければならない。
—	当該のパラメータは適用されない。適用されない当該のパラメータは、オファーやアンサー内に存在してはならない。
c	当該のパラメータの適用は、SDP 処理上の文脈による (conditional)。

B.2.1 映像コーデック

B.2.1.1 MPEG-4 ビデオ

MPEG-4 ビデオ [MPEG-4 Video]を利用する場合にSDPのMedia descriptionに記述するコーデックパラメータと、SDP送受信時の端末動作を付表 B-2 に示す。

付表 B-2/JJ-40.30 SDP 記述コーデックパラメータ (MPEG-4 ビデオ)

コーデック	SDP 記述		信号条件		SDP 送受信時の端末動作	
	行	パラメータ	送信	受信	オファー時	アンサー時
MPEG-4 ビデオ	b=AS		m	m	<ul style="list-style-type: none"> 設定必須。AVSIP プロファイル利用時は表 6-10 に従い値を設定する。 	<ul style="list-style-type: none"> オファーされた帯域値で送受信可能である場合にのみ 200 応答を返し、通信不可である場合は 488 応答を返す。 200 応答の SDP には、オファーされた値と同一値を設定する。
	a=rtpmap	payload type	m	m	<ul style="list-style-type: none"> RFC3016 に従い 96~127 の任意の値を設定する。 	<ul style="list-style-type: none"> オファーされた payload type と同一の値を返す。
		encoding name	m	m	<ul style="list-style-type: none"> RFC3016 に従い"MP4V-ES"を設定する。 	<ul style="list-style-type: none"> RFC3016 に従い"MP4V-ES"を設定する。
		clock rate	m	m	<ul style="list-style-type: none"> RFC3016 に従い 90000 を設定する。 	<ul style="list-style-type: none"> RFC3016 に従い 90000 を設定する。
		encoding parameters	—	o	<ul style="list-style-type: none"> RFC4566 及び RFC3016 に従い、設定しない。 	<ul style="list-style-type: none"> オファーに設定されている場合は 488 応答を返す。 アンサーには設定しない。
	a=fmtp	profile-level-id	m	o	<ul style="list-style-type: none"> 設定必須。AVSIP プロファイル利用時は表 6-10 に従い値を設定する。 	<ul style="list-style-type: none"> profile-level-id パラメータの設定がない場合は、RFC3016 に従い 1 が指定されたと解釈する。 オファーされたプロファイル・レベルで送受信可能である場合にのみ 200 応答を返し、通信不可である場合は 488 応答 (warn-code が 305) を返す。 200 応答の SDP には、オファーされた値と同一値を設定する。
config		m	o	<ul style="list-style-type: none"> ISO/IEC 14496-2 に規定される範囲で profile-level-id パラメータに記述するプロファイル・レベルで許容される DCI (Decoder Configuration Information) 値を記述する。 AVSIP プロファイル利用時は DCI 値に記述される画面サイズ (video_object_layer_width : 水平画素数と video_object_layer_height : 垂直画素数) は表 6-8 及び表 6-10 に従う。 	<ul style="list-style-type: none"> config パラメータの設定がない場合は、profile-level-id パラメータで指定されたプロファイル・レベルの最大能力に対応できるならば 200 応答を返し(注 1)、通信不可である場合は 488 応答を返す。 config パラメータの設定がある場合は、DCI 値として記述される画面サイズ (video_object_layer_width : 水平画素数と video_object_layer_height : 垂直画素数) を用いて送受信可能かつユーザへ表示可能である場合にのみ 200 応答を返し(注 1)、通信不可である場合は 488 応答 (warn-code が 305) を返す。Simple Profile を利用する場合は、config 中の他のパラメータについては、ISO/IEC 14496-2 に規定される範囲の値が指定される限りは、オファーで指定された内容を受け入れデコードが可能でなければならない。 	

	上記以外の a=fmtp 行 のパラメータ	－	o	<ul style="list-style-type: none"> 設定しない (RFC 3016 [RFC 3016]に規定がない)。 	<ul style="list-style-type: none"> 無視する。
	a=framerate	o	o	<ul style="list-style-type: none"> 設定してもよい。AVSIP プロファイル利用時は表 6-10 に従い値を設定する。 	<ul style="list-style-type: none"> a=framerate 行の指定がない場合は、当該 AVSIP プロファイルにおける最大フレームレートがオファーされたものと解釈する。 オファーされたフレームレート以下で通信が可能である場合にのみ 200 応答を返し、通信不可である場合は 488 応答を返す。 200 応答の SDP には、オファーされた値以下の値を設定する(注 2)。

(注 1) オファー側はアンサーで受信した config 値を、アンサー側はオファーで受信した config 値を用いて通信を行う。

(注 2) オファー側・アンサー側、ともにアンサーに記述される a=framerate 行の値を用いて通信を行う。

B.2.1.2 H.263

H.263 [H263]を利用する場合に SDP の Media description に記述するコーデックパラメータと、SDP 送受信時の端末動作を付表 A-3 に示す。

付表 B-3/JJ-40.30 SDP 記述コーデックパラメータ (H.263)

コーデック	SDP 記述		信号条件		SDP 送受信時の端末動作	
	行	パラメータ	送信	受信	オファー時	アンサー時
H.263	b=AS		m	m	<ul style="list-style-type: none"> 設定必須。AVSIP プロファイル利用時は表 6-11 に従い値を設定する。 	<ul style="list-style-type: none"> オファーされた帯域値で送受信可能である場合にのみ 200 応答を返し、通信不可である場合は 488 応答を返す。 200 応答の SDP には、オファーされた値と同一値を設定する。
	a=rtmpmap	payload type	m	m	<ul style="list-style-type: none"> RFC4629 に従い 96~127 の任意の値を設定する。 	<ul style="list-style-type: none"> オファーされた payload type と同一の値を返す。
		encoding name	m	m	<ul style="list-style-type: none"> RFC4629 に従い"H263-1998"を設定する。 	<ul style="list-style-type: none"> RFC4629 に従い"H263-1998"を設定する。
		clock rate	m	m	<ul style="list-style-type: none"> RFC4629 に従い 90000 を設定する。 	<ul style="list-style-type: none"> RFC4629 に従い 90000 を設定する。
		encoding parameters	—	o	<ul style="list-style-type: none"> RFC4566 及び RFC4629 に従い、設定しない。 	<ul style="list-style-type: none"> オファーに設定されている場合は 488 応答を返す。 アンサーには設定しない。
	a=fmtp	QCIF	m	o	<ul style="list-style-type: none"> 設定必須(注 1)。 	<ul style="list-style-type: none"> QCIF で通信を行う場合に限り参照し、送受信可能かつユーザへの表示が可能である場合は 200 応答を返し、それ以外の場合は 488 応答を返す。 AVSIP プロファイル利用時には無視する(注 1)。
		CUSTOM	o	o	<ul style="list-style-type: none"> AVSIP プロファイル利用時は、18 の「PC 画面映像フォーマット」に従い値を設定する。 	<ul style="list-style-type: none"> CUSTOM で指定された値で送受信可能かつユーザへの表示が可能である場合は 200 応答を返し、それ以外の場合は 488 応答を返す。 200 応答の SDP には、オファーされた CUSTOM と同一の画面解像度で、オファー以下のフレームレートを設定する(注 2)。
		PAR	o	o	<ul style="list-style-type: none"> AVSIP プロファイル利用時は、表 6-11 に従い値を設定する。 	<ul style="list-style-type: none"> PAR の指定がない場合は、RFC4629 に従い 12:11 がオファーされたものと解釈する。 オファーされたピクセル縦横比で送受信可能かつユーザへの表示が可能である場合のみ 200 応答を返し、それ以外の場合は 488 応答を返す。 200 応答の SDP には、オファーされた PAR と同一の値を設定する。
		その他 RFC4629 記載のパラメータ	o	o	<ul style="list-style-type: none"> RFC4629 の規定に従い、設定してもよい。 	<ul style="list-style-type: none"> RFC4629 の規定に従い解釈を行い、通信が送受信とも可能である場合にのみ 200 応答を返し、通信不可である場合は 488 応答を返す。

a=framerate	o	o	<ul style="list-style-type: none"> 設定してもよい。設定する場合は、CUSTOM 等のフレームレートを指定する a=fmtp 行のパラメータと矛盾しない値を設定する。 AVSIP プロファイル利用時は表 6-11 に従い値を設定する。 	<ul style="list-style-type: none"> a=framerate 行の指定がある場合は、オファーされたフレームレート以下で通信が可能である場合にのみ 200 応答を返し、通信不可である場合は 488 応答を返す。200 応答の SDP には、オファーされた値以下の値を設定し^(*2)、CUSTOM 等のフレームレートを指定する a=fmtp 行のパラメータと矛盾しない値とする。 a=framerate 行の指定がない場合は、アンサーに a=framerate 行は指定せず、CUSTOM 等のフレームレートを指定する a=fmtp 行のパラメータのみ指定を行う。
<p>(注1) QCIF パラメータは RFC4629 の規定に基づき設定必須であるが、表 6-11 の規定から AVSIP プロファイルを利用するシステムでは参照する必要がない。</p>				
<p>(注2) オファー側・アンサー側、ともにアンサーに記述される a=framerate 行の値を用いて通信を行う。</p>				

B.2.1.3 H.264

H.264 [H264]を利用する場合に SDP の Media description に記述するコーデックパラメータと、SDP 送受信時の端末動作を付表 B-4 に示す。

付表 B-4/JJ-40.30 SDP 記述コーデックパラメータ (H.264)

コーデック	SDP 記述		信号条件		SDP 送受信時の端末動作	
	行	パラメータ	送信	受信	オファー時	アンサー時
H.264	b=AS		m	m	<ul style="list-style-type: none"> 設定必須。AVSIP プロファイル利用時は表 6-12 に従い値を設定する。 	<ul style="list-style-type: none"> オファーされた帯域値で送受信可能である場合にのみ 200 応答を返し、通信不可である場合は 488 応答を返す。 200 応答の SDP には、オファーされた値と同一値を設定する。
	a=rtptime	payload type	m	m	<ul style="list-style-type: none"> RFC6134 に従い 96~127 の任意の値を設定する。 	<ul style="list-style-type: none"> オファーされた payload type と同一の値を返す。
		encoding name	m	m	<ul style="list-style-type: none"> RFC6134 に従い"H264"を設定する。 	<ul style="list-style-type: none"> RFC6134 に従い"H264"を設定する。
		clock rate	m	m	<ul style="list-style-type: none"> RFC6134 に従い 90000 を設定する。 	<ul style="list-style-type: none"> RFC6134 に従い 90000 を設定する。
		encoding parameters	—	o	<ul style="list-style-type: none"> RFC4566 及び RFC6134 に従い、設定しない。 	<ul style="list-style-type: none"> オファーに設定されている場合は 488 応答を返す。 アンサーには設定しない。
	a=fmtsp	profile-level-id	m	o	<ul style="list-style-type: none"> 設定必須。AVSIP プロファイル利用時は表 6-12 に従い値を設定する。 	<ul style="list-style-type: none"> profile-level-id パラメータの設定がない場合は、RFC6134 に従い基本プロファイルのレベル 1 が指定されたと解釈する。 オファーされた値を解析し、profile_idc (1 バイト目)及び profile_iop (2 バイト目) が対応可能で、且つ、level_idc(3 バイト目)の値以下で送受信可能ある場合に 200 応答を返し、通信不可である場合は 488 応答 (warn-code が 305) を返す。 sprop-parameter-sets で明示的な指定がない場合は level_idc の値(3 バイト目)を低く応答することを許容する。
packetization-mode		cl	o	<ul style="list-style-type: none"> デフォルト (すなわち 0=シングル NAL モード) 以外の場合に記述することとし、分割ユニットを使用する場合は明示的に示す(注 1)。 	<ul style="list-style-type: none"> packetization-mode パラメータの設定がない場合は、RFC6134 に従い 0 (シングル NAL モード) が指定されたと解釈する。 オファーされた packetization-mode で送受信可能である場合にのみ 200 応答を返し、通信不可である場合は 488 応答 (warn-code が 305) を返す。 200 応答の SDP には、オファーされた値と同一値を設定する。 	

	sprop-parameter-sets	o	o	<ul style="list-style-type: none"> • RFC6134 の規定に従い、設定してもよい。 	<ul style="list-style-type: none"> • RFC6134 の規定に従い、設定してもよい。
	max-mbps	c2	o	<ul style="list-style-type: none"> • RFC6134 の規定に従い、設定してもよい。 	<ul style="list-style-type: none"> • RFC6134 の規定に従い解釈を行い、通信が送受信とも可能である場合にのみ 200 応答を返し、通信不可である場合は 488 応答を返す。
	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	<ul style="list-style-type: none"> • RFC6134 の規定に従い、設定してもよい。 • AVSIP プロファイル利用時には設定しない(注 2)。 	
その他 RFC6134 記載のパラメータ	c3	o			
a=framerate	o	o	<ul style="list-style-type: none"> • 設定してもよい。AVSIP プロファイル利用時は表 6-12 に従い値を設定する。 	<ul style="list-style-type: none"> • a=framerate 行の指定がない場合は、当該 AVSIP プロファイルにおける最大フレームレートがオファーされたものと解釈する。 • オファーされたフレームレート以下で通信が可能である場合にのみ 200 応答を返し、通信不可である場合は 488 応答を返す。 • 200 応答の SDP には、オファーされた値以下の値を設定する。 	

- c1: シングル NAL モードを使用する場合には、設定は必須ではない。シングル NAL モード以外を使用する場合には、設定必須。
- c2: H.264 の付属資料 A で、profile-level-id で指定したプロファイル・レベルに対応する値以上を使用する場合には、設定必須。
- c3: 使用するコーデックの設定により、RFC6134 の規定により設定が必要な場合は、設定必須。
- (注 1) H.264 のスライスのサイズを制御すれば、広帯域の通信の場合でもシングル NAL モードで対応可能である。ただし、スライスヘッダのオーバーヘッド、スライスを越えた画面内予測と動きベクトル情報の動き予測ができないため、符号化効率が劣化する。スライスのサイズが MTU サイズを超えないように制御できない場合や符号化効率を考慮する場合は、分割ユニットを利用すべきである。また、CABAC 適用時においては、CABAC はスライス長が短くなると圧縮効率は劣化し、スライス先頭における演算負荷が他と比べて大きいという実装上の課題点に配慮が必要であるため、分割ユニットを利用すべきである。
- (注 2) AVSIP プロファイル利用時は、他のパラメータに関しては RFC6134 で規定されるデフォルト値を使用するか、またはプロファイル・レベルの能力範囲内で通信を行うため設定不要である。

B.2.2 音声コーデック

B.2.2.1 G.711 μ -law

G.711 μ -law [G711]を利用する場合に SDP の Media description に記述するコーデックパラメータと、SDP 送受信時の端末動作を付表 B-5 に示す。

付表 B-5/JJ-40.30 SDP 記述コーデックパラメータ (G.711 μ -law)

コーデック	SDP 記述		信号条件		SDP 送受信時の端末動作	
	行	パラメータ	送信	受信	オファー時	アンサー時
G.711 μ -law	a=rtpmap	payload type	m	m	<ul style="list-style-type: none"> • RFC3551 に従い 0 を設定。 	<ul style="list-style-type: none"> • RFC3551 に従い 0 を設定。
		encoding name	m	m	<ul style="list-style-type: none"> • RFC3551 に従い "PCMU" を設定。 	<ul style="list-style-type: none"> • RFC3551 に従い "PCMU" を設定。
		clock rate	m	m	<ul style="list-style-type: none"> • RFC3551 に従い 8000 を設定。 	<ul style="list-style-type: none"> • RFC3551 に従い 8000 を設定。
		encoding parameters	—	o	<ul style="list-style-type: none"> • 設定しない。 	<ul style="list-style-type: none"> • オファーに設定されている場合は 488 応答を返す。 • アンサーには設定しない。
	a=ptime		m	o	<ul style="list-style-type: none"> • TR-1020 に従い 20 を設定。 	<ul style="list-style-type: none"> • ptime 行の設定がない場合は、20 が指定されたものと解釈する。 • オファーされたパケット化周期で送受信可能である場合にのみ 200 応答を返し、通信不可である場合は 488 応答を返す。 • 200 応答の SDP には、オファーされた値と同一値を設定する。

B.2.2.2 G.722

G.722 [G722]を利用する場合に SDP の Media description に記述するコーデックパラメータと、SDP 送受信時の端末動作を付表 B-6 に示す。

付表 B-6/JJ-40.30 SDP 記述コーデックパラメータ (G722)

コーデック	SDP 記述		信号条件		SDP 送受信時の端末動作	
	行	パラメータ	送信	受信	オファー時	アンサー時
G.722	a=rtpmap	payload type	m	m	<ul style="list-style-type: none"> • RFC3551 に従い 9 を設定。 	<ul style="list-style-type: none"> • RFC3551 に従い 9 を設定。
		encoding name	m	m	<ul style="list-style-type: none"> • RFC3551 に従い "G722" を設定。 	<ul style="list-style-type: none"> • RFC3551 に従い "G722" を設定。
		clock rate	m	m	<ul style="list-style-type: none"> • RFC3551 に従い 8000 を設定。 	<ul style="list-style-type: none"> • RFC3551 に従い 8000 を設定。
		encoding parameters	—	o	<ul style="list-style-type: none"> • 設定しない。 	<ul style="list-style-type: none"> • オファーに設定されている場合は 488 応答を返す。 • アンサーには設定しない。
	a=ptime		m	o	<ul style="list-style-type: none"> • TR-1020 に従い 20 を設定。 	<ul style="list-style-type: none"> • ptime 行の設定がない場合は、G.711 μ-law と同様に 20 が指定されたものと解釈する。 • オファーされたパケット化周期で送受信可能である場合にのみ 200 応答を返し、通信不可である場合は 488 応答を返す。 • 200 応答の SDP には、オファーされた値と同一値を設定する。

B.2.2.3 G.711.1

G.711.1 [G711.1]を利用する場合に SDP の Media description に記述するコーデックパラメータと、SDP 送受信時の端末動作を付表 B-7 に示す。

付表 B-7/JJ-40.30 SDP 記述コーデックパラメータ (G711.1)

コーデック	SDP 記述		信号条件		SDP 送受信時の端末動作	
	行	パラメータ	送信	受信	オファー時	アンサー時
G.711.1	a=rtpmap	payload type	m	m	<ul style="list-style-type: none"> • RFC5391 に従い 96~127 の任意の値を設定する。 	<ul style="list-style-type: none"> • オファーされた payload type と同一の値を返す。
		encoding name	m	m	<ul style="list-style-type: none"> • RFC5391 に従い"PCMU-WB"を設定。 	<ul style="list-style-type: none"> • RFC5391 に従い"PCMU-WB"を設定。
		clock rate	m	m	<ul style="list-style-type: none"> • 設定必須。AVSIP プロファイル利用時は表 6-4 に従い設定する。 	<ul style="list-style-type: none"> • オファーされた値で送受信可能である場合にのみ 200 応答を返し、通信不可である場合は 488 応答を返す。 • 200 応答の SDP には、オファーされた値と同一値を設定する。
		encoding parameters	—	o	<ul style="list-style-type: none"> • 設定しない。 	<ul style="list-style-type: none"> • オファーに設定されている場合は 488 応答を返す。 • アンサーには設定しない。
	a=fmtp	mode-set	m	o	<ul style="list-style-type: none"> • 設定必須。AVSIP プロファイル利用時は表 6-4 に従い設定する。 	<ul style="list-style-type: none"> • オファーされたモードで送受信可能である場合にのみ 200 応答を返し、通信不可である場合は 488 応答を返す。 • 200 応答の SDP には、オファーされた値と同一値を設定する。
		上記以外の a=fmtp 行パラメータ	—	o	<ul style="list-style-type: none"> • 設定しない (RFC5391 に規定がない)。 	<ul style="list-style-type: none"> • 無視する。
	a=ptime		m	o	<ul style="list-style-type: none"> • 設定必須。AVSIP プロファイル利用時は表 6-4 に従い設定する。 	<ul style="list-style-type: none"> • ptime 行の設定がない場合は、G.711 μ-law と同様に 20 が指定されたものと解釈する。 • オファーされたパケット化周期で送受信可能である場合にのみ 200 応答を返し、通信不可である場合は 488 応答を返す。 • 200 応答の SDP には、オファーされた値と同一値を設定する。

B.2.2.4 MPEG-4 オーディオ

MPEG-4 オーディオ [MPEG-4 Audio]を利用する場合に SDP の Media description に記述するコーデックパラメータと、SDP 送受信時の端末動作を付表 B-8 に示す。

付表 B-8/JJ-40.30 SDP 記述コーデックパラメータ (MPEG-4 オーディオ)

コーデック	SDP 記述		信号条件		SDP 送受信時の端末動作	
	行	パラメータ	送信	受信	オファー時	アンサー時
MPEG-4 オーディオ	b=AS		m	m	<ul style="list-style-type: none"> 設定必須。AVSIP プロファイル利用時は表 6-5 に従い設定する。 	<ul style="list-style-type: none"> オファーされた帯域値で送受信可能である場合にのみ 200 応答を返し、通信不可である場合は 488 応答を返す。 200 応答の SDP には、オファーされた値と同一値を設定する。
	a=rtmpmap	payload type	m	m	<ul style="list-style-type: none"> RFC3016 に従い 96~127 の任意の値を設定する。 	<ul style="list-style-type: none"> オファーされた payload type と同一の値を返す。
		encoding name	m	m	<ul style="list-style-type: none"> RFC3016 に従い、"MP4A-LATM"を設定。 	<ul style="list-style-type: none"> RFC3016 に従い"MP4A-LATM"を設定。
		clock rate	m	m	<ul style="list-style-type: none"> 設定必須。AVSIP プロファイル利用時は表 6-5 に従い設定する。 	<ul style="list-style-type: none"> オファーされた値で送受信可能である場合にのみ 200 応答を返し、通信不可である場合は 488 応答を返す。
		encoding parameters	—	o	<ul style="list-style-type: none"> 設定しない。 	<ul style="list-style-type: none"> オファーに設定されている場合は 488 応答を返す。 アンサーには設定しない。
	a=fmtp	profile-level-id	m	o	<ul style="list-style-type: none"> 設定必須。AVSIP プロファイル利用時は表 6-5 に従い設定する。 	<ul style="list-style-type: none"> profile-level-id パラメータの指定がない場合は、RFC3016 に従い 30 (Natural Audio Profile/Level 1) が指定されたと解釈する。 オファーされた値で送受信可能である場合にのみ 200 応答を返し、通信不可である場合は 488 応答を返す。
		object	m	o		<ul style="list-style-type: none"> オファーされた値で送受信可能である場合にのみ 200 応答を返し、通信不可である場合は 488 応答を返す。
		config	m	o		
		bitrate	m	o		
		cpresent	c1	o	<ul style="list-style-type: none"> RFC3016 に従い、設定してもよい。AVSIP プロファイル利用時は 1 を設定する。 	<ul style="list-style-type: none"> cpresent パラメータの指定がない場合は、RFC3016 に従い 1 が指定されたと解釈する。 200 応答にはオファーされた値と同一値を設定する。
		上記以外の a=fmtp 行のパラメータ	—	o	<ul style="list-style-type: none"> 設定しない (RFC3016 に規定がない)。 	<ul style="list-style-type: none"> 無視する。
	a=ptime		m	o	<ul style="list-style-type: none"> 設定必須。AVSIP プロファイル利用時は表 6-5 に従い設定する。 	<ul style="list-style-type: none"> ptime 行の設定がない場合は、G.711 μ-law と同様に 20 が指定されたものと解釈する。 オファーされたパケット化周期で送受信可能である場合にのみ 200 応答を返し、通信不可である場合は 488 応答を返す。 200 応答の SDP には、オファーされた値と同一値を設定する。

c1: RTP ペイロード中に config 情報が記述されない場合は、0 を設定しなければならない。

付録I 映像利用時の b=行に指定する値の算出方法

(本付録は参考資料であり、仕様ではない)

I.I はじめに

本付録では、映像を RTP で送信する場合に、SDP の b=行に指定する値を算出するための方法について、一例を記述する。具体的には、SDP の b=行に指定する値にはレイヤ 3、レイヤ 4 のオーバーヘッドを含めた値を指定する必要があるため、エンコーダの出力である映像ストリームのビットレートから、b=行に指定する値を算出する方法について、一例を示す。

I.II 前提条件

算出方法に関する前提条件を下記する。

なお、本前提条件は、計算式を考える上での前提条件であり、本標準の実装を制約するものではない。

- ・ エンコードのモードは、CBR (ConstantBitRate)とする。
- ・ RTP の CSRC、拡張ヘッダは利用しない。
- ・ パケット化サイズ(注 1)のサイズは一定とする。
- ・ MPEG-4 の場合、パケット化サイズ単位で分割された映像ストリームを 1 個の IP パケットに変換する。
- ・ H.264 の場合、パケット化サイズが 1 個の IP パケットに変換可能なサイズであれば、シングル NAL ユニットパケットを利用する。パケット化サイズが 1 個の IP パケットに変換できないサイズの場合は、分割ユニットを利用し複数の IP パケットに変換する。ここでは、分割ユニットとして、FU-A を利用することを前提とする。

(注 1) パケット化サイズとは、MPEG-4 ビデオの場合は VideoPacket サイズであり、H.264 の場合はスライスのサイズである。パケット化サイズを小さくすれば、IP パケット数が増えるため、その分のヘッダのオーバーヘッドにより、必要な帯域が増加する。

I.III パラメータ

算出方法に関するパラメータを下記する。

付表 I-1/JJ-40.30 b=行に指定する値算出のためのパラメータ

パラメータ	変数名	備考
映像ストリームのビットレート	x [bit/s]	
映像ストリームのフレームレート	y [fps]	
パケット化サイズ	z [bit]	
IP ヘッダ長	a [bit]	IPv4 (オプションなし) の場合、 20byte(160bit) IPv6 (オプションなし) の場合、 40byte(320bit)
UDP ヘッダ長	b [bit]	8byte(64bit)
RTP ヘッダ長	c [bit]	CSRC、拡張ヘッダなしの場合、12byte (96bit)
RTP ペイロードヘッダ長	d [bit]	MPEG-4 ビデオ、H.264 (シングル NAL ユニットパケット) の場合は、0bit H.264 (FU-A) の場合は、2byte (16bit)
RTP ペイロード長	e [bit]	RTP パケット 1 個当りに設定可能なペ イロードのデータ長。RTP パケット化 した際の全データ長が MTU サイズ以 下になるような値でなければならない。 MPEG-4 ビデオ、H.264 (シングル NAL ユニットパケット) の場合は、パケッ ト化サイズと同値。
1 フレーム当りの平均ビット量	x / y [bit]	
1 フレーム当りの平均パケット数	((x / y) / z) * (z / e) [個]	小数点以下は、切り上げ。

I.IV 計算式

$$\begin{aligned}
 \text{b=行に指定するビットレート} &= \text{映像ストリームのビットレート} \\
 &+ ((\text{IP ヘッダ長} + \text{UDP ヘッダ長} + \text{RTP ヘッダ長} + \text{RTP ペイロードヘッダ長}) \\
 &\quad * \text{パケット数/秒}) \\
 &= \text{映像ストリームのビットレート} \\
 &+ ((\text{IP ヘッダ長} + \text{UDP ヘッダ長} + \text{RTP ヘッダ長} + \text{RTP ペイロードヘッダ長}) \\
 &\quad * (\text{1 フレーム当りの平均パケット数} * \text{映像ストリームのフレームレート})) \\
 &= x + (a + b + c + d) * ((x / y) / z * (z / e)) * y \\
 &= x + (a + b + c + d) * x / e
 \end{aligned}$$

ただし、計算結果の小数点以下は切り上げとする。

I.V 留意点

- ・ 上記の計算式は映像のバースト性（I フレームは符号量が多いなど）を考慮していないため、エンコーダの特性に合わせて、上記計算式で算出した値に、定数の加算や乗算などを行う必要がある。

I.VI 参考値

上記の計算式を用いた場合における各システムプロファイル利用時に b=行に指定する値の参考値を下表にまとめる。

なお、転送プロトコルは、IPv4 とする。

付表 I-2/JJ-40.30 プロファイルと b=行の値の関係

システムプロファイル	パラメータ	b=行の値
AVSIP-1a (モバイル)	MPEG-4 ビデオ Simple Profile Level 0 QCIF / 15fps / 43kbit/s パケット化サイズ=180byte	48
AVSIP-1.5a (MPEG-4 QVGA)	MPEG-4 ビデオ Simple Profile Level 3 QVGA / 15fps / 365kbit/s パケット化サイズ=1200byte	384
AVSIP-2a (MPEG-4 SD)	MPEG-4 ビデオ Simple Profile Level 4a VGA / 30fps / 1932kbit/s パケット化サイズ=1200byte	2000
AVSIP-2b (H.264 SD)	H.264 Constrained Baseline profile 3.0 4SIF / 30fps / 1932kbit/s パケット化サイズ=1200byte 分割パケットは利用しない	2000
AVSIP-3a , AVSIP-3b (H.264 720p)	H.264 Constrained Baseline profile 3.1 720p / 30fps / 4834kbit/s パケット化サイズ=1200byte 分割パケットは利用しない	5000
AVSIP-4b-i (H.264 1080i)	H.264 High profile 4.0 1080i / 30fps(60field/s) / 9657kbit/s パケット化サイズ=1200byte 分割パケットは利用する	10000
AVSIP-4b-p (H.264 1080p)	H.264 Constrained High profile 4.0 1080p / 30fps/ 9657kbit/s パケット化サイズ=1200byte 分割パケットは利用する	10000

付録II 3D 映像通信の使用例

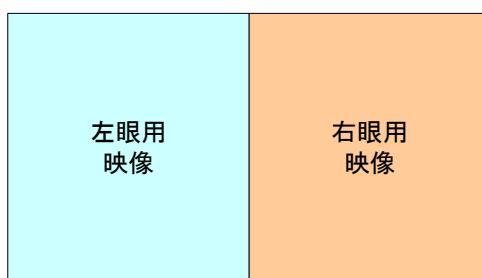
(本付録は参考資料であり、仕様ではない)

II.I はじめに

本付録では、3D 映像を送信する場合の使用例を示し、運用で注意すべき点について説明する。

II.II 3D 映像フォーマット

3D 映像フォーマットは、6.2 節に記載しているように Side-by-Side である。詳細フォーマットは 6.2.1 節所要特性の H.264 の Frame packing arrangement SEI 規定に記載があり、付図 II-1 に示すように、右半分に右眼用映像、左半分に左眼用映像が格納されたフレーム構造をしている。



付図 II-1/JJ-40.30 Side-by-Side 映像のフレーム構造

II.III 3D 映像のディスプレイ表示

II.III.1 一体型でない 3D 映像受信端末の 3D 映像対応ディスプレイでの 3D 映像/2D 映像自動表示

一体型でない 3D 映像受信端末の場合、3D 映像に対応したディスプレイであっても、ディスプレイに Side-by-Side の 3D 映像であることを通知しなければディスプレイが 2D 映像と解釈し、付図 II-1 に示すフレームのままディスプレイに表示される。3D 対応の HDMI 規格を用いてディスプレイと接続する場合には、Side-by-Side の 3D 映像であることを 3D 映像受信端末からディスプレイに通知できるので、3D 映像/2D 映像の自動表示切り替えが実現できる。

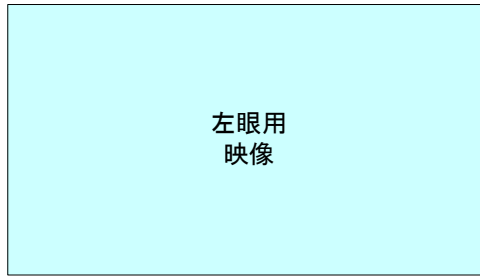
また、3D 対応の HDMI 規格を備えないディスプレイでも、受信側ユーザによるディスプレイ側の操作で 2D 表示と Side-by-Side の 3D 映像表示を切り替えることが可能である。

II.III.2 3D 映像対応ディスプレイでの 3D 映像の 2D 表示

3D 映像対応のディスプレイとしては、眼鏡式ディスプレイと、裸眼ディスプレイの 2 通りがある。後者の場合は常に 3D 映像で表示することも可能だが、前者の場合は眼鏡を着用していない等、場合によっては 3D 映像を 2D 映像としてディスプレイに表示したい場合がある。

そこで、3D 映像受信端末では、受信ユーザの操作によって、Side-by-Side の左眼用映像を付図 II-2 に示すようにフレーム全体に拡大した 2D 映像としてディスプレイに出力する。

また、一体型でない 3D 映像受信端末の 3D ディスプレイの場合は、受信端末の操作ではなく、Side-by-Side 映像のまま 3D ディスプレイに出力し、ディスプレイ側の操作で左眼用映像を付図 II-2 に示すようにフレーム全体に拡大表示することでも 2D 映像表示が可能である。



付図 II-2/JJ-40.30 Side-by-Side 映像を 2D 表示する場合の代表的なフレーム構造

II.III.3 一体型でない 3D 映像受信端末の 3D 映像非対応ディスプレイへの 3D 映像表示

一体型でない 3D 映像受信端末で、3D 映像受信端末に 3D 非対応のディスプレイが接続されうる場合には、3D 映像対応受信機で Side-by-Side の左眼用映像をフレーム全体に拡大した 2D 映像としてディスプレイに出力する機能が必要である。前記機能が無い場合、ディスプレイには付図 II-1 のように表示されることに注意が必要である。

II.IV 3D 映像の送受信例

3D 映像を送受信する場合に、送信端末および受信端末が 3D 映像に非対応の場合がある。また、SIP では 3D 映像対応の有無を調停できない。そこで、各ユースケースについて、送受信の方法について運用の注意事項などを例示する。

II.IV.1 送信端末と受信端末が 3D 映像対応の場合

2D 映像を送信する場合、2D 映像の H.264 ストリームには Frame packing arrangement SEI が含まれないため、3D 映像受信端末は 2D 映像を受信していると判断する。従って、3D 映像受信端末は自動的に 2D 映像を正しく受信し表示できる。

3D 映像を送信する場合、3D 映像非対応受信端末では H.264 ストリームの Frame packing arrangement SEI を解釈し、付図 II-1 の Side-by-Side フレーム映像であることを認識できる。従って、3D 映像受信端末は自動的に 3D 映像を正しく受信し表示できる。なお、一体型でない 3D 映像受信端末で 3D 対応の HDMI 規格を用いてディスプレイと接続する場合には、接続されたディスプレイが 3D 映像対応か非対応かを自動判断し、II.III.1 節、II.III.2 節に記載の方法でディスプレイに適切な表示ができる。

II.IV.2 送信端末が 3D 映像対応で受信端末が 3D 映像非対応の場合

2D 映像を送信する場合、受信側は 2D 映像を正しく受信し表示できる。

しかしながら、3D 映像を送信する場合、3D 映像非対応受信端末では H.264 ストリームの Frame packing arrangement SEI を解釈できないため、付図 II-1 の Side-by-Side フレーム映像をそのままディスプレイに表示してしまう。そこで、受信端末ユーザから音声チャネルを使った口頭指示などの方法で送信端末ユーザに連絡し、送信端末のユーザ操作で 3D 映像ではなく 2D 映像を送信するように切り替えてもらうことで、3D 映像非対応受信端末で 2D 映像として正しく表示できる。

II.IV.3 送信端末が 3D 非対応で受信端末が 3D 映像対応の場合

送信端末から常に 2D 映像が送信される。2D 映像の H.264 ストリームには Frame packing arrangement SEI が含まれないため、3D 映像受信端末は 2D 映像を受信していると判断する。従って、3D 映像受信端末は 2D 映像を自動的に正しく受信し表示できる。

II.IV.4 送信端末と受信端末が 3D 映像非対応の場合

送信端末から常に 2D 映像が送信される。従って、3D 映像非対応受信端末は 2D 映像を自動的に正しく受信し表示できる。