|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Question(s):** | 5/16 | | **Meeting, date:** | | Geneva, 12 - 23 October 2015 | |
| **Study Group:** | 16 | **Working Party:** | 1 | **Intended type of document** (R-C-TD): | | TD |
| **Source:** | Editor H.TPS-AV | | | | | |
| **Title:** | H.TPS-AV “Audio/video parameters for telepresence systems” (New): Output draft Ed. 1.17 (Geneva, 12 - 23 October 2015) | | | | | |
| **Contact:** | Stephen Botzko  Polycom  USA | | | | Tel: +1 (978) 292-5395  Fax:  Email: [[Stephen.Botzko@polycom.com](mailto:Stephen.Botzko@polycom.com)](mailto:Stephen.Botzko@polycom.com) | |
|  | | | | | | |

Summary

This document contains the Editor's output draft (H.TPS-AV Ed. 1.17) for H.TPS-AV "Audio/video parameters for telepresence systems" (New). It incorporates [C.922](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-0922), [C.923](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-0923), [C.949](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-0949), [C.997R1](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-0997), and [C.1000R1](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-1000). It is proposed as output text for the Geneva, October 2015 meeting.

Document History

| Document History | |
| --- | --- |
| **Issue** | **Notes** |
| H.TPS-AV Ed. 0.1 | Initial baseline text. Output from Boston meeting July 2011. Based on AVD-4092, 4075, 4106. |
| H.TPS-AV Ed. 0.2 | Input text to the November/December ITU-T SG16 meeting. It is a copy of the output of the Boston meeting with changes marks accepted. |
| H.TPS-AV Ed. 0.3 | Output text from the November/December ITU-T SG16 meeting. It incorporates: [C.694](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T09-SG16-C-0694), [C.747](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T09-SG16-C-0747), [C.748](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T09-SG16-C-0748), [C.749](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T09-SG16-C-0749), [C.709](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T09-SG16-C-0709) and [C.762](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T09-SG16-C-0762). |
| H.TPS-AV Ed. 0.4 | Input text to the Geneva, Switzerland meeting February 2012. It is a copy of [TD 0645R1/WP2](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T09-SG16-111121-TD-WP2-0645) from the Geneva, Switzerland 21 November – 02 December 2011 meeting with the change marks accepted. |
| H.TPS-AV Ed. 0.5 | Output text from the Geneva meeting February 2012. It incorporates: [AVD-4215](http://wftp3.itu.int/av-arch/avc-site/2009-2012/1202_Gen/AVD-4215.zip), [AVD-4216](http://wftp3.itu.int/av-arch/avc-site/2009-2012/1202_Gen/AVD-4216.zip), [AVD-4217](http://wftp3.itu.int/av-arch/avc-site/2009-2012/1202_Gen/AVD-4217.zip), [AVD-4218](http://wftp3.itu.int/av-arch/avc-site/2009-2012/1202_Gen/AVD-4218.zip), and [AVD-4219](http://wftp3.itu.int/av-arch/avc-site/2009-2012/1202_Gen/AVD-4219.zip). |
| H.TPS-AV Ed. 0.6 | Input text to the SG 16 meeting April 2012. It is a copy of TD-15 from the Geneva, February 2012 meeting with the change marks accepted and editorial cleanup. TD 768/WP2. |
| H.TPS-AV Ed. 0.7 | Output Text from the Geneva Meeting April 2012. It incorporates: [C.899](file:///S:\Q5\Inputs\T09-SG16-C-0899!!MSW-E.doc), [[C.931](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T09-SG16-C-0748)](file:///S:\Q5\Inputs\T09-SG16-C-0931!!MSW-E.docx), [C.940](file:///S:\Q5\Inputs\T09-SG16-C-0940!!MSW-E.doc) and agreed Requirements Terminology. |
| H.TPS-AV Ed. 0.8 | Input text to the Brisbane, Australia meeting September 2012. It is a copy of [TD 0768R1/WP2](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T09-SG16-111121-TD-WP2-0768) from the Geneva, Switzerland 30 April – 11 May 2012 meeting with the change marks accepted. |
| H.TPS-AV Ed. 0.9 | Output text from the Brisbane meeting September 2012. It incorporates: [AVD-4287](http://ftp3.itu.int/av-arch/avc-site/2009-2012/1209_Bri/AVD-4287.zip), [AVD-4289](http://ftp3.itu.int/av-arch/avc-site/2009-2012/1209_Bri/AVD-4289.zip), [AVD-4290](http://ftp3.itu.int/av-arch/avc-site/2009-2012/1209_Bri/AVD-4290.zip), and [AVD-4334](http://ftp3.itu.int/av-arch/avc-site/2009-2012/1209_Bri/AVD-4334.zip) |
| H.TPS-AV Ed. 1.0 | Input text to the Geneva, Switzerland meeting January 2013. It is a copy of [TD-39](http://ftp3.itu.int/av-arch/avc-site/2009-2012/1209_Bri/TD-39.zip) from the Brisbane, Australia September 2012 meeting with the change marks accepted. |
| H.TPS-AV Ed. 1.1 | Output text from the Geneva, Switzerland meeting January 2013. It incorporates [C.90](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-0090), [C.140](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-0140), [C.141](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-0141), [C.142](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-0142), and [TD 45/WP1](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-130114-TD-WP1-0045) |
| H.TPS-AV Ed. 1.2 | Input text to the Oslo, Norway meeting June 2013. It is a copy of [TD-31](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-130114-TD-WP1-0031) from the Geneva, Switzerland January 2013 meeting with the change marks accepted. |
| H.TPS-AV Ed. 1.3 | Output text from the Oslo, Norway meeting June 2013. It incorporates [AVD-4440](http://ftp3.itu.int/av-arch/avc-site/2013-2016/1306_Osl/AVD-4440.zip), [AVD-4441](http://ftp3.itu.int/av-arch/avc-site/2013-2016/1306_Osl/AVD-4441.zip), [AVD-4443](http://ftp3.itu.int/av-arch/avc-site/2013-2016/1306_Osl/AVD-4443.zip), [AVD-4445](http://ftp3.itu.int/av-arch/avc-site/2013-2016/1306_Osl/AVD-4445.zip), [AVD-4461](http://ftp3.itu.int/av-arch/avc-site/2013-2016/1306_Osl/AVD-4461.zip), [AVD-4467](file:///C:\ITU\1306-Oslo\Q5\AVD-4467.docx), and [AVD-4465](http://ftp3.itu.int/av-arch/avc-site/2013-2016/1306_Osl/AVD-4465.zip) |
| H.TPS-AV Ed. 1.4 | Input text to the Geneva Meeting, October 2013. It is a copy of [TD-25a](http://ftp3.itu.int/av-arch/avc-site/2013-2016/1306_Osl/TD-25a.zip) from the Oslo, Norway June 2013 meeting with the change marks accepted and editorial cleanup. |
| H.TPS-AV Ed. 1.5 | Output text from the Geneva Meeting October 2013. It incorporates [C.298](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-0298), [C.299](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-0299), [C.300](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-0300), [C.379](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-0379), and editorial cleanup. |
| H.TPS-AV Ed. 1.6 | Input text to the Geneva Meeting, March 2014. It is a copy of [TD-89](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-131028-TD-WP1-0089) from the Geneva, October 2013 meeting with the change marks accepted and editorial cleanup. |
| H.TPS-AV Ed. 1.7 | Output text from the Geneva Meeting, March 2014. It incorporates [AVD-4543](http://ftp3.itu.int/av-arch/avc-site/2013-2016/1403_Gen/AVD-4543.zip), [AVD-4575](http://ftp3.itu.int/av-arch/avc-site/2013-2016/1403_Gen/AVD-4575.zip), [AVD-4529](http://ftp3.itu.int/av-arch/avc-site/2013-2016/1403_Gen/AVD-4529.zip), [AVD-4530](http://ftp3.itu.int/av-arch/avc-site/2013-2016/1403_Gen/AVD-4530.zip) and [AVD-4531](http://ftp3.itu.int/av-arch/avc-site/2013-2016/1403_Gen/AVD-4531.zip). |
| H.TPS-AV Ed. 1.8 | Input text to the Sapporo Meeting, June 2014. It is a copy of [TD-26](http://ftp3.itu.int/av-arch/avc-site/2013-2016/1403_Gen/TD-26.zip) from the Geneva, March 2014 meeting with the change marks accepted and editorial cleanup. |
| H.TPS-AV Ed. 1.9 | Output Text from the Sapporo Meeting, June 2014. It incorporates [C.509](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-0509), [C.510](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-0510), [C.511](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-0511), [C.513](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-0513), [C.628](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-0628), [C.629](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-0629), and [C.630](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-0630) |
| H.TPS-AV Ed. 1.10 | Input text to the Seoul, Korea meeting, November 2014. It is a copy of [TD-175](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-140630-TD-WP1-0175) from the Sapporo, June 2014 Meeting with the change marks accepted and editorial cleanup |
| H.TPS-AV Ed. 1.11 | Output text from the Seoul, Korea Meeting, November 2014. It incorporates: [AVD-4627](http://ftp3.itu.int/av-arch/avc-site/2013-2016/1411_Seo/AVD-4627.zip), [AVD-4628](http://ftp3.itu.int/av-arch/avc-site/2013-2016/1411_Seo/AVD-4628.zip), [AVD-4630](http://ftp3.itu.int/av-arch/avc-site/2013-2016/1411_Seo/AVD-4630.zip), [AVD-4631](http://ftp3.itu.int/av-arch/avc-site/2013-2016/1411_Seo/AVD-4631.zip), [AVD-4657](http://ftp3.itu.int/av-arch/avc-site/2013-2016/1411_Seo/AVD-4657.zip), and [AVD-4658](http://ftp3.itu.int/av-arch/avc-site/2013-2016/1411_Seo/AVD-4658.zip). |
| H.TPS-AV Ed. 1.12 | Input text to the Geneva Meeting, February 2015. It is a copy of [TD-30](http://ftp3.itu.int/av-arch/avc-site/2013-2016/1411_Seo/TD-30.zip) with the change marks accepted and editorial clean-up. |
| H.TPS-AV Ed. 1.13 | Output text from the Geneva Meeting, February 2015. It incorporates [C.739](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-0739), [C.740](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-0740), [C.741](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-0741), [C.740](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-0740), [C.749](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-0749), and [C.781](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-C-0781) |
| H.TPS-AV Ed. 1.14 | Input text to the Chengdu Meeting, June 2015. It is a copy of [T.247 R1](http://www.itu.int/md/meetingdoc.asp?lang=en&parent=T13-SG16-150209-TD-WP1-0247) with the change marks accepted and editorial clean-up |
| H.TPS-AV Ed. 1.15 | Output text from the Chengdu Meeting, June 2015. It incorporates [AVD-4709](http://ftp3.itu.int/av-arch/avc-site/2013-2016/1506_Che/AVD-4709.zip), [AVD-4710](http://ftp3.itu.int/av-arch/avc-site/2013-2016/1506_Che/AVD-4710.zip), [AVD-4770](http://ftp3.itu.int/av-arch/avc-site/2013-2016/1506_Che/AVD-4770.zip), and [AVD-4772](http://ftp3.itu.int/av-arch/avc-site/2013-2016/1506_Che/AVD-4772.zip). |
| H.TPS-AV Ed. 1.16 | Input text to the Geneva Meeting, October 2015. It is a copy of [TD-25a](http://ftp3.itu.int/av-arch/avc-site/2013-2016/1506_Che/TD-25a.zip) with the change marks accepted and editorial clean-up. |
| H.TPS-AV Ed. 1.17 | Output text from the Geneva meeting October 2015. |

Living List Items:

1. All parameters that correspond to IETF CLUE information elements need to be periodically checked for accuracy and updated as needed.

**CONTENTS**

| ` |
| --- |
| [1 Scope 5](#_Toc430907026)  [2 References 5](#_Toc430907027)  [3 Definitions 7](#_Toc430907028)  [3.1 Terms defined elsewhere 7](#_Toc430907029)  [3.2 Terms defined in this Recommendation 7](#_Toc430907030)  [4 Abbreviations and acronyms 8](#_Toc430907031)  [5 Conventions 8](#_Toc430907032)  [5.1 Requirement terminology 8](#_Toc430907033)  [6 H.323 Telepresence Parameter Usage 8](#_Toc430907034)  [7 Telepresence endpoint parameters 10](#_Toc430907035)  [7.1 Capture Related Parameters 10](#_Toc430907036)  [7.1.1 General Parameters 10](#_Toc430907037)  [7.1.2 Visual parameters 17](#_Toc430907038)  [7.1.3 Audio parameters 25](#_Toc430907039)  [7.1.4 Delay parameters 29](#_Toc430907040)  [7.1.5 Multiple Source Capture Parameters 30](#_Toc430907041)  [7.2 Render related Parameters 32](#_Toc430907042)  [7.2.1 Video parameters 32](#_Toc430907043)  [7.2.2 Audio parameters 32](#_Toc430907044)  [7.3 Telepresence system environment parameters 33](#_Toc430907045)  [7.3.1 Colour and lighting parameters 33](#_Toc430907046)  [7.3.2 Acoustic Parameters 35](#_Toc430907047)  [Appendix I Parameter signalling analysis 38](#_Toc430907048)  [I.1 Capture Related Parameters 38](#_Toc430907049)  [I.1.1 General Parameters 38](#_Toc430907050)  [I.1.2 Visual parameters 39](#_Toc430907051)  [I.1.3 Audio parameters 39](#_Toc430907052)  [I.1.4 Delay parameters 40](#_Toc430907053)  [I.1.4 Multiple Source Capture Parameters 40](#_Toc430907054)  [I.2 Render related Parameters 41](#_Toc430907055)  [I.2.1 Video parameters 41](#_Toc430907056)  [I.2.2 Audio parameters 41](#_Toc430907057)  [I.3 Telepresence system environment parameters 41](#_Toc430907058)  [I.3.1 Colour and lighting parameters 41](#_Toc430907059)  [I.3.2 Acoustic Parameters 41](#_Toc430907060)  [Bibliography 41](#_Toc430907061) |

**List of Tables**

| Page |
| --- |
| **No table of figures entries found.** |

**List of Figures**

| Page |
| --- |
| [Figure 1: Capture area axis (top view) 21](#_Toc411975382)  [Figure 2: Capture area axis (3D view) 22](#_Toc411975383) |

Draft new ITU-T H.??? (ex H.TPS-AV)

Audio/video parameters for telepresence systems

AAP Summary

Telepresence is an interactive audio-visual communications experience between remote locations, where the users enjoy a strong sense of realism and presence between participants by optimizing a variety of attributes. Rec. H.??? (H.TPS-AV) defines the audio/video/environment endpoint parameters for telepresence systems, including those which are used in the media setup phase during capability negotiation, and gives the specific format definitions for these parameters as well. Collectively, these parameters and their associated values can be expected to provide a high quality telepresence experience.

Summary

This Recommendation defines the audio/video/environment endpoint parameters for telepresence systems, including those which are used in the media setup phase during capability negotiation, and gives the specific format definitions for these parameters as well. Collectively, these parameters and their associated values can be expected to provide a high quality telepresence experience. As far as possible, the parameters have been aligned with those specified by the IETF CLUE Working Group [IETF CLUE FW]. The Recommendation provides guidance on the need for signalling these parameters at session initiation and during a session.

# Scope

In a telepresence system, multiple media streams need to be captured, encoded, transported, decoded and rendered. Telepresence systems made by different vendors may have diverse solutions or implementations for similar purposes. These solutions or implementations may use disparate techniques to negotiate, describe and control media streams and cause interoperability problems. Some problems may be solved through translation or transcoding, but this requires additional devices and would introduce unnecessary delay and deteriorated performance.

In order to prevent the above problems the capabilities of the telepresence session can:

1. Be negotiated between the telepresence endpoints (or other intermediaries as required) before media transmission in order to agree on a common set of capabilities.
2. Assume default capabilities based on support of a common profile.
3. A combination of the above.

For example, a receiver could send its audio/video capabilities parameters to the sender through the capability negotiation. The sender would coordinate and send media streams to the receiver according to the receiver's capabilities.

This Recommendation provides details of telepresence endpoint characteristics and how these relate to telepresence in [ITU-T H.323] systems in terms of a number of parameters. Collectively, these parameters and their associated values can be expected to provide a high level of QoS/QoE.

# References

The following ITU-T Recommendations and other references contain provisions, which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published.

The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-R BT.1361] ITU-R Recommendation BT.1361 (02/1998), *Worldwide unified colorimetry and related characteristics of future television and imaging systems*.

[ITU-R BT.2020] ITU-R Recommendation BT.2020 (08/2012), *Parameter values for ultra-high definition television systems for production and international programme exchange*.

[ITU-R BT.601] ITU-R Recommendation BT.601-7 (03/2011), *Studio encoding parameters of digital television for standard 4:3 and wide screen 16:9 aspect ratios.*

[ITU-R BT.709] ITU-R Recommendation BT.709-5 (04/2002), *Parameter values for the HDTV standards for production and international programme exchange*.

[ITU-T G.122] ITU-T Recommendation G.122 (03/1993), *Influence of national systems on stability and talker echo in international connections.*

[ITU-T G.191] ITU-T Recommendation G.191 (03/2010), *Software tools for speech and audio coding standardization.*

[ITU-T H.241] ITU-T Recommendation H.241 (02/2012), *Extended video procedures and control signals for ITU-T H.300-series terminals*.

[ITU-T H.245] ITU-T Recommendation H.245 (05/2011), *Control protocol for multimedia communication*.

[ITU-T H.264] ITU-T Recommendation H.264 (04/2013), *Advanced video coding for generic audiovisual services*.

[ITU-T H.265] ITU-T Recommendation H.265 (04/2013), *High efficiency video coding*.

[ITU-T H.323] ITU-T Recommendation H.323 (12/2009), *Packet-based multimedia communications systems*.

[ITU-T P.56] ITU-T Recommendation P.56 (12/2011), *Objective measurement of active speech level*.

[ITU-T P.300] ITU-T Recommendation P.300 (11/2001), *Transmission performance of group audio terminals (GATs)*.

[ITU-T P.501] ITU-T Recommendation P.501 (01/2012), *Test signals for use in telephonometry.*

[ITU-T Q.931] ITU-T Recommendation Q.931 (05/1998), *ISDN user-network interface layer 3 specification for basic call control*.

[IETF CLUE SIG] IETF draft-ietf-clue-signaling-06, *CLUE Signaling*.

[IETF CLUE DM] IETF draft-ietf-clue-data-model-schema-10, *An XML schema for the CLUE Data Model.*

[IETF CLUE FW] IETF draft-ietf-clue-framework-23, *Framework for Telepresence Multi-Streams*.

[IETF CLUE PROT] IETF [draft-ietf-clue-protocol-04](https://datatracker.ietf.org/doc/draft-ietf-clue-protocol/), *CLUE protocol*.

[IETF RFC 4566] IETF RFC 4566 (2006), *SDP: Session Description Protocol*.

[IETF RFC 4796] IETF RFC 4796 (2007), *The Session Description Protocol (SDP) Content Attribute*.

[IETF RFC 5646] IETF RFC 5646 (2009), *Tags for Identifying Languages*.

[IETF RFC 6184] IETF RFC 6184 (2011), *RTP Payload Format for H.264 Video*.

[IETF RFC 6236] IETF RFC 6236 (2011), *Negotiation of Generic Image Attributes in the Session Description Protocol (SDP)*.

[IETF RFC 6351] IETF RFC 6351 (2009), *xCard: vCard XML Representation*.

[IETF RFC 6464] IETF RFC 6464 (2011), *A Real-time Transport Protocol (RTP) Header Extension for Client-to-Mixer Audio Level Indication*.

[ISO 11664-2] ISO 11664-2 (2007) / CIE S 014-2/E (2006), *Colorimetry - Part 2: CIE standard illuminants*.

[ISO 140-4] ISO 140-4 (1998), *Acoustics - Measurement of sound insulation in buildings and of building elements - Part 4: Field measurements of airborne sound insulation between rooms*.

[ISO 717-1] ISO 717-1 (1996), *Acoustics – Rating of sound insulation in buildings and of building elements – Part 1: Airborne Sound Insulation*.

[ISO 8995] ISO 8995 (2002), *Lighting of indoor work places*.

[P.Supp 16] ITU-T Recommendation Series P Supplement 16 (11/1988), *Guidelines for placement of microphones and loudspeakers in telephone conference rooms and for group audio terminals (GATs)*.

# Definitions

## Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

**3.1.1 camera-left and right** [IETF CLUE FW]: For media captures, camera-left and camera- right are from the point of view of a person observing the rendered media. They are the opposite of stage-left and stage-right.

## Terms defined in this Recommendation

This Recommendation defines the following terms:

**3.2.1** **capability**: An endpoint has a particular capability if it is able to support functionality related to a particular telepresence parameter. This may involve capturing, encoding, transmitting, receiving, decoding, or rendering of a signal related to the parameter.

**3.2.2 chroma:** An adjective specifying that a sample in a digital picture is representing one of the two colour difference signals related to the primary colours.

NOTE – The term chroma is used rather than the term chrominance in order to avoid the implication of the use of linear light transfer characteristics that is often associated with the term chrominance.

**3.2.3 consumer:** An endpoint that receives media streams.

**3.2.4 cycles per degree:** The number of alternating white and black lines that can be captured or rendered per degree of visual field.

**3.2.5 endpoint:** A terminal, Gateway, or MCU. An endpoint can call and be called. It generates and/or terminates information streams.

**3.2.6 luma:** An adjective specifying that a sample in a digital picture is representing the monochrome signal related to the primary colours.

NOTE – The term luma is used rather than the term luminance in order to avoid the implication of the use of linear light transfer characteristics that is often associated with the term luminance.

**3.2.7 provider:** An endpointthat sends media streams.

# Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

|  |  |
| --- | --- |
| 3D | Three dimensional |
| AD | Analogue to Digital |
| CCT | Correlated colour temperature |
| CLUE | Signalling Protocol defined in [IETF CLUE SIG] |
| CPD | Cycles Per Degree |
| CRI | Colour Rendering Index |
| DA | Digital to Analogue |
|  |  |
| MCC | Multiple Content Capture |
| MCU | Multipoint Control Unit |
| PCM | Pulse Code Modulation |
| QoE | Quality of Experience |
| QoS | Quality of Service |
|  |  |
| RTCP | RTP Control Protocol |
| RTP | Real Time Protocol |
| R'w | Weighted apparent sound Reduction index |
| Rw | Weighted sound Reduction index |
| SDP | Session Description Protocol |
| SEI | Supplemental Enhancement Information |
| STC | Sound Transmission Class |
| VUI | Video Usability Information |

# Conventions

## Requirement terminology

In this Recommendation the following conventions are used:

* "Shall" indicates a mandatory requirement.
* "Should" indicates a suggested but optional course of action.
* "May" indicates an optional course of action rather than a recommendation that something take place.

# H.323 Telepresence parameter usage

As shown in clause 9 / [ITU-T H.420], a H.323 based telepresence endpoint may communicate with another endpoint using a "peer to peer" or "multi-point communication" mode. In either of these modes the establishment of a telepresence session follows three main phases:

* Call setup phase: including RAS and [ITU-T Q.931] call signalling.
* Media setup phase: including master-slave determination, capability negotiation and "logical channel" opening through [ITU-T H.245] control signalling.
* Media control phase: including media stream transmission and control using RTP/RTCP and H.245 control signalling.

This Recommendation details parameters that are used to describe telepresence endpoint characteristics, including those which are used in the media setup phase during capability negotiation. Typically, the provider advertises captures with its characteristics and capabilities (what it can provide), e.g. some or all of those listed above. The consumer selects captures with characteristics and capabilities (only those it wants to receive). The provider can then set up media transportation channels based on the consumer's choice. The characteristics may be explicitly signalled or a default (not requiring signalling) may be assumed based on the support of this Recommendation.

A telepresence endpoint may have different capabilities related to audio, visual and environmental characteristics. For example these characteristics may include:

* The number of cameras, spatial relationships of cameras and capture capabilities, such as resolution, frame rate and aspect ratio, etc.
* Viewpoint and field of view for cameras include overlap regions, missing regions or whether images of cameras can be composed to a seamless image in whole depth ranges.
* The number, type, size of displays, spatial relationships of displays and display capabilities, such as resolution, frame rate, aspect ratio, and display border width, etc.
* The number and type of presentation displays, spatial relationships of presentation displays and display capabilities.
* Spatial relationships of microphones and parameters of microphones.
* Spatial relationships of speakers and parameters of speakers.
* System environment parameters, such as colour temperature, etc.

As shown in Figure 10 / [ITU-T H.420] there are a number of interfaces (A1, G1 and G2) where negotiation related to media associated with a telepresence end point may take place. Thus parameters such as those shown in the examples above may need to be transported across these interfaces.

In order that each telepresence endpoint (or intermediate device) can understand what a parameter means it needs to be described in detail. Section 7 below provides a list of parameters related to telepresence endpoints.

By utilising these parameters, the characteristics of the streams may be identified and actions performed for different applications. For example, spatial information for the captured audio and video media can be used to determine the best method for preserving the spatial relationships when rendered on the room equipment.

Multiple parameters can be used together to determine the importance of the streams. Rendering and distribution decisions may be based on such combinations of parameters. For example, in congestion control less important streams may simply be dropped, or alternative captures that use lower resolutions or that use less bandwidth could be substituted.

NOTE: Clause 9/[IETF CLUE PROT] defines CLUE extension procedures that could be used for carriage of any future parameters specified in this recommendation via CLUE.

# Telepresence endpoint parameters

This clause indicates describes the parameters related to a telepresence endpoint in more detail. The parameters are described according to the format below:

|  |  |
| --- | --- |
| **Identity:** | *A unique alphanumeric identifier in this Recommendation.* |
| **Description:** | *A description of the function of the parameter.* |
| **Format:** | *The type of parameter, i.e. Boolean, alpha numeric string, integer, sequence of octets, character or enumeration.* |
| **Possible Values:** | *The parameter must specify which values can be transported.* |
| **Default:** | *Indicate if there is a default that may be assumed.* |
| **Reference:** | *Where a parameter has previously been defined a reference to it may be included.* |
| **Signalling Mechanism:** | *Method(s) for signalling the parameter.* |
| **Supplementary Information:** | *Further information to the use or definition of the parameter.* |

## Capture related parameters

This clause describes parameters related to the capture of media.

### General parameters

#### Captured media type

|  |  |
| --- | --- |
| **Identity:** | mediaType |
| **Description:** | This parameter is a mandatory field specifying the media type of the capture ("audio", "video", "text"...). |
| **Format:** | Enumeration. |
| **Possible Values:** | Audio, Video, Text… |
| **Default:** | None. |
| **Reference:** | See the 'mediaType' attributes in [IETF CLUE DM] for further information. |
| **Signalling Mechanism:** | CLUE |
| **Supplementary Information:** | - |

#### Scene description

|  |  |
| --- | --- |
| **Identity:** | captureScene description |
| **Description:** | This parameter provides a textual description of the overall scene. Multiple languages may be included in the description. |
| **Format:** | List of Strings. |
| **Possible Values:** | A list of <description text> and <language tag> may be provided. |
| **Default:** | None. |
| **Reference:** | See the "Description" attribute in 7.3.1 of [IETF CLUE FW] for further information. |
| **Signalling Mechanism:** | CLUE |
| **Supplementary Information:** | Describing the scene using text in multiple languages may assist in the consumer being able to make a more informed decision about which scenes to choose in the event that multiple scenes are offered. |

#### Capture scene view description

|  |  |
| --- | --- |
| **Identity:** | sceneView description |
| **Description:** | This parameter provides a textual description of the captures that represent a scene. Multiple sets of one or more captures may represent a scene. This description provides text associated with each set. Multiple languages may be included in the description. |
| **Format:** | List of Strings. |
| **Possible Values:** | A list of <description text> and <language tag> may be provided. |
| **Default:** | None. |
| **Reference:** | See the "Description" attribute in 7.3.2 of [IETF CLUE FW] for further information. |
| **Signalling Mechanism:** | CLUE |
| **Supplementary Information:** | Describing the set of captures representing the scene using text in multiple languages may assist in the consumer being able to make a more informed decision about which set of captures to choose in the event that multiple set of captures are offered. |

#### Language

|  |  |
| --- | --- |
| **Identity:** | lang |
| **Description:** | This attribute indicates one or more languages which are used in the content of the media capture. Captures may be offered in different languages in case of multilingual and/or accessible conferences. A Consumer may use this parameter to differentiate between them. |
| **Format:** | Enumeration |
| **Possible Values:** | The values of the 'Subtag' column for the "Type: language" entries in the "Language Subtag Registry" defined in [IETF RFC 5646] |
| **Default:** | None. |
| **Reference:** | See "Language" in [IETF CLUE FW] for further information. |
| **Signalling Mechanism:** | CLUE |
| **Supplementary Information:** | - |

#### Priority

|  |  |
| --- | --- |
| **Identity:** | priority |
| **Description:** | This parameter indicates a relative priority between different Media Captures. The Provider sets this priority, and the Consumer may use the priority to help decide which captures it wishes to receive. |
| **Format:** | Integer |
| **Possible Values:** | 1 upwards. |
| **Default:** | None. |
| **Reference:** | See "Priority" in [IETF CLUE FW] for further information. |
| **Signalling Mechanism:** | CLUE |
| **Supplementary Information:** | - |

#### Embedded text

|  |  |
| --- | --- |
| **Identity:** | embeddedText |
| **Description:** | This parameter indicates that a capture provides embedded textual information. For example the video capture may contain speech to text information composed with the video image. This attribute is only applicable to video captures and presentation streams with visual information. |
| **Format:** | Boolean |
| **Possible Values:** | |  |  | | --- | --- | | True | The capture contains embedded text | | False | The capture does not contain embedded text | |
| **Default:** | False |
| **Reference:** | See "Embedded Text" in [IETF CLUE FW] for further information. |
| **Signalling Mechanism:** | CLUE |
| **Supplementary Information:** | - |

#### Related to

|  |  |
| --- | --- |
| **Identity:** | relatedTo |
| **Description:** | This attribute indicates the capture contains additional complementary information related to another capture. The value indicates the other capture to which this capture is providing additional information. |
| **Format:** | String |
| **Possible Values:** | A capture identity. |
| **Default:** | None. |
| **Reference:** | See "Related to" in [IETF CLUE FW] for further information. |
| **Signalling Mechanism:** | CLUE |

#### Presentation

|  |  |
| --- | --- |
| **Identity:** | Presentation |
| **Description:** | This parameter indicates that the capture originates from a presentation device, that is one that provides supplementary information to a conference through slides, video, still images, data etc. This parameter shall only be sent if the capture is a presentation. |
| **Format:** | Enumeration |
| **Possible Values:** | Presentation optionally with further description information (slides, image). |
| **Default:** | None. |
| **Reference:** | See "Presentation" in [IETF CLUE FW] for further information. |
| **Signalling Mechanism:** | CLUE |
| **Supplementary Information:** | This parameter enables a receiver to uniquely identify captures that contain presentation material. |

#### Person information

|  |  |
| --- | --- |
| **Identity:** | personInfo |
| **Description:** | The person information parameter allows an endpoint to provide specific information regarding the people in a Capture. The endpoint may gather the information automatically or manually from a variety of sources. The xCard [IETF RFC 6351] format is used to convey the information. |
| **Format:** | String |
| **Possible Values:** | As per clause 7.1.1.10/[IETF CLUE FW]. |
| **Default:** | None. |
| **Reference:** | See the 'Person Information' attribute in [IETF CLUE FW] for further information. |
| **Signalling Mechanism:** | CLUE |
| **Supplementary Information:** | - |

#### Person type

|  |  |
| --- | --- |
| **Identity:** | personType |
| **Description:** | The person type attribute indicates the type of person/people contained in the capture in the conference with respect to the meeting agenda. As a capture may include multiple people the attribute may contain multiple values. However values shall not be repeated within the attribute. |
| **Format:** | Sub-list of String |
| **Possible Values:** | As per clause 7.1.1.12/[IETF CLUE FW]. |
| **Default:** | None. |
| **Reference:** | See the 'People Type' attribute in [IETF CLUE FW] for further information. |
| **Signalling Mechanism:** | CLUE |
| **Supplementary Information:** | - |

#### Scene information

|  |  |
| --- | --- |
| **Identity:** | sceneInformation |
| **Description:** | The Scene information parameter provides information regarding the  Capture Scene rather than individual people in the scene. The Provider may gather the information automatically or manually from a variety of sources. The scene information attribute allows a Provider to indicate information such as: organizational or geographic information allowing a Consumer to determine which Capture Scenes are of interest in order to then perform Capture selection. It also allows a Consumer to render information regarding the Scene or to use it for further processing. The xCard format is used to convey this information |
| **Format:** | String |
| **Possible Values:** | As per clause 7.3.1.1/[IETF CLUE FW]. |
| **Default:** | None. |
| **Reference:** | See the 'Scene Information' attribute in [IETF CLUE FW] for further information. |
| **Signalling Mechanism:** | CLUE |
| **Supplementary Information:** | - |

#### General spatial information

The parameters in this clause relate to all types of media captures. It is assumed that all the spatial information parameters from the same endpoint share a common origin.

##### Capture description

|  |  |
| --- | --- |
| **Identity:** | mediaCapture description |
| **Description:** | This parameter provides a textual description of the capture. Multiple languages may be included in the description. |
| **Format:** | List of Strings. |
| **Possible Values:** | A list of <description text> and <language tag> may be provided. |
| **Default:** | None. |
| **Reference:** | See the "Description" attribute in clause 7.1.1.6/[IETF CLUE FW] for further information. |
| **Signalling Mechanism:** | CLUE |
| **Supplementary Information:** | Describing the capture using text in multiple languages may assist in the consumer being able to make a more informed decision about which captures to choose in the event that multiple captures are offered. |

##### Area scale

|  |  |
| --- | --- |
| **Identity:** | captureScene scale |
| **Description:** | This parameter indicates the type of scale that is being applied to the spatial information parameters |
| **Format:** | Enumeration |
| **Possible Values:** | "mm" The metric system with the millimetre unit is being applied.  "unknown" An unknown scale is being used that is the same for all captures in the scene. This may be used to indicate relative distances.  "No scale" This indicates that the scale may be different for each capture in a scene. |
| **Default:** | None. |
| **Reference:** | See "Scale" in clause 6/[IETF CLUE FW] for further information. |
| **Signalling Mechanism:** | CLUE |
| **Supplementary Information:** | - |

##### Mobility of capture

|  |  |
| --- | --- |
| **Identity:** | mediaCapture mobility |
| **Description:** | This parameter indicates whether or not the point of capture, line on point of capture, and area of capture values will stay the same, or are expected to change frequently. This allows a receiver to determine the optimal placement when rendering the capture. |
| **Format:** | Enumeration |
| **Possible Values:** | |  |  | | --- | --- | | Static | The point of capture, line on point of capture, and area of capture values will stay the same for the duration of a conference. | | Dynamic | The point of capture, line on point of capture, and area of capture values may change during a conference. | | HighlyDynamic | The point of capture, line on point of capture, and area of capture values are likely to change frequently during a conference. Periods of constant change may be expected. | |
| **Default:** | Static |
| **Reference:** | See "Mobility of Capture" in [IETF CLUE FW] for further information. |
| **Signalling Mechanism:** | CLUE |
| **Supplementary Information:** | - |

##### View

|  |  |
| --- | --- |
| **Identity:** | mediaCapture view |
| **Description:** | This parameter indicates what type of view the capture relates to. The Consumer can use this information to help choose which Media Captures it wishes to receive. |
| **Format:** | Enumeration |
| **Possible Values:** | |  |  | | --- | --- | | Room | Captures the entire scene | | Table | Captures the conference table with seated participants | | Individual | Captures an individual participant | | Lectern | Captures the region of the lectern including the presenter in a classroom style conference | | Audience | Captures a region showing the audience in a classroom style conference | |
| **Default:** | None. |
| **Reference:** | See "View" in [IETF CLUE FW] for further information. |
| **Signalling Mechanism:** | CLUE |
| **Supplementary Information:** | - |

#### General encoding information

##### Overall maximum bitrate

|  |  |
| --- | --- |
| **Identity:** | maxGroupBandwidth |
| **Description:** | This parameter indicates the maximum number of bits per second relating to all encoded media of a particular type (e.g. audio or video) sent from the endpoint. |
| **Format:** | Numeric |
| **Possible Values:** | 0 upwards. |
| **Default:** | None. |
| **Reference:** | See "maxGroupBandwidth" in [IETF CLUE FW] for further information. |
| **Signalling Mechanism:** | CLUE |
| **Supplementary Information:** | - |

##### Simulcast

|  |  |
| --- | --- |
| **Identity:** | simulcast |
| **Description:** | Telepresence systems may provide multiple encodings for the one capture through a technique known as simulcast. For example, this may be achieved by sending multiple video coding streams with different characteristics to allow a receiving endpoint to choose the stream that meets its needs. |
| **Format:** | - |
| **Possible Values:** | - |
| **Default:** | None. |
| **Reference:** | - |
| **Signalling Mechanism:** | Bearer / media control signalling |
| **Supplementary Information:** | [ITU-T H.323] allows defines the simultaneous capabilities mechanism that allows endpoints to indicate which capabilities they can simultaneously send. In addition the set submode procedure defined in [ITU-T H.241] clause 6.2.5 may be used to request an encoding that better meets its needs. |

### Visual parameters

This clause describes the telepresence parameters related to the visual capabilities.

#### Video colour gamut

|  |  |
| --- | --- |
| **Identity:** | colorGamut |
| **Description:** | This parameter indicates the Colour Gamut used in a Telepresence Video Stream. |
| **Format:** | String |
| **Possible Values:** | BT.709, BT.601, BT.1361, BT.2020 |
| **Default:** | BT.709. |
| **Reference:** | - |
| **Signalling Mechanism:** | H.264 and H.265 SEI |
| **Supplementary Information:** | In the absence of visual usability information (VUI) signalling (or if the VUI signalling indicates "unspecified"), telepresence systems shall process all received video signals per [ITU-R BT.709]. For example: Annex E.2 / [ITU-T H.264] regarding VUI has the "colour primaries" element that allows the signalling of the chromaticity and white point.  If the transmitted video is not conformant with [ITU-R BT.709], senders should always signal the black level and range of the luma and chroma signals, transfer characteristics, the colour primaries, and the matrix coefficients in their video bit streams. Receivers may use this information to optimally render the received pictures.  See clause 7.3.1.4 for more information on the use of colour and lighting parameters. |

#### Video LumaBitDepth

|  |  |
| --- | --- |
| **Identity:** | lumaBitDepth |
| **Description:** | This parameter indicates the bit depth of the luma samples in a digital picture. |
| **Format:** | Integer |
| **Possible Values:** | 8 through 14 |
| **Default:** | 8 |
| **Reference:** | - |
| **Signalling Mechanism:** | H.264 and H.265 SEI |

#### Video ChromaBitDepth

|  |  |
| --- | --- |
| **Identity:** | chromaBitDepth |
| **Description:** | This parameter indicates the bit depth of the chroma samples in a digital picture. |
| **Format:** | Integer |
| **Possible Values:** | 8 through 14 |
| **Default:** | 8 |
| **Reference:** | - |
| **Signalling Mechanism:** | H.264 and H.265 SEI |

#### Effective resolution

Human visual acuity is the ability to distinguish fine detail. It is generally measured in *cycles per degree* (CPD), and therefore measures an angular resolution. People with excellent visual acuity can resolve 50 CPD – 1.2 arc minute per line pair (or 0.35 mm line pair at 1 meter) [b-ImagePro]). Camera resolution is usually measured in *lines* across the entire horizontal or vertical field of view. Black and white lines are alternated. A *cycle* is equivalent to 2 lines of camera resolution – that is, a black+white line pair comprises a cycle. The field of view of a camera is measured in degrees. Therefore the *cycles per degree* for a camera can be computed as *lines*/(2\**field\_of\_view*).

Similarly, display resolution is measured in pixels (width by height). For modern flat panel displays, a *line* is equivalent to a row or column of pixels. The *cycles per degree* value for a display depends on the subtended angle of the display, which in turn depends on the size of the display and the distance of the display surface from the viewer. Therefore, the *cycles per degree* for a display can be computed as *pixels*/(2\**subtended\_angle*).

##### Example

A typical three-display telepresence system uses three cameras with horizontal field of view of 25 degrees. It uses 65 inch 16x9 displays, and is calibrated to show images at actual size.

Each display has a width of 56.65 inches (1439 mm), and the viewers are seated 3.245 meters from the display. At this distance, the horizontal subtended angle is also 25 degrees, matching the camera field of view.

The native resolution of each display is 1920 x1080 pixels (1080p); which corresponds to 1920/(2\*25) cycles per degree – or about 38.4 CPD.

Similarly, 720p corresponds to about 24 CPD. A 65" 4K display at this viewing distance provides about 71 CPD, which exceeds the resolving power of the human eye.

|  |  |
| --- | --- |
| **Identity:** | effectiveResolution |
| **Description:** | This parameter indicates effective resolution of a rendered video stream as perceived by the viewer. |
| **Format:** | Floating Point number |
| **Possible Values:** | Positive numbers in Cycles per Degree |
| **Default:** | None. |
| **Reference:** | - |
| **Signalling Mechanism:** | Not Signalled |

Telepresence systems should provide effective resolutions higher than 25 CPD. Note that humans with normal vision cannot resolve CPD higher than 50, thus the experience isn’t enhanced with higher effective resolutions.

#### Video capture spatial information

This clause describes parameters related to spatial information of a single video capture.

##### Capture area

|  |  |
| --- | --- |
| **Identity:** | captureArea |
| **Description:** | This parameter indicates which area the video capture represents from the overall capture scene in terms of a set of four X, Y and Z Cartesian coordinates. |
| **Format:** | List of Strings |
| **Possible Values:** | Four list elements each with (X, Y, Z) where X, Y and Z are numbers. |
| **Default:** | None. |
| **Reference:** | See "Area of Capture" in [IETF CLUE FW] for further information. |
| **Signalling Mechanism:** | CLUE |
| **Supplementary Information:** | This parameter is used to indicate the area of the video capture, and is also used to determine the spatial relationship between video captures. It is also possible to determine from the capture area information where gaps are present due to display borders etc. which should be taken into consideration by a renderer when rendering video streams. |

###### Example usage of the capture area parameter

Capturers send the Capture Area parameters to renderers, in terms of a set of four X, Y and Z coordinates, respectively. For example, the coplanar capture area is represented by for point bottom left (X1, Y1, Z1), bottom right (X2, Y2, Z2), top left (X3, Y3, Z3) and top right (X4, Y4, Z4). Following the conventions of "Area of Capture" [IETF CLUE FW], as numbers move from lower to higher, the location goes from camera left to right (in the case of the 'x' dimension), audience front to back (in the case of the 'y' dimension) or low to high (in the case of the 'z' dimension), which is shown in figure 1(top view) and figure 2 (3D view).

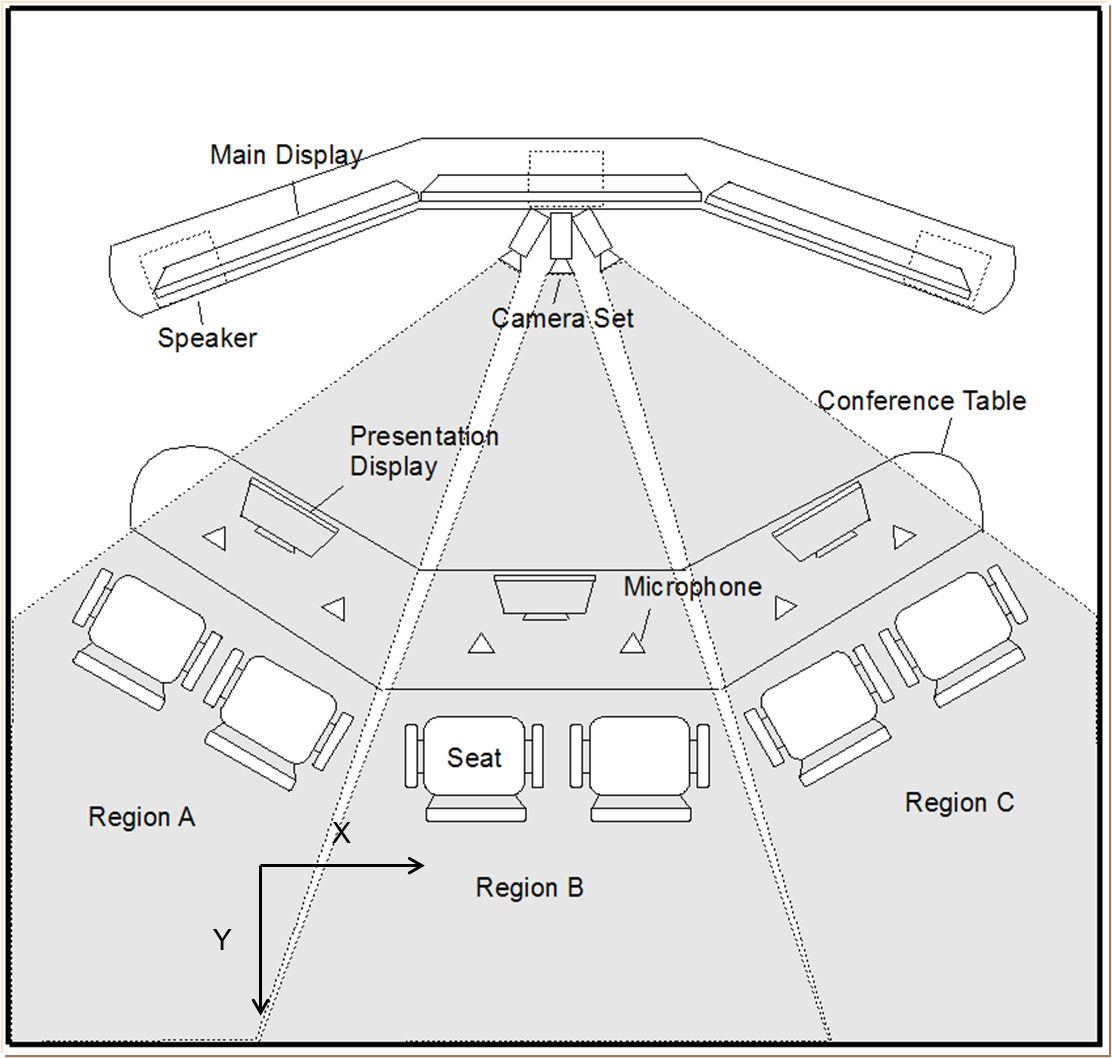


Figure 1: Capture area axis (top view)



Figure 2: Capture area axis (3D view)

NOTE ‑ The "Capture point" and "Point on line of capture" parameters are not considered in the above diagrams.

On reception of the Capture Area information renderers may use this information in several ways to render the received video streams illustrated below. The examples below only discusses the x-axis for simplicity, the calculations may equally be applied to the Y and Z axis. A scale of millimetres is assumed. First, the capture area can be used to indicate the spatial information of the video capture. For example, a capture of Region A with an area of capture:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | bottom left | bottom right | top left | top right |
| Region A | (0,2000,0) | (980,2200,0) | (0,2000,600) | (980,2200,600) |

The capture area above indicates that the X-axis is 980mm. Considering the X coordinates this leads to a planer width of 1000mm. Thus the renderer may use this information to render the pictures in the correct spatial place.

Second, the capture area can be used to describe the spatial relationship between video captures. For example if we consider the above capture of Region A and add a capture for Region B below:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | bottom left | bottom right | top left | top right |
| Region B | (990,2200,0) | (1990,2200,0) | (990,2200,600) | (1990,2200,600) |

By analysing the capture areas of Region A and Region B, respectively, the spatial relationship of the two video captures is that Region A is to the left of Region B and Region B is to the right of Region A. Therefore the renderer may use this information to render the pictures in correct spatial arrangement.

Third, the capture area can be used to determine the gap between video captures. For example, by analysing the respective X coordinates (i.e. 980 and 990) it can be seen that there is a gap of 10mm between Regions A and B. The renderer may use this information to perform processing on the images and render them in a continuous manner. That is, if the width of the gap is larger than the width of the borders between the displays on the receiver side, a virtual display border (For example, a black zone, usually with a width half of the difference between the gap and the borders) may be added to the captured images to achieve visual continuity. If the width of the gap is smaller than the width of the borders between the displays on the receiver side, the captured images may be cropped and/or scaled to achieve a continuous vision.

##### Capture point

|  |  |
| --- | --- |
| **Identity:** | capturePoint |
| **Description:** | This parameter indicates where in the capture scene the video was captured from. |
| **Format:** | String |
| **Possible Values:** | X, Y, Z  Where X, Y and Z are numbers. |
| **Default:** | None. |
| **Reference:** | See "Point of Capture" in [IETF CLUE FW] for further information. |
| **Signalling Mechanism:** | CLUE |
| **Supplementary Information:** | - |

##### Point on line of capture

|  |  |
| --- | --- |
| **Identity:** | lineOfCapturePoint |
| **Description:** | A field with a single Cartesian (X, Y, Z) point value (virtual or physical) which describes a position in space of a second point on the optical axis of the capturing device; the first point being the Point of Capture (see above). |
| **Format:** | String |
| **Possible Values:** | X, Y, Z  Where X, Y and Z are numbers. |
| **Default:** | None. |
| **Reference:** | See the "Point on line of Capture" attribute in [IETF CLUE FW] for further information. |
| **Signalling Mechanism:** | CLUE |
| **Supplementary Information:** | Knowing the optical axis associated with a capture may assist a renderer to apply an appropriate geometric correction when rendering a capture. |

#### Video capture encoding information

This clause describes parameters related to encoding information of a single video capture.

##### Maximum video bitrate

|  |  |
| --- | --- |
| **Identity:** | maxVideoBitrate |
| **Description:** | This parameter indicates the maximum number of bits per second relating to a single video encoding. |
| **Format:** | Numeric |
| **Possible Values:** | 0 upwards. |
| **Default:** | None. |
| **Reference:** | See "max-mbps" in [IETF RFC 6184] and "CustomMaxMBPS" in [ITU-T H.241] for further information. |
| **Signalling Mechanism:** | H.245, SDP |
| **Supplementary Information:** | - |

##### Maximum video resolution width

|  |  |
| --- | --- |
| **Identity:** | maxWidth |
| **Description:** | This parameter indicates the maximum video resolution width in pixels. |
| **Format:** | Numeric |
| **Possible Values:** | 0 upwards. |
| **Default:** | None. |
| **Reference:** | See "horizontal image size" in [IETF RFC 6236] and "CustomPictureFormat" in [ITU-T H.245] for further information. |
| **Signalling Mechanism:** | H.241, SDP. |
| **Supplementary Information:** | *-* |

##### Maximum video resolution height

|  |  |
| --- | --- |
| **Identity:** | maxHeight |
| **Description:** | This parameter indicates the maximum video resolution width in pixels. |
| **Format:** | Numeric |
| **Possible Values:** | 0 upwards. |
| **Default:** | None. |
| **Reference:** | See "vertical image size" in [IETF RFC 6236] and "CustomPictureFormat" in [ITU-T H.245] for further information. |
| **Signalling Mechanism:** | H.241, SDP |
| **Supplementary Information:** | *-* |

##### Maximum video framerate

|  |  |
| --- | --- |
| **Identity:** | maxFrameRate |
| **Description:** | This parameter indicates the maximum video framerate. |
| **Format:** | Numeric |
| **Possible Values:** | 0 upwards. |
| **Default:** | None. |
| **Reference:** | See "framerate" in [IETF RFC 4566] and "MaxFPS" in [ITU-T H.241] for further information. |
| **Signalling Mechanism:** | H.241, SDP |
| **Supplementary Information:** | *-* |

### Audio parameters

This clause describes the telepresence parameters related to the audio capabilities.

#### Audio capture spatial information

This clausedescribes parameters related to spatial information of a single audio capture.

**7.1.3.2.1 Audio capture point**

The "Capture point" parameter in clause 7.1.2.5.2 applies to audio captures where it describes the nominal midpoint of the microphones.

**7.1.3.2.2 Audio point on line of capture**

The "Point on line of capture" parameter in clause 7.1.2.5.3 applies to audio captures where it describes axis of the microphones. It may be used with the sensitivity pattern to infer a volume where sounds can be expected to be picked up.

**7.1.3.2.3 Audio capture sensitivity pattern**

|  |  |
| --- | --- |
| **Identity:** | Audio sensitivityPattern |
| **Description:** | This parameter provides information regarding the nominal sensitivity pattern of the microphone used to capture the audio. |
| **Format:** | Enumeration |
| **Possible Values:** | omni, shotgun, cardioid, hyper-cardioid |
| **Default:** | None. |
| **Reference:** | See “Audio Capture Sensitivity Pattern” in [IETF CLUE FW] for further information. |
| **Signalling Mechanism:** | CLUE |
| **Supplementary Information:** | - |

#### Audio capture encoding information

This clause describes parameters related to encoding information of a single audio capture.

##### Maximum audio bitrate

|  |  |
| --- | --- |
| **Identity:** | maxAudioBitrate |
| **Description:** | This parameter indicates the maximum number of bits per second relating to a single audio encoding. |
| **Format:** | Numeric |
| **Possible Values:** | 0 upwards. |
| **Default:** | None. |
| **Reference:** | See "bandwidth" in [IETF RFC 4566] and "maxBitRate" in [ITU-T H.245] for further information. |
| **Signalling Mechanism:** | H.245,SDP |
| **Supplementary Information:** | - |

#### Nominal audio level

|  |  |
| --- | --- |
| **Identity:** | nominalAudioLevel |
| **Description:** | This parameter indicates the nominal audio level sent in the Telepresence audio stream. |
| **Format:** | Numeric, in dBov |
| **Possible Values:** | 0 to -63 dBov |
| **Default:** | -26 |
| **Reference:** | - |
| **Signalling Mechanism:** | [ITU-T H.245] |

Telepresence systems shall send audio to other telepresence systems with a nominal audio level of NominalAudioLevel dBov as measured by method B of [ITU-T P.56] , which may be implemented using the tool sv-p56.c in the Software Tools Library [ITU-T G.191].

When a telepresence system connects to non-telepresence devices, it should use the NominalAudioLevel expected by that device. If this is unknown, it should use the Telepresence NominalAudioLevel.

#### Dynamic audio level

|  |  |
| --- | --- |
| **Identity:** | dynamicAudioLevel |
| **Description:** | This parameter indicates the actual audio level sent in the Telepresence audio stream as it varies as a function of time. |
| **Format:** | Numeric, in dBov |
| **Possible Values:** | -127 to 0 |
| **Default:** | None |
| **Reference:** | - |
| **Signalling Mechanism:** | RTP |
| **Supplementary Information:** | Dynamic audio level shall be carried as an RTP header extension of the associated audio stream and measured using the root-mean-square method specified in [IETF RFC 6464] on linear PCM samples.  NOTE – [IETF RFC 6464] assumes carriage of monophonic audio. When multiple audio channels are carried in the same RTP stream, the RTP header extension specified in [IETF RFC 6464] may still be used, with the measurement method appropriately adapted. For instance, in the case of a stereophonic stream, the stereophonic audio might be converted to monophonic audio by averaging the channels prior to the measurement. Audio with more than two channels could be treated in a similar fashion. Other adaptations may be used. |

#### Sending loudness rating

|  |  |
| --- | --- |
| **Identity:** | sendLoudnessRating |
| **Description:** | This parameter indicates the acoustic to electric transfer function of the Telepresence audio stream. |
| **Format:** | Numeric, in dB |
| **Possible Values:** | +4 to +16 |
| **Default:** | (+ 13–Fs) |
| **Reference:** | - |
| **Signalling Mechanism:** | *Not signalled* |
| **Supplementary Information:** | The purposes of specifying Sending Loudness Rating for the sound capture are to:   * Set the microphone preamplifiers gain value in order to use the whole dynamic of the A to D converter without saturating it. * Set the nominal electrical sending level from the Telepresence system to be compatible with that from conventional telephones. |

In order to achieve a transmitted nominalAudioLevel of -26 dBov, Telepresence systems shall send audiowith a SendLoudnessRating of (+ 13–Fs) dB, measured in accordance with [ITU-T P.300].

In order to take into account the difference between the reference test positioning and the actual microphone‑talker operating distance (ds), for which the terminal is adjusted, the following correction factor Fs is defined:

 (ds in meters)

NOTE ‑ The formula for Sending Loudness Rating is currently only defined in an audio bandwidth of up to 7 kHz, whereas Telepresence systems may have audio bandwidths up to 20 kHz. However, loudness ratings are intended for speech, where most energy is below 7 kHz, and any difference in loudness level in wider bandwidths is expected to be small. Adaptation of psychoacoustic based loudness algorithms aiming to determine the loudness for all bandwidths ranging from narrowband to full-band audio signals is an on-going work item in ITU-T SG12, and new results will be incorporated in this Recommendation when available.

#### Weighted microphone loudspeaker coupling loss

|  |  |
| --- | --- |
| **Identity:** | weightedMicrophoneLoudspeakerCouplingLoss |
| **Description:** | This parameter indicates the minimum weighted coupling loss between each microphone and loudspeaker pair. In practice, the worst case (ie the one where the coupling loss is minimum) will be the case where the microphone and loudspeaker are physically closest together. |
| **Format:** | Numeric, in dB. |
| **Possible Values:** | ≥ 46 dB nominal volume, ≥ 40 dB maximum volume |
| **Default:** | None. |
| **Reference:** | *-* |
| **Signalling Mechanism:** | *Not Signalled* |
| **Supplementary Information:** | The inherent coupling between loudspeakers and microphones in a telepresence system plays a key role in determining how much additional echo path loss is required to render any echo essentially inaudible. Maximizing the weightedMicrophoneLoudspeaker‌Coupling‌Loss of the system facilitates the operation of the Acoustic Echo Canceller which is required to provide this additional echo path loss, leading to improved full-duplex behaviour.  The audio levels in the system must first be calibrated to their nominal levels, and then for each loudspeaker/microphone pair a signal based on Annex A/[ITU-T P.501] is output through the loudspeaker, and simultaneously the signal picked-up by the microphone is recorded. The coupling loss is computed as the difference in level between these two signals, and is then weighted following B.4/[ITU-T G.122] in the 100 Hz - 16 kHz frequency range. During this measurement, every signal processing module, such as noise suppressor, echo canceller, automatic gain control, must be switched off (except for the frequency equalization filters if any).  NOTE ‑ This can only be measured for static microphones. |

### Delay parameters

This clause describes the telepresence parameters related to delay.

#### End to end video delay

|  |  |
| --- | --- |
| **Identity:** | endtoEndVideoDelay |
| **Description:** | This parameter indicates the one-way end to end delay (camera lens to video display) of the video media sent between two Telepresence endpoints. |
| **Format:** | Integer. |
| **Possible Values:** | Positive numbers in milliseconds. |
| **Default:** | None. |
| **Reference:** | *-* |
| **Signalling Mechanism:** | *Not signalled* |
| **Supplementary Information:** | In order to provide a high QoE telepresence experience to end-users, it is desirable for the end to end video delay to be less than 320 milliseconds. |

#### End to end audio delay

|  |  |
| --- | --- |
| **Identity:** | endtoEndAudioDelay |
| **Description:** | This parameter indicates the one-way end to end delay (mouth to ear) of the audio media sent between two Telepresence endpoints. |
| **Format:** | Integer. |
| **Possible Values:** | Positive numbers in milliseconds. |
| **Default:** | None. |
| **Reference:** | - |
| **Signalling Mechanism:** | *Not signalled* |
| **Supplementary Information:** | In order to provide a high QoE telepresence experience to end-users, telepresence systems, it is desirable for the end to end audio delay to be less than 280 milliseconds. |

#### Audio / video synchronization

|  |  |
| --- | --- |
| **Identity:** | AudioVideoSynchronization |
| **Description:** | This parameter indicates the synchronization between an audio and the corresponding video media stream (EndtoEndVideoDelay-EndtoEndAudioDelay). |
| **Format:** | Integer. |
| **Possible Values:** | Positive numbers in milliseconds indicate that the audio delay is less than the video delay; negative numbers in milliseconds indicate that the audio delay is greater than the video delay. |
| **Default:** | None. |
| **Reference:** | *-* |
| **Signalling Mechanism:** | *Not Signalled* |
| **Supplementary Information:** | In order to provide high QoE telepresence services to end-users, telepresence systems should maintain synchronization within 40 and -60 milliseconds (i.e. synchronization error is less than 40 ms if the audio stream is ahead of the video stream and less than 60 ms if the video stream is ahead of the audio stream). |

### Multiple source capture parameters

#### Multiple content capture

|  |  |
| --- | --- |
| **Identity:** | multiContentCapture |
| **Description:** | This parameter indicates that a capture contains content from one or more media captures sources. Source media captures may or may not be present in the resultant media stream depending on time or space. |
| **Format:** | Boolean |
| **Possible Values:** | Yes / No |
| **Default:** | None. |
| **Reference:** | See the 'Multiple content capture' in [IETF CLUE FW] for further information. |
| **Signalling Mechanism:** | CLUE |
| **Supplementary Information:** | - |

#### Multiple content capture sources

|  |  |
| --- | --- |
| **Identity:** | MCC Sources |
| **Description:** | This parameter is indicates the list of media captures sources that are contained in the media stream. |
| **Format:** | Not applicable |
| **Possible Values:** | References to the identity of media capture sources. |
| **Default:** | None. |
| **Reference:** | See the 'Multiple content capture' in [IETF CLUE FW] for further information. |
| **Signalling Mechanism:** | CLUE |
| **Supplementary Information:** | - |

#### Maximum multiple content capture sources

|  |  |
| --- | --- |
| **Identity:** | MCC maxCaptures |
| **Description:** | This parameter indicates the maximum number of media capture sources that will appear in a media stream at any point in time. It is also possible to indicate whether the number of captures may be less than and/or equal to the maximum number of captures. |
| **Format:** | Integer |
| **Possible Values:** | 1 upwards. |
| **Default:** | None. |
| **Reference:** | See the 'Maximum Number of Captures within a MCC' MCC attribute in [IETF CLUE FW] for further information. |
| **Signalling Mechanism:** | CLUE |
| **Supplementary Information:** | - |

#### Multiple content capture policy

|  |  |
| --- | --- |
| **Identity:** | MCC policy |
| **Description:** | This parameter indicates the policy of how capture sources are chosen to appear in a media stream. For example: a video capture may appear based on a sound level policy or may appear at regular intervals based on a round-robin policy. |
| **Format:** | Enumeration. |
| **Possible Values:** | As per clause 7.2.1.2/[IETF CLUE FW]. |
| **Default:** | None. |
| **Reference:** | See the 'Policy' MCC attribute in [IETF CLUE FW] for further information. |
| **Signalling Mechanism:** | CLUE |
| **Supplementary Information:** | - |

#### Multiple content capture synchronisation

|  |  |
| --- | --- |
| **Identity:** | MCC synchronizationID |
| **Description:** | This parameter provides a linkage between multiple MCC captures to synchronise which captures appear in a media stream at a particular point in time. MCCs with the same synchronisation identity will display media from the same capture sources at a particular point in time. |
| **Format:** | String |
| **Possible Values:** | As per clause 7.2.1.3/[IETF CLUE FW]. |
| **Default:** | None. |
| **Reference:** | See the 'Synchronisation Identity' MCC attribute in [IETF CLUE FW] for further information. |
| **Signalling Mechanism:** | CLUE |
| **Supplementary Information:** | - |

## Render related Parameters

### Audio parameters

#### Receiving Loudness Rating

|  |  |
| --- | --- |
| **Identity:** | receiveLoudnessRating |
| **Description:** | This parameter indicates the electric to acoustic transfer function of the Telepresence audio stream. |
| **Format:** | Numeric, in dB |
| **Possible Values:** | ‑26 to +18 |
| **Default:** | (+ 5–Fr) |
| **Reference:** | - |
| **Signalling Mechanism:** | *Not Signalled* |
| **Supplementary Information:** | * The purpose of specifying Receiving Loudness Rating for the sound rendering is to set the loudspeakers gain value in order to obtain a comfortable acoustic sound level at the conferee's positions (not too low and not too loud), given a nominal incoming electrical receive level. |

When receiving a nominalAudioLevel of -26 dbov, Telepresence systems shall render audio from other telepresence systems with a nominal Receive‌Loudness‌Rating of (+ 5–Fr) dB, measured in accordance with [ITU-T P.300].

In order to take into account the difference between the reference test positioning and the actual loudspeaker‑listener operating distance (dr) for which the terminal is adjusted, the following correction factor Fr is defined:

 (dr in metres)

NOTE 1 ‑ The formula for Receiving Loudness Rating is currently only defined in an audio bandwidth of up to 7 kHz, whereas Telepresence systems may have audio bandwidths up to 20 kHz. However, loudness ratings are intended for speech, where most energy is below 7 kHz, and any difference in loudness level in wider bandwidths is expected to be small. Adaptation of psychoacoustic based loudness algorithms aiming to determine the loudness for all bandwidths ranging from narrowband to full-band audio signals is an on-going work item in ITU-T SG12, and new results will be incorporated in this Recommendation when available.

NOTE 2 ‑ Practical Telepresence systems will usually have some means for adjusting the receive volume according to user preference. However, it is still useful to specify the nominal receive loudness rating as a default starting point for further volume adjustment, and to provide a reference amount of receive gain for echo path loss measurements.

## Telepresence system environment parameters

### Colour and lighting parameters

For a pleasing telepresence experience good lighting in terms of quantity and quality is needed for the environment where the telepresence endpoint is placed. [ISO 8995] defines several criteria for lighting design in indoor environments. It highlights the following parameters as contributing to the luminous environment:

* Luminance distribution
* Illuminance
* Glare
* Directionality of light
* Colour aspect of the light and surfaces
* Flicker
* Daylight
* Maintenance.

These aspects should be considered when designing and commissioning telepresence endpoints.

#### Illuminant

|  |  |
| --- | --- |
| **Identity:** | illuminantType |
| **Description:** | This parameter describes the profile of the visible light at a telepresence endpoint. It is characterised by a CIE standard illuminant value. |
| **Format:** | Enumeration. |
| **Possible Values:** | "Dx" Illuminant representing Daylight (x is in integer, e.g. D65 Default, D50, 55, 75).  "E" Illuminant representing an equal energy radiator.  "FLx" Illuminant representing fluorescent lamp (x is an integer, e.g. FL1, FL2, FL12 etc.).  "FL3.x" Illuminant representing a three band fluorescent lamps (x is an integer, e.g. FL3.1, FL3.14).  "HPx" Illuminant representing a high-pressure lamp spectrum (x is an integer, e.g.:   * HP1: Standard high pressure lamp * HP2: Colour enhanced high pressure sodium lamp * HP3 to HP5: High Pressure metal halide lamps) |
| **Default:** | None. |
| **Reference:** | [b-CIE 15.3:2004] and [ISO 11664-2] |
| **Signalling Mechanism:** | Annex E of [ITU-T H.264] and Annex E of [ITU-T H.265] |
| **Supplementary Information:** | If the CRI value and colour temperature are omitted they are implied by the type. For example, type "A" indicates that the light source is tungsten-filament lighting with CRI of more than 95 and colour temperature of 2856K. |

#### Colour Rendering Index

|  |  |
| --- | --- |
| **Identity:** | illuminantCRI |
| **Description:** | This parameter describes the colour rendering index (CRI) of the visible (ambient) light at the telepresence endpoint. The method of calculation of the CRI is according to [b-CIE 13.3-1995]. |
| **Format:** | Numeric. |
| **Possible Values:** | 0 upwards. |
| **Default:** | None. |
| **Reference:** | [b-CIE 13.3-1995] |
| **Signalling Mechanism:** | Annex E of [ITU-T H.264] and Annex E of [ITU-T H.265] |
| **Supplementary Information:** | If the light source is a standard CIE illuminant, the CRI value may be omitted. |

#### Colour temperature

|  |  |
| --- | --- |
| **Identity:** | illuminantColourTemp |
| **Description:** | This parameter describes the correlated colour temperature (CCT) [b-CIE 17.4-1987] of the visible (ambient) light at the telepresence endpoint. |
| **Format:** | Numeric, degrees Kelvin. |
| **Possible Values:** | 0 upwards. |
| **Default:** | None. |
| **Reference:** | - |
| **Signalling Mechanism:** | Annex E of [ITU-T H.264] and Annex E of [ITU-T H.265] |
| **Supplementary Information:** | If the light source is a standard CIE illuminant, the CCT value can be omitted. |

#### Colour and lighting parameter usage

One of the goals of telepresence is to provide a "being there" experience. One of the aspects of the human vision system is to perform chromatic adaptation based on the ambient lighting conditions so that objects appear to have the same colour under various lighting conditions. Therefore people at the same location would generally perceive scenes in the same way. However since capture devices such as cameras cannot self-adapt, the captured objects will be perceived with different colour in different lighting conditions. In telepresence systems people are at disparate locations with different lighting conditions. Therefore a local scene would be perceived to be a different colour by people at the remote end to the local people.

Capture devices such as cameras can perform "colour correction", such as "white balancing" on the captured images so that the colour more closely correlates to what a person sees. In transforming the image colour the device must estimate the ambient light (illuminant) and its colour characteristics for "colour correction" in order to scale the image colours correctly. The "illuminant", "colour rendering index" and "colour temperature" are local colour and lighting parameters that may be utilised for white balancing colour enhancement.

When an image is transmitted to the remote end and displayed it will in many cases be under a different ambient light. As discussed above, if the colours of the image and luminance of the image are perceived by the remote participants as being different to the local conditions then the "being there" experience is diminished. In order to maximize the "being there" experience the display equipment may transform the received images to ensure consistency with the local colour and lighting parameters (i.e., as if it were captured in the local site). The colour and lighting parameters of the remote site may be based on pre-agreed profile (i.e. the support of [ITU-R BT.709] which includes a D65 illuminant) or through in-band signalling such as the VUI parameters in [ITU-T H.264]. The "Video colour gamut" parameter (clause 7.1.2.1) provides a means for the communication of this information. After obtaining the colour and lighting parameters of the remote site and the ambient light, the renderer can estimate a transform function based on them, and then may perform image correction to achieve a better "being there" experience. The correction function can be a white balance transform, a gamma correction, tone mapping, a colour matrix or a colour space lookup table.

### Acoustic Parameters

#### Reverberation time

Reverberation time should be typically between 0.3 and 0.4 seconds, but at a minimum of 0.2 seconds, when measured in the 125 Hz-4 kHz octave bands. Such values would apply to medium size meeting rooms with a volume of approximately 100 m3 or less. For larger meeting rooms, a slightly higher reverberation time would be acceptable.

This represents somewhat lower reverberation times than it is usual in meeting rooms today, but these values are absolutely attainable in practice.

The audio quality will worsen gradually with increasing reverberation times (T) above 0.4 s.

|  |  |
| --- | --- |
| **Identity:** | reverberationTime |
| **Description:** | This parameter indicates the reverberation time in a telepresence meeting room. |
| **Format:** | Numeric, in seconds |
| **Possible Values:** | between 0.3 and 0.4 |
| **Default:** | None. |
| **Reference:** | See clause 3 / [P.Supp 16] for further information regarding reverberation considerations. |
| **Signalling Mechanism:** | *Not Signalled* |

#### Ambient noise

Care should be taken to minimize the ambient noise level, including both noise from technical installations (air-conditioning, in particular) and any other external noise transmitted into the room. The ambient noise is measured as an Equivalent A-weighted Sound Pressure Level (LAeq).For a high quality experience, the noise level should be less or equal to 35 dBA.

This can be a difficult goal to meet in some buildings, and most smaller telepresence meeting rooms (holding 20 people or less) will work well with a higher ambient noise limit like 40 dBA. However, the quality of experience will in general gradually decrease with increasing noise level.

|  |  |
| --- | --- |
| **Identity:** | ambientNoise |
| **Description:** | This parameter indicates the Equivalent A-Weighted Sound Pressure Level in a telepresence meeting room environment measured at all relevant seating positions. |
| **Format:** | Numeric, in dBA |
| **Possible Values:** | <= 35 |
| **Default:** | None. |
| **Reference:** | See clause 3 / [P.Supp 16] for further information regarding ambient noise. |
| **Signalling Mechanism:** | *Not Signalled* |

#### Sound insulation

Sound insulation between rooms is rated with a "Weighted *apparent* sound Reduction index" (R'w) which can only be measured in the field (finished buildings). Specific building elements are rated with a "Weighted sound Reduction index" (Rw) and can only be measured in laboratories. See [ISO 140-4], [ISO 717-1] for more information.

Recommendations for sound insulation will vary according to the local building plan and external noise sources. However, one universal recommendation for privacy can be stated. To ensure that confidential information cannot be overheard from outside the room, the weighted apparent sound reduction index (R'w) of the room boundaries is recommended to be R'w ≥ 48 dB. See [ISO 140‑4] for more information. This approximately corresponds to a sound transmission class STC ≥ 50.

NOTE ‑ STC is another measure of sound reduction. See [b-ASTM E90] and [b-ASTM E413] for more information.

Doors in the meeting room should have Rw ≥ 35 dB (STC ≥ 35). For walls that include doors, it will be difficult to achieve R'w ≥ 48 dB since the door will be a weak link reducing the R'w of the total construction.

|  |  |
| --- | --- |
| **Identity:** | soundInsulation |
| **Description:** | This parameter indicates the Weighted apparent sound Reduction index (R'w) used between a Telepresence meeting room and adjacent rooms. |
| **Format:** | Numeric, in dB |
| **Possible Values:** | ≥ 48 |
| **Default:** | None. |
| **Reference:** | - |
| **Signalling Mechanism:** | *Not Signalled* |

Appendix I  
  
Parameter signalling analysis

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

## I.1 Capture related parameters

### I.1.1 General parameters

| Parameter | Need for signalling at session initiation | Need for signalling during session | Remarks |
| --- | --- | --- | --- |
| media Type | Y | Y | See the ' CapturedMedia ' attributes in [IETF CLUE DM]. |
| captureScene description | Y | Y | See the “Description" attribute in 7.3.1 of [IETF CLUE FW]. |
| sceneView description | Y | Y | See the “Description" attribute in 7.3.2 of [IETF CLUE FW]. |
| Lang | Y | N | See "Language" in [IETF CLUE FW]. |
| Priority | Y | Y | See "Priority" in [IETF CLUE FW]. |
| embeddedText | Y | Y | See "Embedded Text" in [IETF CLUE FW]. |
| relatedTo | Y | Y | See "Related to" in [IETF CLUE FW]. |
| Presentation | Y | Y | See "Presentation" in [IETF CLUE FW] |
| personInfo | Y | Y | As per clause 7.1.1.10/[IETF CLUE FW]. |
| personType | Y | Y | As per clause 7.1.1.11/[IETF CLUE FW]. |
| sceneInformation | Y | Y | As per clause 7.3.1.1/[IETF CLUE FW]. |
| mediaCapture description | Y | Y | See the “Description" attribute in 7.1.1/[IETF CLUE FW]. |
| captureScene scale | Y | N | See "Scale" in [IETF CLUE FW]. |
| mediaCapture mobility | Y | N | See "Mobility of capture" in [IETF CLUE FW]. |
| mediaCapture view | Y | Y | See "View" in [IETF CLUE FW]. |
| maxGroupBandwidth | Y | N | See "maxGroupBandwidth" in [IETF CLUE FW]. |
| Simulcast | Y | Y | See [ITU-T H.323] and [ITU-T H.241] clause 6.2.5. |

### I.1.2 Visual parameters

| Parameter | Need for signalling at session initiation | Need for signalling during session | Remarks |
| --- | --- | --- | --- |
| colorGamut | Y | N | Signalled as part of the codec information, e.g. in H.264 as described by clause 7.1.2.1. |
| lumaBitDepth | Y | N | This is signalled as part of the codec payload information. |
| chromaBitDepth | Y | N | This is signalled as part of the codec payload information. |
| effectiveResolution | N | N | Not signalled. |
| captureArea | Y | Y | See "Area of Capture" in [IETF CLUE FW]. |
| capturePoint | Y | Y | See "Point of Capture" in [IETF CLUE FW]. |
| lineOfCapturePoint | Y | Y | See the "Point on line of Capture" attribute in [IETF CLUE FW]. |
| maxVideoBitrate | Y | Y | See “max-mbps” in [IETF RFC 6184] and “CustomMaxMBPS” in [ITU-T H.241]. |
| maxWidth | Y | N | See “horizontal image size” in [IETF RFC 6236] and “CustomPictureFormat” in [ITU-T H.245]. |
| maxHeight | Y | N | See “vertical image size” in [IETF RFC 6236] and “CustomPictureFormat” in [ITU-T H.245]. |
| maxFramerate | Y | N | See "framerate" in [IETF RFC 4566] and “MaxFPS” in [ITU-T H.241]. |

### I.1.3 Audio parameters

| Parameter | Need for signalling at session initiation | Need for signalling during session | Remarks |
| --- | --- | --- | --- |
| Audio capturePoint | Y | Y | See "Point of Capture" in [IETF CLUE FW]. |
| Audio lineOfCapturePoint | Y | Y | See the "Point on line of Capture" attribute in [IETF CLUE FW]. |
| Audio sensitivityPattern | Y | Y | See the "Audio Capture Sensitivity Pattern" attribute in [IETF CLUE FW]. |
| maxAudioBitrate | Y | Y | See “bandwidth” in [IETF RFC 4566] and “maxBitRate” in [ITU-T H.245]. |
| nominalAudio Level | Y | Y | [ITU-T H.245] |
| dynamicAudioLevel | N | Y | See [IETF RFC 6464] |
| sendLoudnessRating | N | N | Not signalled. |
| Weighted-Microphone-Loudspeaker Coupling Loss | N | N | Not signalled. |

### I.1.4 Delay parameters

| Parameter | Need for signalling at session initiation | Need for signalling during session | Remarks |
| --- | --- | --- | --- |
| endToEndVideoDelay | N | N | Not signalled. |
| endToEndAudioDelay | N | N | Not signalled. |
| audioVideoSynchronization | N | N | Not signalled. |

### I.1.5 Multiple source capture parameters

| Parameter | Need for signalling at session initiation | Need for signalling during session | Remarks |
| --- | --- | --- | --- |
| multiContentCapture | Y | Y | See the ' Multiple content capture ' in [IETF CLUE FW]. |
| MCC sources | Y | Y | See the ' Multiple content capture ' in [IETF CLUE FW]. |
| MCC maxCaptures | Y | Y | See the ' Maximum Number of Captures within a MCC ' MCC attribute in [IETF CLUE FW]. |
| MCC policy | Y | Y | See the ' Policy ' MCC attribute in [IETF CLUE FW]. |
| MCC synchronizationID | Y | Y | See the ' Synchronisation Identity ' MCC attribute in [IETF CLUE FW]. |

## I.2 Render related parameters

### 

### I.2.1 Audio parameters

| Parameter | Need for signalling at session initiation | Need for signalling during session | Remarks |
| --- | --- | --- | --- |
| receiveLoudnessRating | N | N | Not signalled. |

## I.3 Telepresence system environment parameters

### I.3.1 Colour and lighting parameters

| Parameter | Need for signalling at session initiation | Need for signalling during session | Remarks |
| --- | --- | --- | --- |
| illuminantType | Y | Y | See clause 7.3.1.4 regarding signalling of this information. |
| illuminantCRI | Y | Y | See clause 7.3.1.4 regarding signalling of this information. |
| illuminantColourTemp | Y | Y | See clause 7.3.1.4 regarding signalling of this information. |

### I.3.2 Acoustic parameters

| Parameter | Need for signalling at session initiation | Need for signalling during session | Remarks |
| --- | --- | --- | --- |
| reverberationTime | N | N | Not signalled. |
| ambientNoise | N | N | Not signalled. |
| soundInsulation | N | N | Not signalled. |

Bibliography

[b-ImagePro] John C. Russ, *The Image Processing Handbook*, CRC Press, 2006.

[b-ASTM E90] ASTM E90 - 09 *Standard Test Method for Laboratory Measurement of Airborne Sound Transmission Loss of Building Partitions and Elements*.

[b-ASTM E413] ASTM E413-10, *Classification for Rating Sound Insulation*.

[b-CIE 15.3:2004] CIE 15.3:2004, *Colorimetry*, Third Edition, 2004.

[b-CIE 13.3-1995] CIE 13.3-1995, *Method of Measuring and Specifying Colour Rendering Properties of Light Sources*, 1995.

[b-CIE 17.4-1987] CIE 17.4-1987, *International Lighting Vocabulary*, 1987.

[b-ITU-T G.114] ITU-T Recommendation G.114, *One-way transmission time*, 2003

[b-ITU-R BT.1359-1] ITU-R Recommendation BT.1359-1, *Relative timing of sound and vision for broadcasting,* 1998.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_