|  |
| --- |
| 3GPP TS 26.113 V0.4.0 (2023-4) |
| Technical Specification |
| 3rd Generation Partnership Project;Technical Specification Group Services and System Aspects;Enabler for Immersive Real-time Communication(Release 18) |
|   |
|  |  |
|  |
| The present document has been developed within the 3rd Generation Partnership Project (3GPP TM) and may be further elaborated for the purposes of 3GPP.The present document has not been subject to any approval process by the 3GPPOrganizational Partners and shall not be implemented.This Specification is provided for future development work within 3GPPonly. The Organizational Partners accept no liability for any use of this Specification.Specifications and Reports for implementation of the 3GPP TM system should be obtained via the 3GPP Organizational Partners' Publications Offices. |

|  |
| --- |
|  |
| ***3GPP***Postal address3GPP support office address650 Route des Lucioles - Sophia AntipolisValbonne - FRANCETel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16Internethttp://www.3gpp.org |
| ***Copyright Notification***No part may be reproduced except as authorized by written permission.The copyright and the foregoing restriction extend to reproduction in all media.© 2023, 3GPP Organizational Partners (ARIB, ATIS, CCSA, ETSI, TSDSI, TTA, TTC).All rights reserved.UMTS™ is a Trade Mark of ETSI registered for the benefit of its members3GPP™ is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational PartnersLTE™ is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational PartnersGSM® and the GSM logo are registered and owned by the GSM Association |

Contents

Foreword 4

Introduction 5

1 Scope 6

2 References 6

3 Definitions of terms, symbols and abbreviations 7

3.1 Terms 7

3.2 Symbols 7

3.3 Abbreviations 7

4 System description 8

4.1 High-level architecture 8

4.2 iRTC client in terminal 8

4.3 Web real-time communication 9

5 Functional components 10

5.1 General 10

5.2 Audio 10

5.2.1 Microphone 10

5.2.2 Pre/post-processor 11

5.2.3 Codec 11

5.3 Video 11

5.3.1 Camera 11

5.3.2 Pre/post-processor 11

5.3.3 Codec 11

5.4 Sensor 11

5.3.1 Measure 12

5.5 Transport protocols 12

6 Session management 12

6.1 General 12

6.2 WebRTC functions in 5GS 12

7 Inter-working 12

4.1 General 12

8 Packet-loss handling 12

8.1 General 12

9 Implementor's guide 13

9.1 General 13

Annex A (informative): 14

A.1 General 14

Annex B (informative): 15

B.1 General 15

Annex <X> (informative): Change history 16

# Foreword

This Technical Specification has been produced by the 3rd Generation Partnership Project (3GPP).

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

x the first digit:

1 presented to TSG for information;

2 presented to TSG for approval;

3 or greater indicates TSG approved document under change control.

y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.

z the third digit is incremented when editorial only changes have been incorporated in the document.

In the present document, modal verbs have the following meanings:

**shall** indicates a mandatory requirement to do something

**shall not** indicates an interdiction (prohibition) to do something

The constructions "shall" and "shall not" are confined to the context of normative provisions, and do not appear in Technical Reports.

The constructions "must" and "must not" are not used as substitutes for "shall" and "shall not". Their use is avoided insofar as possible, and they are not used in a normative context except in a direct citation from an external, referenced, non-3GPP document, or so as to maintain continuity of style when extending or modifying the provisions of such a referenced document.

**should** indicates a recommendation to do something

**should not** indicates a recommendation not to do something

**may** indicates permission to do something

**need not** indicates permission not to do something

The construction "may not" is ambiguous and is not used in normative elements. The unambiguous constructions "might not" or "shall not" are used instead, depending upon the meaning intended.

**can** indicates that something is possible

**cannot** indicates that something is impossible

The constructions "can" and "cannot" are not substitutes for "may" and "need not".

**will** indicates that something is certain or expected to happen as a result of action taken by an agency the behaviour of which is outside the scope of the present document

**will not** indicates that something is certain or expected not to happen as a result of action taken by an agency the behaviour of which is outside the scope of the present document

**might** indicates a likelihood that something will happen as a result of action taken by some agency the behaviour of which is outside the scope of the present document

**might not** indicates a likelihood that something will not happen as a result of action taken by some agency the behaviour of which is outside the scope of the present document

In addition:

**is** (or any other verb in the indicative mood) indicates a statement of fact

**is not** (or any other negative verb in the indicative mood) indicates a statement of fact

The constructions "is" and "is not" do not indicate requirements.

# Introduction

This clause is optional. If it exists, it shall be the second unnumbered clause.

# 1 Scope

The present document …

# 2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non‑specific.

- For a specific reference, subsequent revisions do not apply.

- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[x2] W3C Recommendation: "WebRTC 1.0: Real-Time Communication Between Browsers".

[x3] IETF RFC 7478 (2015): "Web Real-Time Communication Use Cases and Requirements".

[x4] 3GPP TS 26.119: "Media Capabilities for Augmented Reality"

[x5] 3GPP TS 26.506: "5G Real-time Media Communication Architecture (Stage 2)"

[x6] 3GPP TS 26.522: "5G Real-time Media Transport Protocol Configurations".

[x7] 3GPP TS 26.114: " IP Multimedia Subsystem (IMS); Multimedia Telephony; Media handling and interaction".

[x8] IETF RFC 8825 (2021): "Overview: Real-Time Protocols for Browser-Based Applications".

[x9] IETF RFC 8835 (2021): "Transports for WebRTC".

[x10] 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".

[x11] Microsoft: "[Microphone Array Geometry Descriptor Format](https://learn.microsoft.com/en-us/windows-hardware/drivers/audio/microphone-array-geometry-descriptor-format)".

[x12] Apple: "[Getting Raw Accelerometer Events](https://developer.apple.com/documentation/coremotion/getting_raw_accelerometer_events)".

[x13] Google: "[Sensor Coordinate System](https://developer.android.com/guide/topics/sensors/sensors_overview)".

[x14] 3GPP TS 36.323: "Evolved Universal Terrestrial Radio Access (E-UTRA); Packet Data Convergence Protocol (PDCP) specification".

[x15] 3GPP TS 37.324: "Evolved Universal Terrestrial Radio Access (E-UTRA) and NR; Service Data Adaptation Protocol (SDAP) specification".

[x16] 3GPP TS 38.300: "NR; NR and NG-RAN Overall description; Stage-2".

[x17] 3GPP TS 38.331: "NR; Radio Resource Control (RRC); Protocol specification".

# 3 Definitions of terms, symbols and abbreviations

## 3.1 Terms

For the purposes of the present document, the terms given in TR 21.905 [1] and the following apply. A term defined in the present document takes precedence over the definition of the same term, if any, in TR 21.905 [1].

**example:** text used to clarify abstract rules by applying them literally.

## 3.2 Symbols

For the purposes of the present document, the following symbols apply:

<symbol> <Explanation>

## 3.3 Abbreviations

For the purposes of the present document, the abbreviations given in TR 21.905 [1] and the following apply. An abbreviation defined in the present document takes precedence over the definition of the same abbreviation, if any, in TR 21.905 [1].

3DoF Three Degrees of Freedom

6DoF Six Degrees of Freedom

API Application Programming Interface

AR Augmented Reality

DRB Data Radio Bearer

DTLS Datagram Transport Layer Security

FFS For Further Study

FoV Field of View

HMD Head-Mounted Display

HTTP Hyper-Text Transfer Protocol

ICE` Interactive Connectivity Establishment

IMU Inertial Measurement Unit

iRTC Immersive Real-Time Communication

LIDAR Light Detection and Ranging

MR Mixed Reality

MNO Mobile Network Operator

NAT Network Address Translation

OTT Over-The-Top

RGB Red-Green-Blue colour space

RGBD Red-Green-Blue-Depth

RTC Real-Time Communication

RTP Real-time Transport Protocol

SCTP Stream Control Transmission Protocol

SLAM Simultaneous Localization And Mapping

SRTCP Secure Real-time Transport Control Protocol

SRTP Secure Real-time Transport Protocol

SSE Server-Sent Events

STUN Session Traversal Utilities for NAT

TLS Transport Layer Security

ToF Time of Flight

TURN Traversal Using Relays around NAT

WebRTC Web Real-Time Communication

XHR XMLHttpRequest

XR Extended Reality

# 4 System description

## 4.1 High-level architecture

The immersive Real-Time Communication (iRTC) system is designed based on the 5G-RTC General Architecture specified in [x5]. Figure 4.1.x illustrates the high-level view of the iRTC system that uses 5GRTC AF and AS for realizing the services. 5G-RTC AF and AS provide the Control Plane (C-Plane) functionalities for setting up and controlling media and data sessions (U-Plane). The functionalities depend on supported scenarios of collaboration, which are described in [x5]. iRTC system shall support at least one scenario.



**Figure 4.1.x: High-level architecture figure showing two iRTC clients in terminals.**

NOTE 1: Only 5GRTC AS exists in the media/data path.

NOTE 2: 5GRTC AF and AS are provided by MNO or 3rd party, depending on adopted collaboration scenarios.

NOTE 3: Operator B is depicted for collaboration scenario 4. In other collaboration scenarios, "Operator B" is replaced with "Operator A", and the boxes representing the same functionality can be provided by an operator.

## 4.2 iRTC client in terminal

The functional components of a terminal including an iRTC client using 3GPP access are shown in figure 4.2.x. The scope of the present document is to specify the handling and interaction of media and data, which includes their capture and generation, pre/post-processing, and compression. Transport of media and data consists of the encapsulation of the coded media and data in one or more transport protocols, which is shown in figure 4.2.x as the "packet-based network interface" and is illustrated in more detail in the user-plane and control-plane protocol stacks shown in figure 5.5.x.



**Figure 4.2.x: Functional components of a terminal including an iRTC client in terminal using 3GPP access**

The 3GPP Layer 2 protocol to be interfaced with iRTC client is PDCP [x14] for EPC and SDAP [x15] for 5GC, which is used on top of PDCP as shown in clause 4.4.1 of [x16]. It is assumed that the SDAP would be configured without header for both directions in the typical iRTC cases, effectively interfacing with PDCP, as SDAP header would be needed only when more than one QoS flows are multiplexed in a DRB or reflective mapping is enabled. An architecture for XR baseline client can be found in [x4].

NOTE 1: Functional components in the grey box, except audio and video pre/post-processors, are within the scope of present document, which also specifies output formats of sensor, microphone, and camera.

NOTE 2: In certain codec types, e.g., avatars, at least the decoders need to be personalized with the information of those who will participate in the communication before or during session setup.

NOTE 3: Device information is assumed to be stored in the UE and loaded to the iRTC client during session setup.

NOTE 4: The iRTC client may exchange media and data with external devices tethered over wired links such as USB-C, 3GPP PC5 [x17], or non-3GPP radio access technologies such as Wi-Fi or Bluetooth.

NOTE 5: Text can be entered via user interface, typically available on display.

## 4.3 Web real-time communication

The iRTC client supports a subset of WebRTC, which enables real-time communication via application programming interfaces (APIs), supporting audio, video, and generic data to be sent between peers [x2]. Functionalities of WebRTC are available as JavaScript APIs for browsers, and libraries for applications [x8]. Information on use cases and requirements of WebRTC can be found in [x3].

# 5 Functional components

## 5.1 General

[Editor’s note: description of each functional component and its operation as specified in clauses 4.1 and 4.2, refer relevant clauses of TS 26.119, 26.522 and other documents for codec & protocol issues]

## 5.2 Audio

## 5.2.1 Microphone

An iRTC client in terminal can be connected to one or more microphones. The outputs of microphones are audio samples in 16-bit uniform Pulse Code Modulation (PCM) format. An iRTC client or audio infra may identify the direction of each microphone with a coordinate system described in figure A.1.x and table A.1.y.



**Figure A.1.x: Microphone array coordinate system**

MicrophoneType, whose default value of 0 indicates an omni-directional microphone, identifies the microphones when other types are used. How to assign a value to each microphone type is left to the discretion of the implementation.

|  |  |  |  |
| --- | --- | --- | --- |
| **Parameter** | **Unit** | **Definition** | **Note** |
| Yaw (𝛹) | int | Direction angle | -31416 < 𝛹 ≦ 31416 |
| Pitch (𝛳) | int | Elevation angle | -31416 < 𝛳 ≦ 31416 |
| Roll (𝛷) | int | Rotation angle | -31416 < 𝛷 ≦ 31416 |
| MicrophoneType | int | A number that uniquely identifies microphone type | May be used for indicating vendor-defined microphone types |

**Table A.1.y: Microphone description parameters**

NOTE 1: The coordinate system and two angles, yaw and pitch, are originally defined in [X11] for computers.

NOTE 2: The positive X-, Y-, Z-axis shown in figure A.1.x correspond to positive Z-, negative Y-, positive Z-axis of a coordinate system commonly used for sensors in mobile operating systems [X12], [X13].

## 5.2.2 Pre/post-processor

An iRTC client in terminal may pre-process the outputs of microphones before they are input to audio encoders, e.g., for limiting bandwidth or converting the output into spatial audio representations. An iRTC client in terminal may post-process the outputs of audio decoders before they are input to speakers, e.g., for acoustically matching the perceived directions or locations of audio with those of video scenes.

## 5.2.3 Codec

Audio codecs for the iRTC client in terminal are specified in [x4], [x7].

## 5.3 Video

## 5.3.1 Camera

An iRTC client in terminal can be connected to one or more color cameras, and/or to one or more depth cameras. Depth cameras in this document typically consist of infrared projectors and infrared cameras that estimate the depth from measured time-of-flight or distortion of projected patterns. Resolutions and frame rates of the cameras are set to meet available bit-rate, complexity, storage, or nature of applications.

The output formats of color cameras, in the form of *Y*, *CR*, *CB* or *R*, *G*, *B* signals, are specified in [x10]. The RGB signals can be input to (2D) video encoders. The output pixel of depth cameras has a value of a 16-bit unsigned number that represents the distance (in millimeters) from the reference point of a depth camera to a point in the captured scene, up to 32.7 meters. The depth signals for a rectangular area (map) can be input to a lossless or lossy encoder, or combined with RGB signals for further processing.

NOTE 1: With infrared-based depth cameras, measurable distance is typically less than several meters.

NOTE 2: When the resolutions or aspect ratios of RGB and depth signals differ, the depth signals, whose resolutions are typically lower than those of RGB, can be interpolated to match the RGB signals.

## 5.3.2 Pre/post-processor

An iRTC client in terminal may pre-process the outputs of cameras before they are input to video encoders, e.g., for limiting bandwidth or converting the outputs into other representations, e.g., point cloud. An iRTC client in terminal may post-process the outputs of video decoders before they are input to displays, e.g., for selecting scenes within FoV or enhancing perceived video quality through appropriate filtering.

## 5.3.3 Codec

Video codecs for the iRTC client in terminal are specified in [x4], [x7].

## 5.4 Sensor

[Editor’s note: description of sensors supporting the operation of a generic iRTC client in terminal, focusing on sensors whose measurements need to be transmitted]

## 5.3.1 Measure

## 5.5 Transport protocols

The iRTC client supports transport protocols used in WebRTC, including the protocols for interaction with intermediate boxes such as firewalls, relays, and NAT boxes [x9]. Figure 5.5.x shows the user-plane and control-plane protocol stack.



**Figure 5.5.x: Protocol stack for a basic iRTC client**

# 6 Session management

## 6.1 General

[Editor’s note: description of integrating WebRTC into 5G system, functionality exposure, QoS realization, etc]

## 6.2 WebRTC functions in 5GS

# 7 Inter-working

## 4.1 General

[Editor’s note: description of inter-working an iRTC client in terminal with another client connected to 3GPP and non-3GPP networks, including tethering]

# 8 Packet-loss handling

## 8.1 General

[Editor’s note: description of measures for link quality degradation, e.g., in poor channel condition, overloading, etc]

# 9 Implementor's guide

## 9.1 General

[Editor’s note: generic guidelines for configuring & operating an iRTC client in terminal]

Annex A (informative):

# [Editor’s note: ]A.1 General

[Editor’s note: ]

Annex B (informative):

[Editor’s note: ]

# B.1 General

[Editor’s note: ]

Annex <X> (informative):
Change history

|  |
| --- |
| Change history |
| Date | Meeting | TDoc | CR | Rev | Cat | Subject/Comment | New version |
|  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |