|  |
| --- |
| 3GPP TS 26.261 V0.1.0 (2024-05) |
| Technical Specification |
| 3rd Generation Partnership Project;Technical Specification Group Services and System Aspects;Terminal audio quality performance requirements for immersive audio servicesRelease 18 |
|   |
|  | 3GPP-logo_web |
|  |  |
|  |  |
|  |  |
|  |  |
| 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.© 2024, 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 6

3.1 Terms 6

3.2 Symbols 6

3.2 Abbreviations 7

4 Interfaces 7

4.1 General 7

4.2 Air interfaces 7

4.3 Acoustical interfaces 7

4.4 Electrical interfaces 8

5 Performance in sending 8

5.1 Applicability 8

5.2 Delay 8

5.3 Loudness 8

5.4 Frequency response 8

5.5 Directional information 8

Performance in receiving 9

6.1 Applicability 9

6.2 Delay 9

6.3 Loudness 9

6.4 Frequency response 9

6.5 Binaural rendering 9

6 Performance in sending+receiving 9

6.1 Applicability 9

6.2 Delay 9

Annex A (informative): Change history 10

# 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, certain modal verbs have the following meanings:

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

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

NOTE 1: The constructions "shall" and "shall not" are confined to the context of normative provisions, and do not appear in Technical Reports.

NOTE 2: 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

NOTE 3: 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

NOTE 4: The constructions "can" and "cannot" shall not to be used as 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

NOTE 5: The constructions "is" and "is not" do not indicate requirements.

# Introduction

TS 26.261 specifies minimum performance requirements for the electro-acoustic characteristics of LTE, NR and WLAN terminals when used to provide immersive services. The performance requirements are specified in the main body of the text; the test methods and considerations are described in TS 26.260.

1 ScopeThe present document is applicable to any terminal capable of supporting wideband, super-wideband or fullband immersive services. The present document specifies minimum performance requirements for the electro-acoustic characteristics of LTE, NR and WLAN terminals.

The set of minimum performance requirements enables a guaranteed level of speech quality while taking possible physical limits of the terminal design into account. Some performance objectives are also defined, if such design limits can be overcome. Care must be taken in applying performance objectives in isolation, not to degrade overall end-user audio quality.

The present document covers both conversational services based on MTSI / telepresence and non-conversational services.

# 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".

[2] 3GPP TS 26.260: "Objective test methodologies for the evaluation of immersive audio systems"

# 3 Definitions of terms, symbols and abbreviations

## 3.1 Terms

For the purposes of the present document, the terms given in 3GPP 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 3GPP TR 21.905 [1].

For the purposes of the present document the terms *wideband,* *super-wideband* and *fullband* refer to signals associated with the corresponding operating codec modes specified in TS 26.260.

The overload point (maximum load capacity) is for the purposes of this document defined as the RMS level of a digital representation of a full-scale pure tone at the input of the speech encoder. The overload point is defined at 3,14 dBm0 for IVAS.

For the purposes of the present document, the term *electrical interface* is defined as an analogue or digital access to a UE, which allows injecting and capturing signals electrically instead of through an acoustical interface. The interface can be either analogue (wired) or digital (wired or wireless). The purpose of this interface is to connect a separate device (typically a headset), which provides a receiver and transmitter.

.

## 3.2 Symbols

Void

## 3.2 Abbreviations

For the purposes of the present document, the abbreviations given in 3GPP 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 3GPP TR 21.905 [1].

DTX Discontinuous Transmission

EVS Enhanced Voice Services

HATS Head and Torso Simulator

IVAS Immersive Voice and Audio Services

POI Point of Interconnection (with PSTN)

RLR Receive Loudness Rating

SLR Send Loudness Rating

WLAN Wireless Local Area Network

# 4 Interfaces

## 4.1 General

The (EVS nteroperable) mono mode of IVAS should be tested according to TS 3GPP TS 26.131 [x1] and TS 26.132 [x2] with a bitrate of [tbd] kbit/s.

The interfaces required to define immersive terminal electro-acoustic characteristics are shown in TS 26.260. These are the air interface and the point of interconnect (POI). The interfaces are shown for different types of immersive formats.

Measurements can be made using the system simulator (SS) described in TS 26.260. For conversational services, MTSI aspects are specified by TS 26.114 [x1].

## 4.2 Air interfaces

## The Air Interfaces for LTE and NR are specified by GSM 05, 3GPP 45, 3GPP 25, 3GPP 36 and 3GPP 38 series specifications, and the Air Interface for WLAN access to EPC is specified by WLAN access to EPC as defined in TS 23.402 [x2] and TS 24.302 [x3]. 4.3 Acoustical interfaces

The following classes of acoustical interface are considered in this specification:

- Headset UE/mode (sending + receiving);

- Table-mounted UE/mode (sending + receiving);

- Hand-held UE/mode (sending + receiving);

- Loudspeaker UE/mode (receiving);

## NOTE: The handset mode is not listed because it may not be fully applicable for immersive communication due to the following reasons. A handset device is typically held close to the user's head, i.e., mouth and (a single) ear. In receiving, monaural listening cannot provide any spatial/immersive audio. In sending, the device is positioned close to the user's mouth, which can limit spatial information into the uplink signal.4.4 Electrical interfaces

The electrical interface UE/mode (sending + receiving) is considered in this specification and details on standardized analogue (wired) and digital (wired and wireless) headset interfaces can be found in TS 26.260. For the electrical interface, the POI in sending / receiving direction is respectively defined as the input / output of the reference coder of the system simulator.

Any of the UE types mentioned in clause 4.3 providing an electrical interface can be considered as Electrical Interface UE.

# 5 Performance in sending

## 5.1 Applicability

tbd

## 5.2 Delay

Tbd

## 5.3 Loudness

Tbd

## 5.4 Frequency response

Tbd

## 5.5 Directional information

Tbd

# Performance in receiving

## 6.1 Applicability

tbd

## 6.2 Delay

Tbd

## 6.3 Loudness

Tbd

## 6.4 Frequency response

Tbd

## 6.5 Binaural rendering

Tbd

# 6 Performance in sending+receiving

## 6.1 Applicability

tbd

## 6.2 Delay

Tbd

Annex A (informative):
Change history

|  |
| --- |
| **Change history** |
| **Date** | **Meeting** | **TDoc** | **CR** | **Rev** | **Cat** | **Subject/Comment** | **New version** |
| 2019-04 | SA4#103 | S4-190426 |  |  |  | Initial version | 0.0.1 |
| 2024-04 | SA4#128 | S4-24 |  |  |  |  | 0.1.0 |