|  |  |
| --- | --- |
| 3GPP TS 26.261 V18.0.0 (2024-06) | |
| Technical Specification | |
| 3rd Generation Partnership Project;  Technical Specification Group Services and System Aspects;  Terminal audio quality performance requirements for immersive audio services  Release 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 address  3GPP support office address  650 Route des Lucioles - Sophia Antipolis  Valbonne - FRANCE  Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16  Internet  http://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 members  3GPP™ is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners LTE™ is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners  GSM® 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 7

3.2 Abbreviations 7

4 Interfaces 7

4.1 General 7

4.2 Air interfaces 7

4.3 Acoustical interfaces 8

4.4 Electrical interfaces 8

5 Performance in sending 8

5.1 Applicability 8

5.2 Delay 8

5.3 Loudness 9

5.4 Frequency response 9

5.4.1 Frequency response for single source 9

5.4.1.1 General 9

5.4.1.2 Acoustical interface 9

5.4.1.3 Electrical interface 9

5.5 Directional information 10

5.5.1 Stereo 10

5.5.2 SBA 10

5.5.3 MASA 10

6 Performance in receiving 10

6.1 Applicability 10

6.2 Delay 10

6.3 Loudness 10

6.4 Frequency response 11

6.4.1.1 General 11

6.4.1.2 Acoustical interface 11

6.4.1.3 Electrical interface 12

6.5 Binaural rendering 12

6.5.1 ILD 12

6.5.2 ITD 12

7 Performance in sending+receiving 13

7.1 Applicability 13

7.2 Delay 13

**Annex A (informative): …………………………………………………………………………………….14**

# 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 test considerations are described in TS 26.260.

# 1 Scope

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

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.

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

[4] 3GPP TS 26.131: "Terminal Acoustic Characteristics for Telephony; Requirements".

[5] ETSI TS 103 739 (2021-10) V1.4.1: " Transmission requirements for wideband mobile wireless terminals (handset and headset) from a QoS perspective as perceived by the user".

[6] ETSI TS 103 740 (2021-10) V1.4.1: "Transmission requirements for wideband mobile wireless terminals (hands-free) from a QoS perspective as perceived by the user".

# 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 test signals associated with the corresponding operating codec modes 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 IVAS encoder. The overload point is defined in TS 26.260.

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.

**Stereo panorama**: The spatial image of a stereo signal, in which the sound source directions lie in a range from -100% (left) to 100% (right).

## 3.2 Symbols

dB Decibel

Hz Unit of frequency (Hertz)

ϕ Azimuth angle (phi)

θ Elevation angle (theta)

TR UE delay in receiving direction

TS UE delay in sending direction

Estimate for a single source direction in the stereo panorama that is physically positioned in direction relative to the capturing device

## 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

ILD Interaural Level Difference

ISM Independent Stream with Metadata

ITD Interaural Time Difference

IVAS Immersive Voice and Audio Services

MASA Metadata-Assisted Spatial Audio

OMASA Objects (ISM) with Metadata-Assisted Spatial Audio

OSBA Objects (ISM) with Scene-Based Audio

POI Point of Interconnection (with PSTN)

RLR Receive Loudness Rating

SBA Scene-Based Audio

SLR Send Loudness Rating

WLAN Wireless Local Area Network

# 4 Interfaces

## 4.1 General

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 [3].

## 4.2 Air interfaces

The same air interfaces as in clause 4.2 of TS 26.131 [4] apply.

## 4.3 Acoustical interfaces

The acoustical interfaces are described in clause 5.4.2 of TS 26.260 [2].

## 4.4 Electrical interfaces

The electrical interface UE is considered in this specification and details on standardized analogue (wired) and digital (wired and wireless) electrical 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

The performance requirements in this clause shall apply when UE is used to provide immersive audio capture.

## 5.2 Delay

The UE delay TS in send direction shall be less or equal to the delay requirements in Table 5.2.

Table 5.2: Requirement on sending UE delay

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Audio format | Subformat | UE type(s) | Maximum delay TS (ms) | |
| Performance Requirement | Performance Objective |
| Stereo | Stereo | All | TBD | TBD |
| Binaural | All | TBD | TBD |
| ISM | ISM1-ISM4 | All | TBD | TBD |
| SBA | FOA | All | TBD | TBD |
| HOA2 | All | TBD | TBD |
| HOA3 | All | TBD | TBD |
| MASA | 1 TC | All | TBD | TBD |
| 2 TC | All | TBD | TBD |
| OSBA | (ISM1-ISM4) x (FOA, HOA2, HOA3) | All | TBD | TBD |
| OMASA | (ISM1-ISM4) x MASA | All | TBD | TBD |
| Multichannel | 5.1 | All | TBD | TBD |
| 7.1 | All | TBD | TBD |
| 5.1.2 | All | TBD | TBD |
| 5.1.4 | All | TBD | TBD |
| 7.1.4 | All | TBD | TBD |

NOTE: This requirement only applies for one-way (sending only) cases where round-delay does not apply.

Compliance shall be checked by the relevant tests described in TS 26.260.

## 5.3 Loudness

The nominal values of SLR to the POI shall be:

SLR = 13 ± 3 dB

NOTE: The value of 13 dB is motivated by the corresponding SLR requirement for desktop hands-free and handheld hands-free in TS 26.131 [4].

Compliance shall be checked by the relevant tests described in TS 26.260.

## 5.4 Frequency response

### 5.4.1 Frequency response for single source

#### 5.4.1.1 General

The sensitivity/frequency characteristics shall be as follows.

The sending sensitivity frequency response, measured from the source direction under test to the SS audio output (digital output of the reference speech decoder of the SS), shall be within a mask, which can be drawn between the points given in the respective requirements table. The mask is drawn with straight lines between the breaking points in the requirements table on a logarithmic (frequency) - linear (dB sensitivity) scale.

#### 5.4.1.2 Acoustical interface

The requirements for acoustical interface testing are given in Table 5.4.1.2.

Table 5.4.1.2: Sending sensitivity/frequency requirements table for acoustical interface

|  |  |  |
| --- | --- | --- |
| **Frequency (Hz)** | **Upper limit (dB)** | **Lower limit (dB)** |
| 100 | TBD |  |
| 200 | TBD | TBD |
| 300 | TBD | TBD |
| 4000 | TBD | TBD |
| 8000 | TBD | TBD |
| 12000 | TBD |  |

Compliance shall be checked by the relevant tests described in TS 26.260.

#### 5.4.1.3 Electrical interface

The requirements for electrical interface testing are given in Table 5.4.1.3.

Table 5.4.1.3: Sending sensitivity/frequency requirements table for electrical interface

|  |  |  |
| --- | --- | --- |
| **Frequency (Hz)** | **Upper limit (dB)** | **Lower limit (dB)** |
| 100 | TBD |  |
| 200 | TBD | TBD |
| 300 | TBD | TBD |
| 6000 | TBD | TBD |
| 8000 | TBD | TBD |
| 12000 | TBD |  |

Compliance shall be checked by the relevant tests described in TS 26.260.

## 5.5 Directional information

### 5.5.1 Stereo

The estimated source directions in the stereo panorama shall be monotonically increasing between -60° and +60°, i.e., where is the source direction index according to Table 5 in TS 26.260. This requirement verifies that the estimated stereo panorama is consistent for the relevant source directions .

The estimated source directions in the stereo panorama shall provide a minimum and symmetric width between and , i.e., and shall be larger than or equal to 60%.

The estimated source directions in the stereo panorama shall be consistent at the edges, i.e., and .

For frontal incidence position (), the estimated sound source direction shall be in the center of the stereo panorama, i.e.,

For all *L* = 7 source directions of Table 5 in TS 26.260, the absolute value of ICTD shall be less than 1.5 ms.

Compliance shall be checked by the relevant tests described in TS 26.260.

### 5.5.2 SBA

The maximum absolute error () shall be less than TBD degree for all directions.

The maximum absolute error of shall be less than TBD degree for all directions.

Compliance shall be checked by the relevant tests described in TS 26.260.

### 5.5.3 MASA

The maximum absolute error () shall be less than TBD degree for all directions.

The maximum absolute error of shall be less than TBD degree for all directions.

Compliance shall be checked by the relevant tests described in TS 26.260.

# 6 Performance in receiving

## 6.1 Applicability

The performance requirements in this clause shall apply when UE is used to provide immersive audio rendering.

## 6.2 Delay

TBD

NOTE: this requirement only applies for one-way (receiving only) cases where round-delay does not apply.

Compliance shall be checked by the relevant tests described in TS 26.260.

## 6.3 Loudness

The nominal value of Loudness Level in Receive (LLR) shall be:

- 75 ± 4 phon for headset UE,

- 75 ± 4 phon for electrical interface UE (including binaural rendering and calibration),

- 67 -4/+8 phon for handheld hands-free UE,

- 71 ± 4 phon for table-mounted hands-free UE,

- 75 ± 4 phon for loudspeaker UE.

In case a user controlled receive volume control is provided, for at least one setting of the control the LLR shall meet the nominal value.

When the control is set to maximum, the LLR shall not be louder than 89 phon for all UE types. With the volume control set to the minimum position the LLR shall not be quieter than 52 phon and should not be quieter than 58 phon for all UE types.

NOTE: The loudness level requirements are motivated by corresponding requirements in ETSI TS 103 739 [5] for desktop hands-free and handheld hands-free and in ETSI TS 103 740 [6] for headset, respectively.

Performance requirements and objectives apply for source position of azimuth 0° and elevation 0° in the test signal. They should also be evaluated for the source positions listed in Table 6.3, and results should be included in the test report.

Table 6.3: Additional source positions for loudness

|  |  |
| --- | --- |
| **Azimuth [°]** | **Elevation [°]** |
| +90 | 0 |
| -90 | 0 |
| 180 | 0 |
| 0 | +90 |

Compliance shall be checked by the relevant tests described in TS 26.260.

## 6.4 Frequency response

#### 6.4.1.1 General

The sensitivity/frequency characteristics shall be as follows.

The receiving sensitivity frequency response, measured either from the digital interface to the DRP with diffuse-field correction or from the SS audio input (analogue or digital input of the reference speech encoder of the SS) to the DRP with diffuse-field correction shall be within a mask, which can be drawn with straight lines between the breaking points in table 2 on a logarithmic (frequency) - linear (dB sensitivity) scale.

#### 6.4.1.2 Acoustical interface

The requirements for acoustical interface testing are given in Table 6.4.1.2.

Table 6.4.1.2: Sending sensitivity/frequency requirements table for acoustical interface

|  |  |  |
| --- | --- | --- |
| **Frequency (Hz)** | **Upper limit (dB)** | **Lower limit (dB)** |
| 100 | TBD |  |
| 200 | TBD | TBD |
| 300 | TBD | TBD |
| 4000 | TBD | TBD |
| 8000 | TBD | TBD |
| 12000 | TBD |  |

Compliance shall be checked by the relevant tests described in TS 26.260.

#### 6.4.1.3 Electrical interface

The requirements for electrical interface testing are given in Table 6.4.1.3.

Table 6.4.1.3: Sending sensitivity/frequency requirements table for electrical interface

|  |  |  |
| --- | --- | --- |
| **Frequency (Hz)** | **Upper limit (dB)** | **Lower limit (dB)** |
| 100 | TBD |  |
| 200 | TBD | TBD |
| 300 | TBD | TBD |
| 6000 | TBD | TBD |
| 8000 | TBD | TBD |
| 12000 | TBD |  |

Compliance shall be checked by the relevant tests described in TS 26.260.

## 6.5 Binaural rendering

#### 6.5.1 ILD

For headset UE or electrical interface UE (if intended for headset usage), the UE shall pass ILD requirements defined in Table 6.5.1.

If the UE supports headtracking, the same requirements in Table 6.5.1 shall be met for each HATS orientation .

NOTE: If the UE supports headtracking and the requirements are not met, the test operator should verify if the failure might be caused by an automatic reset of the reference orientation.

Table 6.5.1: ILD requirements

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| i | Source azimuth (°) | Source elevation (°) | 500 Hz | | 1000 Hz | | 2000 Hz | | 4000 Hz | | 8000 Hz | |
| Lower limit (dB) | Upper limit (dB) | Lower limit (dB) | Upper limit (dB) | Lower limit (dB) | Upper limit (dB) | Lower limit (dB) | Upper limit (dB) | Lower limit (dB) | Upper limit (dB) |
| 1 | 0 | 0 | TBD | TBD | TBD | TBD | TBD | TBD | TBD | TBD | TBD | TBD |
| 2 | 180 | 0 | TBD | TBD | TBD | TBD | TBD | TBD | TBD | TBD | TBD | TBD |
| 3 | 0 | 90 | TBD | TBD | TBD | TBD | TBD | TBD | TBD | TBD | TBD | TBD |
| 4 | 90 | 0 | TBD | TBD | TBD | TBD | TBD | TBD | TBD | TBD | TBD | TBD |
| 5 | -90 (270) | 0 | TBD | TBD | TBD | TBD | TBD | TBD | TBD | TBD | TBD | TBD |

Compliance shall be checked by the relevant tests described in TS 26.260.

#### 6.5.2 ITD

For headset UE or electrical interface UE (if intended for headset usage), the UE shall pass ITD requirements defined in Table 6.5.2.

If the UE supports headtracking, the same requirements in Table 6.5.2 shall be met for each HATS orientation .

NOTE: If the UE supports headtracking and the requirements are not met, the test operator should verify if the failure might be caused by an automatic reset of the reference orientation.

Table 6.5.2: ITD requirements

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| i | Source azimuth (°) | Source elevation (°) | Lower limit (ms) | Upper limit (ms) |
| 1 | 0 | 0 | TBD | TBD |
| 2 | 180 | 0 | TBD | TBD |
| 3 | 0 | 90 | TBD | TBD |
| 4 | 90 | 0 | TBD | TBD |
| 5 | -90 (270) | 0 | TBD | TBD |

Compliance shall be checked by the relevant tests described in TS 26.260.

# 7 Performance in sending+receiving

## 7.1 Applicability

The performance requirements in this clause shall apply when UE is used to provide end to end immersive audio, including both capture and rendering.

## 7.2 Delay

TBD

NOTE: This requirement does not apply for one-way scenarios, delay corresponds here to complete terminal delay.

Annex A (informative):

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Change history** | | | | | | | |
| **Date** | **Meeting** | **TDoc** | **CR** | **Rev** | **Cat** | **Subject/Comment** | **New version** |
| 2019-04 | SA4#103 | S4-190426 |  |  |  | Initial version | 0.0.1 |
| 2024-05 | SA4#128 | S4-241235 |  |  |  | Definition of initial clauses and revision of structure | 0.1.0 |
| 2024-05 | SA4#128 | S4-241333 |  |  |  | Inclusion of preliminary requirements | 0.2.0 |
| 2024-06 | SA4-e (AH) Audio SWG post 128 | SA4aA240032 |  |  |  | Additional requirements, improvements and editorial changes | 0.3.0 |
| 2024-06 |  |  |  |  |  | Version 1.0.0 created by MCC | 1.0.0 |
| 2024-06 |  |  |  |  |  | Version 18.0.0 created by MCC upon TSG approval | 18.0.0 |