ETSI TS 126 255 V18.1.0 (2024-07)

LTE;
5G;
Codec for Immersive Voice and Audio Services
- Error concealment of lost packets
(3GPP TS 26.255 version 18.1.0 Release 18)

Reference
RTS/TSGS-0426255vi10

Keywords
5G,LTE

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - APE 7112B
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° w061004871

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (https://ipr.etsi.org/).

Pursuant to the ETSI Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP™** and **LTE™** are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM**® and the GSM logo are trademarks registered and owned by the GSM Association.

Legal Notice

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

The present document may refer to technical specifications or reports using their 3GPP identities. These shall be interpreted as being references to the corresponding ETSI deliverables.

The cross reference between 3GPP and ETSI identities can be found under https://webapp.etsi.org/key/queryform.asp.

Modal verbs terminology

In the present document "**shall**", "**shall not**", "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Contents

Intellectual Property Rights 2
Legal Notice 2
Modal verbs terminology 2
Foreword 4
1 Scope 6
2 References 6
3 Definitions of terms, symbols and abbreviations 6
3.1 Terms 6
3.2 Symbols 6
3.3 Abbreviations 6
4 General 7
5 Error concealment in the core decoder 8
6 Error concealment per audio format 8
7 SID frame concealment operation 9
8 Error concealment for IVAS split rendering 9
Annex A (informative): Change history 10
History 11

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.

1 Scope

The present document defines a frame loss concealment procedure, also termed frame substitution and muting procedure, which is executed by the Immersive Voice and Audio Services (IVAS) decoder when one or more frames (speech or audio or SID frames) are unavailable for decoding due to e.g. packet loss, corruption of a packet or late arrival of a packet.

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 26.445: "Codec for Enhanced Voice Services (EVS); Detailed Algorithmic Description".

[3] 3GPP TS 26.447: "Codec for Enhanced Voice Services (EVS); Error Concealment of Lost Packets".

[4] 3GPP TS 26.253: "Codec for Immersive Voice and Audio Services - Detailed Algorithmic Description incl. RTP payload format and SDP parameter definitions".

[5] 3GPP TS 26.249: " Immersive Audio for Split Rendering Scenarios; Detailed Algorithmic Description of Split Rendering Functions".

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].

Further IVAS codec specific definitions are found in clause 3.1 of [4].

3.2 Symbols

Void

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].

BFI	Bad Frame Indicator
CPE	Channel Pair Element
DFT	Discrete Fourier Transform

ECU	Error Concealment Unit
EVS	Enhanced Voice Services
HQ MDCT	High Quality MDCT mode
IVAS	Immersive Voice and Audio Services
MASA	Metadata-Assisted Spatial Audio
MC	Multi-channel Audio
MCT	Multi-channel Coding Tool
MDCT	Modified Discrete Cosine Transform
OMASA	Objects and Metadata-Assisted Spatial Audio
OSBA	Objects and Scene-Based Audio
PLC	Packet Loss Concealment
SBA	Scene-Based Audio
SCE	Single Channel Element

4 General

Packet loss concealment serves to ensure the availability of useful audio output when valid packets are unavailable to the decoder. These losses are typically a result of impaired channel conditions like transmission errors or network congestion. The aim is to synthesize a substitution of the decoded audio represented by the lost packet, to prepare for a potential future packet loss, and to handle the transition from the concealment operation back to the decoded audio. The latter is also referred to as recovery operation. An overview of the IVAS codec's decoder operation is given in clause 6.1 of [4], where Figure 6.1-1 shows the functional structure of the decoder. To complement the picture with the packet loss concealment functionality, Figure 1 below shows the packet loss concealment (PLC) operation of the decoder. A major part of the PLC resides in the core decoding tools, where the audio decoding is mainly handled by the core decoder based on EVS [2]. The Single Channel Elements (SCE) decoder comprises one core-decoder, the Channel Pair Elements (CPE) comprises one or two core-decoders and the Multichannel Coding Tool (MCT) comprises joint decoding using multiple core-decoders, all including associated PLC methods. For the Low Frequency Effect (LFE) channel of multichannel audio, an LFE decoder with associated PLC method is available. Spatial metadata including spatial coding parameters are reconstructed in the spatial parameter decoders or by the associated PLC methods of the respective spatial audio formats. Spatial audio output is finally generated by a scene decoder, upmixer and renderer based on the reconstructed transport channels and the reconstructed spatial metadata. In case of a missing or corrupted packet, a bad frame indicator (BFI) is input to the decoding tools, activating the PLC operation. Notably, scene decoding, upmixing and rendering processing are independent of a bad frame indicator.

The IVAS split rendering feature exposes a further interface, the interface between the entity carrying out pre-rendering and encoding into the intermediate audio representation and the end-device doing post-rendering. To cope with potential transmission errors on that interface, the decoder of the intermediate audio representation features packet loss concealment techniques besides the actual decoding scheme. Note that the IVAS specific split rendering functionality including packet-loss concealment is mostly described in TS 26.253 [4] whereas more generic split rendering functionality is specified in TS 26.249 [5].

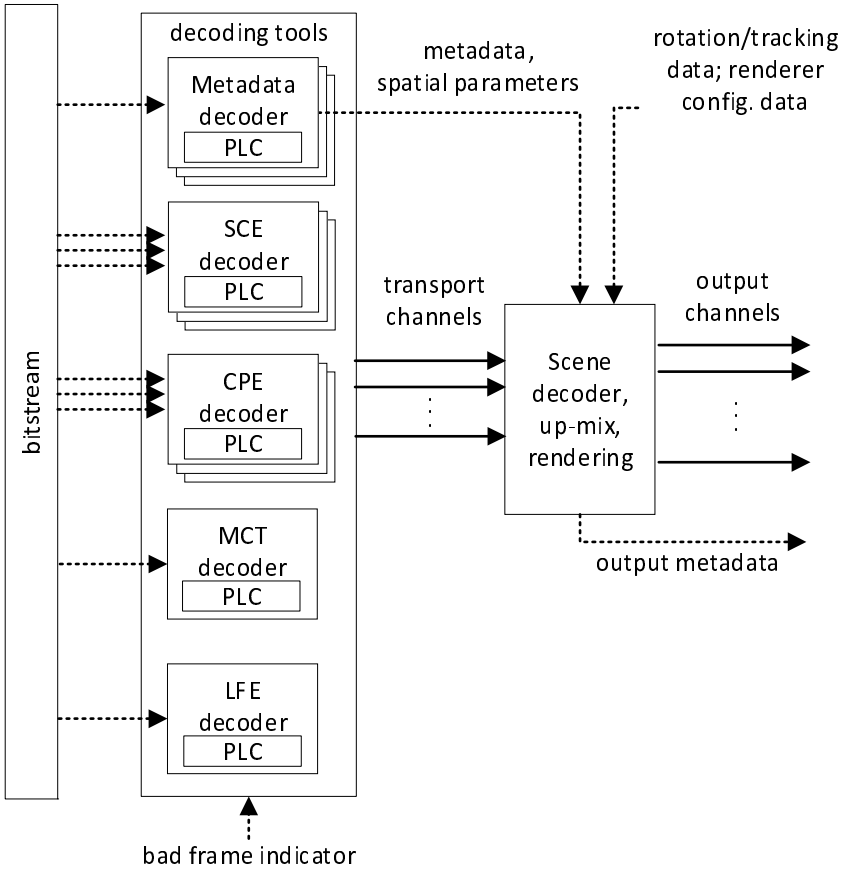

Figure 1: Overview of error concealment operation

5 Error concealment in the core decoder

Since IVAS is based on the Codec for Enhanced Voice Services (EVS) [2], the main functionality of the core-codec is inherited from EVS. This includes the error concealment operations as described in [3], and for mono operation the functionality is implemented in a bit-exact manner. IVAS provides a few enhancements of the core-coder on top of EVS.

The following clauses in [4] describe the enhancements that have been made for the error concealment operation for the IVAS core-coder:

- PLC Method selection in HQ MDCT error concealment can be found in [4], clause 6.2.2.3.4.
- Phase ECU enhancements can be found in [4], clause 6.2.2.3.5.

6 Error concealment per audio format

In addition to the mono operation with EVS compatibility, IVAS supports stereo, Independent Streams with metadata (ISM), multi-channel audio (MC), scene-based audio (Ambisonics or SBA), metadata assisted spatial audio (MASA) and combinations of objects with MASA (OMASA) and combination of objects with scene-based audio (OSBA). To handle the variation in audio formats across the supported range of input audio channels and bit rates, several dedicated encoding and decoding modules are employed. The general principle is that the parameters are recycled from the previously decoded frame, but there may also be further concealment operators performed on the parameters. The following clauses in [4] describe the error concealment operations within each of modules decoding the various audio formats.

- MCT PLC can be found in [4], clause 6.2.3.4.10.

- DFT-based stereo parameter error concealment can be found in [4], clause 6.3.2.3.10.
- PLC in MDCT-based stereo can be found in [4], clause 6.3.3.7.
- PLC in the SBA format decoder can be found in [4], clause 6.4.8.
- PLC in the MASA format decoder can be found in [4], clause 6.5.5.
- PLC in the ISM format decoder can be found in [4], clause 6.6.5.
- LFE channel PLC within the MC format decoder can be found in [4], clause 6.7.1.7.
- McMASA mode PLC within the MC format decoder can be found in [4], clause 6.7.2.5.
- ParamMC mode PLC within the MC format decoder can be found in [4], clause 6.7.3.7.
- Discrete MC mode PLC within the MC format decoder can be found in [4], clause 6.7.5.2.
- PLC in the OSBA format decoder can be found in [4], clause 6.8.3.
- PLC in the OMASA format decoder can be found in [4], clause 6.9.8.

7 SID frame concealment operation

In the case of the loss of an SID frame, the comfort noise will be generated based on the last received SID frame.

8 Error concealment for IVAS split rendering

The intermediate audio format of the IVAS split rendering feature comprises coded pose correction metadata and coded binaural audio. The binaural audio may be encoded using the LCLD coding format or the LC3plus coding format. The respective PLC schemes are described in [4] as follows:

- PLC for pose correction metadata is described in [4], clause 7.6.3.5.
- PLC for LCLD binaural audio coding is described in [4], clause 7.6.4.4.
- PLC for LC3plus binaural audio coding is described in [4], clause 7.6.4.6.

NOTE: In case binaural audio is transmitted using the PCM interface, no PLC method is provided. An implementor needs to make sure that suitable corresponding methods are available.

Annex A (informative): Change history

Change history							
Date	**Meeting**	**TDoc**	**CR**	**Rev**	**Cat**	**Subject/Comment**	**New version**
2023-11	SA4#126	S4-231870	-			First version presented to SA4 for information.	v0.0.1
2023-11	SA4#126	S4-231968	-			Submitted to SA4 plenary for information.	v0.1.0
2023-12	SA#102	SP-231298				Version 1.0.0 created by MCC	v1.0.0
2024-01	SA4#127	S4-240225				Version 1.0.1 presented to SA4 for agreement	v1.0.1
2024-01	SA4#127	S4-240378				Version 1.1.0 presented to SA4 plenary for agreement	v1.1.0
2024-03	SA#103	SP-240031				Version 2.0.0 created by MCC	2.0.0
2024-03							18.0.0
2024-06	SA#104	SP-240693				Adding ISAR track-a split rendering feature	18.1.0
2024-06	SA#104					Change of spec title as approved by TSG SA in SP-240917	18.1.0

History

Document history		
V18.0.0	May 2024	Publication
V18.1.0	July 2024	Publication