

ETSI TS 126 094 V13.0.0 (2016-01)



**Digital cellular telecommunications system (Phase 2+);
Universal Mobile Telecommunications System (UMTS);
LTE;**

**Mandatory speech codec speech processing functions;
Adaptive Multi-Rate (AMR) speech codec;
Voice Activity Detector (VAD)
(3GPP TS 26.094 version 13.0.0 Release 13)**



Reference

RTS/TSGS-0426094vd00

Keywords

GSM,LTE,UMTS

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 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

The present document can be downloaded from:

<http://www.etsi.org/standards-search>

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the only prevailing document is the print of the Portable Document Format (PDF) version kept on a specific network drive within ETSI Secretariat.

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other ETSI documents is available at

<http://portal.etsi.org/tb/status/status.asp>

If you find errors in the present document, please send your comment to one of the following services:

<https://portal.etsi.org/People/CommiteeSupportStaff.aspx>

Copyright Notification

No part may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm except as authorized by written permission of ETSI.

The content of the PDF version shall not be modified without the written authorization of ETSI.

The copyright and the foregoing restriction extend to reproduction in all media.

© European Telecommunications Standards Institute 2016.

All rights reserved.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are Trade Marks of ETSI registered for the benefit of its Members. **3GPP™** and **LTE™** are Trade Marks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners.

GSM® and the GSM logo are Trade Marks registered and owned by the GSM Association.

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is 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 IPR Policy, no investigation, 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.

Foreword

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, UMTS identities or GSM identities. These should be interpreted as being references to the corresponding ETSI deliverables.

The cross reference between GSM, UMTS, 3GPP and ETSI identities can be found under <http://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
Foreword.....	2
Modal verbs terminology.....	2
Foreword.....	4
1 Scope	5
2 References	5
3 Technical Description of VAD Option 1.....	5
3.1 Definitions, symbols and abbreviations.....	5
3.1.1 Definitions	5
3.1.2 Symbols	5
3.1.2.1 Variables	5
3.1.2.2 Constants.....	6
3.1.2.3 Functions.....	7
3.1.3 Abbreviations.....	7
3.2 General	7
3.3 Functional description	7
3.3.1 Filter bank and computation of sub-band levels	8
3.3.2 Pitch detection	10
3.3.3 Tone detection	10
3.3.4 Correlated Complex Signal Analysis (and detection).....	11
3.3.5 VAD decision	11
3.3.5.1 Hangover addition.....	12
3.3.5.2 Background noise estimation	14
4 Technical Description of VAD Option 2.....	16
4.1 Definitions, symbols and abbreviations.....	16
4.1.1 Definitions	16
4.1.2 Symbols	16
4.1.2.1 Variables	16
4.1.2.2 Constants.....	17
4.1.2.3 Functions.....	17
4.1.3 Abbreviations.....	18
4.2 General	18
4.3 Functional description	18
4.3.1 Frequency Domain Conversion	19
4.3.2 Channel Energy Estimator	19
4.3.3 Channel SNR Estimator.....	20
4.3.4 Voice Metric Calculation.....	20
4.3.5 Frame SNR and Long-Term Peak SNR Calculation.....	20
4.3.6 Negative SNR Sensitivity Bias	21
4.3.7 VAD Decision	21
4.3.8 Spectral Deviation Estimator	22
4.3.9 Sinewave Detection	22
4.3.10 Background Noise Update Decision	23
4.3.10 Background Noise Estimate Update	23
5 Computational details.....	24
Annex A (informative): Change history	25
History	26

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.

1 Scope

The present document specifies two alternatives for the Voice Activity Detector (VAD) to be used in the Discontinuous Transmission (DTX) as described in [3]. Implementors of mobile station and infrastructure equipment conforming to the AMR specifications can choose which of the two VAD options to implement. There are no interoperability factors associated with this choice.

The requirements are mandatory on any VAD to be used either in User Equipment (UE) or Base Station Systems (BSS)s that utilize the AMR speech codec.

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 TS 26.073: "Adaptive Multi-Rate (AMR); ANSI C source code".

[2] 3GPP TS 26.090: "Transcoding functions".

[3] 3 GPP TS 26.093: "Source Controlled Rate operation".

[4] ITU, The International Telecommunications Union, Blue Book, Vol. III, Telephone Transmission Quality, IXth Plenary Assembly, Melbourne, 14-25 November, 1988, Recommendation G.711, Pulse code modulation (PCM) of voice frequencies.

3 Technical Description of VAD Option 1

3.1 Definitions, symbols and abbreviations

3.1.1 Definitions

For the purposes of the present document, the following terms and definitions apply:

frame: time interval of 20 ms corresponding to the time segmentation of the speech transcoder

3.1.2 Symbols

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

3.1.2.1 Variables

bckr_est[n]	background noise estimate
burst_count	counts length of a speech burst, used by VAD hangover addition
hang_count	hangover counter, used by VAD hangover addition
complex_hang_count	hangover counter, used by CAD hangover addition