ETSI TS 101 376-6-6 V1.1.1 (2001-03)

Technical Specification

GEO-Mobile Radio Interface Specifications;
Part 6: Speech coding specifications;
Sub-part 6: Vocoder: Voice Activity Detection (VAD);
GMR-1 06.032

Reference

DTS/SES-001-06032

Keywords

DTX, GMR, GSM, GSO, interface, MES, mobile, MSS, radio, satellite, S-PCN, VAD

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

Individual copies of the present document can be downloaded from:
http://www.etsi.org

The present document may be made available in more than one electronic version or in print. In any case of existing or perceived difference in contents between such versions, the reference version is the Portable Document Format (PDF). In case of dispute, the reference shall be the printing on ETSI printers of the 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://www.etsi.org/tb/status/

If you find errors in the present document, send your comment to:
editor@etsi.fr

Contents

Intellectual property rights

The information pertaining to essential IPRs 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 (http://www.etsi.org/ipr).

The attention of ETSI has been drawn to the Intellectual Property Rights (IPRs) listed below which are, or may be, or may become, Essential to the present document. The IPR owner has undertaken to grant irrevocable licences, on fair, reasonable and non-discriminatory terms and conditions under these IPRs pursuant to the ETSI IPR Policy. Further details pertaining to these IPRs can be obtained directly from the IPR owner.

The present IPR information has been submitted to ETSI and 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.

IPRs:

Project	Company	Title	Country of Origin	Patent n°	Countries Applicable
TS 101 376 V1.1.1	Digital Voice Systems Inc		US	US 5,226,084	US
TS 101 376 V1.1.1	Digital Voice Systems Inc		US	US 5,715,365	US
TS 101 376 V1.1.1	Digital Voice Systems Inc		US	US 5,826,222	US
TS 101 376 V1.1.1	Digital Voice Systems Inc		US	US 5,754,974	US
TS 101 376 V1.1.1	Digital Voice Systems Inc		US	US 5,701,390	US

IPR Owner: Digital Voice Systems Inc
One Van de Graaff Drive Burlington,
MA 01803
USA

Contact: John C. Hardwick
Tel.: +1 781-270-1030
Fax: +1 781-270-0166

Project	Company	Title	Country of Origin	Patent n°	Countries Applicable
TS 101 376 V1.1.1	Ericsson Mobile Communication	Improvements in, or in relation to, equalisers	GB	GB 2 215 567	GB
TS 101 376 V1.1.1	Ericsson Mobile Communication	Power Booster	GB	GB 2 251 768	GB
TS 101 376 V1.1.1	Ericsson Mobile Communication	Receiver Gain	GB	GB 2 233 846	GB
TS 101 376 V1.1.1	Ericsson Mobile Communication	Transmitter Power Control for Radio Telephone System	GB	GB 2 233 517	GB

IPR Owner: Ericsson Mobile Communications (UK) Limited
The Keytech Centre, Ashwood Way
Basingstoke
Hampshire RG23 8BG
United Kingdom

Contact: John Watson
Tel.: +44 1256 864821

Project	Company	Title	Country of Origin	Patent n°	Countries Applicable
TS 101 376 V1.1.1	Hughes Network Systems		US	Pending	US

IPR Owner: Hughes Network Systems
11717 Exploration Lane
Germantown, Maryland 20876
USA

Contact: John T. Whelan
Tel: +1 301-428-7172
Fax: +1 301-428-2802

Project	Company	Title	Country of Origin	Patent n°	Countries Applicable
TS 101 376 V1.1.1	Lockheed Martin Global Telecommunic. Inc	2.4-to-3 KBPS Rate Adaptation Apparatus for Use in Narrowband Data and Facsimile Communication Systems	US	US 6,108,348	US
TS 101 376 V1.1.1	Lockheed Martin Global Telecommunic. Inc	Cellular Spacecraft TDMA Communications System with Call Interrupt Coding System for Maximizing Traffic ThroughputCellular Spacecraft TDMA Communications System with Call Interrupt Coding System for Maximizing Traffic Throughput	US	US 5,717,686	US
TS 101 376 V1.1.1	Lockheed Martin Global Telecommunic. Inc	Enhanced Access Burst for Random Access Channels in TDMA Mobile Satellite System	US	US 5,875,182	
TS 101 376 V1.1.1	Lockheed Martin Global Telecommunic. Inc	Spacecraft Cellular Communication System	US	US 5,974,314	US
TS 101 376 V1.1.1	Lockheed Martin Global Telecommunic. Inc	Spacecraft Cellular Communication System	US	US 5,974,315	US
TS 101 376 V1.1.1	Lockheed Martin Global Telecommunic. Inc	Spacecraft Cellular Communication System with Mutual Offset High-argin Forward Control Signals	US	US 6,072,985	US
TS 101 376 V1.1.1	Lockheed Martin Global Telecommunic. Inc	Spacecraft Cellular Communication System with Spot Beam Pairing for Reduced Updates	US	US 6,118,998	US

IPR Owner: Lockheed Martin Global Telecommunications, Inc.
900 Forge Road
Norristown, PA. 19403
USA

Contact: R.F. Franciose
Tel.: +1 610.354.2535
Fax: +1 610.354.7244

Foreword

This Technical Specification (TS) has been produced by ETSI Technical Committee Satellite Earth Stations and Systems (SES).

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

Version 1.m.n

where:

- the third digit (n) is incremented when editorial only changes have been incorporated in the specification;
- the second digit (m) is incremented for all other types of changes, i.e. technical enhancements, corrections, updates, etc.

The present document is part 6, sub-part 6 of a multi-part deliverable covering the GEO-Mobile Radio Interface Specifications, as identified below:

Part 1: "General specifications";

Part 2: "Service specifications";

Part 3: "Network specifications";

Part 4: "Radio interface protocol specifications";

Part 5: "Radio interface physical layer specifications";

Part 6: "Speech coding specifications";

Sub-part 1: "Speech Processing Functions; GMR-1 06.001";

Sub-part 2: "Vocoder: Speech Transcoding; GMR-1 06.010";

Sub-part 3: "Vocoder: Substitution and Muting of Lost Frames; GMR-1 06.011";

Sub-part 4: "Vocoder: Comfort Noise Aspects; GMR-1 06.012";

Sub-part 5: "Vocoder: Discontinuous Transmission (DTX); GMR-1 06.031";

Sub-part 6: "Vocoder: Voice Activity Detection (VAD); GMR-1 06.032";

Part 7: "Terminal adaptor specifications".

Introduction

GMR stands for GEO (Geostationary Earth Orbit) Mobile Radio interface, which is used for mobile satellite services (MSS) utilizing geostationary satellite(s). GMR is derived from the terrestrial digital cellular standard GSM and supports access to GSM core networks.

Due to the differences between terrestrial and satellite channels, some modifications to the GSM standard are necessary. Some GSM specifications are directly applicable, whereas others are applicable with modifications. Similarly, some GSM specifications do not apply, while some GMR specifications have no corresponding GSM specification.

Since GMR is derived from GSM, the organization of the GMR specifications closely follows that of GSM. The GMR numbers have been designed to correspond to the GSM numbering system. All GMR specifications are allocated a unique GMR number as follows:

GMR-n xx.zyy

where:

xx.0yy (z=0) is used for GMR specifications that have a corresponding GSM specification. In this case, the numbers xx and yy correspond to the GSM numbering scheme.

xx.2yy (z=2) is used for GMR specifications that do not correspond to a GSM specification. In this case, only the number xx corresponds to the GSM numbering scheme and the number yy is allocated by GMR.

n denotes the first (n=1) or second (n=2) family of GMR specifications.

A GMR system is defined by the combination of a family of GMR specifications and GSM specifications as follows:

- If a GMR specification exists it takes precedence over the corresponding GSM specification (if any). This precedence rule applies to any references in the corresponding GSM specifications.

NOTE: Any references to GSM specifications within the GMR specifications are not subject to this precedence rule. For example, a GMR specification may contain specific references to the corresponding GSM specification.

- If a GMR specification does not exist, the corresponding GSM specification may or may not apply. The applicability of the GSM specifications is defined in GMR-1 01.201 [6].

1 Scope

The present document defines the VAD algorithm that is used in the GMR-1 system to facilitate discontinuous transmission (DTX) as described in GMR-1 06.031 [3]. In addition, the present document includes test methods that must be used to verify that any VAD algorithms used in the GMR-1 system are compliant with 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 and/or edition number or version number) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies.

[1] GMR-1 01.004 (ETSI TS 101 376-1-1): "GEO-Mobile Radio Interface Specifications; Part 1: General specifications; Sub-part 1: Abbreviations and acronyms; GMR-1 01.004".

[2] GMR-1 06.012 (ETSI TS 101 376-6-4): "GEO-Mobile Radio Interface Specifications; Part 6: Speech coding specifications; Sub-part 4: Vocoder: Comfort Noise Aspects; GMR-1 06.012".

[3] GMR-1 06.031 (ETSI TS 101 376-6-5): "GEO-Mobile Radio Interface Specifications; Part 6: Speech coding specifications; Sub-part 5: Vocoder: Discontinuous Transmission (DTX); GMR-1 06.031".

[4] GMR-1 05.008 (ETSI TS 101 376-5-6): "GEO-Mobile Radio Interface Specifications; Part 5: Radio interface physical layer specifications; Sub-part 6: Radio Subsystem Link Control; GMR-1 05.008".

[5] GMR-1 06.001 (ETSI TS 101 376-6-1): "GEO-Mobile Radio Interface Specifications; Part 6: Speech coding specifications; Sub-part 1: Speech Processing Functions; GMR-1 06.001".

[6] GMR-1 01.201 (ETSI TS 101 376-1-2): "GEO-Mobile Radio Interface Specifications; Part 1: General specifications; Sub-part 2: Introduction to the GMR-1 Family; GMR-1 01.201".

3 Definitions and abbreviations

3.1 Definitions

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

Voice Activity Detection (VAD): method of classifying short segments of speech as either "voice" or "background noise." The decision is based upon comparing the current level and spectral characteristics of the input signal with that of a typical level and spectral characteristics

Comfort Noise Insertion (CNI): method of synthesizing low-level noise on the receive side during breaks in voice transmission. To increase the perceived voice quality, the synthesized noise has characteristics that are similar to the background noise present on the transmit side

Forward Error Correction (FEC): method of introducing redundancy to binary data that allows for the detection and/or correction of errors introduced during transmission of that data

V/UV(Voiced/Unvoiced): each spectral band is declared either "voiced" or "unvoiced", depending upon the amount of periodic energy in that band. This voicing decision is frequently referred to as a V/UV decision

frame: data representing a full 40 msec of continuous data input to or output from the vocoder. The frame data may consist of model parameters, quantized bits, FEC encoded channel data, or speech samples at various points in the vocoder

subframe: data representing 10 msec of continuous data input to or output from the vocoder, or the result of processing that data through various points in the vocoder. For example, "The second subframe of model parameters is passed to the quantizer" is a valid use of the term as is "The decoder outputs one subframe of 8 kHz speech samples"

subframe number: each frame is composed of four consecutive subframes that are each assigned a subframe number. The first, second, third, and fourth subframes within a frame are assigned subframe numbers 0, 1, 2, and 3 respectively

quantizer-frame: data representing the 20 msec of continuous vocoder data that is formed by combining subframes 0 and 1 or subframes 2 and 3

quantizer-frame number: each frame is composed of two consecutive quantizer-frames that are each assigned a quantizer frame number. The first and second quantizer-frames within a frame are assigned quantizer-frame numbers 0 and 1 respectively

voice frame: 40-msec frame that contains some voice data but no tone data. It may also contain comfort noise data

SID frame: (Silence Descriptor): 20-msec frame that contains only comfort noise data. No voice or tone data may be present in a SID frame

tone frame: 40-msec frame that contains tone data. It may also contain voice data or comfort noise data

dbm0: power in dBm referred to or measured at a zero transmission level point (0TLP)

3.2 Abbreviations

Abbreviations used in the present document are listed in GMR-1 01.004 [1].

4 General

The typical signal input to a vocoder has periods of speech and periods without speech. The periods that are lacking speech typically contain background noise. The function of the VAD algorithm is to indicate whether or not each 40-msec frame that is output by the voice encoder contains voice data. The VAD algorithm outputs a binary flag for each frame that indicates either case.

5 Functional description of VAD

5.1 VAD overview

The VAD algorithm must distinguish between frames that contain speech and frames that contain only noise, which is a relatively easy task when the input signal is clean speech (i.e. not noisy). When the signal-to-noise ratio is decreased it becomes more difficult to distinguish between speech frames and noise frames. Low signal-to-noise ratios are commonly encountered in mobile environments.

The multifrequency VAD algorithm embedded in the vocoder operates by maintaining an adaptive model of the spectral characteristics of the background noise. The algorithm assumes that the background noise is somewhat stationary. The background noise estimates are updated on a frame-by-frame basis when background noise is present. The VAD algorithm also maintains an estimate of the average voice energy, which is updated only when voice is present.

A spectral error metric, ε_D, is computed by comparing the spectral characteristics of the adaptive noise model with those of the current frame. The spectral error metric, average voice energy, and the total energy of the current frame are then used to distinguish between "voice" frames and "noise" frames. In addition there is a fixed threshold, ε_{in}, which may be set at the user interface. If the input signal exceeds this level for any frame it will be declared a "voice" frame. The nominal setting for ε_{in} is -25 dBm0.

To prevent clipping of low-level speech, the VAD algorithm uses a VAD holdover counter that ensures that frames for a short period after a voice burst are also declared voice. The integrated VAD algorithm takes advantage of the vocoder's inherent delay to provide an early indication of the onset of voice activity.

The flag output by the VAD algorithm is used by the DTX (see GMR-1 06.031 [3]) to reduce the channel rate during periods that have no voice activity.

During periods in which the VAD flag is deactivated, the vocoder outputs a SID frame, which is used for comfort noise insertion (see GMR-1 06.012 [2]) that models the background noise characteristics present at the transmit end. The DTX (see GMR-1 06.031 [3] and GMR-1 05.008 [4]) controls how SID frames are transmitted.

5.2 Vocoder interface notes relevant to VAD

This clause discusses some implementation and interface issues that relate to VAD operation.

The vocoder outputs the voice_active VAD flag twice for each frame (once per quantizer-frame) output. The voice_active flag for quantizer-frames 0 and 1 must be logically OR-ed together to obtain a VAD decision for the overall frame. All frames having voice activity in either quantizer-frame or both quantizer-frames must be transmitted.

The VAD level threshold, ε_{in}, is configurable at the vocoder interface. A recommended value is -25 dBm0.

History

Document history		
V1.1.1	March 2001	Publication