# SLOVENSKI STANDARD
# SIST EN 300 730 V8.0.1:2003

**01-december-2003**

---

**Digitalni celiļ ni telekomunikacijski sistem (faza 2+) ï Detektor govornih dejavnosti (VAD) v kanalih za izboljġani polnohitrostni (EFR) govorni promet (GSM 06.82, razliļ ica 8.0.1, izdaja 1999)**

Digital cellular telecommunications system (Phase 2+) (GSM); Voice Activity Detector (VAD) for Enhanced Full Rate (EFR) speech traffic channels (GSM 06.82 version 8.0.1 Release 1999)



**Ta slovenski standard je istoveten z:** **EN 300 730 Version 8.0.1**

---

**ICS:**

| | | |
|---|---|---|
| 33.070.50 | Globalni sistem za mobilno telekomunikacijo (GSM) | Global System for Mobile Communication (GSM) |

**SIST EN 300 730 V8.0.1:2003** **en**

# ETSI EN 300 730 V8.0.1 (2000-11)

*European Standard (Telecommunications series)*

# Digital cellular telecommunications system (Phase 2+); Voice Activity Detector (VAD) for Enhanced Full Rate (EFR) speech traffic channels (GSM 06.82 version 8.0.1 Release 1999)



GLOBAL SYSTEM FOR MOBILE COMMUNICATIONS

Reference
REN/SMG-110682Q8

Keywords
Digital cellular telecommunications system,
Global System for Mobile communications (GSM)

***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

***Copyright Notification***

# Contents

Intellectual Property Rights ..... 4

Foreword ..... 4

1 Scope ..... 5

2 References ..... 5

3 Definitions, symbols and abbreviations ..... 5
3.1 Definitions ..... 5
3.2 Symbols ..... 6
3.2.1 Variables ..... 6
3.2.2 Constants ..... 6
3.2.3 Functions ..... 7
3.3 Abbreviations ..... 7

4 General ..... 7

5 Functional description ..... 7
5.1 Overview and principles of operation ..... 7
5.2 Algorithm description ..... 8
5.2.1 Adaptive filtering and energy computation ..... 9
5.2.2 ACF averaging ..... 9
5.2.3 Predictor values computation ..... 9
5.2.4 Spectral comparison ..... 10
5.2.5 Information tone detection ..... 10
5.2.6 Threshold adaptation ..... 11
5.2.7 VAD decision ..... 13
5.2.8 VAD hangover addition ..... 13
5.2.9 Periodicity detection ..... 13

6 Computational description overview ..... 14
6.1 VAD modules ..... 14
6.2 Pseudo-floating point arithmetic ..... 14

**Annex A (informative): Simplified block filtering operation ..... 16**

**Annex B (informative): Pole frequency calculation ..... 17**

**Annex C (informative): Change Request History ..... 18**

History ..... 19

# 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 (http://www.etsi.org/ipr).

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 European Standard (Telecommunications series) has been produced by ETSI Technical Committee Special Mobile Group (SMG).

The present document specifies the Voice Activity Detector (VAD) to be used in the Discontinuous Transmission (DTX) for Enhanced Full Rate (EFR) speech traffic channels within the digital cellular telecommunications system.

The contents of the present document is subject to continuing work within SMG and may change following formal SMG approval. Should SMG modify the contents of the present document it will be re-released with an identifying change of release date and an increase in version number as follows:

Version 8.x.y

where:



8 indicates Release 1999 of GSM Phase 2+.



x the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.

y the third digit is incremented when editorial only changes have been incorporated in the specification.

| National transposition dates | |
|---|---|
| Date of adoption of this EN: | 3 November 2000 |
| Date of latest announcement of this EN (doa): | 28 February 2001 |
| Date of latest publication of new National Standard or endorsement of this EN (dop/e): | 31 August 2001 |
| Date of withdrawal of any conflicting National Standard (dow): | 31 August 2001 |

# 1 Scope

The present document specifies the Voice Activity Detector (VAD) to be used in the Discontinuous Transmission (DTX) as described in GSM 06.81 [5] Discontinuous transmission (DTX) for Enhanced Full Rate (EFR) speech traffic channels.

The requirements are mandatory on any VAD to be used either in GSM Mobile Stations (MS)s or Base Station Systems (BSS)s that utilize the enhanced full-rate speech traffic channel.

# 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.
- A non-specific reference to an ETS shall also be taken to refer to later versions published as an EN with the same number.
- For this Release 1999 document, references to GSM documents are for Release 1999 versions (version 8.x.y).



[1] GSM 01.04: "Digital cellular telecommunications system (Phase 2+); Abbreviations and acronyms".



[2] GSM 06.53: "Digital cellular telecommunications system (Phase 2+); ANSI-C code for the GSM Enhanced Full Rate (EFR) speech codec".

[3] GSM 06.54: "Digital cellular telecommunications system (Phase 2+); Test vectors for the GSM Enhanced Full Rate (EFR) speech codec".

[4] GSM 06.60: "Digital cellular telecommunications system (Phase 2+); Enhanced Full Rate (EFR) speech transcoding".

[5] GSM 06.81: "Digital cellular telecommunications system (Phase 2+); Discontinuous transmission (DTX) for Enhanced Full Rate (EFR) speech traffic channels".

# 3 Definitions, symbols and abbreviations

## 3.1 Definitions

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

**noise:** signal component resulting from acoustic environmental noise.

**mobile environment:** any environment in which mobile stations may be used.

## 3.2 Symbols

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

### 3.2.1 Variables

| | |
|---|---|
| aav1 | filter predictor values, see clause 5.2.3 |
| acf | the ACF vector which is calculated in the speech encoder (GSM 06.60 [4]) |
| adaptcount | secondary hangover counter, see clause 5.2.6 |
| av0 | averaged ACF vector, see clause 5.2.2 |
| av1 | a previous value of av0, see clause 5.2.2 |
| burstcount | speech burst length counter, see clause 5.2.8 |
| den | denominator of left hand side of equation 8 in annex B, see clause 5.2.5 |
| difference | difference between consecutive values of dm, see clause 5.2.4 |
| dm | spectral distortion measure, see clause 5.2.4 |
| hangcount | primary hangover counter, see clause 5.2.8 |
| lagcount | number of subframes in current frame meeting periodicity criterion, see clause 5.2.9 |
| lastdm | previous value of dm, see clause 5.2.4 |
| lags | the open loop long term predictor lags for the two halves of the speech encoder frame (GSM 06.60 [4]) |
| num | numerator of left hand side of equation 8 in annex B, see clause 5.2.5 |
| oldlagcount | previous value of lagcount, see clause 5.2.9 |
| prederr | fourth order short term prediction error, see clause 5.2.5 |
| ptch | Boolean flag indicating the presence of a periodic signal component, see clause 5.2.9 |
| pvad | energy in the current filtered signal frame, see clause 5.2.1 |
| rav1 | autocorrelation vector obtained from av1, see clause 5.2.3 |
| rc | the first four unquantized reflection coefficients calculated in the speech encoder (GSM 06.60 [4]) |
| rvad | autocorrelation vector of the adaptive filter predictor values, see clause 5.2.6 |
| smallag | difference between consecutive lag values, see clause 5.2.9 |
| stat | Boolean flag indicating that the frequency spectrum of the input signal is stationary, see clause 5.2.4 |
| thvad | adaptive primary VAD threshold, see clause 5.2.6 |
| tone | Boolean flag indicating the presence of an information tone, see clause 5.2.5 |
| vadflag | Boolean VAD decision with hangover included, see clause 5.2.8 |
| veryoldlagcount | previous value of oldlagcount, see clause 5.2.9 |
| vvad | Boolean VAD decision before hangover, see clause 5.2.7 |



### 3.2.2 Constants

| | |
|---|---|
| adp | number of frames of hangover for secondary VAD, see clause 5.2.6 |
| burstconst | minimum length of speech burst to which hangover is added, see clause 5.2.8 |
| dec | determines rate of decrease in adaptive threshold, see clause 5.2.6 |
| fac | determines steady state adaptive threshold, see clause 5.2.6 |
| frames | number of frames over which av0 and av1 are calculated, see clause 5.2.2 |
| freqth | threshold for pole frequency decision, see clause 5.2.5 |
| hangconst | number of frames of hangover for primary VAD, see clause 5.2.8 |
| inc | determines rate of increase in adaptive threshold, see clause 5.2.6 |
| lthresh | lag difference threshold for periodicity decision, see clause 5.2.9 |
| margin | determines upper limit for adaptive threshold, see clause 5.2.6 |
| nthresh | frame count threshold for periodicity decision, see clause 5.2.9 |
| plev | lower limit for adaptive threshold, see clause 5.2.6 |
| predth | threshold for short term prediction error, see clause 5.2.5 |
| pth | energy threshold, see clause 5.2.6 |
| thresh | decision threshold for evaluation of stat flag, see clause 5.2.4 |

### 3.2.3 Functions

| | |
|---|---|
| + | addition |
| - | subtraction |
| * | multiplication |
| / | division |
| \| x \| | absolute value of x |
| AND | Boolean AND |
| OR | Boolean OR |
| b | |
| MULT(x(i)) | the product of the series x(i) for i=a to b |
| i=a | |
| b | |
| SUM(x(i)) | the sum of the series x(i) for i=a to b |
| i=a | |

## 3.3 Abbreviations

For the purposes of the present document, the following abbreviations apply:

| | |
|---|---|
| ACF | Autocorrelation function |
| ANSI | American National Standards Institute |
| DTX | Discontinuous Transmission |
| LTP | Long Term Predictor |
| TX | Transmission |
| VAD | Voice Activity Detector |



For abbreviations not given in this clause, see GSM 01.04 [1].

# 4 General



The function of the VAD is to indicate whether each 20 ms frame produced by the speech encoder contains speech or not. The output is a Boolean flag (vadflag) which is used by the Transmit (TX) DTX handler defined in GSM 06.81 [5].

The present document is organized as follows.

Clause 5 describes the principles of operation of the VAD. Clause 6 provides an overview of the computational description of the VAD. The computational details necessary for the fixed point implementation of the VAD algorithm are given in the form of ANSI C program contained in GSM 06.53 [2].

The verification of the VAD is based on the use of digital test sequences which are described in GSM 06.54 [3].

# 5 Functional description

The purpose of this clause is to give the reader an understanding of the principles of operation of the VAD, whereas GSM 06.53 [2] contains the fixed point computational description of the VAD. In the case of discrepancy between the two descriptions, the description in GSM 06.53 [2] will prevail.

## 5.1 Overview and principles of operation

The function of the VAD is to distinguish between noise with speech present and noise without speech present. This is achieved by comparing the energy of a filtered version of the input signal with a threshold. The presence of speech is indicated whenever the threshold is exceeded.

The detection of speech in a mobile environment is difficult due to the low speech/noise ratios which are encountered, particularly in moving vehicles. To increase the probability of detecting speech the input signal is adaptively filtered (see clause 5.2.1) to reduce its noise content before the voice activity decision is made (see clause 5.2.7).

The frequency spectrum and level of the noise may vary within a given environment as well as between different environments. It is therefore necessary to adapt the input filter coefficients and energy threshold at regular intervals as described in clause 5.2.6.

## 5.2 Algorithm description

The block diagram of the VAD algorithm is shown in figure 1. The individual blocks are described in the following clauses. The variables shown in the block diagram are described in table 1.

**Table 1: Description of variables in figure 1**

| Var | Description |
|---|---|
| acf | The ACF vector which is calculated in the speech encoder (GSM 06.60 [4]). |
| av0 | Averaged ACF vector. |
| av1 | A previous value of av0. |
| lags | The open loop long term predictor lags for the two halves of the speech encoder frame (GSM 06.60 [4]). |
| ptch | Boolean flag indicating the presence of a periodic signal component. |
| pvad | Energy in the current filtered signal frame. |
| rav1 | Autocorrelation vector obtained from av1. |
| rc | The first four reflection coefficients calculated in the speech encoder (GSM 06.60 [4]). |
| rvad | Autocorrelation vector of the adaptive filter predictor values. |
| stat | Boolean flag indicating that the frequency spectrum of the input signal is stationary. |
| thvad | Adaptive primary VAD threshold. |
| tone | Boolean flag indicating the presence of an information tone. |
| vadflag | Boolean VAD decision with hangover included. |
| vvad | Boolean VAD decision before hangover. |



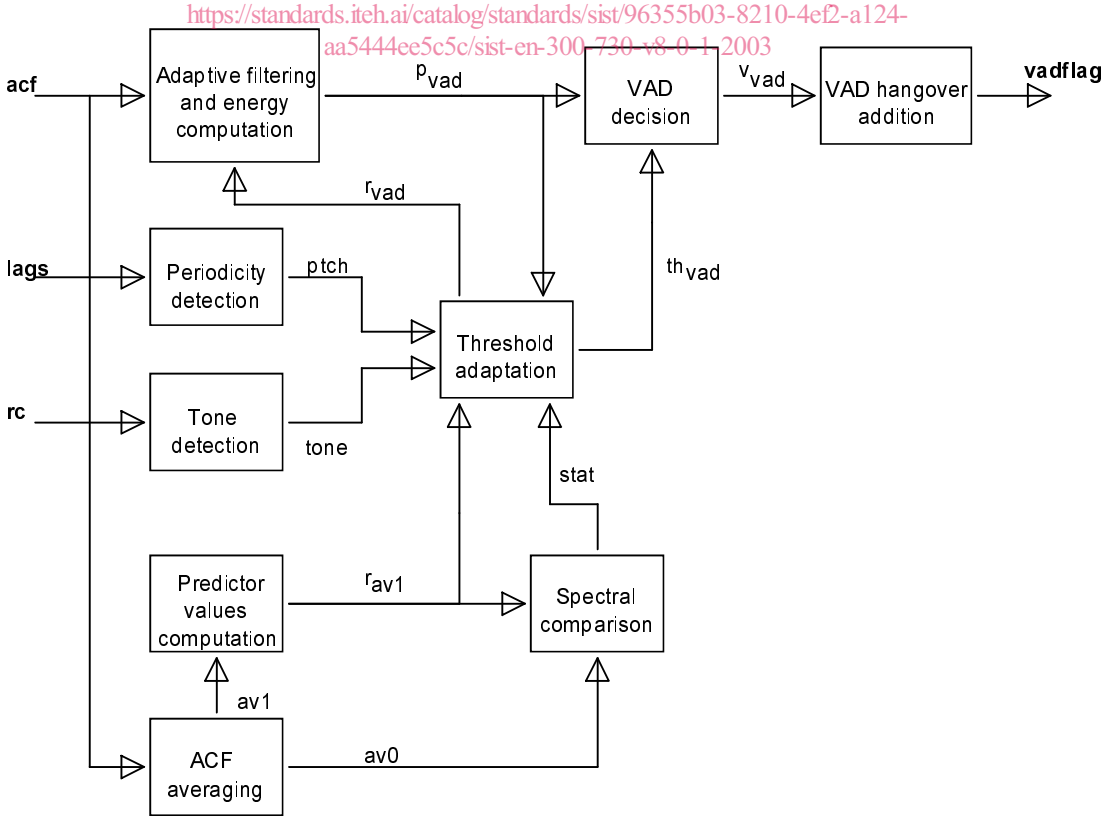


**Figure 1: Functional block diagram of the VAD**

### 5.2.1 Adaptive filtering and energy computation

The energy in the current filtered signal frame (pvad) is computed as follows:

$$pvad = rvad[0] * acf[0] + 2 * \sum_{i=1}^{8} (rvad[i] * acf[i]) \quad (1)$$

This corresponds to performing an 8th order block filtering on the filtered input samples to the speech encoder. This is explained in annex A.

### 5.2.2 ACF averaging

Spectral characteristics of the input signal have to be obtained using blocks that are larger than one 20 ms frame. This is done by averaging the ACF (autocorrelation function) values for several consecutive frames. The averaging is given by the following equations:

$$av0\{n\}[i] = \sum_{j=0}^{frames-1} (acf\{n-j\}[i]) \quad ; i = 0..8 \quad (2)$$

$$av1\{n\}[i] = av0\{n-frames\}[i] \quad ; i = 0..8 \quad (3)$$

where (n) represents the current frame, (n-1) represents the previous frame. The values of constants are given in table 2.

**Table 2: Constants and variables for ACF averaging**

| Constant | Value | Variable | Initial value |
|---|---|---|---|
| frames | 4 | previous ACF's, av0 & av1 | All set to 0 |



### 5.2.3 Predictor values computation



The filter predictor values aav1 are obtained from the autocorrelation values av1 according to the equation:

$$a = R^{-1} p \quad (4)$$

where:

$$R = \begin{bmatrix} av1[0] & av1[1] & av1[2] & av1[3] & av1[4] & av1[5] & av1[6] & av1[7] \\ av1[1] & av1[0] & av1[1] & av1[2] & av1[3] & av1[4] & av1[5] & av1[6] \\ av1[2] & av1[1] & av1[0] & av1[1] & av1[2] & av1[3] & av1[4] & av1[5] \\ av1[3] & av1[2] & av1[1] & av1[0] & av1[1] & av1[2] & av1[3] & av1[4] \\ av1[4] & av1[3] & av1[2] & av1[1] & av1[0] & av1[1] & av1[2] & av1[3] \\ av1[5] & av1[4] & av1[3] & av1[2] & av1[1] & av1[0] & av1[1] & av1[2] \\ av1[6] & av1[5] & av1[4] & av1[3] & av1[2] & av1[1] & av1[0] & av1[1] \\ av1[7] & av1[6] & av1[5] & av1[4] & av1[3] & av1[2] & av1[1] & av1[0] \end{bmatrix}$$