SLOVENSKI STANDARD
SIST EN 300 730 V6.0.1:2003

01-december-2003

Digitalni celiļ ni telekomunikacijski sistem (faza 2+) ï Detektor govornih dejavnosti (VAD) v kanalih za izboljģani polnohitrostni (EFR) govorni promet (GSM 06.82, razliļ ica 6.0.1, izdaja 1997)

Digital cellular telecommunications system (Phase 2+) (GSM); Voice Activity Detector (VAD) for Enhanced Full Rate (EFR) speech traffic channels (GSM 06.82 version 6.0.1 Release 1997)

Ta slovenski standard je istoveten z: **EN 300 730 Version 6.0.1**

ICS:

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

SIST EN 300 730 V6.0.1:2003 **en**

EN 300 730 V6.0.1 (1999-06)

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 6.0.1 Release 1997)

GLOBAL SYSTEM FOR MOBILE COMMUNICATIONS

Reference

DEN/SMG-020682Q6 (6rc0300o.PDF)

Keywords

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

ETSI

Postal address

F-06921 Sophia Antipolis Cedex - FRANCE

Office address

650 Route des Lucioles - Sophia Antipolis
Valbonne - 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

Internet

secretariat@etsi.fr
Individual copies of this ETSI deliverable
can be downloaded from
http://www.etsi.org
If you find errors in the present document, send your
comment to: editor@etsi.fr

Contents

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 SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available **free of charge** 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 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 the 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 6.x.y

where:

- 6 indicates Release 1997 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.

Proposed national transposition dates	
Date of adoption of this EN:	05 June 1999
Date of latest announcement of this EN (doa):	30 September 1999
Date of latest publication of new National Standard or endorsement of this EN (dop/e):	31 March 2000
Date of withdrawal of any conflicting National Standard (dow):	31 March 2000

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.

[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 definitions apply:

noise: The 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 subclause 5.2.3

acf the ACF vector which is calculated in the speech encoder (GSM 06.60 [4])

adaptcount secondary hangover counter, see subclause 5.2.6

av0 averaged ACF vector, see subclause 5.2.2

av1 a previous value of av0, see subclause 5.2.2

burstcount speech burst length counter, see subclause 5.2.8

den denominator of left hand side of equation 8 in annex B, see subclause 5.2.5

difference difference between consecutive values of dm, see subclause 5.2.4

dm spectral distortion measure, see subclause 5.2.4

hangcount primary hangover counter, see subclause 5.2.8

lagcount number of subframes in current frame meeting periodicity criterion, see subclause 5.2.9

lastdm previous value of dm, see subclause 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 subclause 5.2.5

oldlagcount previous value of lagcount, see subclause 5.2.9

prederr fourth order short term prediction error, see subclause 5.2.5

ptch Boolean flag indicating the presence of a periodic signal component, see subclause 5.2.9

pvad energy in the current filtered signal frame, see subclause 5.2.1

rav1 autocorrelation vector obtained from av1, see subclause 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 subclause 5.2.6

smallag difference between consecutive lag values, see subclause 5.2.9

stat Boolean flag indicating that the frequency spectrum of the input signal is stationary, see subclause 5.2.4

thvad adaptive primary VAD threshold, see subclause 5.2.6

tone Boolean flag indicating the presence of an information tone, see subclause 5.2.5

vadflag Boolean VAD decision with hangover included, see subclause 5.2.8

veryoldlagcount previous value of oldlagcount, see subclause 5.2.9

vvad Boolean VAD decision before hangover, see subclause 5.2.7

3.2.2 Constants

adp number of frames of hangover for secondary VAD, see subclause 5.2.6

burstconst minimum length of speech burst to which hangover is added, see subclause 5.2.8

dec determines rate of decrease in adaptive threshold, see subclause 5.2.6

fac determines steady state adaptive threshold, see subclause 5.2.6

frames number of frames over which av0 and av1 are calculated, see subclause 5.2.2

freqth	threshold for pole frequency decision, see subclause 5.2.5
hangconst	number of frames of hangover for primary VAD, see subclause 5.2.8
inc	determines rate of increase in adaptive threshold, see subclause 5.2.6
lthresh	lag difference threshold for periodicity decision, see subclause 5.2.9
margin	determines upper limit for adaptive threshold, see subclause 5.2.6
nthresh	frame count threshold for periodicity decision, see subclause 5.2.9
plev	lower limit for adaptive threshold, see subclause 5.2.6
predth	threshold for short term prediction error, see subclause 5.2.5
pth	energy threshold, see subclause 5.2.6
thresh	decision threshold for evaluation of stat flag, see subclause 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

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 subclause, 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 subclause 5.2.1) to reduce its noise content before the voice activity decision is made (see subclause 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 subclause 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 subclauses. 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.

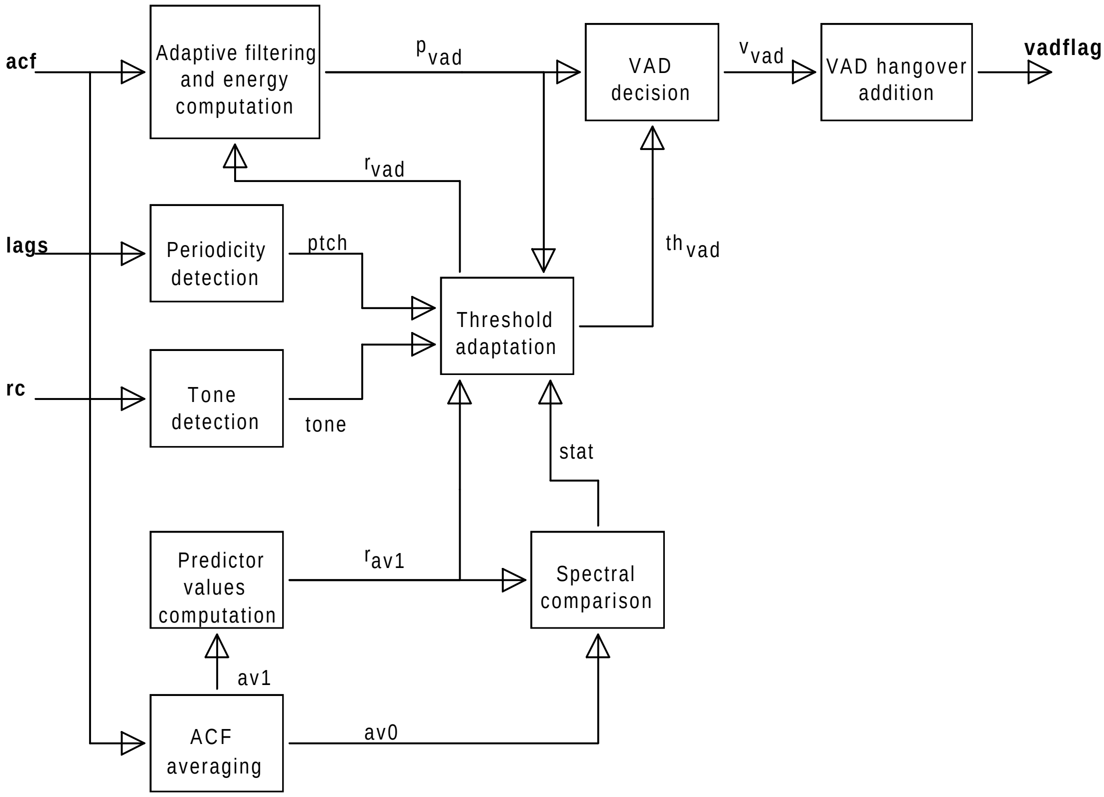

Figure 1: Functional block diagram of the VAD

5.2.1 Adaptive filtering and energy computation

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

```
                                 8
pvad = rvad[0] * acf[0] + 2 * SUM (rvad[i] * acf[i])                                (1)
                                i=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:

```
         frames-1
av0{n}[i] = SUM  (acf{n-j}[i])          ; i = 0..8                                 (2)
            j=0

av1{n}[i] = av0{n-frames}[i]          ; i = 0..8                                   (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