



SLOVENSKI STANDARD
SIST EN 300 969 V7.0.1:2003
01-december-2003

8 [[]HJb]`WV] b]`hY`ca i b]_UW`g_]`g]ghYa `fZUnU&ŽL`È; cj cf`g`dc`cj] bc` \]f`cg`h`c`È
DfY_cX]fUb`Y`df]`dc`cj] b]` \]f`cg`h`c`cj] GA` \$` *` &`ž`f`Un`]]WU+` '\$` %`ž`nXU`U`%` - , Ł

Digital cellular telecommunications system (Phase 2+) (GSM); Half rate speech; Half rate speech transcoding (GSM 06.20 version 7.0.1 Release 1998)

iTeh STANDARD PREVIEW
(standards.iteh.ai)

Ta slovenski standard je istoveten z: **EN 300 969 Version 7.0.1**
<https://standards.iteh.ai/catalog/standards/sist/e0d69577-ebb4-46eb-91c1-9089bcccf796/sist-en-300-969-v7-0-1-2003>

ICS:

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

SIST EN 300 969 V7.0.1:2003 **en**

iTeh STANDARD PREVIEW
(standards.iteh.ai)

[SIST EN 300 969 V7.0.1:2003](https://standards.iteh.ai/catalog/standards/sist/e0d69577-cb64-46eb-91c1-9089bccccf796/sist-en-300-969-v7-0-1-2003)

<https://standards.iteh.ai/catalog/standards/sist/e0d69577-cb64-46eb-91c1-9089bccccf796/sist-en-300-969-v7-0-1-2003>

ETSI EN 300 969 V7.0.1 (2000-01)

European Standard (Telecommunications series)

**Digital cellular telecommunications system (Phase 2+);
Half rate speech;
Half rate speech transcoding
(GSM 06.20 version 7.0.1 Release 1998)**

iTeh STANDARD PREVIEW
(standards.iteh.ai)

GSM®
GLOBAL SYSTEM FOR
MOBILE COMMUNICATIONS

[SIST EN 300 969 V7.0.1:2003](https://standards.iteh.ai/catalog/standards/sist/e0d69577-cb64-46eb-91c1-9089bccc796/sist-en-300-969-v7-0-1-2003)

<https://standards.iteh.ai/catalog/standards/sist/e0d69577-cb64-46eb-91c1-9089bccc796/sist-en-300-969-v7-0-1-2003>



ReferenceREN/SMG-110620Q7

Keywords

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

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

Important notice

This ETSI deliverable 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.

Copyright Notification

No part may be reproduced except as authorized by written permission.
The copyright and the foregoing restriction extend to reproduction in all media.

© European Telecommunications Standards Institute 2000.
All rights reserved.

Contents

Intellectual Property Rights	5
Foreword.....	5
1 Scope	7
2 References	7
3 Definitions, symbols and abbreviations	7
3.1 Definitions	7
3.2 Symbols	9
3.3 Abbreviations	10
4 Functional description of the GSM half rate speech codec	11
4.1 GSM half rate speech encoder	11
4.1.1 High-pass filter	13
4.1.2 Segmentation	14
4.1.3 Fixed Point Lattice Technique (FLAT).....	14
4.1.4 Spectral quantization.....	15
4.1.4.1 Autocorrelation Fixed Point Lattice Technique (AFLAT)	15
4.1.5 Frame energy calculation and quantization.....	17
4.1.6 Soft interpolation of the spectral parameters.....	17
4.1.7 Spectral noise weighting filter coefficients	18
4.1.8 Long Term Predictor lag determination.....	19
4.1.8.1 Open loop long term search initialization.....	20
4.1.8.2 Open loop lag search.....	21
4.1.8.3 Frame lag trajectory search (Mode $\neq 0$).....	26
4.1.8.4 Voicing mode selection	28
4.1.8.5 Closed loop lag search	28
4.1.9 Harmonic noise weighting	29
4.1.10 Code search algorithm	31
4.1.10.1 Decorrelation of filtered basis vectors.....	32
4.1.10.2 Fast search technique	33
4.1.11 Multimode gain vector quantization	34
4.1.11.1 Coding GS and P0.....	34
4.2 GSM half rate speech decoder.....	37
4.2.1 Excitation generation	37
4.2.2 Adaptive pitch prefilter	38
4.2.3 Synthesis Filter.....	38
4.2.4 Adaptive spectral postfilter	38
4.2.5 Updating decoder states	40
5 Homing sequences.....	40
5.1 Functional description	40
5.2 Definitions	40
5.3 Encoder homing.....	40
5.4 Decoder homing	41
5.5 Encoder home state.....	41
5.6 Decoder home state	41
Annex A (normative): Codec parameter description	42
A.1 Codec parameter description.....	42
A.1.1 MODE.....	42
A.1.2 R0	42
A.1.3 LPC1 - LPC3.....	43
A.1.4 LAG_1 - LAG_4	43
A.1.5 CODE _x _1 - CODE _x _4.....	43
A.1.6 GSP0_1 - GSP0_4.....	43

A.2 Basic coder parameters	43
Annex B (normative): Order of occurrence of the codec parameters over Abis	44
Annex C (informative): Bibliography.....	45
Annex D (informative): Change Request History	46
History	47

iTeh STANDARD PREVIEW (standards.iteh.ai)

[SIST EN 300 969 V7.0.1:2003](https://standards.iteh.ai/catalog/standards/sist/e0d69577-cb64-46eb-91c1-9089bccccf796/sist-en-300-969-v7-0-1-2003)

<https://standards.iteh.ai/catalog/standards/sist/e0d69577-cb64-46eb-91c1-9089bccccf796/sist-en-300-969-v7-0-1-2003>

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 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 speech codec to be used for the GSM half rate channel for the digital cellular telecommunications system. The present document is part of a series covering the half rate speech traffic channels as described below:

- | | |
|------------------|--|
| GSM 06.02 | "Digital cellular telecommunications system (Phase 2+); Half rate speech; Half rate speech processing functions". |
| GSM 06.06 | "Digital cellular telecommunications system (Phase 2+); Half rate speech; ANSI-C code for the GSM half rate speech codec". |
| GSM 06.07 | "Digital cellular telecommunications system (Phase 2+); Half rate speech; Test sequences for the GSM half rate speech codec". |
| GSM 06.20 | "Digital cellular telecommunications system (Phase 2+); Half rate speech; Half rate speech transcoding". |
| GSM 06.21 | "Digital cellular telecommunications system (Phase 2+); Half rate speech; Substitution and muting of lost frames for half rate speech traffic channels". |
| GSM 06.22 | "Digital cellular telecommunications system (Phase 2+); Half rate speech; Comfort noise aspects for half rate speech traffic channels". |
| GSM 06.41 | "Digital cellular telecommunications system (Phase 2+); Half rate speech; Discontinuous Transmission (DTX) for half rate speech traffic channels". |
| GSM 06.42 | "Digital cellular telecommunications system (Phase 2+); Half rate speech; Voice Activity Detector (VAD) for half rate speech traffic channels". |

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 7.x.y

where:

- 7 indicates Release 1998 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:	31 December 1999
Date of latest announcement of this EN (doa):	31 March 2000
Date of latest publication of new National Standard or endorsement of this EN (dop/e):	30 September 2000
Date of withdrawal of any conflicting National Standard (dow):	30 September 2000

**iTeh STANDARD PREVIEW
(standards.iteh.ai)**

[SIST EN 300 969 V7.0.1:2003](https://standards.iteh.ai/catalog/standards/sist/e0d69577-cb64-46eb-91c1-9089bccccf796/sist-en-300-969-v7-0-1-2003)

<https://standards.iteh.ai/catalog/standards/sist/e0d69577-cb64-46eb-91c1-9089bccccf796/sist-en-300-969-v7-0-1-2003>

1 Scope

The present document specifies the speech codec to be used for the GSM half rate channel. It also specifies the test methods to be used to verify that the codec implementation complies with the present document.

The requirements are mandatory for the codec to be used either in GSM Mobile Stations (MS)s or Base Station Systems (BSS)s that utilize the half rate GSM 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 1998 document, references to GSM documents are for Release 1998 versions (version 7.x.y).

- [1] GSM 06.02: "Digital cellular telecommunications system (Phase 2+); Half rate speech; Half rate speech processing functions"
- [2] GSM 06.06: "Digital cellular telecommunications system (Phase 2+); Half rate speech; ANSI-C code for the GSM half rate speech codec"
- [3] GSM 06.07: "Digital cellular telecommunications system (Phase 2+); Half rate speech; Test sequences for the GSM half rate speech codec".

3 Definitions, symbols and abbreviations

3.1 Definitions

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

adaptive codebook: The adaptive codebook is derived from the long term filter state. The lag value can be viewed as an index into the adaptive codebook.

adaptive pitch prefilter: In the GSM half rate speech decoder, this filter is applied to the excitation signal to enhance the periodicity of the reconstructed speech. Note that this is done prior to the application of the short term filter.

adaptive spectral postfilter: In the GSM half rate speech decoder, this filter is applied to the output of the short term filter to enhance the perceptual quality of the reconstructed speech.

allowable lags: The set of lag values which may be coded by the GSM half rate speech encoder and transmitted to the GSM half rate speech decoder. This set contains both integer and fractional values (see table 3).

analysis window: For each frame, the short term filter coefficients are computed using the high pass filtered speech samples within the analysis window. The analysis window is 170 samples in length, and is centered about the last 100 samples in the frame.

basis vectors: A set of M, M1, or M2 vectors of length Ns used to generate the VSELP codebook vectors. These vectors are not necessarily orthogonal.

closed loop lag search: A process of determining the near optimal lag value from the weighted input speech and the long term filter state.

closed loop lag trajectory: For a given frame, the sequence of near optimal lag values whose elements correspond to each of the four subframes as determined by the closed loop lag search.

codebook: A set of vectors used in a vector quantizer.

Codeword (OR Code): An M, M1, or M2 bit symbol indicating the vector to be selected from a VSELP codebook.

Delta (LAG) code: A four bit code indicating the change in lag value for a subframe relative to the previous subframe's coded lag. For frames in which the long term predictor is enabled (MODE 1, 2, or 3), the lag for subframe 1 is independently coded using eight bits, and delta codes are used for subframes 2, 3, and 4.

direct form coefficients: One of the formats for storing the short term filter parameters. All filters which are used to modify speech samples use direct form coefficients.

fractional lags: A set of lag values having sub-sample resolution. Note that not every fractional lag value considered in the GSM half rate speech encoder is an allowable lag value.

frame: A time interval equal to 20 ms, or 160 samples at an 8 kHz sampling rate.

harmonic noise weighting filter: This filter exploits the noise masking properties of the spectral peaks which occur at harmonics of the pitch frequency by weighting the residual error less in regions near the pitch harmonics and more in regions away from them. Note that this filter is only used when the long term filter is enabled (MODE = 1, 2 or 3).

high pass filter: This filter is used to de-emphasize the low frequency components of the input speech signal.

integer lags: A set of lag values having whole sample resolution.

interpolating filter: An FIR filter used to estimate sub-sample resolution samples, given an input sampled with integer sample resolution.

lag: The long term filter delay. This is typically the pitch period, or a multiple or sub-multiple of it.

long term filter: This filter is used to generate the periodic component in the excitation for the current subframe. This filter is only enabled for MODE = 1, 2 or 3.

LPC coefficients: Linear Predictive Coding (LPC) coefficients is a generic descriptive term for describing the short term filter coefficients.

open loop lag search: A process of estimating the near optimal lag directly from the weighted speech input. This is done to narrow the range of lag values over which the closed loop lag search shall be performed.

open loop lag trajectory: For a given frame, the sequence of near optimal lag values whose elements correspond to the four subframes as determined by the open loop lag search.

reflection coefficients: An alternative representation of the information contained in the short term filter parameters.

residual: The output signal resulting from an inverse filtering operation.

short term filter: This filter introduces, into the excitation signal, short term correlation which models the impulse response of the vocal tract.

soft interpolation: A process wherein a decision is made for each frame to use either interpolated or uninterpolated short term filter parameters for the four subframes in that frame.

soft interpolation bit: A one bit code indicating whether or not interpolation of the short term parameters is to be used in the current frame.

spectral noise weighting filter: This filter exploits the noise masking properties of the formants (vocal tract resonances) by weighting the residual error less in regions near the formant frequencies and more in regions away from them.

subframe: A time interval equal to 5 ms, or 40 samples at an 8 kHz sampling rate.

vector quantization: A method of grouping several parameters into a vector and quantizing them simultaneously.

GSP0 vector quantizer: The process of vector quantization, its intermediate parameters (GS and P0) for the coding of the excitation gains β and γ .

VSEL P codebook: Vector-Sum Excited Linear Predictive (VSEL P) codebook, used in the GSM half rate speech coder, wherein each codebook vector is constructed as a linear combination of the fixed basis vectors.

zero input response: The output of a filter due to all past inputs, i.e. due to the present state of the filter, given that an input of zeros is applied.

zero state response: The output of a filter due to the present input, given that no past inputs have been applied, i.e. given the state information in the filter is all zeroes.

3.2 Symbols

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

$A(z)$	Short term spectral filter.
α_i	The LPC coefficients.
$b_L(n)$	The output of the long term filter state (adaptive codebook) for lag L.
β	The long term filter coefficient.
$C(z)$	Second weighting filter.
$e(n)$	Weighted error signal
$f_j(i)$	The coefficients of the j^{th} phase of the 10th order interpolating filter used to evaluate candidate fractional lag values; i ranges from 0 to P_f-1 .
$g_j(i)$	The coefficients of the j^{th} phase of the 6th order interpolating filter used to interpolate C's and G's as well as fractional lags in the harmonic noise weighting; i ranges from 0 to P_g-1 .
γ	The gain applied to the vector(s) selected from the VSEL P codebook(s).
H	A M2 bit code indicating the vector to be selected from the second VSEL P codebook (when operating in mode 0).
I	A M or M1 bit code indicating the vector to be selected from one of the two first VSEL P codebooks.
L	The long term filter lag value.
L_{max}	142 (samples), the maximum possible value for the long term filter lag.
L_{min}	21 (samples), the minimum possible value for the long term filter lag.
M	9, the number of basis vectors, and the number of bits in a codeword, for the VSEL P codebook used in modes 1, 2, and 3.
M1	7, the number of basis vectors, and the number of bits in a codeword, for the first VSEL P codebook used in mode 0.
M2	7, the number of basis vectors, and the number of bits in a codeword, for the second VSEL P codebook used in mode 0.
MODE	A two bit code indicating the mode for the current frame (see annex A).
N_A	170, the length of the analysis window. This is the number of high pass filtered speech samples used to compute the short term filter parameters for each frame.
N_F	160, the number of samples per frame (at a sampling rate of 8 kHz).
N_p	10, the short term filter order.
N_s	40, the number of samples per subframe (at a sampling rate of 8 kHz).
P1	6, the number of bits in the prequantizer for the r1 - r3 vector quantizer.
P2	5, the number of bits in the prequantizer for the r4 - r6 vector quantizer.
P3	4, the number of bits in the prequantizer for the r7 - r10 vector quantizer.
P_f	The order of one phase of an interpolating filter used to evaluate candidate fractional lag values. P_f equals 10 for $j \neq 0$ and equal to 1 for $j = 0$.
P_g	The order of one phase of an interpolating filter, $f_j(n)$, used to interpolate C's and G's as well as fractional lags in the harmonic noise weighting, P_g equals 6.
pitch	The time duration between the glottal pulses which result when the vocal chords vibrate during speech production.
Q1	11, the number of bits in the r1 - r3 reflection coefficient vector quantizer.
Q2	9, the number of bits in the r4 - r6 reflection coefficient vector quantizer.
Q3	8, the number of bits in the r7 - r10 reflection coefficient vector quantizer.

R0	A five bit code used to indicate the energy level in the current frame.
$r(n)$	The long term filter state (the history of the excitation signal); $n < 0$
$r_L(n)$	The long term filter state with the adaptive codebook output for lag L appended.
$s'(n)$	Synthesized speech.
$W(z)$	Spectral weighting filter.
λ_{hnw}	The harmonic noise weighting filter coefficient.
ξ	The adaptive pitch prefilter coefficient.
$\lceil x \rceil$	Ceiling function: the largest integer y where $y < x + 1,0$.
$\lfloor x \rfloor$	Floor function: the largest integer y where $y \leq x$.
$\sum_{i=j}^K x(i)$	Summation: $x(j)+x(j+1)+\dots+x(K)$.
$\prod_{i=j}^K x(i)$	Product: $x(j)(x(j+1))\dots(x(K))$
$\max(x,y)$	Find the larger of two numbers x and y.
$\min(x,y)$	Find the smaller of two numbers x and y.
$\text{round}(x)$	Round the non-integer x to the closest integer y: $y = \lfloor x + 0,5 \rfloor$; $y = x + 0,5$.

3.3 Abbreviations

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

AFLAT	Autocorrelation Fixed point Lattice Technique
CELP	Code Excited Linear Prediction
FLAT	Fixed Point Lattice Technique
LTP	Long Term Predictor
SST	Spectral Smoothing Technique
VSELP	Vector-Sum Excited Linear Prediction

<https://standards.iteh.ai/catalog/standards/sist/e0d69577-cb64-46eb-91c1-9089bccc796/sist-en-300-969-v7-0-1-2003>

4 Functional description of the GSM half rate speech codec

The GSM half rate codec uses the VSELP (Vector-Sum Excited Linear Prediction) algorithm. The VSELP algorithm is an analysis-by-synthesis technique and belongs to the class of speech coding algorithms known as CELP (Code Excited Linear Prediction).

The GSM half rate codec's encoding process is performed on a 20 ms speech frame at a time. A speech frame of the sampled speech waveform is read and based on the current waveform and the past history of the waveform, the codec encoder derives 18 parameters that describe it. The parameters extracted are grouped into the following three general classes:

- energy parameters (R0 and GSP0);
- spectral parameters (LPC and INT_LPC);
- excitation parameters (LAG and CODE).

These parameters are quantized into 112 bits for transmission as described in annex A and their order of occurrence over Abis is given in annex B.

The GSM half rate codec is an analysis-by-synthesis codec, therefore the speech decoder is primarily a subset of the speech encoder. The quantized parameters are decoded and a synthetic excitation is generated using the energy and excitation parameters. The synthetic excitation is then filtered to provide the spectral information resulting in the generation of the synthesized speech (see figure 1).

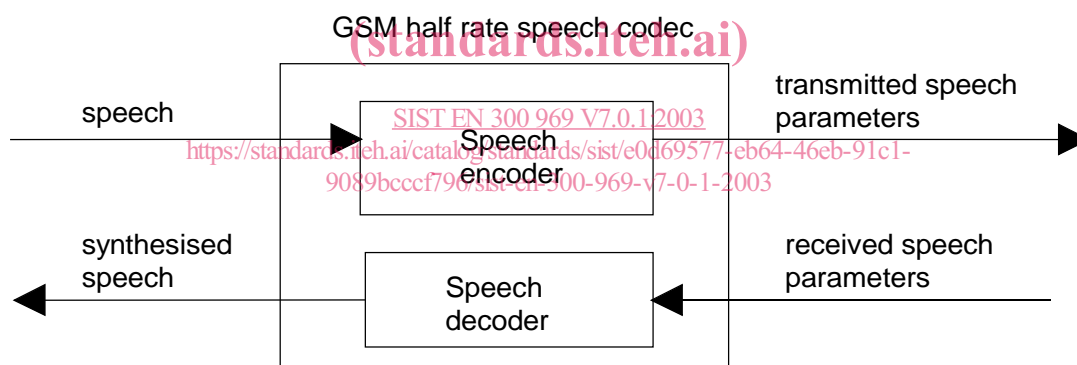


Figure 1: Block diagram of the GSM half rate speech codec

The ANSI-C code that describes the GSM half rate speech codec is given in GSM 06.06 [2] and the test sequences in GSM 06.07 [3] (see clause 5 for the codec homing test sequences).

4.1 GSM half rate speech encoder

The GSM half rate speech encoder uses an analysis by synthesis approach to determine the code to use to represent the excitation for each subframe. The codebook search procedure consists of trying each codevector as a possible excitation for the Code Excited Linear Predictive (CELP) synthesizer. The synthesized speech $s'(n)$ is compared against the input speech and a difference signal is generated. This difference signal is then filtered by a spectral weighting filter, $W(z)$, (and possibly a second weighting filter, $C(z)$) to generate a weighted error signal, $e(n)$. The power in $e(n)$ is computed. The codevector which generates the minimum weighted error power is chosen as the codevector for that subframe. The spectral weighting filter serves to weight the error spectrum based on perceptual considerations. This weighting filter is a function of the speech spectrum and can be expressed in terms of the α parameters of the short term (spectral) filter.