ETSI/TC SMG
Released by : ETSI/PT 12
Release date: February 1992

RELEASE NOTE

Recommendation GSM 06.01

Sppech processing functions: General description

Previously distributed version : 3.0.0 (Release 1/90)
New Released version February 92 : **3.0.0 (Release 92, Phase 1)**

---

1. Reason for changes

No changes since the previously distributed version.

ETSI-GSM

Technical

Specification

GSM 06.01

Version 3.0.0

European digital cellular telecommunication system (phase 1);

Speech Processing Functions:

General Description

ETSI

PREFATORY NOTE

ETSI has constituted stable and consistent documents which give specifications for the implementation of the European Cellular Telecommunications System. Historically, these documents have been identified as "GSM recommendations".

Some of these recommendations may subsequently become Interim European Telecommunications Standards (I-ETSs) or European Telecommunications Standards (ETSs), whilst some continue with the status of ETSI-GSM Technical Specifications. These ETSI-GSM Technical Specifications are for editorial reasons still referred to as GSM recommendations in some current GSM documents.

The numbering and version control system is the same for ETSI-GSM Technical Specifications as for "GSM recommendations".

Recommendation: GSM 06.01

Title: Speech processing functions: General description

Date: February, 1992

<u>Language of original</u>: English

<u>Number of pages</u>: 8

Detailed list of contents
****

1. INTRODUCTION
============

This recommendation is an introduction to the 06 series of the GSM recommendations dealing with the speech processing functions in the GSM system. A general overview of the speech processing parts is given with reference to the recommendations where each part is specified in detail.

Fig 1-1 presents a reference configuration where the various speech processing functions are identified. In this figure, the relevant recommendations for each function are also indicated.

In fig 1-1, the audio parts including analogue to digital and digital to analog conversion are included to show the complete speech path between the audio input/output in the Mobile Station and the digital interface to the PSTN. The detailed specification of the audio parts are, however, contained in rec GSM 11.15 and also rec GSM 03.50 (Transmission plan aspects of the speech service in the Pan-European PLMN system). These aspects are only considered in the GSM 06-series to the extent that the performance of the audio parts affect the performance of the speech transcoder.

2. FULL RATE SPEECH TRANSCODING (REC GSM 06.10)
===

As shown in fig 1-1, the speech encoder takes its input as a 13 bit uniform PCM signal either from the audio part of the mobile station or on the network side, from the PSTN via an 8 bit/A-law to 13 bit uniform PCM conversion. The encoded speech at the output of the speech encoder is delivered to the channel coding function defined in rec GSM 05.03 to produce an encoded block consisting of 456 bits leading to a gross bit rate of 22.8 kbit/s.

In the receive direction, the inverse operations take place.

Recommendation GSM 06.10 describes the detailed mapping between input blocks of 160 speech samples in 13 bit uniform PCM format to encoded blocks of 260 bits and from encoded blocks of 260 bits to output blocks of 160 reconstructed speech samples. The sampling rate is 8000 sample/s leading to an average bit rate for the encoded bit stream of 13 kbit/s. The coding scheme is the so-called Regular Pulse Excitation - Long Term prediction - Linear Predictive Coder, here-after referred to as RPE-LTP.

The recommendation describes the codec down to the bit level, thus enabling the verification of compliance to the recommendation to a high degree of confidence by use of a set of digital test sequences. These test sequences are also described and are available on floppy disks.
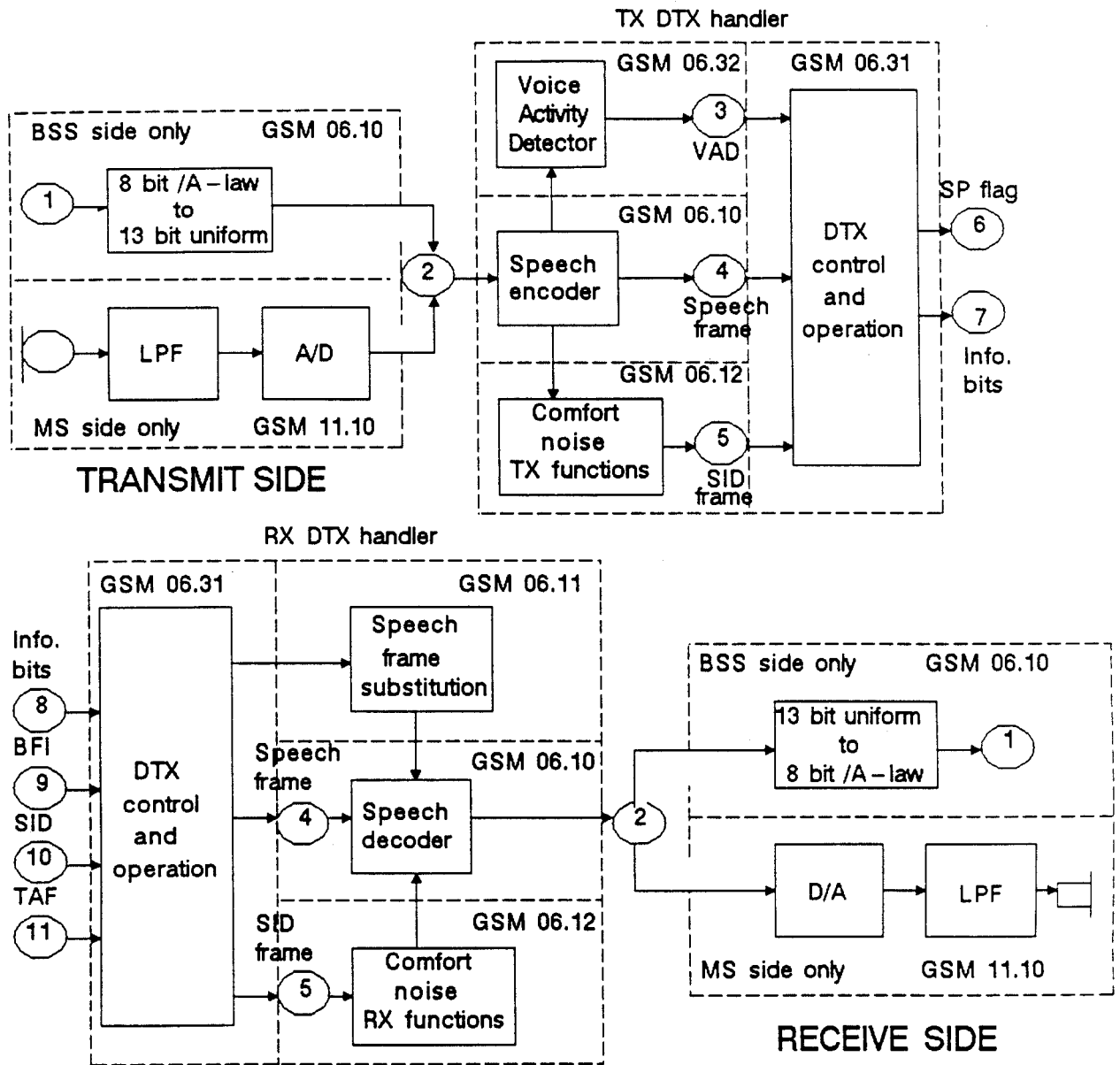
(1) 8 bit /A−law PCM (CCITT rec G.711), 8000 samples/s

(2) 13 bit uniform PCM, 8000 samples/s

(3) Voice activity flag

(4) Encoded speech frame, 50 samples/s, 260 bits/frame

(5) Silence Descriptor (SID) frame, 260 bits/frame

(6) Speech flag, indicates whether information bits are speech or SID information

(7) Information bits delivered to the radio subsystem

(8) Information bits received from the radio subsystem

(9) Bad Frame Indication (BFI) flag

(10) Silence Descriptor (SID) flag

(11) Time Alignment Flag (TAF), marks the position of the SID frame within the SACCH multiframe

Fig 1 − 1. Overview of audio processing functions

3. FULL RATE DISCONTINUOUS TRANSMISSION (DTX) (REC GSM 06.31)
===

During a normal conversation, the participants alternate so that, on the average, each direction of transmission is occupied about 50 % of the time. Discontinuous transmission (DTX) is a mode of operation where the transmitters are switched on only for those frames which contain useful information. This may be done for the following two purposes:

1) In the Mobile Station, battery life will be prolonged or a smaller battery could be used for a given operational duration.

2) The average interference level on the "air" is reduced, leading to better spectrum efficiency.

The overall DTX mechanism is implemented in the DTX handlers (TX and RX) described in recommendation GSM 06.31 and requires the following functions which are described in separate recommendations:

- A Voice Activity Detector on the transmit side

- Evaluation of the background acoustic noise on the transmit side, in order to transmit characteristic parameters to the receive side.

- Generation on the receive side of a similar noise, called comfort noise, during periods where the radio transmission is cut.

The transmission of comfort noise information to the receive side is achieved by means of a special frame (Silence descriptor=SID). This frame is transmitted at the end of speech bursts and serves as an end of speech marker for the receive side. In order to update the comfort noise characteristics at the receive side, SID frames are transmitted at regular intervals also during speech pauses. This also serves the purpose of improving the measurement of the radio link quality by the radio subsystem.

For the overall DTX functionality, the DTX handlers interwork via various flags with the Radio Subsystem, which is in control of the actual transmitter keying on the TX side and which performs various pre-processing functions on the RX side. This is also described in rec GSM 06.31.

A common terminology used throughout the GSM 06-series of recommendations is also defined in this recommendation.

4. FULL RATE VOICE ACTIVITY DETECTION (VAD) (REC GSM 06.32)
==

The input to the VAD is a set of parameters computed by the full-rate speech encoder defined in Recommendation GSM 06.10. The VAD uses this information to decide whether each 20 ms speech coder frame contains speech or not. Note that the VAD flag is an input to TX DTX handler and does not control the transmitter keying directly.

The recommendation describes the VAD algorithm down to the bit level. The conventions used in the bit-exact specification are the same as those used in recommendation GSM 06.10. The verification of compliance to the recommendation is achieved by use of digital test sequences applied to the same interface as the test sequences for the speech codec. These test sequences are also described and are available on floppy disks.

5. FULL RATE COMFORT NOISE INSERTION (REC GSM 06.12)
==

When switching the transmission on and off during DTX operation, the effect would be a modulation of the background noise at the receiving end if no precautions were taken. When transmission is on, the background noise is transmitted together with the speech to the receiving end. As the speech burst ends, the connection is off and the perceived noise would drop to a very low level. This step modulation of noise is perceived as very annoying and may reduce the intelligibility of speech if presented to a listener without modification.

This so-called noise contrast effect is reduced in the GSM system by inserting an artificial noise, termed comfort noise, at the receiving end when speech is absent.

Rec GSM 06.12 deals with the detailed aspects of the comfort noise process:

- the evaluation of the acoustic background noise in the transmitter,

- the noise parameter encoding (SID frames) and decoding

- and the generation of comfort noise in the receiver.

The algorithm for updating the noise parameters during speech pauses is also defined.

The comfort noise mechanism is based on the full rate speech codec defined in rec GSM 06.10.

6. FULL RATE LOST SPEECH FRAME SUBSTITUTION AND MUTING (REC GSM
   06.11)
   ======

In the receiver, frames may be lost due to transmission errors or
frame stealing. Recommendation GSM 06.11 describes the actions to
be taken in these cases, both for lost speech frames and for lost
SID-frames in DTX operation.

In order to mask the effect of an isolated lost frame, a scheme is
used, where the lost speech frame is substituted by a predicted
frame based on previous frames. Insertion of silence frames is not
allowed. For several lost frames in a row, some muting technique
must be used to indicate to the customer that transmission is
interrupted.

7. SPEECH TRANSMISSION FUNCTIONS OF THE HALF RATE TRAFFIC CHANNEL
   ===

Provisions have been made for introduction at a later stage of a
so called half rate speech transcoder. The gross bit rate of the
half rate traffic channel will be 11.4 kbit/s. The algorithm for
the half rate speech transcoder as well as the associated DTX
functions are for further study.