INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

Supplement 32

(Series G)

(03/93)

TRANSMISSION SYSTEMS AND MEDIA

TRANSMISSION PLAN ASPECTS OF SPECIAL CIRCUITS AND CONNECTIONS USING THE INTERNATIONAL TELEPHONE CONNECTION NETWORK

TRANSMISSION ASPECTS OF DIGITAL MOBILE RADIO SYSTEMS

Supplement 32 to

ITU-T Series G Recommendations

(Previously "CCITT Recommendations")

FOREWORD

The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of the International Telecommunication Union. The ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Conference (WTSC), which meets every four years, established the topics for study by the ITU-T Study Groups which, in their turn, produce Recommendations on these topics.

Supplement 32 to ITU-T Series G Recommendations was prepared by the ITU-T Study Group XII (1988-1993) and was approved by the WTSC (Helsinki, March 1-12, 1993).

NOTES

1 As a consequence of a reform process within the International Telecommunication Union (ITU), the CCITT ceased to exist as of 28 February 1993. In its place, the ITU Telecommunication Standardization Sector (ITU-T) was created as of 1 March 1993. Similarly, in this reform process, the CCIR and the IFRB have been replaced by the Radiocommunication Sector.

In order not to delay publication of this Recommendation, no change has been made in the text to references containing the acronyms “CCITT, CCIR or IFRB” or their associated entities such as Plenary Assembly, Secretariat, etc. Future editions of this Recommendation will contain the proper terminology related to the new ITU structure.

2 In this Recommendation, the expression “Administration” is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

CONTENTS

Supplement 32

TRANSMISSION ASPECTS OF DIGITAL MOBILE RADIO SYSTEMS

(Helsinki, 1993)

(referred to in Series G Recommendations)

This Supplement is an introduction to the transmission characteristics of specific public land mobile networks.

1 Transmission aspects of the Pan-European Digital Mobile Radio System

1.1 Introduction

A series of the Groupe Special Mobile (GSM) Recommendations deal with the speech processing functions, transmission plan aspects and mobile termination conformity tests in the GSM system.

A general overview of the speech processing parts is given in [1] with reference to the Recommendations where each part is specified in detail. The transmission plan aspects are covered in [2]. The detailed specification of the audio parts are contained in [2] and [3].

1.2 Speech processing functions

1.2.1 Full rate speech transcoding

The detailed description of the full rate speech transcoding process is given in [4]. The scope of this GSM Recommendation is the conversion between a digital speech signal with 13 bit uniform PCM signal sampled at 8 kHz, and an encoded signal having an average bit rate of 13 kbit/s. The requirements of the transcoder are based on three parameters:

- transmission quality;
- delay;
- power consumption (complexity).

The method used is the Regular Pulse Excitation/Linear Predictive Coding using Long Term Prediction, RPE-LTP. The encoded bit stream consists of frames of 260 bits, corresponding to 20 ms of speech. The speech frame is delivered to the channel coding function defined in [9] to produce an encoded block consisting of 456 bits leading to a gross bit rate of 22.8 kbit/s.

Recommendation [4] describes the codec down to the bit level, thus enabling the verification of compliance with the Recommendation to a high degree of confidence by use of a set of digital test sequences, which are also described in the Recommendation.

For the subjective characterization tests on the RPE-LTP codec, several experiments were performed which characterized the effects of speech input level, listening level, carrier-to-interference (C/I) ratio, transcoding, audio input circuitry and environmental noise.

The codec showed satisfactory performance at all input and listening levels. The performance of the codec has been found to be substantially unaffected down to a C/I ratio of 10 dB and may be considered to have acceptable performance down to 7 dB. Smaller C/I ratios produce unacceptable degradation of speech performance (see Figure 1-1) and should be avoided.

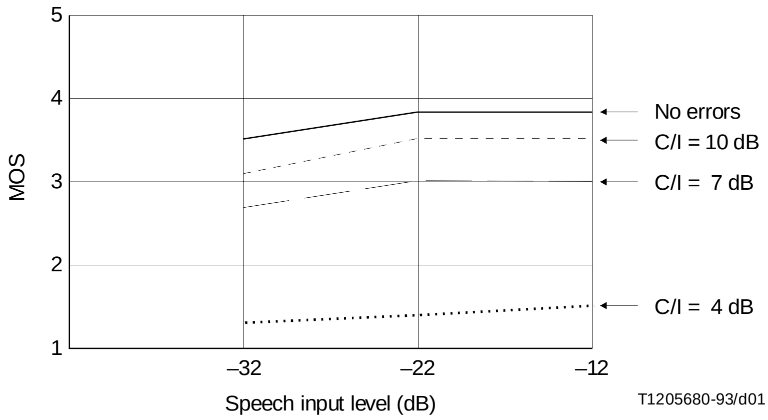

FIGURE 1-1

MOS versus speech input level for different C/I ratio

Under error-free transmission conditions the perceived quality of the codec is lower than both codecs conforming to Recommendations G.711 and G.721. The qdu value for the RPE-LTP codec is about 7-8 and should be used for planning purposes. At present there are no specific rules for determining qdu's for encoding below 32 kbit/s. A qdu value of 4.6 is measured under error-free conditions and nominal input level.

When connected in tandem with a codec conforming to Recommendation G.721, the same performance was obtained no matter whether the RPE-LTP codec came before or after the G.721 codec (see Figure 1-2). Two RPE-LTP codecs connected in tandem are significantly worse than a single codec.

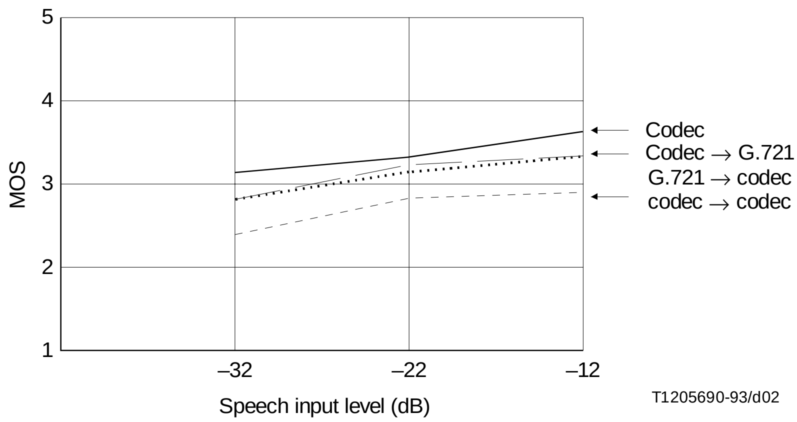

FIGURE 1-2

MOS versus speech input level for different transcodings

The theoretical minimum delay of the RPE-LTP codec is 20 ms. However, practical realizations may have an additional processing time in the order of 3-8 ms.

It should be noted that the RPE-LTP codec is an adaptive system which has been optimized for speech inputs. Great care must be taken when making measurements with non-speech signals because the normal assumptions of time invariance and linearity cannot be made.

1.2.2 Half-rate speech transcoding

Provisions have been made for the introduction of a so-called half-rate speech transcoder at a later stage. The gross bit rate of the half-rate transcoder will be 11.4 kbit/s. The algorithm for the half-rate speech transcoder is for further study.

1.2.3 General strategy for discontinuous transmission (DTX) operation

During a normal conversation, the participants alternate so that, on average, each transmission direction is occupied about 50% of the time. If the transmitters are switched on only for those frames that contain speech, this will result in a gain in two respects:

- in the mobile station, battery life will be prolonged or a smaller battery could be used for a given operational duration;
- the average interference in the "air" is reduced leading to better spectrum efficiency.

The overall process for the various actions using this mode of operation (discontinuous transmission, DTX) and the occurrence of missing frames are described in Recommendation [7]. This Recommendation defines two-state diagrams, one for the transmitting side and one for the receiving side. The Recommendation controls the functions defined in [5], [6] and [8].

1.2.4 Voice activity detection

Recommendation [8] describes the performance objectives for the voice activity detector (VAD) function in the GSM system. The VAD is part of the system for discontinuous transmission (DTX) which is defined in Recommendation [7]. The VAD generates an active/non-active signal for each 20 ms frame of speech.

The basic objective of the VAD is:

- to avoid unacceptable impairment of voice performance; and
- to achieve the required percentage of speech pause detection.

Voice activity detection is achieved by monitoring and analysing a number of signal features that effectively distinguish between speech (or speech + noise) and noise. Examples of such features are signal level, zero crossing rate, auto-correlation coefficients, etc. The detection may be enhanced by use of additional input information such as an additional microphone or information available in the speech encoder.

In order to achieve maximum separability of speech and noise, the signal properties of both signals have to be taken into account when selecting signal features and in the design of a detection algorithm. For the verification of the performance, the VAD device will be tested with actual recordings of speech and noise where the correct on-off patterns of speech have been marked on a reference file.

The basic performance criteria of the VAD are:

- front end clipping of speech bursts;
- mid-speech clipping;
- noise removal performance (noise detected as speech).

1.2.5 Comfort noise insertion

When the active/non-active signal produced by the VAD is used to switch the transmitter on and off, the effect will be a modulation of the background noise at the receiving end. In the "active speech" situation, the background noise is transmitted together with the speech to the receiving end. As the speech burst ends, the connection is interrupted, and the received noise drops to a very low level. This step modulation of noise is perceived as very annoying and may reduce the intelligibility of speech. The noise contrast effect in the GSM system is reduced by inserting noise at the receiving end.

Recommendation [6] deals with all detailed aspects of the overall comfort noise generation process, i.e. the evaluation in the transmitter, the noise parameter encoding (SID frames) and decoding and the insertion in the receiver. The algorithm for updating the noise parameters during speech pauses is also defined.

The generation of comfort noise is based on the speech codec Recommendation.

1.2.6 Speech frame extrapolation

In the receiver, frames may be missing due to one of three causes:

– DTX operation;

– transmission errors (erasure); and

– frame stealing.

In order to mask the effect of missing frames, a speech frame extrapolation scheme must be used: as a minimum, a missing frame should be replaced by repetition of the previous frame, silence frames are not allowed. For subsequent missing frames, some muting must be used. Manufacturers are encouraged to implement more effective extrapolation schemes.

1.3 Mean one-way propagation time of the GSM system

For the whole GSM system with a full rate speech transcoder between the acoustic interface and the network connection point, the propagation time (delay) needs to be considered. Table 1-1 lists the inherent and implementation-dependent delays of the GSM system.

TABLE 1-1

Delays in the GSM system

RPE-LTP codec delay	20.0 ms
Channel coding delay	37.5 ms
Processing delay	ca. 27.5 ms
Transmission system (BS to MSC link)	ca. 5.0 ms
Expected overall delay	ca. 90.0 ms

1.4 References

[1] GSM Recommendation 06.01 *Speech Processing Functions, General Description*.

[2] GSM Recommendation 03.50 *Transmission Planning Aspects of the Speech Service in the GSM PLMN System*.

[3] GSM Recommendation 11.10 *Mobile Station Conformity Specification*.

[4] GSM Recommendation 06.10 *13 kbit/s Regular Pulse Excitation-Long Term Prediction-Linear Predictive Coder for use in the Pan-European Digital Mobile Radio System*.

[5] GSM Recommendation 06.11 *Speech Frame Extrapolation*.

[6] GSM Recommendation 06.12 *Comfort Noise Receiver Functions*.

[7] GSM Recommendation 06.31 *DTX Control and Operation*.

[8] GSM Recommendation 06.32 *Voice Activity Detector.*

[9] GSM Recommendation 05.03 *Channel Coding*.

2 Transmission aspects of the Japanese digital mobile radio system

2.1 Introduction

This is a preliminary report on the transmission aspects of the Japanese digital mobile radio system, which is currently undergoing standardization. Transmission performance shown in this Supplement was obtained by RCR (Research and development Center for Radio systems) through the standardization of the full-rate speech coding method. All of the data are subject to updating after completing the standardization procedure.

2.2 Transmission performance

2.2.1 Speech quality

Speech quality performance shown below was evaluated using 11.2 kbit/s real-time working speech CODECs through subjective listening tests. In the tests, MNRU reference samples were evaluated at the same time to translate MOS values into opinion equivalent Q values and qdu values.

1) *Transmission error*

   Speech quality versus transmission error is shown in Figure 2-1. Here speech input level is adjusted to –18 dBm0 (21 dB lower than the overload level) and no background noise is added.

   Excellent performance of approximately 4 qdu is achieved in an error-free channel. In error-prone channels, speech quality becomes worse than 5 qdu. However, this performance is better than that of conventional analogue FM systems.

2) *Speech input level*

   Speech quality versus speech input level is shown in Figure 2-2, where no transmission error and no background noise are added.

3) *Background noise*

   Speech quality versus background noise is shown in Figure 2-3, where speech input level is adjusted to –18 dBm0 and no transmission error is added.

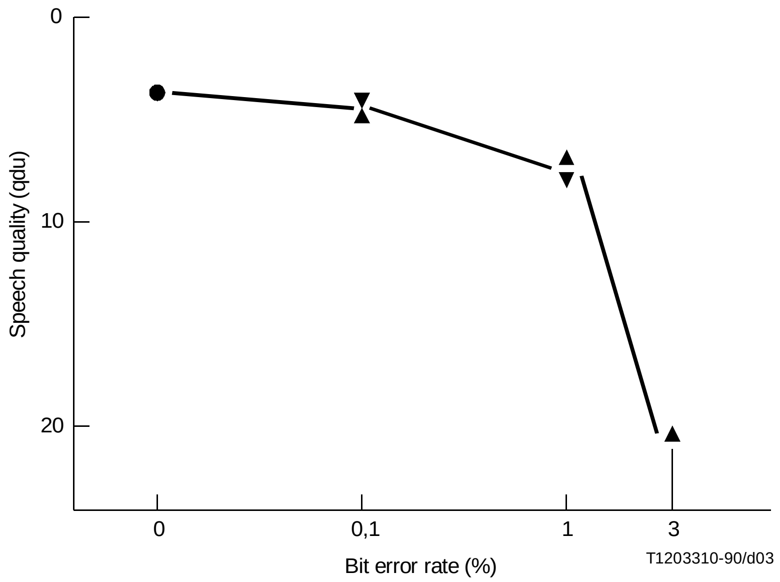

FIGURE 2-1

Speech quality versus transmission error

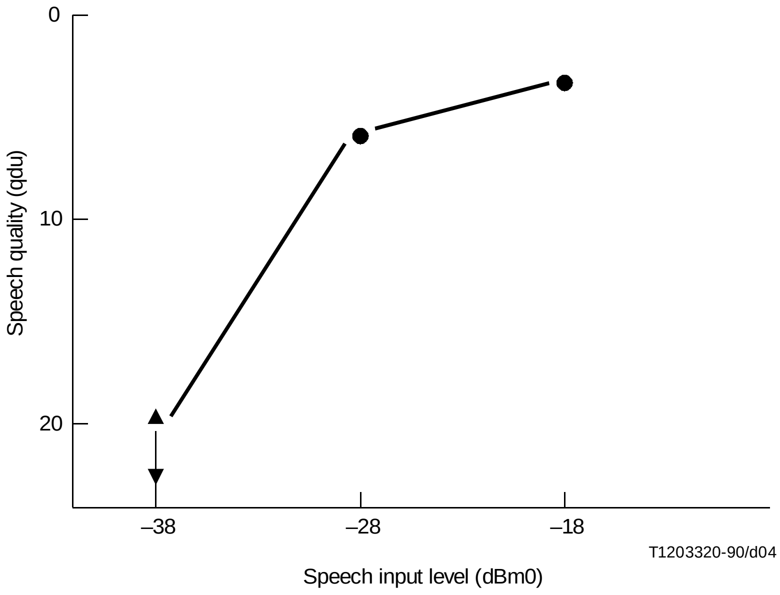

FIGURE 2-2

Speech quality versus speech input level

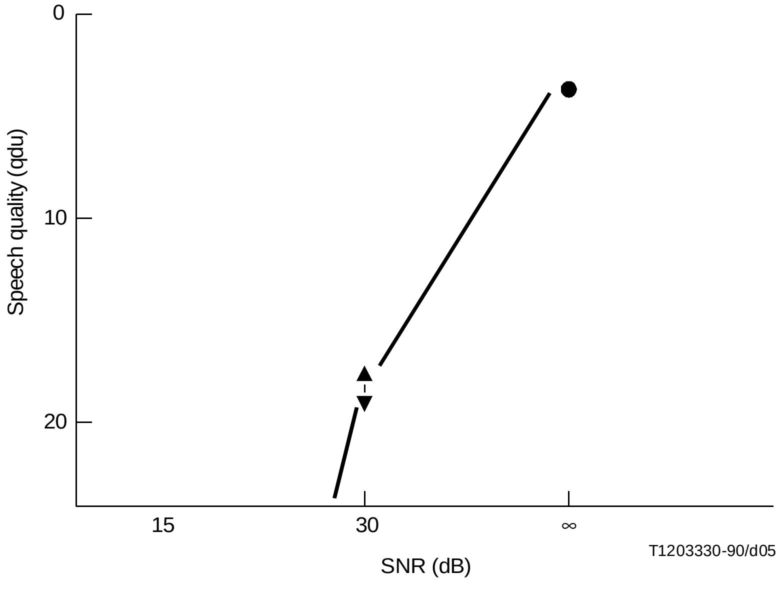

FIGURE 2-3

Speech quality versus SNR (background noise)

2.2.2 One-way transmission delay

One-way transmission delay time of the whole system can be defined as a summation of:

– algorithmic delay of speech encoding/decoding, including error protection;

– processing delay of the real-time working CODEC;

– transmission delay of the TDMA radio link;

– transmission delay between BS and MSC.

Approximate values of the above items are 27-55 ms, 6.7 ms and 20 ms. Thus, the expected overall one-way delay is approximately 73.7-101.7 ms.