

# An Improved Speech and Channel Coding for GSM System

Dariusz Godyń, Dominik Rutkowski

*Motorola Polska Software Center, ul. Jodłowa 3, 30-252 Kraków, Poland*

e-mail: [Dariusz.Godyn@motorola.com](mailto:Dariusz.Godyn@motorola.com)

*Department of Radio Communication Systems, Technical University of Gdańsk*

e-mail: [nick@pg.gda.pl](mailto:nick@pg.gda.pl)

Key words: GSM system, channel error protection

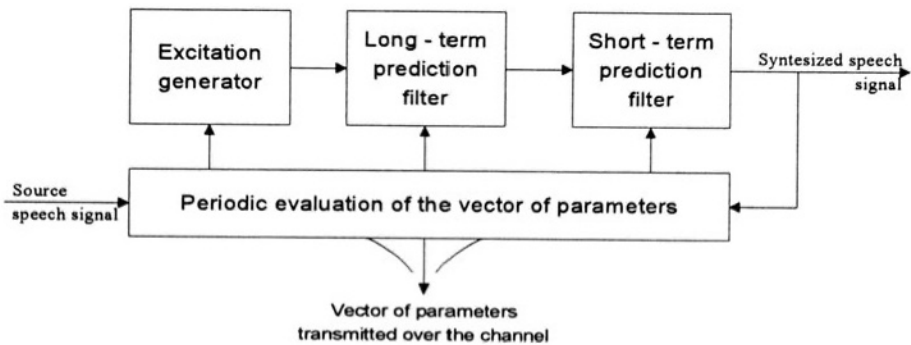
Abstract: In the paper an improved speech and channel coding/decoding for the GSM system has been described. It enables to decrease the source rate and to increase the channel error protection of the speech and thus to achieve better performance of the reconstructed speech than the original speech and channel coding/decoding in the presence of fading and noise.

## 1. INTRODUCTION

The major task of the source coding/decoding of speech is to reduce the redundancy of speech and the corresponding bit rate in order to confine the required channel bandwidth that is a valuable resource in any radio communication system. On the other hand, the channel coding/decoding that introduces some redundancy is used to increase the error protection of speech data transmitted over the channel with fading, interference and noise to improve the performance of the reconstructed speech at the receiver. Despite of that protection one may observe the loss of speech frames in GSM system at the poor channel conditions and, of course, a drastic degradation of performance. However, if one looks more closely on the operation of source coder in GSM system one can notice the possibility of additional reduction of bit rate at its output that combined with more robust error protection on the channel can improve the overall performance of speech reconstruction at the output of the receiver if compared with the standard solution.

## 2. FULL-RATE CODING WITH REDUCED BIT RATE

The source coding of speech in GSM system is based on the digital model of a natural mechanism of speech signal generation. The essence of source coding consists in periodic evaluation of specific parameters characterizing the digital model provided that the segment of source signal is taken in the sufficiently short time interval in which it can be considered as quasistationary. The set of these parameters called a vector is used in the source decoder at the receiver to the synthesis of speech signals. The digital model of speech signal generation shown in Fig. 1 consists of the excitation generator, long- and short-term prediction filters as well as a unit that is approximating each source signal segment by the synthesized speech signal segment to obtain the weighted error minimization. All the parameters characterizing the excitation generator and the prediction filters are quantized usually with the nonlinear quantizers and periodically transmitted over the channel.



*Figure 1. General model of the evaluation of parameters characterizing source speech signal used in GSM system*

The excitation generator is modelling a flow of air stream through the vocal tract of a human speech organ. The linear long-term prediction filter is modelling the pitch period corresponding to the vibrations of vocal cord. The linear, short-term prediction filter is modelling the short-term spectral envelope of speech.

The number of bits delivered periodically at the output of speech encoder depends on the number of parameters describing the model of the source signal and the range of their values that in turn depend on the model complexity and the required performance of synthesized speech.

The standard speech encoder employed in GSM system is based on the RPE- LTP- LPC algorithm [1]. It produces the bit rate of 13 kb/s which when

compared with the input bit rate of 104 kb/s gives eightfold compression at the acceptable quality of the reconstructed speech in terms of MOS value (MOS  $\sim 3.6 \div 4.0$ ) [2], [3]. This result is obtained at the moderate complexity in terms of operations per second and short delay due to processing. Speaking more precisely, the GSM speech encoder delivers for each 20ms segment of speech samples, 8 parameters of short-term prediction filter that are represented by 36 bits and 2 parameters (gain and lag) of long-term prediction filter that are taken every 5ms and coded by 2 and 7 bits, respectively which yields another 36 bits. The parameters of excitation generator are calculated 4 times within one speech segment. These are 13 excitation pulses and the initial time phase of their appearance coded by 2 bits. The amplitudes of excitation pulses are first normalized with respect to the largest of them and then the largest amplitude is coded by 6 bits and the other ones by 3 bits. Summing up, the parameters of excitation pulses are represented by 188 bits every 20 ms and the speech encoder produces 260 bits of speech frame for each segment of source speech signal. The position of each bit in that frame depends on its weight i.e. its influence on the intelligibility of reconstructed speech. According to that weight 260 bits of speech frame are subdivided into 3 classes as it is shown in Fig.2 with the most important class 1a of 50 bits, less important class 1b of 132 bits and the least important class 2 of 78 bits. There is different channel coding scheme used for bits class 1a and 1b.

The speech decoder at the receiver is synthesizing the speech signal on the basis of the received speech frame corrupted by noise, fading and interference on the channel.

As it turns out the number of 260 bits contained in the speech frame can be reduced by using a procedure for finding out optimized quantizers of parameters for the source signal model [4]. The characteristics of quantizers have been obtained with the use of Max – Lloyd algorithm and the kernel estimator of the probability density functions for different parameters.

The investigations carried out with respect to the subjective audio tests of reconstructed speech and the analysis of the average parameter-to-quantization noise ratio (PQR) for each standard quantizer enabled to find some new optimum quantizers. They concern 8 parameters of short-term prediction filter which require 5; 5; 5; 4; 4; 3; 2; and 2 bits, respectively. Apart from this a new 5 bit quantizer of maximum amplitude of excitation pulse has been derived. The number of quantization levels for the remaining quantizers have been unchanged. However, the optimization of their characteristics has been carried out.

Indeed, the obtained reduction of the number of bits in the speech frame has caused the decrease by 2.6 dB of the average PQR, however, the quality of synthesized speech turned out to be sufficiently good. The analysis of the average signal-to-noise ratio [2] of the reconstructed speech for the test

speech signals indicates only small decrease by 0,7 dB. However, taking into account the average log spectral distortion of a speech frame a small increase by 0,1 dB is observed.

The comparable audio tests of the standard and modified speech encoder for 2 women voices and 2 men voices have shown that subjective quality of modified speech encoder is fully acceptable as one can see in table 1.

Table 1. Results of the subjective quality of speech reconstruction. Notation: (+) – higher quality, (-) – lower quality, (=) – the same quality.

Evaluating person	Woman 1 voice		Woman 2 voice		Man 1 voice		Man 2 voice	
	GS	PRO	GS	PRO	GS	PRO	GS	PRO
	M	JECT	M	JECT	M	JECT	M	JECT
D.G.	+	-	+	-	-	+	+	-
P.C.	-	+	=	=	=	=	=	=
R.S.	+	-	+	-	+	-	=	=
J.S.	=	=	=	=	=	=	=	=
M.B.	=	=	=	=	=	=	=	=
S.K.	-	+	-	+	+	-	-	+

The most important result obtained with respect to the modified speech encoder is the reduction of the speech data rate from 13 kb/s to 12,5 kb/s. Precisely speaking, the modified speech encoder is generating 250 bits instead of 260 bits per speech frame of the standard speech encoder.

In table 2 one can find the number of bits needed for the parameters of standard encoder and the modified one.

Table 2. The number of bits required for parameter coding of standard and modified encoder.

Parameters	Number of bits (GSM)	Number of bits (PROJECT)
Parameters of short term prediction filter	36	30
Initial time phase of excitation pulses	8	8
Maximum amplitude of excitation pulses	24	20
Amplitudes of remaining excitation pulses	156	156
Gain of long-term prediction filter	8	8
Lag of long-term prediction filter	28	28
Total	260	250

### 3. IMPROVED CHANNEL CODING

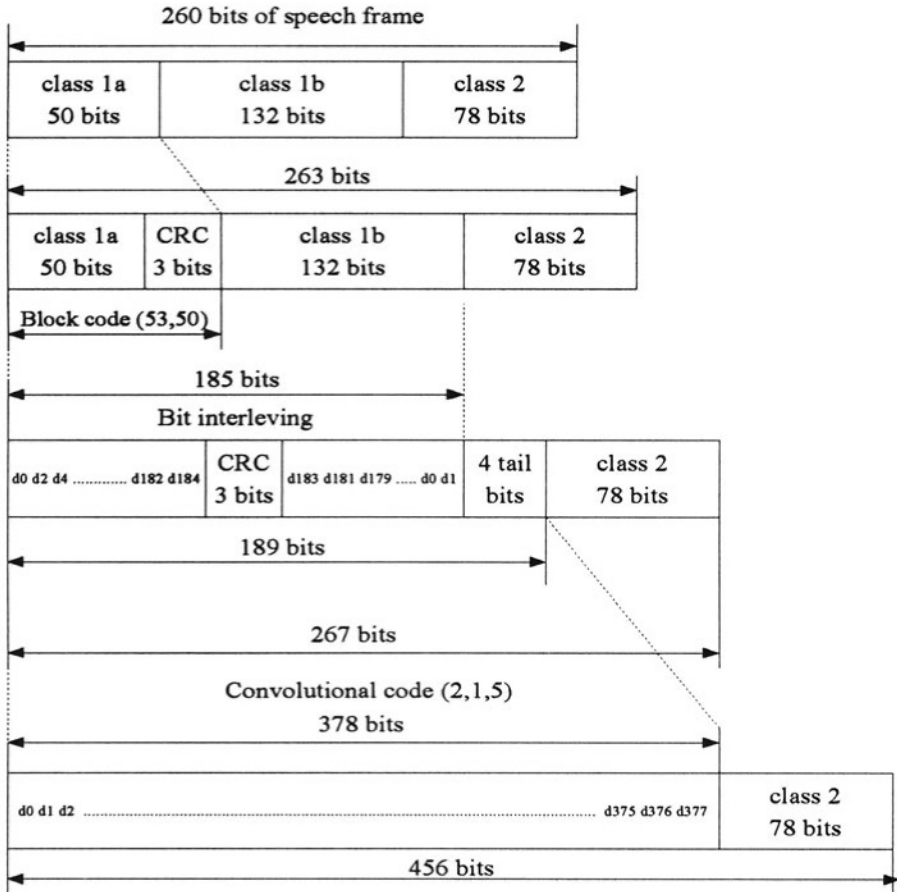


Figure 2. Standard channel coding of speech frame

The standard speech frame containing 260 bits is further undergoing channel coding. First, it is encompassing error detection coding of class 1a bits, by the use of CRC adding 3 parity bits. Next, to the obtained block together with the class 1b bits additional four 0 bits are added to reset the convolution coding employed as error correction coding. Further details of standard coding are shown in Fig 2.

The total number of bits in the speech frame reduced by 10 bits can enable to design a new channel coding with enhanced error protection and to improve the performance of reconstructed speech at the poor though still acceptable channel conditions.

Thus, instead of a simple error detection coding concerning the class 1a bits of standard coding a new error correction coding has been implemented. As a result the rate of discarded speech frames has been reduced and the performance of reconstructed speech has been improved. The code that has been used is BCH (63, 57). It has the error correction capability of 1 error and its generating polynomial is given by

$$g(u) = u^6 + u + 1 \quad (1)$$

To obtain some error detection capability an extended BCH code has been employed with one added parity bit. A new error protection coding is shown in Fig. 3. One can see in Fig.3 that 7 zero bits have extended the length of class 1a bits to carry out block error coding. Of course, on the channel only the sequence of 50 class 1a bits and 6 parity bits is transmitted. The new error coding of speech frame is shown in Fig. 4. As one can see the number of bits allocated to class 1b and class 2 has been slightly changed if compared with the standard channel coding. This change is motivated by convenient arrangement of BCH code parity bits and the format similar to the standard one. Thus, the total number of coded bits in the speech frame remains unchanged.

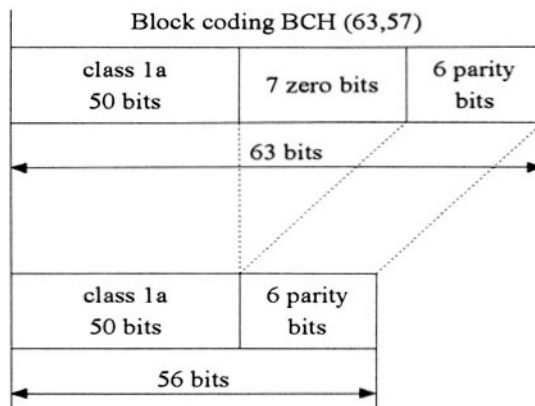


Figure 3. Error correction coding of class 1 a bits

#### 4. AN ALGORITHM OF CHANNEL DECODING

In the standard decoding algorithm the received speech frame is discarded if an error is detected in the sequence of class 1a bits. A new channel coding

requires some changes in the decoding algorithm (see [5]) that are shown in Fig.5.

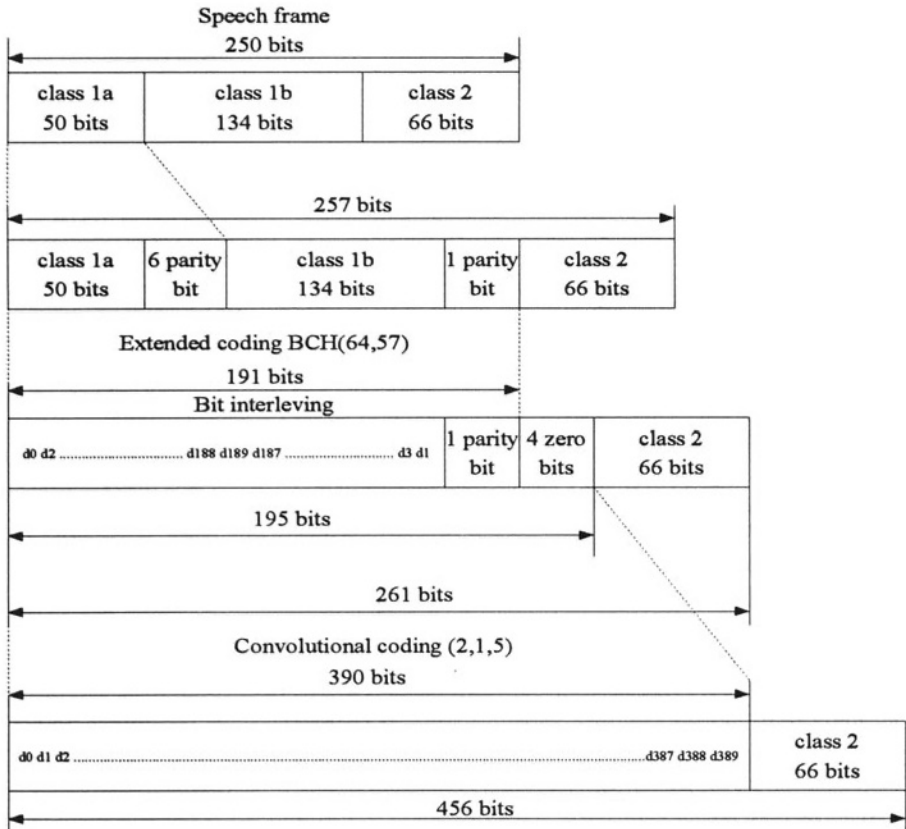


Figure 4. Channel coding of speech frame containing 250 bits

First of all, the syndrome of received class 1a bits must be calculated. This requires to add 7 zero bits (see Fig. 3). If there are no errors an acceptance decision is undertaken. If there are some errors, the speech frame is discarded provided that the error appears on the forbidden positions (7 positions), otherwise the error correction is carried out and the parity bit is calculated. In the case of identical values of parity bit calculated and received, the speech frame is accepted. In another instance it is discarded and the procedure of previous frame substitution is initiated.

The improved speech and channel coding/decoding has been tested through simulation in order to compare the results with that for the standard channel coding/decoding provided that signals are transmitted on the channel with fading, noise and at different mobile station speeds. The evaluation was based on the indication by the listeners, which one out of two reconstructed speech

test signals with the use of standard and improved speech and channel decoder is better in quality. Some of the results (see [5]) for urban propagation environment are shown in tables 3 and 4.

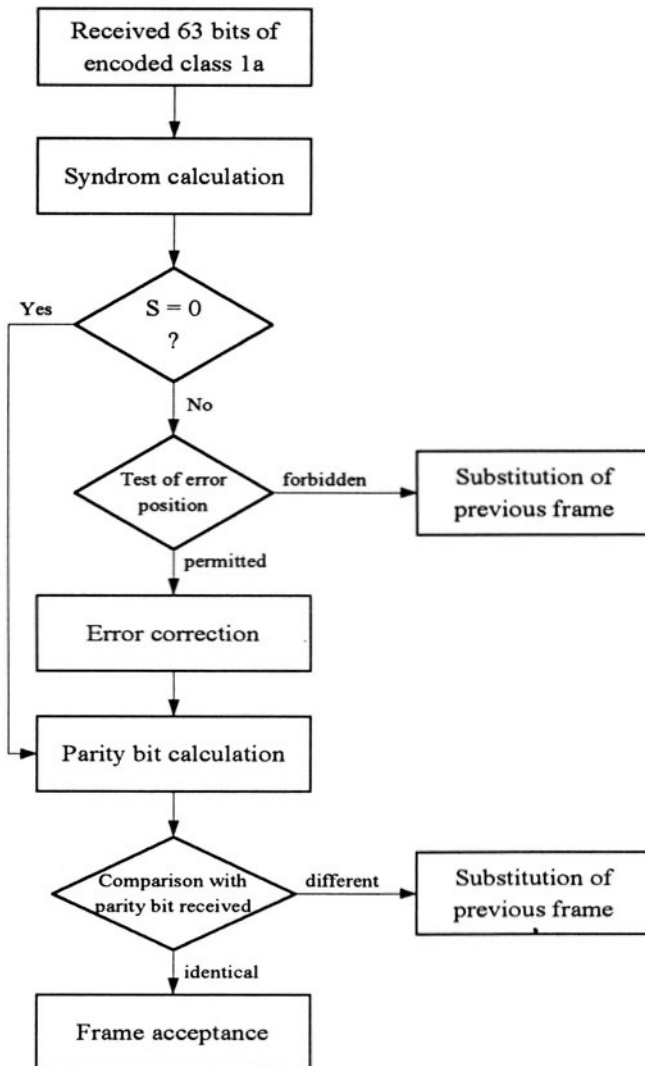


Figure 5. A decoding algorithm of speech frame for new channel coding

As one can see the improved speech and channel coding/decoding enables to obtain better performance particularly at the poor channel conditions i.e. at low  $E_b/N_0$ .



Table 3. Simulations results of the subjective quality of speech reconstruction of a woman voice for the standard and improved speech and channel coding/decoding. Propagation environment: TU 50. Notation: (+) – higher quality, (-) – lower quality, (=) – the same quality

Evaluating person	$E_b/N_0$ [dB]									
	4,5		5,5		6,5		8,5		10,5	
	GSM	PROJECT	GSM	PROJECT	GSM	PROJECT	GSM	PROJECT	GSM	PROJECT
R.S.	-	+	-	+	=	=	-	+	-	+
D.G.	-	+	-	+	-	+	-	+	-	+
J.S.	-	+	-	+	-	+	-	+	+	-
S.K.	-	+	-	+	-	+	-	+	-	+
P.C.	-	+	-	+	-	+	-	+	-	+

Table 4. Simulations results of the subjective quality of speech reconstruction of a man voice for the standard and improved speech and channel coding/decoding. Propagation environment: TU 50. Notation: (+) – higher quality, (-) – lower quality, (=) – the same quality

Evaluating person	$E_b/N_0$ [dB]									
	4,5		5,5		6,5		8,5		10,5	
	GSM	PROJECT	GSM	PROJECT	GSM	PROJECT	GSM	PROJECT	GSM	PROJECT
R.S.	-	+	-	+	-	+	-	+	+	-
D.G.	-	+	-	+	-	+	-	+	-	+
J.S.	-	+	-	+	=	=	=	=	=	=
S.K.	-	+	-	+	-	+	-	+	-	+
P.C.	-	+	-	+	+	-	-	+	-	+

## 5. CONCLUSIONS

The results obtained indicate that the improvement of performance of GSM system is possible particularly in the areas where poor channel conditions exist. This concerns the areas at the cell borders and the areas of propagation shadow. By improving the performance of reception the degree of coverage by the system can also be increased.

**REFERENCES**

- [1] ETSI, *European digital cellular telecommunications system* (Phase 2), (Sophia Antipolis CEDEX - FRANCE), 1995.
- [2] R.A. Salami, L.Hanzo, R.Steele, K.H.J.Wong, and I.Wassell, *Mobile Radio Communication*, pp. 186-346. Pentech Press, 1 st ed., 1992.
- [3] J.Stachurski, *A Pitch Pulse Evolution Model for Linear Predictive Coding of Speech*, Ph.D. dissertation, McGill University, Montreal, Canada, 1997.
- [4] D.Godyń, *A Choice of Characteristics of Nonlinear Quantizers for Spectrum Efficient Speech Coding in GSM System*, Proc. Nat. Comm. Conf. in Polish, KST'98, vol. A, pp. 179-185, Bydgoszcz, Poland, 1998.
- [5] D.Godyń, *Investigation of Spectrum Efficient Techniques of Coding for GSM System*, Ph.D. dissertation, Technical University of Gdańsk, Gdansk, Poland, 1999.