Toll Quality Codec for the GSM-System

Hansen, H. B.; Haagen, J.; Mikkelsen, K. B.; Nielsen, H.; Larsen, Knud J.

[Link back to DTU Orbit](#)

Toll Quality Codec for the GSM-System

Henrik B. Hansen, Jesper Haagen, Keld B. Mikkelsen,
Henrik Nielsen, and Knud J. Larsen[1].

Tele Danmark Research,
Lyngsø Allé 2, DK-2970 Hørsholm, Denmark.

[1] Institute of Telecommunication,
Technical University of Denmark, DK-2800 Lyngby, Denmark.

ABSTRACT

This paper describes the basic performance requirements for a new enhanced GSM codec and a proposal for this application based on the ITU 8 kb/s coder standard. A powerful channel coding scheme using more than half the gross bit rate is designed and the performance of the codec is formally evaluated using MOS testing. The overall system is found to provide toll quality speech with only a graceful degradation at increasing bit error rates, and it meets the preliminary quality requirements for the new standard.

1 Introduction

In Europe, a new codec for the GSM wireless communication system is required. Currently two codecs have been standardized: the GSM full-rate codec [1] at 22.8 kb/s gross bit rate, and the GSM half-rate codec at 11.4 kb/s [2]. Since the introduction of the GSM full-rate codec, the advances in speech coding and VLSI technology have made high quality coding for the full-rate channel feasible. The high quality codec is targeted mainly at new markets like the evolving PCS (Personal Communication Systems), but it is envisaged that an enhanced full-rate codec could be included as an option in multi-mode GSM terminals supporting both full-rate and half-rate coding schemes. In any case, the growing demand for mobile communication stresses the need for speech coders capable of providing toll quality speech on wireless connections at least at relatively low error rates.

Since a relatively tight time schedule is foreseen for the development of the high quality codec (less than one year), it would be convenient if an existing codec could be adapted to function in this wireless environment. In general, high quality codecs are more sensitive to channel errors than the existing GSM full-rate (FR) codec. This means that a new enhanced speech codec should offer better quality at somewhat lower bit rate allowing additional bits to be allocated to channel coding.

In this paper we describe a channel coding scheme which can be used in conjunction with the new toll-quality ITU 8 kb/s codec [3] and the experiments carried out to test the robustness of the codec combination in the GSM environment. The following issues will be addressed:

- GSM enhanced full-rate (EFR) coding performance requirements
- channel coder design
- error detection and concealment techniques
- algorithmic delay

2 Performance requirements

The essential requirements [4] for the high quality codec include speech quality, complexity, and delay. The speech quality should be comparable to that of G.728 [5] for error free conditions and conditions with relatively few channel errors. Under severe error conditions the quality should be significantly better than that of the GSM full-rate codec.

Concerning the complexity of the codec it should not exceed significantly that of the GSM half-rate codec, i.e. 4 times the full-rate codec complexity.

The coding delay, including time-slot interleaving, is limited to that of the GSM full-rate system, i.e. 90 ms end-to-end delay. However, in order to mitigate the echo cancellation problems it is a desirable requirement to limit the coding delay to 40 ms.

3 Channel coding

The channel coding scheme is very similar to those already implemented in the GSM system. It is basically a low constraint length convolutional coding scheme without tail bits transmitted. The low speech coder bit rate allows a channel coding rate of almost 1/3 for all bits in each 10 ms frame. The speech coder bits are divided into two classes containing 54 bits and 25 bits, respectively, excluding the parity bit which is treated differently as explained below. The two classes are encoded at rates $R = 6/17$ and $R = 1/3$, respectively. Even though the tail bits are not transmitted, a low constraint length $CL = 4$ corresponding to 8 coding states is chosen. Preliminary experiments have shown that performance does not really benefit from increasing the constraint length for the convolutional code. A low constraint length code recovers faster from the occasional overloading during channel fades, and it also ensures a fairly low channel coder complexity.

3.1 Error detection

Error detection is based on soft decision information. Both input channel soft decision and channel decoder output soft decision information is used to produce the bad frame indication. The parity check included in the speech coder

59

itself is not very useful in this context because single errors rarely occur during fades. Therefore the parity bit is used for channel coding and after the channel decoding parity is set to either check-or fail depending on the channel decoder output. This technique actually adds some flexibility to the error concealment because in case of a parity error only the pitch parameters are replaced in the decoder. This means that a parity error may be utilised to signal a slightly corrupted frame which should not be totally discarded. The speech decoder also recognises bad frames flagged by the error detector. These frames are discarded and the speech decoder extrapolates the output signal.

4 Delay

The algorithmic delay depends on the speech coder frame size and the channel burst interleaving depth. If two speech coder frames are merged into one 20 ms frame and then interleaved over eight bursts, the configuration is similar to that of the FR system. However, due to the speech coder sample look-ahead the delay limit will be exceeded slightly. A reduction in the total delay can be obtained by reducing the interleaving depth. For a frame size of 10 ms and interleaving depths less than eight the algorithmic delay is below the limit. Some of the possible configurations of frame size

candidate	frame size	interleaving depth	delay
EFR208	20 ms	8	FR + 5 ms
EFR106	10 ms	6	FR - 35 ms
EFR104	10 ms	4	FR - 55 ms

Table 1. *Algorithmic round-trip delays*

and interleaving depth are shown in Table 4 and these candidates are covered by the test described below.

5 Subjective test

The speech coder itself is already characterised during the standardisation process. The purpose of this MOS (Mean Opinion Score) test is to assess the quality of the coding system under bad channel conditions. In order to measure the effect of a delay reduction, three different interleaving depths are included in the experiment as separate conditions. The G728 and GSM FR standard coders are included as references. The G728 is tested in error-free condition only while both the EFR candidates and the GSM FR are subjected to three different error patterns in addition to the error-free condition.

Figure 1 shows the subjective test results. The dotted line represents the reference quality level corresponding to the measured G728 quality on clear channel. The listeners have consistently given rather low scores throughout the experiment. The reason for this could be that the experiment is not balanced sufficiently well because a major part of the speech samples is almost equal in quality.

6 Conclusion

In this paper the application of the ITU 8 kb/s speech coder in a GSM EFR coding system is addressed. A channel coder
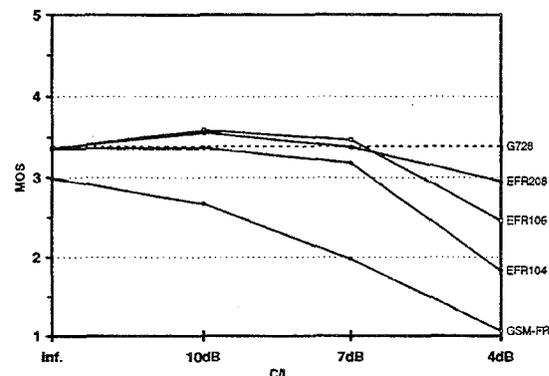

Figure 1. *Subjective test results.*

to accompany the speech coder is described and the performance of the coding system in bad channel conditions is investigated by means of subjective testing. The test results show that by applying a powerful channel coding scheme it is possible to retain a quality level close to toll quality for the bad channel situations.

Besides speech quality, the amount of delay introduced by the codec is an important factor when the goal is to achieve performance approaching that of a wireline network. As the speech coder has a fairly low algorithmic delay of 15 ms, a significant reduction in total delay could be achieved if the burst interleaving is removed. However, since the error correction benefits very much from the burst interleaving, this requires very strong channel coding if speech quality is to be retained. Other aspects like background noise performance and overall complexity need to be investigated further before a final conclusion on applicability to GSM EFR can be reached. So far the results suggest that a solution offering the requested quality with some reduction in delay can be found.

References

[1] ETSI, "GSM Recommendation 06.10: Full-Rate Speech Transcoding," Sep 1988.

[2] ETSI, "GSM Recommendation 06.20: Speech Coding for The GSM Half-Rate Channel," draft, Jan 1995.

[3] ITU-TSS Study Group 15, Draft Recommendation G.729, June 1995.

[4] SMG1, "Enhanced Full-Rate Speech Codec Study Phase Report," ETSI SMG-TC, Dec 1994.

[5] ITU-T, Recommendation G.728.

[6] H. B. Hansen, K. J. Larsen, H. Nielsen, and K. B. Mikkelsen, "Channel Coding Schemes for the GSM Half-Rate System," in *Speech and Audio Coding for Wireless and Network Applications*, edited by Bishnu S. Atal, Vladimir Cuperman, and Allen Gersho, pp. 171-179, Kluwer Academic Publishers, 1993.

[7] K. B. Mikkelsen, K. J. Larsen, H. B. Hansen, H. Nielsen, and H. Jacobsen, "The Danish Candidate for the GSM Half-Rate Speech Channel," *SIGNAL PROCESSING VII, Proc. of EUSIPCO-94*, pp. 995-998, EURASIP 1994.