

A FULL-RATE GSM-AMR CANDIDATE

C. Sriratanaban, A. Kondo

Centre for Communication Systems Research (CCSR)
University of Surrey, Guildford, Surrey, GU2 5XH, United Kingdom
C.Sriratanaban@surrey.ac.uk, A.Kondo@surrey.ac.uk

ABSTRACT

This paper describes a full-rate (FR) candidate speech codec for a new Adaptive Multi-Rate (AMR) codec standard. The proposed codec is a multi-rate codec, operating in various modes that have different bit-rate partitionings between source and channel coding. The source coding is based on Algebraic Code Excited Linear Predictive Coding (ACELP). Convolutional coding is used for the channel coding. Subjective tests show that at poor channel error conditions with low C/I-ratios of 7 dB, the lowest mode of the proposed codec produces speech quality comparable to G.728 in clean channel condition. The codec also provides substantial improvement on error robustness over the Enhanced Full Rate (EFR) GSM codec, by selecting an optimal coding rate to suit the channel condition. The codec performance passes twenty-six out of the twenty-nine test conditions, conducted according to [2]. Therefore, the proposed codec has passed the ETSI qualification test.

1. INTRODUCTION

The Adaptive Multi-Rate (AMR) codec is a new generation codec for Global System for Mobile Communication (GSM). It has been initiated by the European Telecommunications Standards Institute (ETSI) in 1996. AMR was intended to enhance the quality of speech as well as the capacity of the existed GSM system. The AMR system can be operated at both half rate (11.4 kbit/s) and full rate (22.8 kbit/s). It adapts to local radio channel conditions and selects the optimum speech and channel coding rate to deliver the best speech quality.

The GSM system firstly used the FR codec which is based on Regular Pulse Excited Long-term Predictive (RPE-LTP) Coding algorithm. Later on, the number of mobile users rapidly increased. Therefore, HR codec, based on Vector Sum Excited Linear Prediction (VSELP), has been developed to cope with the market needs by using half of the FR channel bandwidth i.e. 11.4 kbit/s. The EFR codec [3], based on ACELP, has now been established to provide higher speech quality to the subscribers through the full-rate channel. Moreover, the EFR codec proves to perform better than FR and HR

codecs in both error-free and typical channel error conditions.

As for the speech quality aspect, the AMR codec aims to provide better speech quality and improved error robustness over not only the FR and HR codecs but also the EFR codec. In addition, it also aims to achieve wireline quality combined with capacity advantages when operating in half rate mode. AMR codec can deliver up to twice as much capacity over the EFR codec and maintain higher quality than the existing FR when HR channel mode is used.

This paper focuses on an adaptive multi-rate codec operating in full-rate channel mode for the AMR standard. The multi-rate codec is based on ACELP coding algorithm and a convolutional channel coding algorithm. These algorithms are also used in the existing GSM-EFR. The proposed codec operates at a fixed gross bit-rate of 22.8 kbit/s. It maintains better speech quality than wire-line quality (G.728 or better) even in poor channel conditions.

2. CODING DESIGN

EFR operates at a gross bit-rate of 22.8 kbit/s which consists of 12.2 kbit/s for source coding and 10.6 kbit/s for channel coding. Since the bit partitions are fixed, they have to be compromised between speech coding bits and error protection bits. Therefore, the speech quality can be degraded through the poor channel when the channel coding bits are not sufficient.

The multi-rate codec which is proposed in this paper is able to adapt its rate to provide the highest quality in clean channel. In addition, it provides better error protection by allocating more channel coding bits in poorer channel conditions. Therefore four versions have been designed to provide different levels of error protection and output quality. All four versions have the same short term and long term prediction. Only their excitation signals are different. The proposed codec also supports Tandem Free Operation (TFO) and performs seamless codec mode bit-rate changes.

2.1 SOURCE CODING

The proposed codec is based on ACELP structure. It produces 4 different source coding bit-rates of 8.25, 10.05, 11.65 and 13.65 kbit/s.

It operates on 20 ms frame basis, with a 4.4 ms (35 samples) look ahead for LPCs (Linear Prediction Coefficients) calculation. Long Term Prediction (LTP) lag and gain are computed on every 5 ms sub-frame i.e. 4 sub-frames in a frame.

Short term parameters, 10th order LPCs, are LSF transformed and split-vector quantised using 28 bits. LPCs calculation is performed on every sub-frame by interpolating between the current and the previous set of LSFs. RMS (Root Mean Squared) energy is calculated and quantised using 7 bits for each frame

LTP parameters are obtained via an open-loop pitch search algorithm followed by a closed-loop refinement. Integer lag is provided by the open-loop search. Then fractional lag is searched around the open-loop with 1/6 resolution. The LTP-lags are coded with 9 bits for the first and the third sub-frames and delta coded with 6 bits for the second and the last sub-frame. 4 bits are used to code the LTP gains covering only the positive range.

Table 1 : Bit allocation of the proposed full-rate AMR speech candidate codec

	Rate1	Rate 2	Rate 3	Rate 4
LSF	28			
Frame RMS Energy	7			
LTP-Lag	30			
LTP-Gain	16			
Excitation	68	88	120	160
Excitation Gain	16	32	32	32
Speech Coding bits	165	201	233	273

Speech Coding [kbit/s]	8.25	10.05	11.65	13.65
Channel Coding [kbit/s]	13.95	12.15	10.55	8.55
Signaling bits [kbit/s]	0.6	0.6	0.6	0.6

The codebook excitation is an improved ACELP structure. The codebook in the proposed codec aims to achieve bit-rate reduction and quality improvement over the EFR codec. The excitation sequence in the lowest rate is obtained by a single search stage and contains 4 non-zero samples (+1 or -1). In the lowest rate, the sequence is searched from 625 out of 8192 possibilities.

In the three higher rates, the excitation vectors are computed in two stages. The search routine in the first stage is similar to that in the lowest rate. After computing the excitation sequence and its gain, its contribution is subtracted from the reference signal and then a second stage excitation computation is performed. The total

searched indices for each rate in each search stage are shown in Table 2.

Table 2 : Number of searched indices for each operation mode

Codec Mode	Rate1	Rate 2	Rate 3	Rate 4
Stage 1	625	512	343	320
Stage 2	-	121	343	320

In all rates, the excitation gain of each subframe is normalised to the overall RMS energy of the current frame of original speech. This reduces the dynamic range of the secondary codebook gain. Hence, the gain is then quantised more effectively. However, no prediction is used during the parameter quantisation stage i.e. LSF and excitation gain parameters are quantised with no reference to past history.

There are 81 common bits per frame that are allocated as follow: 28-bit LSF, 30-bit LTP-lag, 16-bit LTP-gain and 7-bit frame RMS. The excitation bit allocations, including excitation gain are 84, 120, 152 and 192 bits for 8.25, 10.05, 11.65 and 13.65 kbit/s respectively.

2.2 CHANNEL CODING

Unlike the fixed 10.6 kbit/s channel coding in EFR, the proposed AMR codec provides 4 different levels of error protection as shown in Figure 1. The error protection level depends on the availability of the channel coding bits. Therefore better error protection can be provided using reduced amounts of speech coding bits.

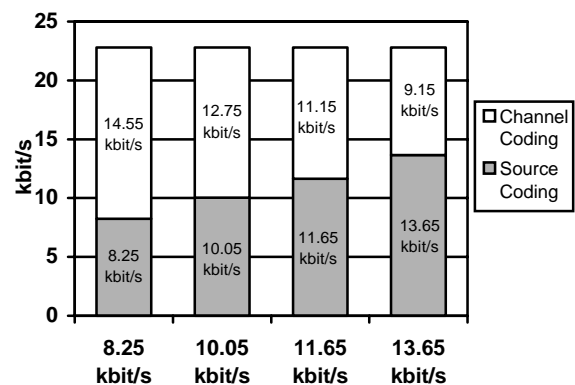


Figure 1 : Bit partitionings for speech and channel coding

The coded speech bits in each frame are grouped into three and four classes for the two lower rates and the two higher rates respectively. A 6-bit tailing sequence is added after the Cyclic Redundancy Check (CRC) bits in order to flush the memory of the encoder.

The total number of speech coding bits which are shown in Table 1 are then half-rate convolutionally encoded. In

the two higher rates, Class-4 bits are not protected and puncturing is applied in order to eliminate the extra bits. The resulting bits are either duplicated or punctured depending on the rate to add up to 22.8 kbit/s.

Table 3 : The channel coding bit allocations per frame in each operation mode.

	Rate 1		Rate 2		Rate 3		Rate 4	
		CRC		CRC		CRC		CRC
Class-1	81	5	81	5	81	5	81	5
Class-2	32	5	52	4	56	4	64	6
Class-3	52	-	68	-	74	-	68	-
Class-4	-	-	-	-	22	-	60	-
Tailing seq.	6		6		6		6	
Total	181		216		248		290	

Error protection is based on a Cyclic Redundancy Check (CRC) and half-rate convolutional coding with a constraint length of seven. CRC codes are used to detect bad frames. Bad/missing data is extrapolated to recover the current frame from the previously decoded frames.

3. RATE ADAPTATION

Rate adaptation mechanism is used to ensure that the AMR system uses the most appropriate rate at any time. It is based on 4 state machines: two operating in the network and the other two operating in the mobile terminal. There are two possible modes. The mode control may either be distributed (i.e. located on both MS and network sides) or centralised (only on the network side), which is required by ETSI. Each pair of state machines share transmit and receive commands but operate independently. Therefore the up-link and down-link can operate at different rates.

For the up-link, the network monitors the link quality and issues the rate command to the mobile terminal. For the down-link, the mobile terminal transmits the estimated link quality to the network and the network issues the appropriate rate if it has to be changed. This in-band signalling can be different for up-link and down-link. In each frame, 12 bits are allocated for a rate control bit stream. These signalling bits include 6 protected controlled bits. The first three bits of the signalling data indicate the status of the state machine.

In the full AMR system, out of band signalling is required for channel mode control to switch the AMR codec between full-rate channel mode for the best speech quality and half-rate channel mode for higher capacity in a clean communication channel.

The overall system performance mainly depends on the speed of the rate adaptation scheme. Since the channel

variation must be tracked accurately for determining the optimum source and channel coding rates.

4. SUBJECTIVE TEST PERFORMANCE

Subjective tests were conducted according to the AMR qualification test plans [2]. The first two tests evaluate the performance of the codec on input speech with and without background noise respectively. Furthermore, the proposed codec is also tested for the effect of switching, speech input level and tandeming under clean speech conditions. Twenty-nine male and female English speaking sentences were used. The original speech signals with 8-bit A-law Modified IRS (MIRS) and Modulated Noise Reference Unit (MNRU) signals at 5, 10, 15, 20, 25, 30 dB were included in the test database. Each test was carried out with different set of twenty-four native English listeners using telephone handsets. The original speech sentences with and without background noise were processed using the standard codecs and a full-rate GSM-AMR codec through a channel interface. The channel interface introduces channel errors at required levels so as to comply with all the test conditions in the ETSI AMR test plans document [2]. These processed speech files were played randomly through out the test. Evaluation of the switching and tandeming performance was included in the tests.

4.1 Clean Speech Performance

The first test was the performance test for clean speech without background noise in clean channel (EC0) and various channel error conditions i.e. C/I ratio of 10dB (EC10), 7dB (EC7) and 4dB (EC4). The test employed Absolute Category Rating (ACR) MOS with a 5-point scale. The reference codecs were G.728 with no errors, EFR and FR. The speech coding, without channel coding of G.728, EFR and FR are 16 kbit/s, 12.2 kbit/s and 12.2 kbit/s respectively.

The subjective test results of the proposed codec with clean speech input in comparison with the reference codecs are shown in Figure 2.

According to the AMR study phase report [1], the full-rate codec candidate is expected to perform similarly to G.728 with no error in the channels, which has C/I ratio of equal or more than 13dB. At C/I ratio less than 13 dB, the codec should be able to perform better than or at least comparable to EFR. Figure 2 shows that at C/I ratio more than 13 dB, the speech quality of the proposed codec operating in 3 lower rates is comparable to G.728 (no error). Moreover, its performance is similar to G.728 at C/I ratio higher than 7 dB when it operates in the lowest mode.

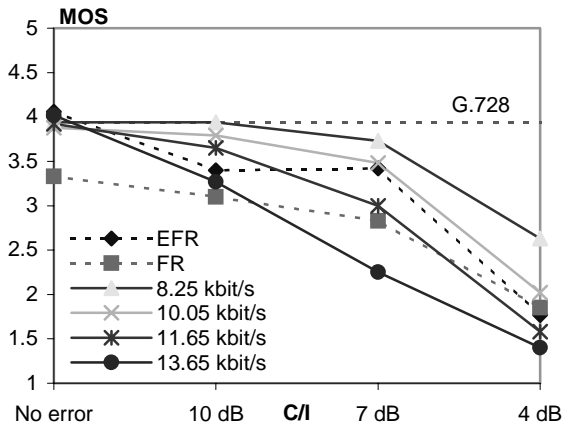


Figure 2 : Subjective results (clean speech).

Furthermore, the proposed codec when operates in two lower rate modes provides substantial improvement to error robustness over the GSM-EFR for all test channel conditions. It also performs better than GSM-FR for error-free and most of the channel error conditions in all rates.

4.2 Background Noise Performance

The second test is the performance test for speech with background noise in clean channel (EC0) and various channel error conditions i.e. C/I ratio of 16dB (EC16), 10dB (EC10) and 4dB (EC4). Degradation Category Rating (DCR) with 5-point scale of degradation was used. The tests are comparison tests of processed speech against direct speech with background noise. They aim to evaluate the degradation of the codec performance under noisy atmosphere.

Both street noise and vehicle noise are used as added background noise at S/N = 15dB. The reference codecs are G.729 (no errors) and the full-rate GSM codec. The speech coding rate, without channel coding of G.729 and FR are 8 kbit/s and 12.2 kbit/s respectively. G.729 codec (8 kbit/s) is used as a reference to low source coding bit-rates.

According to the AMR study phase report [1], the full-rate codec candidate is expected to perform similarly to G.729 and GSM-FR (no error) in channels which have a C/I ratio of 10 dB. Figure 3 shows that the proposed codec performs better than FR and G.729 (no errors) at C/I ratio greater than or equal to 10 dB. At rates of 10.05 kbit/s and 8.25 kbit/s, it provides substantial improvements over GSM-FR at all channel conditions.

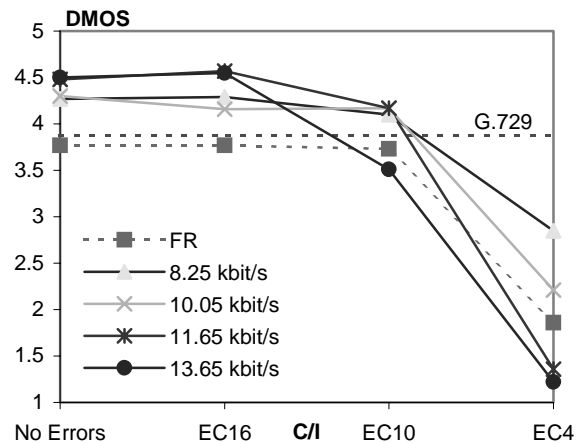


Figure 3 : Subjective result (speech + background street noise input).

Finally, the proposed codec was tested under clean speech conditions on the effect of switching from one rate to another. It provides seamless mode switching and causes no audible glitches.

In tandem operation at nominal level, using the highest speech coding rate, the codec performance is better than EFR. In the proposed solution, the network is able to force all links to use the desired rate. Therefore Tandem Free Operation (TFO) is supported. Hence it satisfies all conditions in this test category [4].

5. CONCLUSION

The proposed codec achieves higher performance than the existing fixed-rate codec standard by selecting an optimal coding rate according to the channel error condition. The subjective tests show that this proposed full-rate GSM-AMR gives improved performance compared to GSM-EFR and GSM-FR.

6. REFERENCES

- [1] ETSI TC SMG, "Adaptive Multi-Rate (AMR) Study Phase Report, Version 1.0", SMG#23 (Tdoc 97-740), Budapest, 13 October 1997.
- [2] ETSI SMG11 AMR#9, "AMR Permanent Document#7, Test plans for the Qualification Phase", Tdoc SQ 31/98R, Versailles, France, March 23-25, 1998.
- [3] ETSI TC-SMG, "Digital Cellular Telecommunications System; Performance Characterization of the GSM Enhanced Full Rate (EFR) Speech Codec (GSM 06.60)", May 1996.
- [4] Mobile VCE, "AMR Speech and Channel Coder", University of Surrey, May 1998.