SPLIT BAND LPC BASED

ADAPTIVE MULTI-RATE GSM CANDIDATE

S.Villette, M.Stefanovic, A.Kondoz

Centre for Communication Systems Research
University of Surrey, Guildford GU2 5XH, Surrey, United Kingdom.
S.Villette@ee.surrey.ac.uk

ABSTRACT

The European Telecommunications Standards Institute (ETSI) has launched a competition for a new mobile communications standard designed to provide better performance than the current GSM standard. This standard is to be called AMR for Adaptive Multi-Rate: the source and channel coding rates can be adapted depending on the state of the channel, thus providing optimal balance between them at any time. The University of Surrey has submitted a candidate for this competition through the Mobile VCE. This candidate was the only one amongst eleven to use a vocoder in the half-rate GSM channel instead of a CELP based coder. The testing which took place as part of the first stage of the competition has shown that this candidate was among the best. This paper presents the system submitted for the half-rate channel as well as the results of the testing.

1. INTRODUCTION

The design requirements for the AMR competition were that the proposed system should be able to operate both in the current full-rate and half-rate GSM channel to allow the reuse of the existing infrastructures [3]. All the candidates opted for a CELP based system similar to the EFR system, to be used in the full-rate channel at 22.8 kb/s. Two to four rates were used in the various systems submitted. The full-rate GSM channel makes it the best choice to use such a system.

However, in the half-rate channel, only 11.4 kb/s are available. The system is expected to maintain good speech quality even at a C/I ratio of 4 dB. In order for the channel coding to keep the number of corrupt frames to a manageable level, we felt desirable to have a source rate below 4 kb/s.

Hence the choice for using a vocoder based system, which would provide better speech quality at 4 kb/s than a CELP based system, and hopefully give good quality at higher source bit rates when the channel is clean. All the other candidates have selected a CELP based system in the half-rate channel.

This paper will focus on the half-rate system only.

The AMR competition consists of several phases. The first phase required each proponent to carry out internal tests on their candidate against various existing standards, following strict guidelines given by ETSI. After this qualification phase, each proponent submitted its solution along with the results. The results of our candidate are presented here.

The vocoder used is an improved version of the Split-Band LPC Vocoder which has been developed at the University of Surrey. It is briefly described in Section 2, together with the improvements made to the model. In Section 3 the complete system is detailed. The results of the qualification phase testing are shown in Section 4. Finally, we summarize the results and present our future work in Section 5.

2. SPLIT BAND LPC VOCODER PRINCIPLES

2.1 General characteristics

The Split-Band LPC Vocoder has been presented in detail in [1,2]. Some improvements have been made, including a varying length windowing, new quantisers, a revised pitch algorithm as well as a new voicing algorithm. The main features have been retained, in particular the voicing quantiser which assumes all harmonic bands below a certain cut-off frequency to be voiced and the rest unvoiced.

The diagrams for the revised encoder and decoder are given in Figure 1 and 2.

2.2 Improvements to the model

The basic model has seen several improvements. In order to improve the pitch and voicing determination under noisy background conditions, the level of background noise is estimated. This is then used to bias the pitch and voicing determination to obtain optimum performance in both clean and noisy background. Results indicate that the voicing algorithm performs very well even at 0dB SNR. At such a signal-to-noise ratio (typically in military applications) only a few pitch errors amounting to less than 0.1% have been detected which were due to significant and sudden increase in the background noise coinciding with either speech on-sets or off-sets.

Variable length windowing has also been introduced. The chosen analysis window should cover at least two pitch cycles, which is the minimum required for the correct analysis. A longer analysis window can affect the accuracy in determining the parameters.

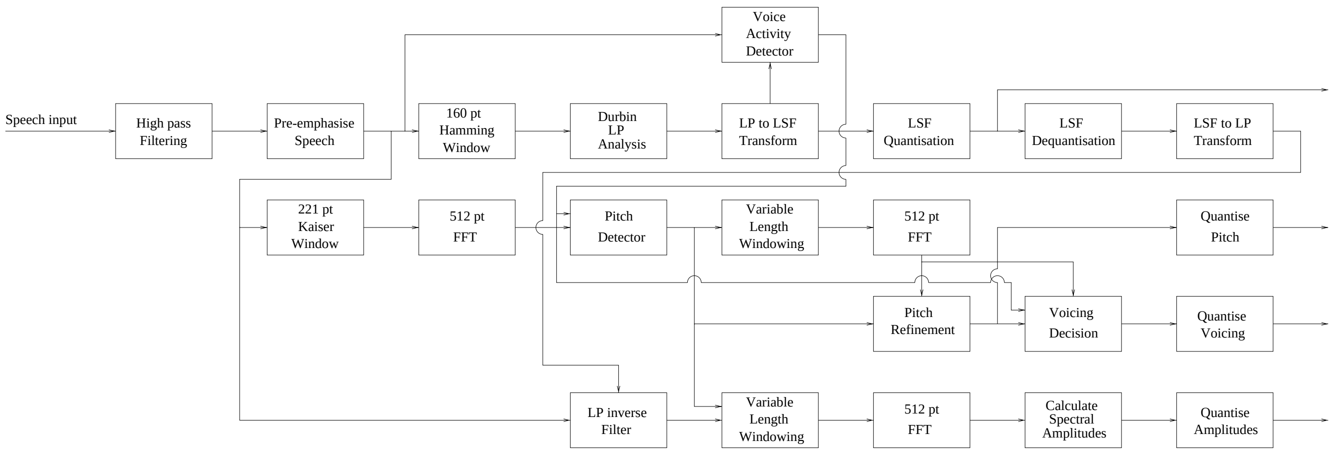

Figure 1: Block diagram of the Split-Band LPC encoder

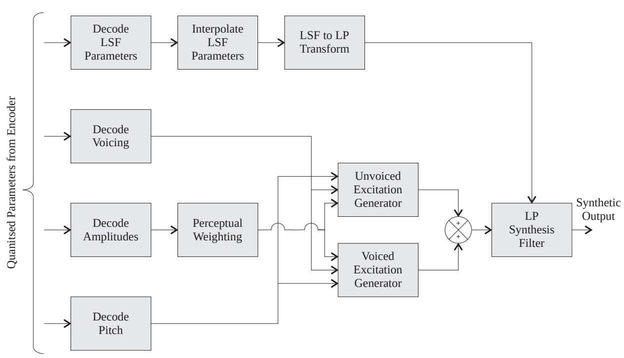

Figure 2: Block diagram of the Split-Band LPC decoder

Finally, new quantisers have been designed for the LPC coefficients and the spectral amplitudes.

3. SUBMITTED SYSTEM FOR THE HALF-RATE CHANNEL

3.1 Operating rates

The original Split-Band LPC Vocoder had been developed to operate at a bit rate of 2.4 kb/s with a frame size of 20 ms. Although very high for this bit rate, the quality of the speech was not high enough to reach a MOS score of 4 which was needed to fulfill the requirements of the AMR competition.

Moreover, as the system is supposed to operate in the half-rate GSM channel, the overall bit rate including channel coding and control bits has to be 11.4 kb/s. This enables us to use higher source bit rates.

Therefore three versions with higher bit rate and higher quality have been designed: 3.9, 5.2 and 6.8 kb/s. The four versions differ only by the quantisation methods used and the update rate of the parameters. Whereas in the 2.4 kb/s version, all parameters were estimated every 20 ms, in the higher bit rate version some are estimated every 10 ms. The bit allocations and update rates are described in Table 1.

Bit Rate	2.4 kb/s		3.9 kb/s		5.2 kb/s		6.8 kb/s	
Update rate (in ms)	20		20		20		20	
	10	10	10	10	10	10	10	10
LPC	28		28		28		28	28
Pitch	7		7	5	7	5	7	7
Voicing	3		4	4	4	4	5	5
RMS energy	6		7	7	7	7	7	7
Spectral amplitudes	4		8	8	21	21	21	21
Total for 20 ms:	48		78		104		136	

Table 1: Bit allocation for the different rates of the Split-Band LPC Vocoder

The quantisers used do not rely on any type of differential coding or inter-frame prediction in order to limit the effect of channel errors. Only the second pitch in two of the rates is encoded differentially using the first pitch to save bits.

3.3 Channel coding

Preliminary studies have shown that the best suited rates to be used in the GSM Half-Rate channel were 3.9, 5.2 and 6.8 kb/s rates, as there was no gain obtained from the 2.4 kb/s version in this application.

A specific channel coder has been designed for each rate. For each rate the bits have been prioritized and CRC checks added to the parameters. Then 6 bits have been added at the beginning for the rate adaptation scheme and a trailing sequence added to flush the bits. This is finally passed through a convolutional encoder of rate ½ and length 7. The resulting bits are either punctured or duplicated depending on the rate to add up to 11.4 kb/s.

The three modes described will be referred as "Sblpc 1,2 or 3" as shown in table 2. All rates add up to 11.4 kb/s, which is the capacity of the GSM half-rate channel.

	source coding	channel coding + rate adaptation
Sblpc 1	3.9 kb/s	7.5 kb/s
Sblpc 2	5.2 kb/s	6.2 kb/s
Sblpc 3	6.8 kb/s	4.6 kb/s

Table 2: Bit rates selected for the AMR competition

3.4 Rate adaptation scheme

In order to always use the most appropriate rate at any time, on the up- and on the down-link, a rate adaptation scheme has to be designed.

It has to synchronize the transmitting rate with the reception rate to allow communication between the base station (BS) and the mobile station (MS) and to decide based on an estimation of the channel quality which rate is to be used. As both links are independent and the estimated bit error rate (EBER) of a link is obtained at the receiving end, EBER and the rate switching decisions have to be exchanged between the MS and the BS.

This is accomplished using 4 state-machines, 2 in the MS and 2 in the BS. They work in pairs, each controlling one of the links, which are independent. Therefore the up- and down-link can operate at different rates. The network is always the master and the mobile is the slave.

For the downlink, the mobile transmits the estimated channel state to the network and the network makes the decision whether the rate is to be changed. This is communicated to the mobile station which follows the network recommendation. For the uplink the network monitors the link quality and issues the rate change command which the mobile acknowledges. These commands are transmitted using 12 bits (including protection) in each frame and over more than one frame. The meaning of the bits vary with respect to the first three bits of the signalling data which indicate the state of the state machines.

The performance of the complete system depends heavily on the speed of the rate adaptation scheme. If it is too slow and a sudden increase in BER occurs, the system might not be able to change rates towards a more channel robust rate, hence losing frames and degrading the quality. In this system, a total delay of 160 ms is needed for a rate adaptation to take place.

4. PERFORMANCE

4.1 Conditions of testing

The testing has been carried out internally following ETSI's Recommendations [3]. It consisted of four parts, each testing a particular aspect of the complete proposed system.

The speech material used was coming from the NTT database. The subjects were naïve listeners with good hearing and were all native English speakers. The testing took place in a dedicated listening room arranged according to the ETSI recommendations [3].

For each of the experiments, many conditions have been tested and only the most significant ones are presented here.

Each experiment required the speech samples to be processed in a specific way: experiment 1 and 4 used M-IRS filtered speech, while experiment 2 and 3 used FLAT filtered speech. DMOS was used in experiment 2, whereas AMOS was used in the other three experiments. Some cases also include processing through a G711 codec. Therefore results from different experiments cannot be compared.

The limited number of subjects (only around 30 for each experiment, the number being fixed by the ETSI testing procedure [3]) leads to a fairly large inaccuracy of the results, which explains why some scores are slightly higher than others where they should in theory be slightly lower.

4.2 Experiment 1: The effect of errors under clean speech conditions

Coder	Sblpc 1	Sblpc 2	Sblpc 3	EFR	FR
Bit rate	11.4	11.4	11.4	22.8	22.8
clean	3.67	3.96	3.85	4.29	N/A
19 dB	3.63	3.67	4.04	N/A	N/A
16 dB	3.81	3.69	3.77	N/A	N/A
13 dB	3.63	3.71	3.38	N/A	3.35
10 dB	3.81	3.58	2.67	4.08	3.40
7 dB	3.25	2.85	1.65	3.69	2.98
4 dB	2.10	1.31	1.00	2.00	2.00
1 dB	1.10	1.02	1.06	N/A	N/A

Table 3: MOS score for the Experiment 1

In this experiment, the AMR candidate performance in the half-rate channel (11.4 kb/s) is compared to EFR and FR GSM standards performance under various error conditions corresponding to a C/I ranging from 1 to 19 dB, plus a clean channel condition. It uses MOS scores.

The EFR and FR are complete with their own channel coder and frame reconstruction techniques, operating in the full rate channel (22.8 kb/s) that is twice the bit rate of the tested AMR candidate.

The fact that Sblpc 3 scores more with a 19 dB channel than with a clean channel is an example of the effect of the limited number of subjects used in the experiments.

4.3 Experiment 2: The effect of background noise for static conditions

This experiment is made up of two parts: street noise and vehicle noise. It used DMOS scores.

The subjects were asked to mark the degradation perceived between the original noisy speech and the processed version of the original. The processing can also include channel errors, in the same conditions as for Experiment 1.

Coder	Sblpc 1	Sblpc 2	Sblpc 3	FR	G729
clean	3.98	3.79	3.86	3.77	3.88
13 dB	3.57	3.86	3.71	3.88	N/A
7 dB	3.36	3.05	2.04	3.57	N/A

Table 4: DMOS scores for Experiment 2, street noise.

Coder	Sblpc 1	Sblpc 2	Sblpc 3	FR	G729
clean	3.84	3.82	3.91	3.88	3.81
13 dB	3.57	3.46	3.49	3.71	N/A
7 dB	3.26	2.49	1.50	3.45	N/A

Table 5: DMOS scores for Experiment 2, car noise.

4.4 Experiment 3: The effect of switching, speech input level and tandeming under clean speech conditions

The AMR candidate should be able to cope with various input level. Part of this experiment consisted of checking the proper operation of the candidate with an input level 10 dB higher or lower than the nominal input level. The candidate was designed to have its nominal level at -26 dB from overload. The scaling of the speech samples was performed using the tools provided by the ETSI for this purpose.

coder	Sblpc1	Sblpc2	Sblpc3	FR	G728	G729
-16 dB	3.69	3.79	3.85	N/A	4.02	3.85
-26 dB	3.23	3.33	3.31	3.23	3.52	3.67
-36 dB	2.50	2.60	2.75	N/A	3.06	2.88
tandem	2.40	2.60	2.92	3.04	3.46	3.10

Table 6: MOS scores for Experiment 3

The effect of switching was also tested. It is important that no artifact should occur when switching from one rate to the other in order to adapt to the channel conditions. The test has shown that switching between rates does not produce any such artifact.

4.5 Experiment 4: The effect of dynamic error patterns

The aim of this experiment is to check the performance of the complete system and to validate the concept of Adaptive Multi-Rate. The complete system is simulated using the dynamic error patterns provided on both links and the rate adaptation scheme controls the switching to optimize the performance. Five different scenarios (called Dynamic Error Condition, DEC) are used, all representative of typical mobile channels.

Coder	Sblpc, 11.4 kb/s	GSM FR, 22.8 kb/s
DEC 1	3.63	3.67
DEC 2	3.57	3.67
DEC 3	3.63	3.66
DEC 4	2.98	2.77
DEC 5	2.82	2.81

Table 7: MOS scores for Experiment 4

4.6 Overall performance

There were eleven candidates for the qualification phase, all of whom were well established companies with only one University: the University of Surrey/Mobile VCE Ltd candidate. After checking that all design constraints had been met by the proponents, the results of the listening tests for the qualification phase were tabulated and the proponents ranked to select the promising ones.

According to the initial figure of merit, our candidate was placed third overall, with the best figure of merit for the half-rate GSM channel.[3]

5. SUMMARY

In this paper we have presented the candidate submitted by the University of Surrey through Mobile VCE Ltd to the ETSI AMR competition. We have presented the speech coder, as well as the channel coder and the rate adaptation scheme.

The results of the ETSI/AMR qualification stage have shown that using a vocoder instead of a CELP coder in the half-rate channel was a viable approach, as our candidate was ranked amongst the best.

An ETSI meeting took place, to decide which candidates should be allowed to go forward to the selection phase, and which ones should be selected. Unfortunately, although among the most promising solutions, the University of Surrey/ Mobile VCE Ltd was not one of the five candidates voted for by the others to go through.

6. REFERENCES

[1] I.Atkinson, S.Yeldener, A.Kondoz, “High Quality Split-Band LPC Vocoder Operating at Low Bit Rates” *ICCASP 97 Proceedings*, Volume 2, pp 1559

[2] S.Villette, M.Stefanovic, I.Atkinson, A.Kondoz “High Quality Split Band LPC Vocoder and its Fixed Point Real Time Implementation” *Eurospeech 97 Proceedings*, Volume3, pp 1243-1246

[3] ETSI AMR Qualification Phase Documentation, 1998