Algebraic Code-Excited Linear Predictive Coding Mechanism for 3GPP Codec

Ganashree T.S, Asst Professor 1, Dr. Josephin Prem Kumar, HOD, 2, and Vidyashree, 7th Sem, ISE 2
1 East Point College of Engineering / ISE, Bangalore, India
Email: ganashreets@yahoo.co.in
2 East Point College of Engineering for Women / CSE & ISE, Bangalore, India
Email: d_prem_k@yahoo.com, varadhya93@gmail.com

Abstract—A variety of ways has been established to detect a new embedded speech coding. The speech coding structure is based on the Adaptive Multi-Rate Wideband+ (AMR−WB+) standard codec. The declared coding scheme consists of three different types of bitrates where the two lower bitrates are embedded into the highest one. The aim of modifying the algebraic codebook search procedure is to achieve an embedded bit stream and the algebraic codebook search procedure adopted for the AMR−WB+ codec. The proposed method offered the advantage of scalability due to the embedded bit stream and some additional calculation complexity for acquiring two different code vectors of the higher bitrate modes. The embedded coder presented highly improved speech qualities for two higher bitrates modes with a slightly increased bitrate generated by the decreased coding efficiency of the algebraic codebook.

Index Terms—wideband speech, Speech coding, embedded speech coder, AMR-WB+.

I. INTRODUCTION

The algebraic code-excited linear predictive (ACELP) coding algorithm is to be adopted for wideband speech codecs like the Adaptive Multi-Rate Wideband (AMR−WB+) standard as well as for many narrow band standard speech codecs like as G.729, AMR-WB+ was standardized in 2004 for streaming and multimedia messaging services in Global System for Mobile communications (GSM). Enhanced Full Rate, Enhanced Variable Rate Coding, and the AMR codec. Presently, most speech codecs used in mobile communication systems have regulated on a narrow bandwidth limited to 200 to 3400 Hz. Moreover, as AMR−WB + enlarges the audio bandwidth from 50 to 7000 Hz, it becomes possible to achieve high quality speech signals both in intelligibility and naturalness. AMR−WB + is a hybrid codec that switches between a time-domain coding model and a transform-domain coding model. The time domain coding model is absolutely the AMR−WB + 3GPP mandatory standard for wideband speech communication [1] and has been selected as the new ITU − T G.722.2 standard [2].

The AMR−WB + codec is a multi-rate codec which consists nine different bitrates between 6.6 and 23.85 kbps. Even though it operates on nine different bit rate, the bit allocation for each codec is very similar because each codec is based on the ACELP algorithm. In spite of that, multimedia communications such as ATM teleconferencing and other various Internet applications are being more widely used. Depending on the application, audio codecs have to face various channel impairments that typically translate into lost frames,
packets or bit errors. Packetized speech communication has become a more important issue. In this packetization process (also called multiplexing) packet losses can occur by the network congestion, exceeding delay constraints, or a buffer overflow. To avoid a large distortion of output speech obtained by packet losses is to utilize an embedded coding scheme [5] – [8]. An embedded coding structure basically consists of a base-layer and an enhancement-layer as shown in Fig. 1.

![Typical embedded coding structure](image)

**Fig. 1: Typical embedded coding structure**

Various Quality levels have been introduced for different version of synthetic speech. The base-layer bit stream guarantees the reformation of the original signals with a less acceptable quality, on the other hand, the enhancement-layer bit stream enables a rather enhanced decoded signal quality. Accordingly, an embedded coding structure usually depends upon an additional functional unit to implement the enhancement-layer.

In this paper, we propose an embedded ACELP speech coding structure based on the AMR-WB+ standard codec without introducing any additional functional unit. Here, embedded speech coder has introduced three different bitrates, and the embedded bit stream can be achieved by modifying the algebraic codebook search method used in the AMR-WB+ codec.

**II. PROPOSED EMBEDDED CODING STRUCTURE (AMR-WB+ BIT STREAM)**

The AMR-WB+ is a multi-rate codec with nine different bitrates. The AMR-WB+ coder works at multiple bitrates, the bit stream for each bitrate is a fixed one. However, if the transmitted bit stream is classified in such a way that the lower bit rate mode can be embedded into the higher bitrate mode, the receiver can reform speech with the available bit stream even when some parts of the highest bit rate bit stream are missed.

The AMR-WB+ codec conduct at nine different bitrates, the bit allocations for the seven highest bitrates, from 12.65 to 23.85 kbps, bit allocation structure is same for each codec; the omission are the algebraic codebook index and the high frequency band energy, which is only used at 23.85 kbps [1]. From this we can conclude that if the coder is able to enhance the algebraic code vector in an embedded coding manner while keeping the other parameters unchanged, an embedded coding structure can be obtained without introducing any additional enhancement unit. Accordingly, we implemented the embedded coding structure by manipulating these similarities in the bit allocation of the AMR-WB+ coder. On the other hand, we propose a way of evaluating the code vectors with the different resolutions by simply modifying the algebraic codebook search procedure.

So, using multi-excitation for bitrates scalability, the proposed coding scheme is similar to general multi-stage agitation coding based on an embedded coding approach which is proposed here [5]. Therefore, the
The proposed coder can be manipulated with a very simple embedded scheme by modifying the codebook search procedure which is adopted in the AMR-WB+ codec. The algebraic codebook of the AMR-WB+ coder, the sub frame is divided into some tracks and each pulse is occupied in these tracks for active modeling of the excitation signal of the sub frame. For the 23.85 kbps mode, the 64 positions contains six pulses which is used to find out the optimal pulse positions, the algebraic codebook is searched by maximizing the search criterion.

\[ Q_k = \frac{(d^t C_k)^2}{C_k^t \Phi C_k} \]

Where \( C_k \) is the code vector with index \( k \), \( d \) is the correlation vector between the target signal and the impulse response of the weighted synthesis filter, and \( \Phi \) is the correlation matrix of the impulse response. The AMR-WB+ coder at the 12.65 kbps mode, the base-layer codec of our embedded coding structure, a best 8-pulse code vector implements a two-track-based sequential search method. While testing the possible merger of two pulses, a defined number of potential positions of the first pulse are tested for better complexity reduction. However, first two pulses are set into the positions according to the maximum values in each track. The remaining pulses are searched in pairs by searching continuously each of the pulse in the consecutive tracks.

The procedure to find the three optimum code vectors for the proposed structure is shown in Fig. 2. In the present proposed embedded coder, the algebraic codebook is obtained by using the depth-first tree search method adopted for the 12.65 kbps mode AMR-WB+ [1]. the embedded coder can get three code vectors instantly by enhancing the search procedure of the depth-first tree search process. The bitrate scalability is achieved by these three code vectors.

A. Iteration

we can summarize the whole process of Fig. 2 for getting the three code vectors as follows a) To get the total of eight pulses per frame, here the two pulses are placed in each tracking for the lowest bitrate mode in the codebook at the same time to finding the code vector consisting eight pulses b) two pulses that are searching each sub frame are divided into four tracks where each track to be searched should be in tracks T0 and T1, T2 and T3, or T4 and T0.

![Fig. 2: Block diagram for generating three code vectors](image-url)
Hence, in this case search tree is having four levels. Pulses I0 and I1 are allotted to tracks T0 and T1 in first case. No search is executed in these levels and the position of both pulses is set to the two maxima of the prespecified cited signal in each track. This iteration of process will be repeated until the positions of all eight pulses are found out. If the two pulses are assigned to different tracks, then this process of assigning track will be done for four times.

After getting the position of eight pulses the search iteration will be continued as previously mentioned method for finding the position of additional eight pulses. Therefore, by assigning four pulses in each track we can have the position of 16 pulses. Again doing the search iteration with four levels according to the Fig. 2, we can obtain the code vector which is having six pulses in every track. For finding the additional 16-pulses for higher bit rate mode we did not take in consideration the iteration procedure to reduce the computational complexity. At same bit rate, the embedded coding structure will need same complexity for searching the mid-rate and lowest-bit-rate modes as AMR-WB+. The embedded coding structure need more computation than AMR-WB+ for high bit rate mode. Now we have found the three types of code vector, hence it becomes easier to organize the bit stream.

III. EVALUATION AND EXPERIMENTAL RESULTS

By this observation, the proposed embedded coder has the advantage of generating scalability at the cost of slightly enhanced computation and encoding the efficiency at a greater bitrate. At the time of encoding the pulses in a track some redundancy occur for the case of more than one pulse in a track. Therefore, in proposed embedded coder this can be encoded at once for bit stream scalability. As a result, coding efficiency for encoding the pulses is decreased in our embedded scheme.

Moreover, the required bits for encoding the pulses are the same for the AMR-WB+ which is shown in Table 1. In this proposed coder, the compression ratio is different for every audio sample and the Algebraic code vectors for the higher bitrate are achieved by searching from wider searching space than that of the ARM-WB+.

However, the length are different for every sample and the ARM-WB+ of sample 1, 2 and 3 is 192000, after a certain period it is changed. The better quality of ARM-WB+ at 52.8% compared to the other audio sample.

<table>
<thead>
<tr>
<th>Audio sample</th>
<th>Length</th>
<th>ARM-WB+ (Before)</th>
<th>ARM-WB+ (After)</th>
<th>Compression Ratio (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample 1</td>
<td>0.1s</td>
<td>192000</td>
<td>101376</td>
<td>52.8</td>
</tr>
<tr>
<td>Sample 2</td>
<td>0.4s</td>
<td>192000</td>
<td>107776</td>
<td>56.1333</td>
</tr>
<tr>
<td>Sample 3</td>
<td>0.5s</td>
<td>192000</td>
<td>127744</td>
<td>66.533</td>
</tr>
</tbody>
</table>

IV. CONCLUSION

In this paper, we proposed an embedded coding structure based on the standard AMR-WB+ coder. The proposed coder has three different bitrate modes and the embedded bitstream was implemented by modifying the algebraic codebook search method used in the AMR-WB+ coder. Thereafter, the proposed coder can provide the advantage of scalability, which is implemented without any additional functional unit for calculating compression ratio of each audio sample. The experimental results show that our coder provides good scalability and maintaining the same or good quality for all different audio sample with a slightly increased bitrates due to decreased coding efficiency of the proposed coder.

REFERENCES

