An Investigation into the Performance Enhancement and Salient Features of AMR-WB Speech Codec Used in GSM Communication

Anwar Ahmad[1], M. T. Beg [2]

[1](Associate Professor, Dept. of Electr. & Comn. Engg. Jamia Millia Islamia, New Delhi-110025, INDIA)
[2](Professor, Dept. of Electr. & Comn. Engg., Jamia Millia Islamia, New Delhi-110025, INDIA)

ABSTRACT : *The 3rd Generation Partnership Project (3GPP) along with European Telecommunication Standard Institute (ETSI) recommended new speech coding scheme, called Adaptive Multi Rate (AMR) in 2000 and 2001. The proposed scheme has two variants: AMR-Narrowband (AMR-NB) and AMR-WB (AMR-Wideband). The speech quality of AMR-WB in GSM systems is remarkably high compared with that of AMR-NB speech codec because it encompasses wider bandwidth of the speech signal, 50-7000 Hz, instead of 200-3400 Hz in AMR-NB. But there are always possibilities for improvement. This paper investigates the salient features of AMR-WM codec and various possibilities for further enhancement in its speech quality. Two algorithms are recommended for incorporation into the standard AMR-WB to achieve this goal.*

Keywords – *ACELP, AMR, AMR-NB, AMR-WB, GSM, Speech Codecs, Speech Quality*

I.   INTRODUCTION

In cellular mobile communication systems the main challenges are to provide high quality speech even under the worst conditions and to design a bandwidth efficient system. To meet these challenges highly efficient and robust speech codec is required. The 3rd Generation Partnership Project (3GPP) along with European Telecommunication Standard Institute (ETSI) standardized two variants of a new speech transmission system, called Adaptive Multi Rate (AMR) speech codec. These two variants are: AMR-NB [1] and AMR-WB [2].

The speech quality in mobile communication depends upon channel conditions. The AMR codec employs variable bit-rate operation depending upon the channel conditions instead of fixed-rate operation. The variable bit-rates are identified as modes of operation. The AMR codec dynamically switches to specific mode depending upon the channel conditions. The low bit-rate modes are adapted in poor channel conditions whereas higher bit-rate modes are adapted in good channel conditions. This technique provides a means to vary the bit-rate of the channel coder to provide adequate bit protection to the transmitted speech data. The channel coder enhances its bit-rate under poor channel conditions and reduces its bit-rate under good channel conditions. Thus, depending upon the channel performance, AMR and channel codecs dynamically tune their bit-rates to provide acceptable speech quality under all channel conditions.

Each of the speech codec modes is associated with a corresponding channel codec mode to provide the gross transmission bit-rate. The gross bit-rate is, therefore, shared dynamically between speech and channel coders as per the channel conditions [7].

Another advantage of AMR is dynamic trade-off between speech quality and system capacity. When traffic load of the system is high, the system may transfer connections which operate under good channel conditions to Half-Rate (HR) channel modes. The speech quality in HR modes is compromised but due to availability of more spectrum, the system capacity is enhanced. On the other hand, the speech quality is enhanced when the traffic demand is low by switching more connections to Full-Rate (FR) mode [7].

In this paper we investigate the salient features and possibilities of enhancing the speech quality of AMR-WB source codecs.

II.   AMR-WB: GENERAL DESCRIPTION

AMR-NB speech codec operates within a narrow speech bandwidth limited to 200-3400 Hz whereas to enhance the quality and naturalness in the delivered speech, a wider bandwidth of 50-7000 Hz is used in AMR-WB. AMR-WB is superior to AMR-NB in terms of better fricative differentiation and therefore higher intelligibility due to its wider bandwidth operation [5]. But the operational features of both the codecs are more or less the same under dynamic channel conditions. The typical AMR behavior is the trade-off between speech quality and error robustness along with the channel quality [5].

AMR-WB includes a set of fixed rate speech and channel codec modes, a Voice Activity Detector (VAD), Discontinuous Transmission (DTX), in-band signaling for codec mode transmission and Link Adaptation to control the mode selection.

The AMR technology is based upon the Analysis-by-Synthesis (AbS) method of source coding. This method utilizes the linear prediction (LP) model and a perceptual distortion measure to reproduce only those characteristics of input speech that are determined to be most important [8]. The AMR codec employs Algebraic Code-Excited Linear Prediction (ACELP) method for speech analysis-by-synthesis. The block diagrams of encoding and decoding algorithms are recommended in [3].

AMR-WB codec operates at 16 kHz sampling rate. The speech bandwidth is divided into two frequency bands: low frequency band (50-6400 Hz) and high frequency band (6400-7000 Hz). These frequency bands are coded separately to reduce the complexity and to allocate more bits to subjectively most important frequency range [8].

The input speech signal is down-sampled to 12.8 kHz and pre-processed using high pass filter. The ACELP algorithm is then applied to down-sampled and pre-processed signal. The LP analysis is performed once per 20ms frame and LP parameters are converted to Immitance Spectrum Pairs (ISP). The ISP are then vector quantized using split multi-stage vector quantization with 46 bits. The speech frame is divided into 4 sub-frames of 5ms each (64 Samples). The adaptive codebook and fixed codebook parameters are transmitted every sub-frame. The pitch lag is encoded with 9 bits in odd sub-frames and relatively encoded with 6 bits in every sub-frame. One bit per frame is used to determine low pass filter applied to past excitation. The pitch and codebook gains are jointly quantized using 7 bits per sub-frame [8].

III.   CODECS MODES AND SELECTION

1. Codec Modes

There are speech coding modes in AMR-WB with bit-rates of 23.85, 23.05, 19.85, 18.25, 15.85, 14.25, 12.65, 8.85 and 6.6 kbps. There is another mode that is used in DTX operation with a bit-rate of 1.75 kbps. The 12.65 kbps and higher modes are used in good channel conditions and offer high quality wideband speech. The lowest modes 8.85 and 6.65 kbps are used in poor channel conditions or during high traffic load [5].

2. Mode Selection and In-band Signaling

The codec mode selection is based upon the current performance of the channel measured in terms of Carrier-to-Interference (C/I) ratio. The channel estimation is carried out at the decoder and this information is sent to the encoder through reverse link for codec mode selection.

The signaling of all information needed for codec mode selection is done in-band using some of the bits normally available for source and channel coding. Adaption requires two kind of information: a codec mode command sent through downlink and channel estimation parameters via uplink [9].

For uplink transmission, the base station monitors the channel conditions and decides which codec mode the mobile station should use. For downlink transmission, the mobile station computes a downlink measurement, which is quantized and transmitted to base station via uplink channel. This is done in-band using one bit delta modulation. The base station then decides which codec mode it will use for the downlink transmission of the next frame [9].

IV.   QUALITY OF SSPECH

The quality of speech in GSM system is the ultimate parameter and determines the quality of the system. Due to this reason, there are continuous efforts to improve the speech quality further. The quality of speech may be improved by adaptive gain re-quantization of fixed codebook (FCB) [10]. This approach is called Sub-band CELP (SB-CELP). The input speech signal is split into two unequal sub-bands. The lower SB is 0-6 kHz whereas the upper SB is 6-7 kHz.

A variable rate, multi-mode ACELP codec is applied to the lower SB. The various bit-rates are integrated in a common structure where the scalability is realized by exchanging fixed excitation codebook while leaving other codec parameters invariant. The upper SB is coded using a very low bit-rate representation based on bandwidth extension technique [10].

The SB-CELP encoder used a method to determine the FCB gain factor that helps to overcome the unwanted gain fluctuations as they appear in CELP coders when coding pure background noise segment or speech in background noise [10].

Another approach to improve the speech quality is the adaptive thresholds for codec mode operation of AMR-WB [11]. The codec mode operation depends upon the channel performance measured in terms of Quality Indicator (QI) that is equivalent to C/I ration [4]. The QI is compared with the pre-defined thresholds associated with each codec mode and the codec mode for the current QI factor is decided. The codec moves to the mode indicated by the QI.

Once a codec mode is adapted, it remains fixed during a call in standard GSM operation. In case the channel conditions vary during the call, the speech quality suffers as the coder cannot change its mode. But the system has liberty to change the code whenever required. This feature of AMR codec is utilized to apply the adaptive thresholds for the selection of the codec mode during a call, if required.

The algorithm proposed in [11] estimates the speech quality on the receiving link i.e. at the mobile station, and compares the estimate against given speech quality for each codec mode. If the estimated speech quality for a given mode is outside its limits, it indicates that the associated threshold for switching to the appropriate codec mode is sub-optimal as per the current channel performance. The algorithm then modifies the thresholds of all the codec modes accordingly. It is necessary to modify the thresholds of all codec modes to maintain the thresholds in a consistent order and avoid overlapping. The adaptive threshold approach can improve the objective speech quality significantly [11].

V.   CONCLUSION

AMR-WB codec delivers a superior speech quality due to its wider speech bandwidth up to 7 kHz, use of ACELP and dynamic codec mode adaptation as per the current channel performance. In this paper, we presented the salient features of AMR-WB codec and investigated the possibilities to further improve performance for GSM system. One of the techniques we recommended is the SB-CELP algorithm that makes the use of adaptive gain re-quantization of fixed codebook [10]. The other recommendation to incorporate is the use of adaptive thresholds for various codec modes as per the current speech quality [11].

REFERENCES

[1]   *"AMR Speech Codec: General Description", 3GPP TS 26.071.*
[2]   *"AMR Wideband Speech Codec: General Description", 3GPP TS 26.171.*
[3]   *"AMR Wideband Speech Codec: Transcoding", 3GPP TS 26.190.*
[4]   *"AMR Wideband Speech Codec: Link Adaption", 3GPP TS 45.009.*
[5]   B. Bessette, R. Salami, R. Lefebvre, M. Jelinek, J. Rotola-Pukkila, J. Vainio, H. Mikkola, and K. Jarvinen, "The Adaptive Multi-rate Wideband Speech Codec (AMR-WB)", *IEEE Transaction on Speech and Audio Processing, Vol 10, No 8, 2002, pp. 620-636.*
[6]   C. Erdmann, P. Vary, K. Fischer, Wen Xu, M. Marke, T. Fingscheidt, I. Varga, M. Kaindl, C. Quinquis, B. Kovesi, and D. Massaloux,"A Candidate Proposal For 3GPP Adaptive Multi-Rate Wideband Speech Codec", *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), (Salt Lake City, Utah, USA), Vol. 2, May 2001, pp. 757–760.*
[7]   A. Uvliden, S. Bruhn, and R. Hagen," Adaptive Multi-Rate: A Speech Service Adapted to Cellular Radio Network Quality", *Thirty-Second Asilomar Conference on Signals, Systems and Computers, California, 1998, pp. 343-347.*
[8]   R. Salami, B. Bessette, R. Lefebvre, M. Jelinek, J. Rotola-Pukkila, J. Vainio, H. Mikkola, and K. Jarvinen, "The Adaptive Multi-rate Wideband Codec: History and Performance", *Speech Coding, 2002, IEEE Workshop Proceedings, Oct. 2002, pp. 144-146.*
[9]   E. Paksoy, D. Martin, A. McCree, G. Gerlach, A. Anandakumar, W. Lai, and V Viswanathan," An Adaptive Multi-Rate Speech Coder For Digital Cellular Telephony", *IEEE International Conference on Acoustics, Speech, and Signal Processing, Phoenix, AZ, Vol.1,1999, pp. 193 - 196.*
[10]   C. Erdmann, P. Vary, K. Fischer, J. Stegmann, C. Quinquis, , D. Massaloux, and B. Kovesi,"An Adaptive Multi Rate Wideband speech codec With adaptive Gain Re-Quantization", *Proceedings of IEEE Workshop on Speech Coding (SCW), Delavan, Wisconsin, USA, September 2000, pp. 145-147.*
[11]   T. Lundberg, P. Bruin, S. Bruhn, S. Hakansson, and S. Craig," Adaptive Thresholds for AMR Codec Mode Selection", *Vehicular Technology Conference, Vol. 4, 2005, pp. 2325 – 2329.*