International Journal of Current Engineering and Technology
E-ISSN 2277 – 4106, P-ISSN 2347 - 5161

Available at http://inpressco.com/category/ijcet

Research Article

Steganography Approach of Weighted Speech Analysis with and without Vector Quantization using Variation in Weight Factor

Nikita Atul Malhotra[A*] and Nikunj Tahilramani[A]

[A]Electronics and Communication, C.G.P.I.T, U.T.U Bardoli, Gujarat India,

Accepted 10 May 2014, Available online 01 June 2014, **Vol.4, No.3 (June 2014)**

Abstract

This Paper basically focuses on the analysis and synthesis of the speech signal which takes reference from the speech coding technique of Conjugate-Structure Algebraic Code Excited Linear Prediction (CS-ACELP) and also the concept of vector quantization. The proposed technique analyses the variation in weighted speech by the changes in the weight factor and also by implementing it with and without vector quantization. This algorithm is based on CELP coding technique which works on analysis by synthesis principle with the frame size of 80 samples (10 ms). Steganography is implemented in the LSB of the input cover speech which contributes to the high frequency component of the cover speech. The weighted speech signal is generated using both quantized and unquantized LP parameters, which is being tested on various grounds to basically observe and analyze its overall performance. At the end an analytic comparison between the original and the weighted speech signal both subjectively as well as objectively is done. The values of the two weight factors produce a considerable change in the weighted speech quality.

Keywords: *CS-ACELP, quantized LP parameters, steganography, unquantized LP parameters, vector quantization*

1. Introduction

This steganography technique is based on CELP coding algorithm which works on analysis and synthesis principle i.e. It has decoder embedded within the encoder. The LP parameters of the coder are selected in such a manner that the mean square error between the input speech as well as the weighted speech is as less as possible. A Block diagram of a synthesis-by-analysis codec is shown in the figure 1. Main functional parts include the following:

1.1Frame segmentation

The input speech is fragmented into frames of N blocks of samples per frame. The number of samples per frame usually varies from 80 to 160 (ITU-T, CS-ACELP *et al* 01/2007) (with the frame length of10 ms -20 ms respectively).

1.2 Pre-emphasis

Pre-processing of the input involves the decimation of the input speech to 8 kHz and then the two pre-emphasis operations involving signal scaling and high pass filtering (ITU-T, CS-ACELP *et al* 01/2007) is applied before the encoding process. These two pre-emphasis functions are implemented by the filter given below (ITU-T, CS-ACELP *et al* 01/2007)

$$H(Z) = \frac{0.46363718-0.92724705z^{-1}+0.46363718z^{-2}}{1-1.9059465z^{-1}+0.9114024z^{-2}} \quad (1)$$

*Corresponding author: **Nikita Atul Malhotra**

1.3 Speech parameterisation

For each of the input frame LP parameters are calculated using the Levinson Durbin algorithm which uses the auto-regressive models and calculates the LP parameters from the auto-correlation values depending on the filter order(in our case it is 10). Then these obtained LP coefficients are transformed to the LSP domain for the purpose of quantization and interpolation. These values of interpolated LSP for each sub frame are used to get back the unquantized LP filter coefficients which are being used to reconstruct the weighted speech using the synthesis (weighting) filter.

LPC-analysis and encoder

The 10^{th} order linear prediction filter is used for analysis and synthesis. The filter order of 10 is decided based on the human vocal tract model (Tamanna Islam et al, April 2000). The synthesis filter is given by equation below (ITU-T, CS-ACELP *et al* 01/2007):

$$\frac{1}{A(z)} = \frac{1}{1+\sum_{i=1}^{10} a_i z^{-i}} \quad (2)$$

Where a_i is the quantised LP parameters where i ranges from 1 to 10 as depicted from 10^{th} order LP filter. These LP parameters are calculated once per frame from the auto correlation values.

1.4 LP to LSP conversion

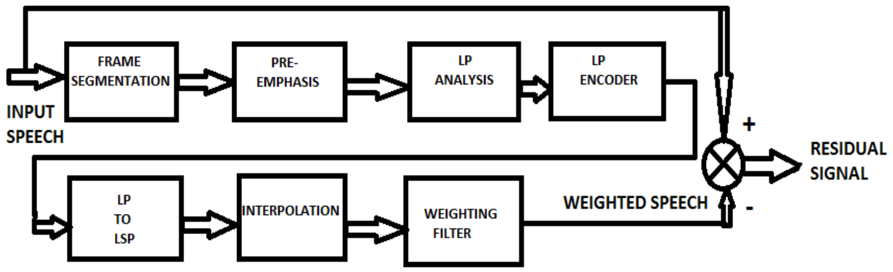

Fig. 1 Block Diagram of Encoder

Normally the obtained LP parameters are converted to LSF (line spectral frequency) representation due to its following properties which makes it a desired choice for LP parameters to be converted to that domain including (Tamanna Islam *et al*, April 2000)

- Bounded range
- Sequential ordering of the parameters
- Simple stability check for the filter (its roots lie on the unit circle)
- The LSF representation is a frequency-domain representation which helps to the exploit and model the human auditory perception.

1.4.1 Quantization of the LSP coefficients

In order to define sum and difference polynomials as given below, the inverse filter polynomial is used (Tamanna Islam *et al*, April 2000)

$$P(z) = A(z) + z^{-(M+1)}\, A(z^{-1}) \quad (3)$$
$$Q(z) = A(z) - z^{-(M+1)}\, A(z^{-1}) \quad (4)$$

Where M is the LP filter order

The LSP (line spectral pair) coefficients are defined the roots of the sum and difference polynomials P (z) and Q (z)

1.4.2 LSF representation

The LSF representation ω_i lies in the range [0, π] which are quantised using the given equation: that is (ITU-T, CS-ACELP *et al* 01/2007)

$$\omega = cos^{-1}(q) \quad (5)$$

Where q denotes the LSP coefficients

1.5 Vector quantisation

Vector quantisation is a form lossy data compression also called as block quantisation.

The input speech is segmented into frames with each frame of K-dimension. The search engine chooses the nearest matching vector from a set of n-dimensional vectors of the random or fixed codebook. All possible combinations of the n-dimensional vector form the input code-book to which all the quantized vectors belong. So instead of transmitting the quantized LSP parameters only the index of the codeword in the codebook is transmitted.

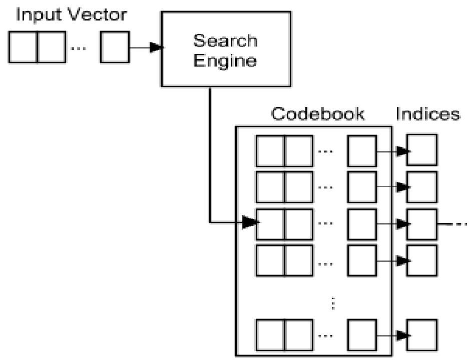

Fig. 2 Vector quantization

1.6 Interpolation of the LSP coefficients

Interpolation guarantees a smooth and improved quality of synthetic speech due to the fact that it smoothens the variation of the filter coefficients as a function of time without having the need to transmit any additional information.

In order to balance the trade-off between frame size and bit rate, a frame is usually divided into sub-frames and this interpolation is usually done at the sub-frame basis (Tamanna Islam *et al*, April 2000). If the interpolation is directly implemented on LP coefficients it does not assure the filter stability, so it is thereby converted to LSF representation.

Interpolation is done on the LSP coefficients which work in the cosine domain. For the first sub-frame the LP coefficients are obtained by linear interpolation in the adjacent sub-frames, whereas for the second frame the quantized LP coefficients are obtained from the previous frame.

The interpolated LSP coefficients for each of the two sub-frames are given by [ITU-T, CS-ACELP *et al* 01/2007]:

$$Subframe\ 1{:}\ q_i^1 = 0.5 * q_i^{previous} + 0.5 * q_i^{current} \quad (6)$$

$$Subframe\ 2{:}\ q_i^2 = q_i^{current} \quad (7)$$

Where i =1,...., 10

1.7 Weighting filter

The weighting filter is built taking into consideration the human auditory masking system. It basically distributes the error energy in the high energy region so that the noise energy is masked and is inaudible to the human listener. The spectrum of the weighted speech should as closely resemble the spectrum of input speech. The commonly used weighting filter is (ITU-T, CS-ACELP *et al* 01/2007):

$$s_w(n) = s(n) + \sum_{i=1}^{10} a_i \gamma_1^i s(n-i) + \beta s_w(n-1) \quad (8)$$

Where n= 0 to L-1
s(n) is the input speech
a represents the LP coefficients of that frame
β=0.68
γ_1=0.4 to 0.7

2. Results

This technique is carried out on music files taken from the http://www.repository.voxforge1.org/downloads/SpeechCorpus/Trunk/Audio/Original/16kHz_16bit
2.1 Subjective analysis

MOS rating

	quantized	unquantized
Raai0002.wav	2.5603	2.5607
Raai0006.wav	1.8734	1.7488
Raai0009.wav	1.9125	1.3721

2.2 Objective analysis

PSEQ
β=0.68; γ_1=0.61

	quantized	Unquantized
Raai0002.wav	2.2603	2.2407
Raai0006.wav	1.5934	1.4388
Raai0009.wav	1.5725	1.0325

Seg_SNR

	Quantized	Unquantized
Raai0002.wav	1.8534	1.8141
Raai0006.wav	1.9929	1.9805
Raai0009.wav	2.0741	2.0654

2.3 Variation of PSEQ on values of weighting factor
or the raai0002.wav file, the PSEQ values are calculated for various values of γ

	Quantized	Unquantized
γ_1 = 0.48	2.1562	2.1562
γ_1 =0.56	2.3117	2.2366
γ_1= 0.58	2.3125	2.2378
γ_1 =0.6	2.3120	2.2401
γ_1 =0.61	2.2603	2.2407
γ_1 =0.64	2.2273	2.1787

Conclusion

The proposed technique analyses the role of variation in weight factor on the perceived speech quality and its overall performance. The values of weight factor ranges from 0.4 to 0.7 with the optimum speech quality obtained in the mid-range rather than selecting the lower bound or upper bound. This research motivates to explore the domain of steganography by selecting the optimum speech as the cover speech which would result in a high quality stego speech without introducing much noticeable distortion. Nevertheless vector quantization adds to the complexity of the coding algorithm still it enhances the speech quality and also enables to hide data in lower number of bits.

References

Samir Kumar(June 2012), BandyopadhyayBarnali, Gupta Banik , LSB Modification and Phase Encoding Technique of Audio Steganography Revisited, *International Journal of Advanced Research in Computer and Communication Engineering* Vol. 1, Issue 4

WU Zhi-jun GAO Wei, YANG Wei(december,2009), LPC parameters substitution for speech information hiding ,*The Journal of China Universities of Posts and Telecommunications*, 16(6): 103–11

Fatiha Djebbar, Beghdad Ayad Karim , Abed Meraim and Habib Hamam(2012), Comparative Study of Digital Audio Steganography Tech-niques, *EURASIP journal on audio, speech and music processing, springe*

Jayaram P, Ranganatha H R, Anupama H S(August 2011), Information Hiding Using Audio Steganography – A Survey, *The International Journal of Multimedia & Its Applications (IJMA)* Vol.3, No.

Xiao Bo, Huang Yongfeng, ShanyuTang (2008), An Approach to Information Hiding in Low Bit-rate Speech Stream, *IEEE GLOBECOM proceedings*.

Song-bin LI, Huai-zhou Tao, Yong-feng Huang(2012), Detection of quantization index modulation steganography in G.723.1 bit stream based on quantization index sequence analysis, *Journal of Zhejiang University-SCIENCE C (Computers & Electronics)*, 13(8):624-634

Siwar Rekik, Driss Guerchi, Sid-Ahmed Selouani and Habib Hamam, (2012) Speech steganography using wavelet and Fourier transforms, *EURASIP Journal on Audio, Speech, and Music Processing*,2012:20

Jin Liu and Ke Zhou and Hui Tian (2011) Least-significant-digit Steganography in Low Bitrate Speech, *National Basic Research Program (973 Program) of China under Grant No. 2011CB302305, Natural Science Foundation of Fujian Province of China under Grant No. 2011J05151, and Scientific Research Foundation of National Huaqiao University under Grant No. 11BS210*

Driss Guerchi and Fatiha Djebbar(2009), Narrowband Speech Hiding using Vector Quantization, *International Journal of Information and Communication Engineering* 5:8

Zaidoon Kh. AL-Ani, A.A.Zaidan, B.B.Zaidan and Hamdan.O.Alanazi(2010), Overview: Main Fundamentals for Steganography, *Journal of Computing*, Volume 2, Issue 3, March 2010

Nedeljko Cvejic (2004), Algorithms For Audio Watermarking And Steganography, *thesis presented in University of Oulu*

Mean opinion score, http://en.wikipedia.org/wiki/Mean_opinion_score

ASCOM *(2009)*, The PESQ Algorithm as the Solution for Speech Quality Evaluation on 2.5G and 3G Network Technical Paper

Premalatha P and Amritha P (2009) *Optimally Locating for Hiding Information in Audio Signal*, *International Journal of Computer Applications* (0975 – 8887) Volume 65– No.14, March 2013

C. Vanmathi, S. Prabu (2013), A Survey of State of the Art techniques of Steganography, C. Vanmathi et al. / *International Journal of Engineering and Technology (IJET)*, Vol 5 No 1 Feb-Mar 2013

Ashwini Mane.,Gajanan Galshetwar,.Amutha Jeyakumar(2012), Data Hiding Technique: Audio Steganography Using Lsb Technique, *International Journal of Engineering Research and Applications (IJERA)*, Vol. 2, Issue 3, May-Jun 2012, pp.1123-1125

K. Gopalan, Cepstral Domain Modification of Audio Signals for Data Embedding – Preliminary Results, *thesis Department of Engineering Purdue University Calumet Hammond*, IN 46323

Adel Almohammad (2010), Steganography-Based Secret and Reliable Communications: Improving Steganographic Capacity and Imperceptibility, *thesis submitted for the degree of Doctor of Philosophy in Brunel University*, August, 2010

Itu - T Cs-Acelp(01/2007) G.729 Telecommunication Standardization Sector Of It

Itu - T G.722 Telecommunication Standardization Sector of Itu, 07/2003

http://www.repository.voxforge1.org/downloads/SpeechCorpus/Trunk/Audio/Original/16kHz_16bit

L.R Rabiner and R.W. (1978) Schfer, Digital processing of speech signals , Prentice-hal

Wai C. Chu.(2003) Speech Coding Algorithms, Foundations & Evolution of Standardized coders, Willey Publications

Tamanna Islam(April 2000) Interpolation of Linear Prediction Coefficients for Speech Coding, Department of Electrical Engineering McGill University Montreal, Canada

Johnson Ilyeh Agbinya (1996) Discrete wavelet transform techniques in speech processing, *IEEE TECON-digital signal processing applications*