ERROR CONFIGURATIONS IN BIT PATTERNS OF COMPRESSED DIGITAL SPEECH BASED ON CASCADED APPROACH

Nazish Nawaz Hussaini , Asadullah Shah , Abdul Wahab Ansari

Abstract :
The research challenges in digital speech processing remain in the traditionally identified area like speech compression. The main purpose to compress speech is to reduce the load of voice traffic for transmission, maintaining the quality of speech. In the initial stage of proposed research a cascaded approach is designed to reduce the bit rate and to the maintain speech quality. This coding technique is developed upon a PCM coded speech by dropping few of the least significant bits (LSB) of every byte and randomly substituting these dropped bits at the decoding end from their respective fixed Codebooks. To analyze the speech quality Mean Opinion Score is conducted, which is a subjective quality analysis. Experiments are conducted to refine the strategy by configuring the LSB bits after substitutions from their respective fixed codebooks but of reduced sizes. It is observed that as the codebooks containing the less number of bit patterns with lesser numeric values like 0 and 1, when substituted, their bytes have no changes in the valuable least significant bits.

Index Terms: Least Significant Bits, Bit Configurations, Fixed Codebook, Mean opinion Score

I. INTRODUCTION

Compression techniques are developed to cope with the problems of limited memory, and deal with bandwidth requirements. A telephonic speech is specified in International Telecommunication Union as standard G.711, producing 64Kbps speech [1]. G.711 is pulse code modulation (PCM) standard produces output at 64 kilobits per second (Kbps) [2]. Differential PCM encodes the differences of PCM values; this type of encoding reduces to the number of bits required to store a sample of about 25% lesser as compared to PCM [3]. Adaptive Differential PCM (ADPCM) is defined in ITU−T G.726 standard. ADPCM sets the level for quantization to the size of the input difference signal at a bit rate reduced to 32 kbps for voice transmission [4]. Most widely used speech coding algorithm proposed yet is Code Excited Linear Prediction and its variations, these are providing significantly better quality of speech than other low bit-rate algorithms [5].

Institute of Mathematics & Computer Science, Sindh University Jamshoro, Hyderabad, 76080, Pakistan, Email: nazish_hussaini@hotmail.com,
nazish.nawaz@usindh.edu.pk
Department of Computer Science, IBA Sukker,
dr_asadullahshah@hotmail.com
Department of Computer Science, Isra University
Hala Naka Road, Hyderabad,Pakistan,
wahabansari@hotmail.com

CELP coding scheme is described in federal standard 1016, providing good quality, at 4.8 Kbits per second, some of its variants are like G. 723.1 CELP based, at 5.3 and 6.4 Kbits/sec and G.729 CELP based, at 8 Kbps [6].

II. LITERATURE REVIEW

The progress of multimedia depends on the solution of technical problems of encoding, storage, transmission, distribution, security and privacy, recognition, understanding, indexing and searching, also subjective quality analysis of multimedia material. An appropriate model of source produces effective transformation of an input stream to an output stream. Performance evaluation of compression algorithms includes metrics like efficiency, compression ratio, bit rate achieved after compression, percentage of compression, scale of perception like MOS and others. Speech can be compressed in telephones with A-law or u-law, with a reduced bit rate of 64kbps. Speech compression can be achieved through parametric or source method like CELP, Waveform method like ADPCM, producing compression rates above 32 kbps. Wavelet and fractal coding are examples of transform method, producing compression rates 2.4kbps or even lower. Code Excited Linear Predictive (CELP) is one of the better compression schemes belonging to parametric scheme and compress speech to 4.8 kbps [8].

Compression systems like CELP, VSELP, GSM 06.10 and Artificial Neural Networks ANN, focus on the fact of voiced and unvoiced sounds. The main cause is the speaker dependency that hinders the standardization of new voice compression systems [9]. Another proposed technique uses lossy compression algorithm that consider perceptual and rate distortion criteria. The achiever bit rate is 54 to 64 kbps. The decoder implements this algorithm effectively in real time [10].

It is analyzed in a research on protecting real time speech signal over Internet, (based on Forward Error Correction) that the performance of G.729 decoder fails to cover up the loss of voiced and unvoiced frames at an unvoiced/voiced transition. The impact of frame loss within a speech signal on the quality and gained the knowledge that the “loss of voiced frames at the beginning segment leads to a significant degradation in speech quality while the loss of other frames covered up well by the G.729 decoder’s concealment algorithm” [13]. A research on “Digitizing

Speech Recordings for Archival Purposes", described digitizing speech recordings for archival purpose. Analyzing frequency response, dynamic range, noise, psychoacoustic and perceptual quality, spectral evaluation of recordings was used to develop digitization best practices. Digitization with sampling rate: 96KHz; on 24-bits; The WAV file format is recommended for speech recordings and are easy to process into a variety of streaming formats [5].

An approach used for the digital speech is to drop 3 least significant bits from every byte at the source end, and substitute randomly the dropped bits at the destination end, from the respective codebook. The codebook for the substitution of 3 dropped bits is of 2^3=8 bit patterns size [14].The research is also carried out on Sindhi and Urdu languages to compare the speech quality of both languages after compression [15]. The strategy [14] is then refined by reducing the codebook size or splitting the codebooks in two equal halves. Byte errors are analyzed to observe error free and error occurring bytes [16].

This refined strategy is further explored to see the impact on each and every byte of digital speech. Bytes after substitution of 3 bits are checked to know the number of changed bits. Bits are configured to see whether out of 3 substituted bits, mostly how many bytes are having changes in their 0th bit only, or changes occurred in 1st bit or 2nd bit individually. Configuration of bits also evaluates the bytes having changes occurred only in 0th and 1st bit, 0th and 2nd bit, 1st and 2nd bit i.e., changes in consecutive bits. This strategy focuses on the bytes having no changes in their least significant bits too i.e.; before and after substitutions they remain same or the bytes after substitution changed totally i.e.; all substituted bit are change bits.

Bit errors occurred in each byte of the .pcm speech from the codebooks of original size and reduced sizes in the refined strategy are shown in Fig-2.1 shows an abstraction for the 3 bit substituted bytes:

A. Codebook containing 8 bit patterns

The Codebook of 8 bit patterns size is used in the first experiment (with respect to 3 dropped bits) i.e., {0,1,2,3,4,5,6,7} or {000, 001, 010, 011, 100, 101, 110, 111}. When bit patterns are picked randomly from that codebook and substituted as least significant bits at the destination end, then it is observed that in all the bytes of the compressed speech file, various bytes are having 50% changes in their 0th, 1st, and 2nd least significant bits respectively.

The file named as F8 (created with codebook of 8 bit patterns size), and the bit configurations are shown in table-2.1. In this experiment F8-file compressed after substituting 8 bit patterns, scored 4.22 MOS (Mean Opinion Score) out of 5.

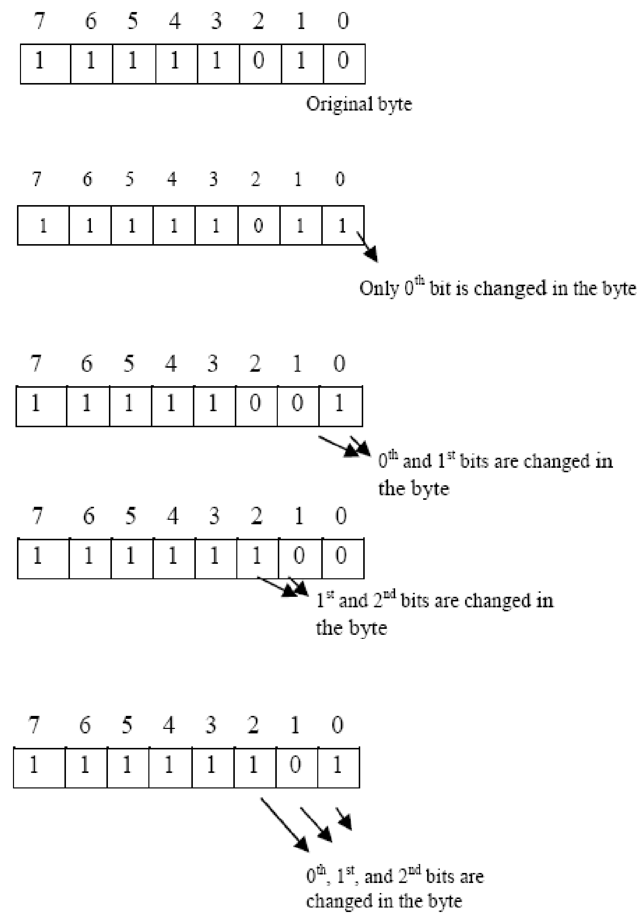

Fig: 2.1. Abstraction for the 3 bit substituted bytes

B. Codebooks containing 4 bit patterns

In the second experiment the Codebook when reduced to two halves, then the first half at left hand side contains 4 bit patterns {0,1,2,3} and the second half at right hand side contains rest of other 4 bit patterns {4,5,6,7}. The compressed speech files with respect to the corresponding codebooks are named as FL48 (File substituted with codebook of 4 bit patterns from the left side values) and FR48 (File substituted with codebook of 4 bit patterns from the right side values).

After substitution of these patterns having lesser numeric values to all the bytes, various bytes are having 50% changes in their 2nd least significant bits respectively i.e., 0th and 1st bit out of 3 least significant bits and 2nd bit of these bytes remained unchanged at the decoding end. The Codebook containing rest of other 4 bit patterns {4,5,6,7}, when substituted as the 3 least significant bits, 50% various bytes found error occurring in their 0th and 1st least significant bit respectively while all the 100% bytes contain error in their 2nd least significant bit at the decoding end. See the bit configurations in table-2.1. As the bit configuration status proceeds change in the bit causes more noticeable error.

. In this experiment FL48-file compressed with the respective 4 bit patterns at left side (lesser numeric values) scored 3.99 and FR48-file compressed with the respective 4 bit patterns at Right side (greater numeric values) scored 3.5 out of 5.

C. Codebooks containing 2 bit patterns

In the third experiment, the Codebooks containing 2 bit patterns each {0,1},{2,3},{4,5},{6,7} named as FLL28, FLR28, FRL28, and FRR28.

After substitution it is observed that from these four fixed reduced codebooks the codebook having {0,1} bit patterns gives better results to listeners and from all the bytes, 50% bytes are having changes in their 0th least significant bit only and no changes found in their other bits, also rest of 50% Bytes found error free. See the bit configurations for this and rest of files in table-2.1. The codebook containing {2,3} bit patterns having 0th least significant bit changes in 50% various bytes while 100% bytes found error in 1st least significant bit, but the 2nd least significant bit of all the bytes of the speech file remained same at the decoding end.

The codebook having {4,5} bit patterns made changes in 50% various bytes in their 0th bit from all the bytes while 100% bytes found change in their 2nd least significant bit, it is observed that 1st bit of all the bytes of the speech file remained as they were before substitution. The codebook having {6, 7} bit patterns caused changes in all the 100% bytes in their 0th, 1st , and 2nd bits of the speech file after substitution.

. In this experiment the files having 2 bit patterns each as codebooks to substitute are FLL28, FLR28, FRL28, and FRR28. These contains {0,1},{2,3},{4,5},{6,7} bit patterns, the MOS scored is 4.22, 4.10,4.10, 3.45 out of 5 respectively.

III. RESULTS AND DISCUSSIONS

Our proposed research work is based on simple strategy to compress data and can compete to the standard coding schemes. The approach drops 3 least significant bits in different experimentations and substitutes these bits with their respective codebooks. The refined strategy uses codebooks of reduced sizes. The strategy also focuses on the codebook sizes and

also values of the bit patterns they have. On substituting 3 least significant bits with 3 bit patterns out of 8 patterns, very good speech quality is achieved with 37.5% compression. The research work also focused on bit configurations i.e., changes in a byte with respect to their 0th , 1st , and 2nd least significant bits for 3 bits substitutions. It is concluded that codebooks of reduced sizes having less number (values) of bit patterns with numeric values like 0 and 1, when substituted, 50% of bytes found unchanged after substitutions.

Table-2.1. Bit Error Configurations after 3 Bit Substitutions Radio recorded Speech

Compressed files (After 3 bit substitution)		**Bit Error Configurations**					
		0th Bit		1st Bit		2nd Bit	
File Name	**Bit Patterns**	No. of Bytes	Error %	No. of Bytes	Error %	No. of Bytes	Error %
F8	{0,1,2,3,4,5, 6,7};	40018	49.80	40082	49.88	40237	50.07
FL48	{0,1,2,3};	40309	50.16	40016	49.80	0	0.00
FR48	{4,5,6,7};	40150	49.97	40257	50.10	80355	100.00
FLL28	{0,1};	40101	49.90	0	0.00	0	0.00
FLR28	{2,3};	40324	50.18	80355	100.00	0	0.00
FRL28	{4,5};	40153	49.97	0	0.00	80355	100.00
FRR28	{6,7};	40179	50.00	80355	100.00	80355	100.00

REFERENCES

[1] http://en.wikipedia.org/wiki/G.711, August 17, 2007.

[2] http://whatis.techtarget.com/definition/0,,sid9_gci1251936 00.html Apr26, 2007.

[3] http://en.wikipedia.org/wiki/Pulse-code_modulation, 19 August 2007.

[4] http://www.cisco.com/warp/public/788/signalling/ waveform_coding.html, Feb 02, 2006.

[5] http://en.wikipedia.org/wiki/CELP, 30 June 2007.

[6] http://www.otolith.com/otolith/olt/lpc.html, October 95

[7] Shijo M Joseph, Firoz Shah A, Babu Anto P, Spoken Digit Compression: A Comparative Study between Discrete Wavelet Transforms and Linear Predictive Coding, International Journal of Computer Applications (0975 – 8887), Volume 6– No.6, September 2010.

[8] W. Kinser, Compression and its Metrics for Multimedia, Proceedings of the First IEEE International Conference on Cognitive Informatics (ICCI'02).

[9] Javeir Bustos and Alejandro Bassi, Voice Compression Systems for Wireless Telephony, XXI International Conference of the Chilean Computer Science Society (SCCC'01), pp. 0041, November 2001.

[10] Rodrigo Capobianco Guido, Lucimar Sasso Vieira, Fabrycio Lopes Sanchez, Jan Frans Willem Slaets, Lyrio Onofre Almeida, Adilson Gonzaga, Marcelo Bianchi , A Matched FIR (Finite Impulse Response) Filter Bank for Audio Coding, pp. 796-801, Proceedings of the Seventh IEEE International Symposium on Multimedia (ISM'05).

[11] J Light, Performance Analysis of Audio Codecs over Real-Time Transmission Protocol (RTP) for Voice Services over Internet protocol, Proceedings of the Second Annual Conference on Communication Networks and Services Research (CNSR'04).

[12] http://whatis.techtarget.com/definition/0,,sid9_gci896692,%20% 2000.html, Apr 03, 2007.

[13] Henning Sanneck and Nguyen Tuong Long Le, Speech Property-Based FEC for Internet Telephony Applications, to appear in Proceedings of the SPIE/ACM SIGMM Multimedia Computing and Networking Conference 2000 (MMCN 2000), San Jose, CA, January 2000.

[14] Nazish nawaz Hussaini and Prof. Dr. Asadullah Shah, "Binary codebook Based Digital Speech Compression", SZABIST Journal of Independent Studies and Research, Vol. 5, No. 2, July 2007.

[15] Nazish nawaz Hussaini and Prof. Dr. Asadullah Shah, "Affects upon Urdu and Sindhi Language Perceptions Using Codebook Based Digital Speech Compression", proceedings of Shaikh Ayaz International conference on Language and Literature, Sindh University, Jamshoro, January 2007.

[16] Nazish nawaz Hussaini and Prof. Dr. Asadullah Shah, "Byte Error Analysis Based on Cascaded Approach for Digital Speech Compression", Bahria University Journal of Information and Communication Technology (BUJICT)-Karachi, Vol. 2, Issue I, ISSN 19994974, December 2009.