



Discussion of Effective Speech Communication under Different Compression Approaches

¹Amit, ²Sunita

¹Student, M.Tech, ²Asstt. Professor,
Deptt. of Computer Sc. & App., College,
Shri Baba Mastnath Engineering College,
Rohtak, Haryana, India

Abstract: Speech is most interactive way of communication to interact with a person directly or distance person. The distance communication is low cost and does not require any specialized device. The only drawback of speech communication is the size of communicating data. As the data size increases, the communication criticality increases. To provide the effective network voice communication, some compression approach is required that will reduce the communication size without affecting the quality of voice. In this paper, a discussion of compression and encoding approaches is done. The paper also discussed some of the effective compression based techniques so that the effective and reliable communication will be drawn.

Keywords: Speech Compression, Encoding, DWT, Noise Reduction.

I. INTRODUCTION

Speech is one of the most interactive communication medium to perform the distance mode communication. Lot of work is done under different applications of speech processing so that the efficient and reliable communication will be drawn based on speech. Speech processing comes under the signal processing. That requires the encoding of the signal along with different associated applications. Speech processing one of the fundamental as well as critical application area because of its real time involvement. There are many applications that uses the speech as the biometric feature to perform the person identification. The application areas associated with speech processing includes the speech recognition, speaker recognition etc. These application areas are shown in figure 1.

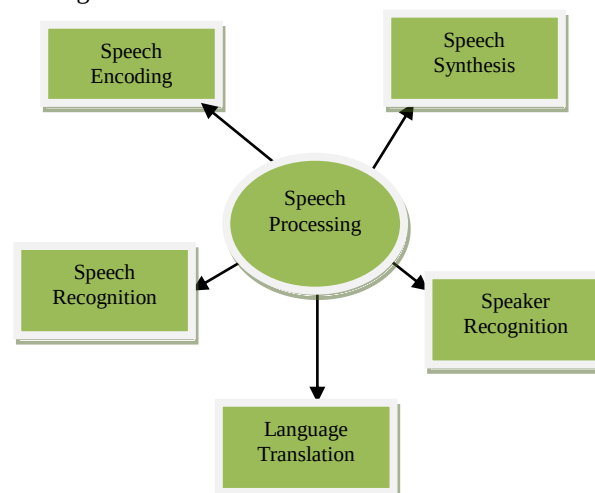


Figure 1: Application Areas of Speech Processing

As the raw speech is collected it is in the form of acoustic signals. To perform different operations and processing, this speech is transformed so that the effective communication as well as processing will be performed over it. The speech processing also depends on multiple communication vectors. These vectors are shown in figure 2. The communication reliability depends on these all factors so that to ensure the effective processing and decisions, the reliable communication vector is drawn. While performing the speech processing, the memory requirement and the processor requirements are high as compared to some traditional applications.

These parameters affect the speech taken from the real time environment. As the speech is collected, the first factor affects the person or the speaker itself. The speech processing outcome different person to person. It is possible that the application is running on the person voice but not working for other. If a person record his voice multiple times, the results can differ. The difference in the speech can be obtained in terms of pitch or the speech. Pitch represents the speech frequency or the intensity and the speed criticality.

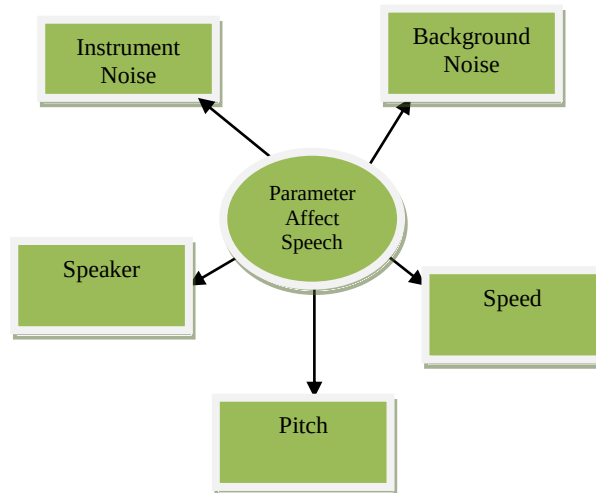


Figure 2: Parameters that affects Speech Processing

In this paper, a study on different compression approaches on speech is discussed. In this section, a study on the speech processing its application and the criticality parameters is defined. In section II, the work done by the earlier authors is discussed in same application area. In section III, the basic compression techniques are discussed. In section IV, the conclusion obtained from the work is presented.

II. RELATED WORK

Lot of work is already done by different researchers in the area of speech processing and the compression techniques. Some of the work done by earlier researchers in same area is discussed in this section. A work on effective compression technique was done by H.B. Kekre[2] in year 2008. Author used the vector quantization method to compress the speech signal and to obtain the effective compression ratio. The work is simulated by the researcher and the obtained results shows the effective outcome is driven. The outcome is presented under different performance parameters. The parameters included in this research work are SNR evaluation, compression ration and absolute error identification. Another work on compression approaches is done using Fractional change analysis. This work is defined by S. Kumar[9] in year 2010. Author combined the vector quantization along with fractional analysis so that effective compression ratio for speech signal will be obtained. Author attain the effective compression ratio and maintained the quality of the signal.

A DWT based work is defined by M.A Najih for speech compression. Author applied a series of wavelet filters so that the best filter for the speech processing will be elected and low bit rate transmission will be performed. Author[10][11] defined the work under different schemes so that the communication reliability will be obtained. The encoding techniques adapted by the author includes DWT, Thresholding, Quantization, Huffman Encoding, Speech signal reconstruction etc. The comparative analysis is performed to generate the signal and to perform the filtration. Author obtained the effective compression ratio as well as bit error rate over the signal is reduced. The parameters considered by the author include the frequency based analysis as well as time domain based analysis. This analysis provided the pitch prediction approach so that high quality speech signal will be composed. Author claims the effective compression ratio so that the signal reliability is attained. The quality signal obtained from the work also depend on the codec used in the signal. A work based on codec processing was defined by Jean-Marc Valin[13]. Author presented the work on different frequency ranges as well as low and high quality signals. Author gain the algebraic vector quantization so that the frequency domain oriented pitch prediction will be obtained as well as the signal process will be done effectively.

Another DWT based compression was defined by H. Elavdi[14]. This compression technique is lossy compression that provided the high compression ratio but degraded the speech quality. The compression was performed using DWT under the bandwidth limits. Author defined the comparison using speech coder so that the effective coefficient vector based encoding will be performed.

K L Neville [15] worked on DWT on speech with different coefficient vectors. Neville focused on analysis of wavelet compression effects. Two wavelet compression techniques, thresholding and low-subband filtering have been utilized. K T Talele in 2012 [16], proposed a simple speech compression algorithm using sub-band

division and ADPCM algorithm. Although speech data are stored in a semiconductor memory device, its capacity and the available network capacity are limited. Therefore, it is necessary to compress the data as much as possible.

S Vatsa did speech compression based on DCT and DWT. The main objective of work is to achieve the compression with encoding approach [17]. In this paper DWT and DCT (discrete cosine transform) based speech compression techniques are implemented with run length encoding, Huffman encoding and their combination M Suman in [18] dealt with multistage vector quantization technique used for coding of narrow band speech signals. The parameter used for coding of speech signals are the line spectral frequencies, so as to ensure filter stability after quantization.

III. SPEECH COMPRESSION TECHNIQUES

The Speech compression techniques are broadly classified as either lossless or lossy, depending, respectively, on whether or not an exact replica of the original Speech could be reconstructed using the compressed Speech. In addition to using the spatial and spectral redundancies, lossy techniques also take advantage of the way people see to discard data that are perceptually insignificant. Lossy schemes provide much higher compression ratios than lossless schemes. Lossless compression is used only for a few applications with stringent requirements such as medical imaging. Lossy schemes are widely used since the quality of the reconstructed Speech is adequate for most applications. Taxonomy of Speech compression techniques is given here under

A. Run length encoding

This encoding technique replaces sequences of identical symbols (pixels), called *runs* by shorter symbols. This technique is usually used as a post-processing step after applying a lossy technique to the Speech and obtaining a set of data values that are suitably reordered to get long runs of similar values

B. Huffman coding

This is a general technique for coding symbols based on their statistical occurrence frequencies (probabilities). The pixels in the Speech are treated as symbols. The symbols that occur more frequently are assigned a smaller number of bits, while the symbols that occur less frequently are assigned a relatively larger number of bits. Huffman code is a *prefix code*. This means that the (binary) code of any symbol is not the prefix of the code of any other symbol. Most Speech coding standards use lossy techniques in the earlier stages of compression and use Huffman coding as the final step.

C. Lempel-Ziv coding

This approach is based on storing frequently occurring sequences of symbols (pixels) in a dictionary (table). Such frequently occurring sequences in the original data (Speech) are represented by just their indices into the dictionary.

D. Bit-plane encoding

This encoding approach is defined under the binary representations of the values of the pixels in the Speech are considered. The corresponding bits in each of the positions in the binary representation form a binary Speech of the same dimensions as the original Speech. This is called a *bit plane*. Each of the bit planes can then be efficiently coded using a lossless technique. The underlying principle is that (in most Speeches) the neighboring pixels are correlated. That means the values of the neighboring pixels differ by small amounts. They can be captured by the representation of pixel values in *gray code* so that the values of neighboring bits in the bit planes are similar. This makes the individual bit planes amenable for good compression.

All known lossy Speech compression techniques take advantage of how we see things. We derive the spatial frequencies of an Speech and suitably allocate more bits for those frequency components that have more visual impact. We then allocate fewer bits, or even discard the insignificant components. The resulting Speech is represented with fewer bits and reconstructed with a better closeness to the original.

D. Block truncation coding

This approach uses the concept to divide an Speech is into non-overlapping blocks of pixels. For each block, *threshold* and *reconstruction* values are determined. The threshold is usually the mean of the pixel values in the block. Then a bitmap of the block is derived by replacing all pixels whose values are greater than or equal (less than) to the threshold by a 1 (0). Then for each segment (group of 1s and 0s) in the bitmap, the reconstruction value is determined. This is the average of the values of the corresponding pixels in the original block.

The *lossy DPCM* is very similar to the lossless version. The major difference is that in lossy DPCM, the pixels are predicted based on the “reconstructed values” of certain neighboring pixels. The difference between the predicted value and the actual value of the pixels is the differential (residual) Speech. It is much less correlated than the original Speech. The differential Speech is then quantized and encoded.

E. Subband coding

This approach analyzes the Speech to produce the components containing frequencies in well defined bands, the *subbands*. Subsequently, quantization and coding is applied to each of the bands. The advantage of this scheme is that the quantization and coding well suited for each of the subbands can be designed separately.

IV. CONCLUSIONS

In this paper, a discussion on different encoding or the compression mechanism is defined for speech encoding. The paper has discussed the speech processing along with different limitations and dependencies. The paper also discussed different encoding mechanisms adapted by speech processing.

REFERENCES

- [1] H Hermansky, "Perceptual linear predictive (PLP) analysis of speech", Journal of the Acoustical Society of America, pp 1738-1752, 1990.
- [2] H. B. Kekre, Tanuja K. Sarode, "Speech data compression using vector quantization", Int'l Journal of Computer and Information Engineering, Vol. 2, No.8, pp 535-538, 2008.
- [3] M H Johnson, Abeer Alwan, "Speech coding: fundamentals and applications", Wiley Encyclopedia of Telecom., pp 1-20, 2003.
- [4] David Suendermann, Harald Höge, and Alan Black, "Challenges in speech synthesis", Speech Technology, Springer Science-Business Media, pp 19-32, 2010.
- [5] Bhupinder Singh, Neha Kapur, Puneet Kaur, "Speech recognition with Hidden- Markov model: A review", Int'l Journal of Advanced Research in Computer Sc and Software Engineering, Vol. 2, No. 3, 2012.
- [6] D A Reynolds, "An overview of automatic speaker recognition technology", Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference, Vol 4, pp 4072-4075, 2002.
- [7] N.S. Bansod, S Kawathekar, S.B. Dabhade, "Review of different techniques for speaker recognition system", Advances in Computational Research, Vol. 4, No. 1, pp 57-60, 2012. (Journal)
- [8] Y Zhang, "Survey of current speech translation research", Carnegie-Mellon Univ, 2003. (Paper)
- [9] Chaudhari, V.K. ; Srivastava, M. ; Singh, R.K. ; Kumar, S. "A new algorithm for voice signal compression and analysis suitable for limited storage devices using Matlab", Int'l Journal of Computer and Electrical Engg. Vol. 1, No. 5, pp 656-665, 2009.
- [10] M A Najih, Ramli, A.R. ; Prakash, V. ; Syed, A.R. "Speech compression using discreet wavelet transform", Telecommunication Technology, 2003. NCTT 2003 Proceedings, pp 1-4, 2003.
- [11] M A Najih, Ramli, A.R. ; Ibrahim, A. ; Syed, A.R., "Comparing speech compression using wavelets with other speech compression schemes", Research and Development, 2003. SCORED 2003., pp 55-58, 2003
- [12] V Radha, "Comparative analysis of compression techniques for Tamil speech datasets", IEEE-Int'l conference on recent trends in Information Technology, pp 712-716, Chennai, 2011.
- [13] J M Valin, T.B. ; Montgomery, C. ; Maxwell, G., "A high-quality speech and audio codec with less than 10 ms delay", IEEE Trans. on Audio, Speech and Language Processing, Vol. 18, No. 1, pp 58-67, 2010.
- [14] H Elaydi, "Speech compression using wavelets", Electrical & Computer Engineering department-Islamic University of Gaza, Palestine, 2010. (Report)
- [15] Katrina L Neville, Zahir M Hussain, "Effects of wavelet compression of speech on its Mel-Cepstral coefficients", Int'l Conference on Communication, Computer and Power (ICCCP'09), pp 387-390, Muscat, 2009.
- [16] KT Talele and S T Gandhe, "Speech compression using ADPCM", IJCA Proceedings on International Conference in Computational Intelligence, Vol 8, 2012 New York, USA.
- [17] S Vatsa, O. P. Sahu, "Speech compression using discrete wavelet transform and discrete cosine transform", International Journal of Engineering Research & Technology, Vol. 1, No. 5, 2012
- [18] M.Suman,, Habibulla Khan,, M. Madhavi Latha, D. Aruna Kumari, "Dimensions of performance in compressed speech signals and its enhancement", International Journal of Engineering Sciences Research, pp 87-93,,2011.
- [19] Preet Kaur, Pallavi Bah, "Comparative analysis between DWT and WPD techniques of speech compression", IOSR Journal of Engineering, Vol. 2, No. 8, pp 120-128, 2012.