Scientific Journal Impact Factor (SJIF): 1.711

e-ISSN: 2349-9745
p-ISSN: 2393-8161

International Journal of Modern Trends in Engineering and Research
www.ijmter.com

# FREQUENCY SPARSE AUDIO SIGNAL USING BY DCT AND COMPRESSIVE SAMPLING

Ms. J.N.Phasale[1], Mr. C.S.Patil[2], U.S.Jathar[3]

[1,2] *E&TC S.G.D.C.O.E.Jalgaon.* [3]S.N.D.Yeola.

**Abstract—**The discrete cosine transform (DCT) and the compressive sampling (CS) are two signal processing techniques with many applications on a great number of engineering fields. In this paper, we advise to apply both techniques to the compression of audio signals. Using spectral analysis and the properties of the DCT, we can treat audio signals as thin signals in the frequency domain. This is particularly true for sounds representing tones. On the other hand, CS has been traditionally used to acquire and compress certain sparse images. We propose the use of DCT and CS to obtain an efficient representation of audio signals, especially when they are sparse in the frequency.

By using the DCT as signal preprocessor in order to obtain a sparse representation in the frequency domain, we show that the successive application of CS represent our signals with less information than the well-known sampling theorem. This means that our results may possibly be the basis for a new compression method for audio and speech signals.

**Key words—**Audio signals, Compressive sampling, DCT, PSNR, SNR,Sparsity.

## I. INTRODUCTION

Speech is very basic way for humans to express information. The main purpose of Speech is communication. Speech can be defined as the response of vocal track to one or more excitation signal. vast amount of data transmission is very difficult both in conditions of transmission and storage. Speech Compression is a method to translate human speech into an encoded form in such a way that it can later be decoded to get back the original signal. Compression is basically to eliminate redundancy between nearest samples and between adjacent cycles. Major aim of speech compression is to represent signal with lesser number of bits. The decrease of data should be done in such a way that there is tolerable loss of quality.

**TYPES OF COMPRESSION -**

### 1.1. Lossless Compression

It is a group of data compression algorithm that allows the accurate original data to be reconstructed from the exact original data to be reconstructed from the compressed data. It is mostly used in cases where it is main that the unique signal and the decompressed signal are almost same or indistinguishable. Examples of lossless compression are Huffman coding.

### 1.2. Lossy Compression

It is a data programming method that compresses data by removing some of them. The plan of this technique is to reduce the amount of data that has to be transmitted. They are generally used for multimedia data compression. The rest of the paper is planned as follow; section 2 gives the Theoretical situation about the speech compression schemes. The speech compression techniques are described in section 3& Section 4 evaluates the presentation of the proposed system followed by the termination.

## II. BLOCK DIAGRAM

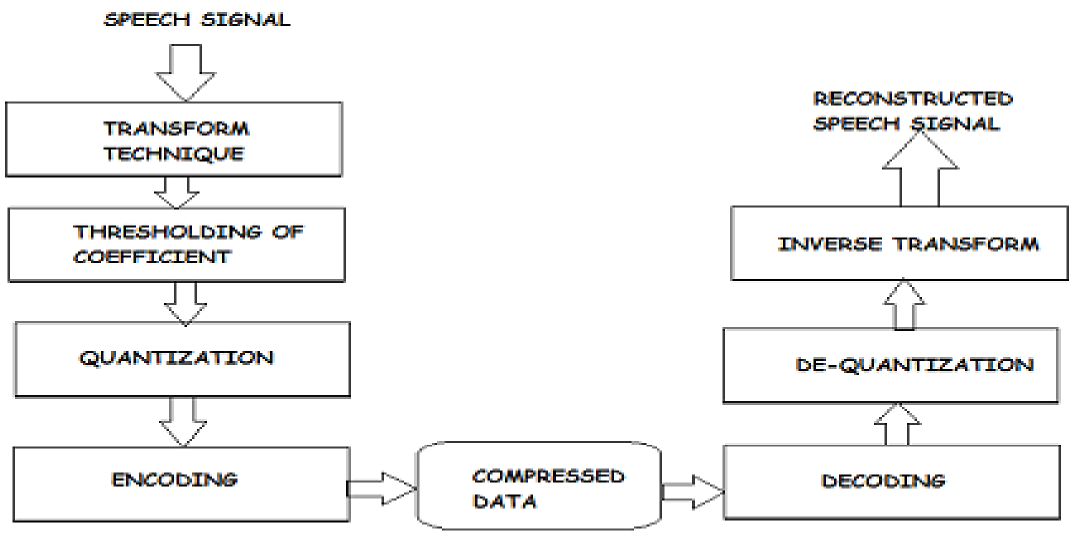


*Fig-1: Block diagram of speech compression*

### *2.1. TRANSFORM TECHNIQUE-*

DCT methods are used on speech signal. Using DCT, reconstruction of signal be capable of be done very accurately; this property of DCT is worn for data compression. Localization characteristic of wavelet along with time frequency motion property makes DWT very suitable for speech compression. The most important idea after signal compression using wavelets is connected mainly to the comparative shortage of the wavelet domain demonstration of signal.

### *2.2. DISCRETE COSINE TRANSFORM*

Discrete Cosine Transform can be used for speech compression because of high link in adjoining coefficient. We can reconstruct a series very accurately from very few DCT coefficients. This property of DCT helps in valuable reduction of data. **3.2.1 The One-Dimensional DCT**

The most common DCT definition of a 1-D sequence of length N is

$$u(t) = \alpha(u)\sum_{x=0}^{N-1} f(x)cos\{\Pi(2x+1)|2N\} \qquad \text{--------(1)}$$

for $u = 0,1,2,\ldots,N-1$. Similarly, the inverse transformation is defined as

$$f(x) = \sum_{u=0}^{N-1} \alpha(u)C(u)cos\{\Pi(2x+1)u|2N\} \qquad \text{----------(2)}$$

for $x = 0,1,2,\ldots,N-1$. In both equations **(1)** and **(2)**〈*(u)* is defined as

 

$$\alpha(u) = \begin{cases} \sqrt{\frac{1}{N}} for\ u = 0 \\ \sqrt{\frac{2}{N}} for\ u \neq 0 \end{cases} \quad \text{------------ (3)}$$

It is clear from (1) that u=0,

$$c(u=0) = \sqrt{\frac{1}{N}} \sum_{x=0}^{N-1} f(x)$$

### *2.3. THRESHOLDING*

After the coefficients are established from dissimilar transforms, thresholding is done. Very few DCT coefficients signify 99% of signal energy; hence Thresholding is designed and applied to the coefficients. Coefficients having ideals less than threshold values are detached.

### *2.4. QUANTIZATION*

It is a method of mapping a set of nonstop valued data to a set of discrete valued data. The plan of quantization is to reduce the information establish in threshold coefficients. This process makes confident that it produces minimum errors. We mainly execute uniform quantization process.

### *2. 5. ENCODING*

We use dissimilar encoding technique like Run Length Encoding and Huffmann Encoding. Encoding technique is used to eliminate data that are frequently occurring. In encoding we can also decrease the number of coefficients by removing the unneeded data. Encoding can use any of the two compression techniques, lossless or lossy. This helps in dropping the bandwidth of the signal hence compression can be achieved. The compressed speech signal can be reconstructed to form the original signal by DECODING followed by DEQUANTIZATION and then the stage the INVERSE-TRANSFORM methods. This would repeat the original signal.

## III. PROPOSED SYSTEM

### 3.1 DCT BASED COMPRESSION TECHNIQUE

The given sound file is read. The vector is divided into smaller frames and set into matrix form. DCT operation is performed on the matrix. DCT operation is performed and the elements are sorted in their matrix form to find components and their indices. The elements are arranged in descending order. After the arrangement has been done, a Threshold value is decided. The coefficients below the threshold values are discarded.

 

Fig.shows the DCT compression and decompression process.

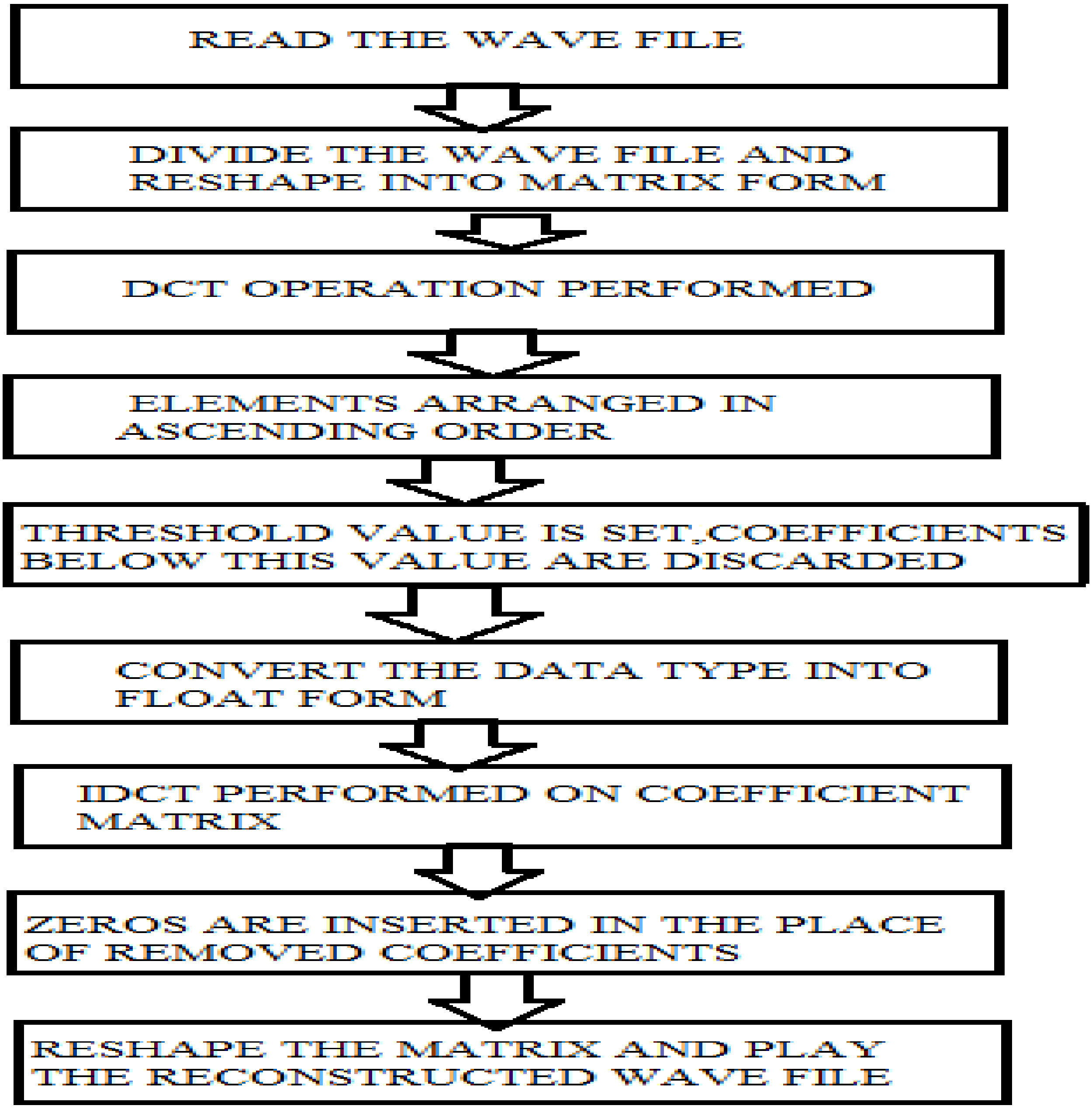


***Fig.2: DCT based speech coder design flow***

## IV. SPARSITY IN SPEECH

The successful models of generating speech and audio signals have been (i) linear system model for speech and (ii) sinusoidal model(AM-FM) for both speech and music. [8] Both are parametric models and parsimoniously represent the time-varying nature of these signals. (For high (transparent) quality reconstruction, such as for music, production models along with a residual signal model is used.) Because of the time-varying nature, we need to do sensing and compressing of a short duration of the signal. It is known that the perceptually significant features of spectral resonances (formants) and the harmonicity due to periodic excitation are the most signal is known to be sparse in a vector space.

## V. SPARSE SIGNAL RECONSTRUCTION

It has been shown that redundant dictionaries are useful for the success of MP; hence, in (4) need not be N×N but N×K, K > N [6], inducing a higher dimensionality for the sparse vector **s**. Thus, $\mathbf{\Phi} \cdot \mathbf{\Psi}$ would be a M × K matrix withM <N <K. For speech, a sparse model based on DCT or DFT is directly usable in eqs.(3) or (4) with the substitution of $\mathbf{\Psi} = \mathbf{C}{-}\mathbf{1}$ or $\mathbf{\Psi}$.

## VI. COMPRESSIVE SAMPLING

CS was first proposed in the literature of Information Theory and Approximation Theory in an conceptual general setting. One measures a small number of random linear combinations of the signal values much smaller than the number of signal samples technically defining it. The signal is reconstructed with good accuracy from these measurements by a nonlinear method. I look at a particular case of CS, where the sample linear combination are basically individual Fourier coefficients (k-space samples). In that situation, CS is claimed to be able to make accurate reconstructions from a small division of k-space, relatively than an entire k-space network. The CS approach requires that:

(a) The preferred signal have a sparse version in a known transform domain (i.e., is compressible),

(b) The aliasing artifacts due to k-space under sampling be incoherent (noise like) in that transform domain.

(c) A non-linear reconstruction is used to enforce both sparsity of the image representation and stability with the acquired data.

To help, keep in mind these ingredients this depicts relationships among some of these main concepts. It shows the image, k-space and the transform domains, the operators connecting these domains and the requirements for CS. Equispaced k-space under sampling and reconstruction by zero-filling results in logical aliasing, a superposition of shifted replica of the signal. In this case, there is an inherent uncertainty; it is not possible to distinguish between the original signal and its replica, as they are all to be uniformly liable.

Random under sampling results in a very different position. The zero-filling Fourier reconstruction exhibit incoherent artifact that really perform much like additive random noise. even though appearance, the artifacts are not noise; rather, under sampling cause leakage of energy away from each individual nonzero coefficient of the original signal. This energy appears in other reconstructed signal coefficients, including those which had been zero in the original signal. It is possible, if all the fundamental original signal coefficients are known, to calculate this leakage systematically. Subtract the interference of the strong mechanism reduces the total interference level and enables improvement of weaker, previously flooded components. By iteratively repeating this procedure, one can pick up the rest of the signal components. A recovery procedure along these lines was planned by Donoho et. al (Sparse Solution of Underdetermined Linear Equations by Stage wise Orthogonal Matching Pursuit, 2006, Stanford University, Statistics Department, technical report #2006-02) as a fast approximate algorithm for CS reconstruction.

The essential observation is that one can design capable sensing or sampling protocols that capture the useful information satisfied embedded in a sparse signal and condense it into a small amount of data. These protocols are nonadaptive and simply need correlating the signal with a small number of fixed waveforms that are incoherent with the basis offer a short description of the signal. (Otherwise there is no dependence of the measurement process on the signal itself.) Further, there is a way to use numerical optimization to reconstruct the full-length signal from the small amount of collected data. In other words, CS is a very simple and capable signal acquisition protocol which samples in a signal independent fashion at a low rate and later uses computational power for reconstruction from what appears to be an imperfect set of measurements.

The basic CS theory which emerged in the works [1] [2] and [3], present the key mathematical ideas underlying this theory, and survey a couple of important results in the field. Our goal here is to explain this theory as obviously as possible, and so our paper is mainly of a tutorial nature. One of the ornaments of this theory is that it draws from various sub disciplines within the applied mathematical sciences, most especially probability theory. In this review, I have decided to highlight this aspect and especially the fact that randomness can be used as a very effective sensing method. Finally, I will discuss significant implication and explain why CS is a concrete protocol for sensing and compressing data simultaneously (this is the origin of the name), and conclude my tour by reviewing important applications.

## VI. CONCLUSIONS

Input data is packed into few coefficients in DCT speech signal representation. This helps quantizer to remove coefficients with smaller amplitudes without generating audio distortion in reconstructed signal. Compressive sampling can be used for compression of audio signal and we can achieve good results by preprocessing the audio signal. This technique can achieves a significant reduction in number of samples required to represent certain audio Signal and it reduces required number of bytes for encoding. As we will compression becomes part of the representation or *coding* scheme which have become popular audio, image and video formats. We will first study basic compression algorithms and then go on to study some actual coding formats

## REFERENCES

[1] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," IEEE Trans. Inform. Theory, vol. 52, pp.489–509, Feb. 2006.
[2] D. L. Donoho, "Compressed sensing," IEEE Trans. Inform. Theory, vol. 52, pp. 1289–1306, July 2006.
[3] M. Lustig, D.L. Donoho, J. Santos, and J. Pauly, "Compressed sensing MRI," IEEE Signal Processing Magazine, March 2008.
[4] E. Candes and J. Romberg, "Robust signal recovery from incomplete observations," in Proc. ICIP, 2006, pp. 1281–1284.
[5] E. Candes and J. Romberg, "Sparsity and incoherence in compressive sampling," Inverse Problems, vol. 23, pp. 969–985, June 2007.
[6] E.Candes and J. Romberg, "Practical signal recovery from random projections," 2005, preprint. [Online]. Available :www.acm.caltech.edu/emmanuel/papers/PracticalRecovery.pdf
[7] Sayed Ali Khayam, "The Discrete Cosine Transform (DCT): Theory and Application" Department of Electrical & Computer Engineering Michigan State University, 2003
[8] l1-magic. [Online]. Available: http://www.l1-magic.org Oct 2010
[9] John G Proakis and Dimitris G Manolakis, "Digital Signal Processing,",ed., Pearson, 2007
[10] R. C. Gonzalez and P. Wintz, "Digital Image Processing," ed., Prentice Hall., 2008.
[11] T. V. Sreenivas and W. B. Kleijn, "Compressive sensingfor sparsely excited speech signals," in Proc. IEEEInt. Conf. Acoust., Speech, Signal Processing, 2009, pp.4125–4128.

[12] D. Giacobello, M. G. Christensen, J. Dahl, S. H. Jensen,and M. Moonen, "Sparse linear predictors for speechprocessing," in Proc. Interspeech, 2008.

[13] R.G. Moreno-Alvarado, Maurico Martinez-Garcia, "DCT-compressive Sampling of Frequency-sparse Audio Signals," IEEE Trans. Inform. Theory, vol. II WCE 2011, July 6-8, 2011, London, U.K.

[14] Moshe Mishali and Yonina C. Eldar," Xampling: Compressed Sensing of Analog Signals", Xiv:1103.2960v1 [cs-IT],15 March 2011.

[15] Amir M. Abdulghani, Alexander J. Casson and Esther Rodriguez-Villegas, "Quantifying the performance of compressive sensing on scalp EEG signals".

[16] Anthony Griffin and Panagiotis Tsakalides, "COMPRESSED SENSING OF AUDIO SIGNALS USING MULTIPLE SENSORS", Marie Curie TOK-DEV "ASPIRE" grant within the 6th European Community Framework Program.

[17] Mads Grasboll Christensen, Jan Ostergaardz, and Soren Holdt Jensen," ON COMPRESSED SENSING AND ITS APPLICATION TO SPEECH AND AUDIO SIGNALS".

[18] T.V. Sreenivas and W. Bastiaan Kleijn," COMPRESSIVE SENSING FOR SPARSELY EXCITED SPEECH SIGNALS', ICASSP 2009 978-1-4244-2354-5/09.

[19] Starck J, Elad M, Donoho D. Image decomposition via the combination of sparse representations and a variational approach. IEEE Trans. On Image Processing 2005; 14:1570–1582.

[20] Taubman DS, Marcellin MW, "JPEG 2000: Image Compression Fundamentals, Standards and Practice.". Kluwer International Series in Engineering and Computer Science., 2002.

[21] Lustig M, Santos JM, Donoho DL, Pauly JM.,"SPARSE: High frame rate dynamic MRI exploiting spatio-temporal sparsity. In: Proceedings of the 13th Annual Meeting of ISMRM", Seattle, 2006. p. 2420.

[22] L. Yaroslavsky and Ye Wang," DFT, DCT, MDCT, DST AND SIGNAL FOURIER SPECTRUM ANALYSIS".

[23] Massimo Fornasier and Holger Rauhut," Compressive Sensing", On Image Processing 2005; 14:1570–1582.

[24] M. Unser. "Sampling 50 Years after Shannon", Proceedings of the IEEE, 88(4):569–587, 2000.

[25] Haykin, Simon (1991), "Advances in Spectrum Analysis and Array Processing". Prentice-Hall. ISBN 0-13-007444-6.

[26] E. Candes and Justi Romera,"l-magic, recovery of sparse signal via convex programming", oct. 2005

[27] http://users.ece.gatech.edu/~sasif/homotopy.

[28] K. Brandenburg and H. Popp, An introduction to mpeg layer-3, tech. report, Fraunhofer Institut fr Integrierte Schaltungen (IIS), Erlangen, Germany.

[29] D. S. Taubman and M. W. Marcellin, "JPEG 2000: Image compression

.

 