RESEARCH ARTICLE OPEN ACCESS

Compressive Sensing in Speech from LPC using Gradient Projection for Sparse Reconstruction

Viral Modha[1], Ravi Butani[2], Ankur Undhad[3]
PG Student[1], Asst. Professor[2, 3]
MEFGI, Rajkot, Gujarat, India

Abstract--
This paper presents compressive sensing technique used for speech reconstruction using linear predictive coding because the speech is more sparse in LPC. DCT of a speech is taken and the DCT points of sparse speech are thrown away arbitrarily. This is achieved by making some point in DCT domain to be zero by multiplying with mask functions. From the incomplete points in DCT domain, the original speech is reconstructed using compressive sensing and the tool used is Gradient Projection for Sparse Reconstruction. The performance of the result is compared with direct IDCT subjectively. The experiment is done and it is observed that the performance is better for compressive sensing than the DCT.

Keywords- Compressive sensing, LPC, GPSR.

I. INTRODUCTION

Compressive Sensing is a signal processing technique for reconstructing a signal, [1] by finding solution to undetermined linear equations. Compressive sensing is a new type of sampling theory, which predicts that sparse signals and speech can be reconstructed from what was previously believed to be incomplete information**.**

As our modern world technology-driven civilization acquires and exploits ever-increasing amounts of data, everyone, now knows that most of the data we acquire can be thrown away with almost no perceptual loss. Instead of sensing all information and throwing away later compressive sensing is a method to sense only the required information. Compressive Sensing sample the signal with lower rate than sampling frequency fs. This Compressive measurement can be small and still contain all the useful information. Compressive sensing is used in mobile phone camera sensor, MRI scanning sessions, holography, astronomy etc.

II. SPEECH COMPRESSION USING COMPRESSIVE SENSING

The proposed speech compression algorithm using compressive sensing is illustrated in Fig. 1. Of the various speech processing front ends, gamma tone filter bank will be utilized due to its resemblance to the shape of human auditory filters [6]. Of the various transforms available to sparsify speech signal, DCT will be chosen due to its simplicity and its good decorrelation property [6].

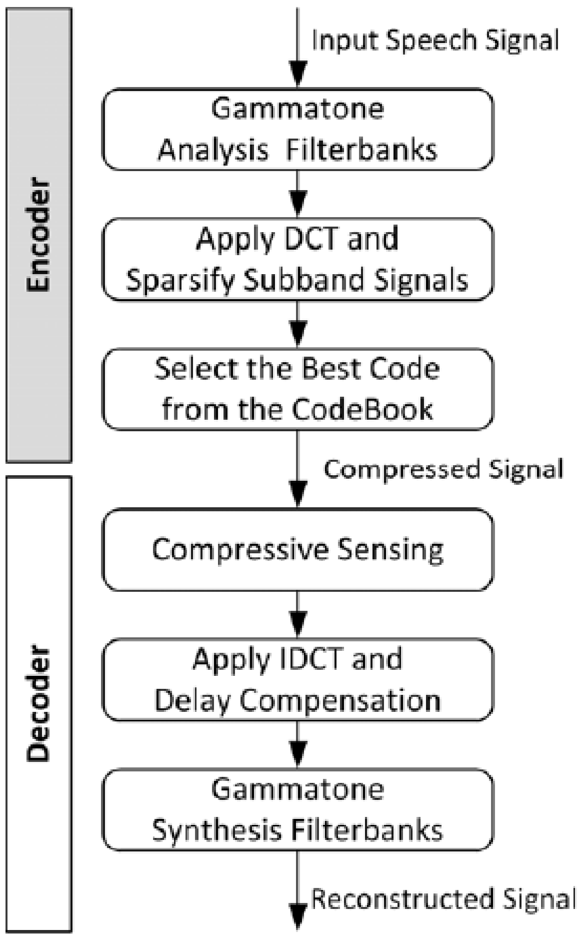

Fig. 1 Proposed Speech Compression Algorithm using Compressive Sensing [6]

On the decoder part, for solving convex optimization of compressive sensing, the gradient projection for sparse reconstruction (GPSR) algorithm [6] was utilized due to its high accuracy and low complexity.

As mentioned in [6], the reconstruction quality can be traded with the available processing power at the decoder side. The higher the processing power or the longer available to solve convex optimization problem, the higher the reconstructed signal quality. Furthermore, IDCT and delay compensation were applied to the compressed signal. The amount of filter delay accumulated by each sub band is different and without compensating for this delay, the reconstruction of sub band signal will lead to an incoherent output signal, i.e. lower quality signal.

III. THE LINEAR PREDICTIVE CODING MODEL

In the linear predictive coding method, a system consisting of only poles and no zeroes is used to model the human vocal tract.[5] All the vocal tract parameters are represented in a set of LPC coefficients, which are calculated automatically from the natural speech signals. The number of coefficients is typically 10 to 20 i.e. 2 for each formants in the 4-5 kHz bandwidth of the speech signal, plus a few for modelling spectral zeros and the source of vocal tract excitation. Speech synthesis model based on LPC is as shown in Fig. 1

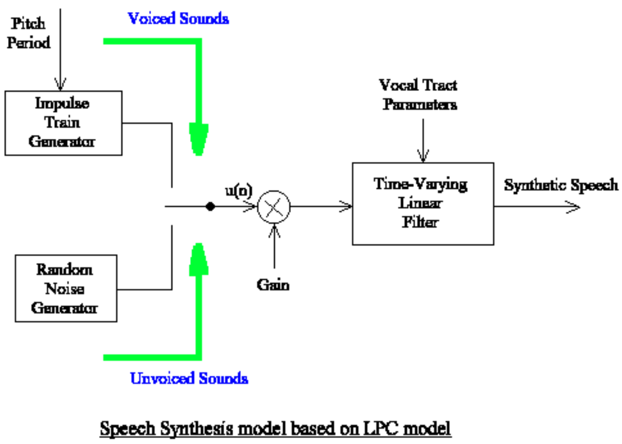

Fig. 2 Speech Synthesis model based on LPC [5]

The LPC synthesis filter is excited by either a periodic (typically impulsive) source or a noise source, depending on whether the analysed speech is estimated to be voiced or not [5]. The excitation is specified by a gain factor, a voiced/unvoiced (periodicity) bit, and (if voiced) a fundamental frequency value. [5]

A. Proposed system

The block diagram of the system is shown in Fig. 2. This is essentially a sequence of steps, each of which can be considered a module as follows:

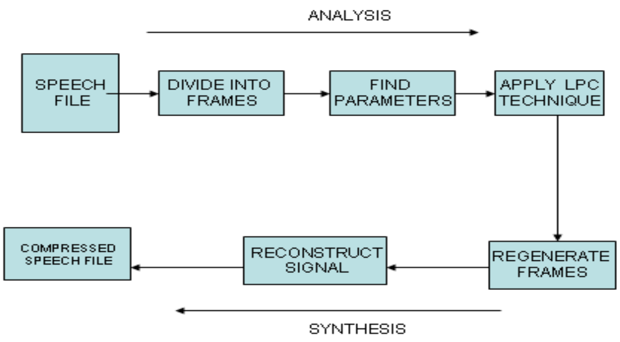

Fig. 3 Proposed System [5]

IV. GPSR (GRADIENT PROJECTION OF SPARSE RECONSTUCTION) [3]

Many problems in signal processing and statistical inference involve finding sparse solution to under-determined or ill conditioned linear system of equations. GPSR algorithm is one of the best techniques to reconstruct the sparse signal. A standard approach consists in minimizing an objective function which includes a quadratic (squared) error term combined with a sparseness-inducing regularization term.

There has been considerable interest in solving the convex unconstrained optimization problem

$$\min_x \frac{1}{2}\,\|y - Ax\|_2^2 + \tau\|x\|_1 \qquad (1)$$

Where$x \in \mathbb{R}^n$,$y \in \mathbb{R}^k$, A is an $k \times n$ matrix, τ is a nonnegative parameter. [3]GPSR is able to solve a sequence of problems efficiently for a sequence of values of τ. Once a solution has been obtained for a particular τ, it can be used as a “warm-start” for a nearby value. Solutions can therefore be computed for a range of τ values for a small multiple of the cost of solving for a single τ value from a “cold start.”

This assumes x is sparse in time domain. If x is sparse in a transform domain $x = Bs$ whereas B is also a N x N basis matrix and s is sparse then the objective function becomes

$$\min_x \frac{1}{2}\|y - ABs\|_2^2 + \tau\|s\|_1 \qquad (2)$$

Here A is called sensing matrix and B is called sparsifying matrix and it is proved that A and B to be mutually non coherent for better reconstruction of x.

In this section an alternative interpretation of above transformation matrix A and B are considered. These matrices constitute the kernel functions (e.g. twiddle factor in DFT) of the transform selected and for our convenience of implementation let us assume them as an operator A and B that process on x to produce the transformed output. Similarly the inverse operators A^{-1} and B^{-1} are also available. Large scale implementation of CS algorithm requires implementation of these operators which are used to iteratively solve the optimization problem in (2). This helps the implementation to avoid representation of complex and large A and B matrices and also helps to use the existing fast blocks to determine the transforms.

V. CONCLUSION

An analytical study, adopting compressive sensing technique and direct Inverse discrete cosine transform methods to reconstruct a sparse speech in case of a defined mask function, results identical. However, compressive sensing technique adopted for sparse speech reconstruction proves to be of higher prudence to random mask functions on. Comparisons to direct inverse discrete cosine transform method.

REFERENCES

[1]. T.V. Sreenivas1 and W. Bastiaan Kleijn, "Compressive Sensing For Sparsely Excited Speech Signals," *IEEE Transl*, 2009, 4125-4128.

[2]. Candès.E.J., & Wakin.M.B,"An introduction to compressive sampling."*IEEE Signal Proces Mag,* 2009.

[3]. M´ario A. T. Figueiredo, Robert D. Nowak, Stephen J. Wright, "Gradient Projection for Sparse Reconstruction: Application to Compressed Sensing and Other Inverse Problems", *IEEE Communication Magazine*, 2007.

[4]. Christenseny.M.G, Østergaardz.J & Jensenz.S.H,"On Compressed Sensing and Its Application To Speech And Audio Signals," *IEEE Conference Publications in Statistical Signal Processing,* 2009.

[5]. G Sawant, A Singh, K Kadam, R Mazarello, P Dumane," Speech Synthesis using LPC", *International Conference and Workshop on Emerging Trends in Technology*,2010.

[6]. T.S. Gunawan, O.O. Khalifa, A.A. Shafie, "Speech Compression using Compressive Sensing on a Multicore System", *4th International Conference on Mechatronics (ICOM)*, 17-19 May 2011.

[7]. Jort Florent Gemmeke et al, "Compressive Sensing for Missing Data Imputation In Noise Robust Speech Recognition",*IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING*, VOL. 4, NO. 2, APRIL 2010, 272- 287.

[8] Dalei Wu et al,"A Compressive Sensing Method for Noise Reduction of Speech and Audio Signals", *Department of Electrical and Computer Engineering, Concordia University,* 2011.

[9] Guangzhao Bao et al., "A Compressed Sensing Approach to Blind Separation of Speech Mixture Based on a Two-Layer Sparsity Model", *IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING*, VOL. 21, NO. 5, MAY 2013,899-906.

[10] Miss Urvashi P Shukla et al., "A survey on recent advances in speech compressive Sensing", *IEEE, 2013, 276*-280.

[11] Tao Xu and Wenwu Wang,"A COMPRESSED SENSING APPROACH FOR UNDERDETERMINED BLIND AUDIO SOURCE SEPARATION WITH SPARSE REPRESENTATION", *IEEE/SP 15th Workshop on Statistical Signal Processing*, 2009, 493-496.

[12] Siddhi Desai, "Compressive Sensing in Speech Processing: A Survey Based on Sparsity and Sensing Matrix", *International Journal of Emerging Technology And Advanced Engineering,* Volume 3, Issue 12, December 2013, 18-23.

[13] Tzu-Chuen Lu, "A Survey of VQ Codebook Generation", *Journal of Information Hiding and Multimedia Signal Processing,* Volume 1, Number 3, July 2010, 190-203.