RECOVERY AND IDENTIFICATION OF AUDIO SIGNALS WITH COMPRESSED SENSING

Syed Yasir Abbas[1,*], Farhat Iqbal[1]

[1]Electrical Engineering Department, University College of Engineering and Technology,University of Sargodha

[*] Corresponding Author, E-mail: yasirabbas65@yahoo.com, [1]farhat_ieee@yahoo.com

***ABSTRACT**: In modern era audio compression has become one of the essential multimedia technologies. Selection of an competent compression scheme for that kind of signals that is capable to preserve the signal quality while providing a high compression ratio that is desirable in the different principles globally and also proved scientifically since the optimization techniques in sampling. Compressed sensing generally describes a system in which the sampling and compression of an audio signals can be done concurrently. This means that, when dealing with analog signals such as audio,image, etc. the signal must first pass through an ADC(analog to digital convertor) and then be sampled at a uniform rate size foe efficiently acquiring reconstructed signals. In this paper, we describe a mathematical compressive date algorithm based on the l_1-minimizationc(least square regression) method to decide audio signals into various iterations with uniform rate size and then comparing these audio signals with the input audio signals already recorded in the GUI or database which must be user dependent and data size can be changed manually. Since the quality of audio signals is also resolving factor to get a better audio compression ratio at the output so adaptive filters makes a good impact in this area because these are important to get mean square calculations at every minimization step also to accomplish better and to preserve signal quality. This scheme helps the user to record the various audio signals and then helps one to identify their voices based on their pitch detection and estimation. One possible extension is to incorporate the proposed approach in various speech applications such as voice activity detection and signal resolving calculations.The sound signals recovered if not according to data rate size decided earlier the it must be passed through a adaptive filter for obtaining better signal quality.*

Keywords: Audio signals, Compressed sensing, l_1-minimization, Pitch detection

1. INTRODUCTION

Compressed sensing for audio signals has become a revolution in research discoveries which are used in the field of digital signal processing as well as image processing. Use of these audio signals in the field of signal processing the nature made a change/shift from the basic sampling, quantization and coding schemes to the most effective compression based on Linear Prediction Coding (LPC) and Code Excited Linear prediction coding (CELP) and other model based systems for speech signal which depends on the vocal territory model **[1]** which are the basic conditions for converting analog signals into digital signals**.** Recently, sparseness of audio signal has become an aim of achieving with uniform sample rate size than the current compression techniques used in the multimedia coding standards **[2].** Sparse signal exemption is an active research area that has recently witnessed most important research trends especially in compressive sensing(sampling) depending upon the uniform data rate size. Compressed sensing made an impact on the world by the work of Cand`es and Donoho **[3,4]** which gave the universal framework for random sampling and achieving sparse signals using linear measurements at a sampling rate much smaller than the Nyquist rate (comparable with the fundamental frequency) which is the frequency distribution phenomanon. The removal of incoherent noise in sampling intervals of audio signals is the main objective behind that work **[5]** and the use of distribution matrix which is highly sparse gives the random samples, the original audio signal can be recovered at the user side.. The main objective behind the work is that random sampling will introduce disjoint noise signals which must be removed to preserve better signal quality. In this paper we construct the database of the audio signals for different systems based on their pitch estimation to calculate linear minimization steps at each frame of the speech signals and recognize that signals after applying least square regression (robust and noise free) techniques to identify sound signals**[6].** Also, this paper gives us a great view to generalize audio signals either they are instrument's vocal quality or a general human voice must be recognized at the decoder side by analizing identical mathematical techniques which must be analyzed digitally.

The original signal must be recovered at the decoder side.

.

2. COMPRESSIVE SENSING PROCESS

Initially we are having a vector matrix $\mathbf{x} \in \mathbf{R}^N$ of real numbers, which can have another matrix called transform matrix $\mathbf{\Psi}$ (sometimes called as Wavelet Transform (WT)). There exists another input signal/matrix having the representation of $\mathbf{x} = \mathbf{\Psi}\, \mathbf{f}$. where $\mathbf{\Psi}$ is *MX1* matrix(must be a column matrix with each row has the one eigen value) whose columns are orthogonal/orthonormal based functions and the matrix diagonals are the combinations of eigon values, and **f** is the coefficient vector to be transformed after reconstruction and data rate size must be same as decided earlier. **x** is called K-sparse when there exists K non zero co-efficients in $\mathbf{\Psi}$ domain[11]. According to the Compressed Sensing theory signal **x** obtained through a number M (called as sparseness factor) which can be obtained by equation $y_{CS} = \phi x$ where **y** is an MX1- two dimensional linear measurement vector in the linear domain. This technique is a non-linear sampling theorem (A=bx) which must be realistic and used for most beneficial signal recovery using uniform uncertainty principle. Number of non-linear solutions always has a sparse level actually no. of zeros that will lead to give unclear iterations. The uncertainty principle will be represented by $S = K\log(\frac{N}{K})$. For recovery and reconstruction of (mx1) matrix M, the rule of thumb is to be applied which says that for incoherence the sampling measurements should be about 4 times to the level of sparsity[12]. The size of framing shold be taken 1 initially with respect to iterations. This fact is known as four-to-one practical rule (mostly applied to image processing in real world for fixing the size of pixels)**[11].** This analogue will help the users to reconstruct the M dimensional real matrix with fixed data size. The main feature for the user is that reconstructed sound signals are obtained with uniform and estimated pitch size by the user..

3. RECONSTRUCTION

The reconstruction method of audio signals for the process of compressed sensing is possible only by using two methods of least square regression(L1) and least absolute deviation regression(L2) which is the main feature of Compressive sensing and must be operated with uniform data rate size. L1 method is robust and the other one is non robust but both features are user dependent and must be operated manually. Since the technique of least square regression or L1 minimization is robust and uniform(gives quick frames with least deviation from results)[9]. Since all norms are equivalent, i.e., there exists a pair of real numbers $0 < C1 \leq C2$(in frame size) such that, for all of these $x \in V$, the inequality in real frames or real numbers for matrix M holds when A=b.

$$C_1\|x\|_b \leq \|x\|_a \leq C_2\|x\|_b.$$

That means, we can replace L1-Norm with a constant factor times[10],[12] before selecting data size.

Consider the optimization problem

$$\min \|Ax\|^1 + \lambda\|x - b\|^2$$

where A∈R(in domain and range), x&b∈R , and λ is positively greater than 0. (This problem is closely resembled to the "lasso" problem in basis algorithm detection.) Can anything be said about the value of λ (the eigen value in matrix M) for which Ax=b is sparsest or not? Clearly some values are bad in real time domain: for example, if λ is larger and b is not clear then it is doubtful that Ax=b will be very sparse in real domain with uniform data rate size and the eigen values may not occur in real time domain. In other words: among all λ>0 there is at least one value λ such that $\|Ax\star(\lambda)\|0$ is minimized. Are there, say, limitations on λ in terms of A and b. By choosing fixed size of sparseness factor K the recovery of signal can be possible at the decoder side with suitable values of x.

4. AUDIO COMPRESSION METHODOLOGY

Methodology for compressed sensing of audio signals for both the sparsed sensing based algorithm and the estimated pitch detection algorithm (frequency domain in real time) analysis is presented in this section. The compressed sensing process algorithm is shown in Fig 1 which shows that the overall process must also be divided into two sequences, the first one consisting of algorithm till compression and the second one continues until the reconstruction. Each frame is distributed into the frequency domain using one of three uniform trusted transform domains, namely, wavelet transform, least square regression technique (L-1 minimization) and compressive data size which must the methodology for selecting appropriate data size **[11].** The data size gives the values which must be compared from Gaussian distribution matrix. At the decoder side, original signals can be recovered (must consider the size of sparseness factor in database also) from the sparse coefficients**[13].** The rate of reconstruction must be uniform at the decode side to obtain better signal quality.

It is significance noting that both sparse frames and sparse signals which must be improved at the decoder side follows the steps as under.

1- Reconstruction of audio signals and its frames and construction of a database to initialize audio signals data for the frame segmentation.

2- Applying the algorithm to remove K-sparse coefficients of the recovered signals in the frequency domain or wavelet//freuency transform domain(re-occurrence).

3- At the receiver the recovered signals obtained with fixed data size follows the rule of Gaussian distribution matrix which helps the compressed frames to be recovered with rate of iterations. All the rest of the values after framing is fixed is set to zero

4- Frame by frame investigation of audio signals for comparing the data from database to verify the recorded signals and original signals so that original signal can be generated at receiver side with proper pitch estimation techniques.

5- Comparison of the recovered audio signal with the input database signals with uniform pitch detection technique.

The selection of these results after pitch estimations reports with detection algorithms helps the users to record various voices based on vocal territory model. The sound signals recovered if not according to data rate size decided earlier the it must be passed through adaptive filter for obtaining better signal quality

5. ANALYSIS OF SIMULATIONS AND RESULTS

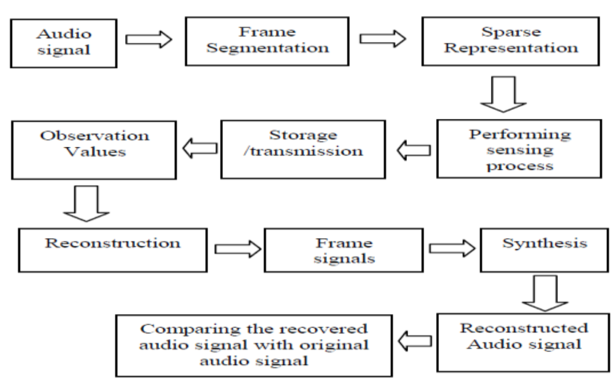


Fig 1: Algorithm for Compressed sensing based audio signals

In this section we present the use of compressed sensing transformation (CST) in order to obtain sparse representation of human voice or audio signal in the frequency domain or in real time interval. Also the use of filtering algorithms for compression of audio signal appropriately considering framing with suitable size.

The real time tested audio signals are formerly stored in a database system and each signal is a twenty (20) second music piece or less or a human stored or generated voice and sampling is done at rate of 44100 sample/s, the recovered samples of vocal territory model used in frequency domain are divided into samples and each sample contained 1024(1K) samples. By applying least square regression or robust techniques we obtain a sparse depiction of signal frames. Fig.3(a,b,c) shows the frequency domain database system having the original signals with fixed frame sizes.

The recovered signal must be obtained at the decoder side with uniform data rate apporiate for about 44100 samples per second recorded most significantly. It must have pitch detection technique at uniform data rate size. It must be obtained at the decoder side

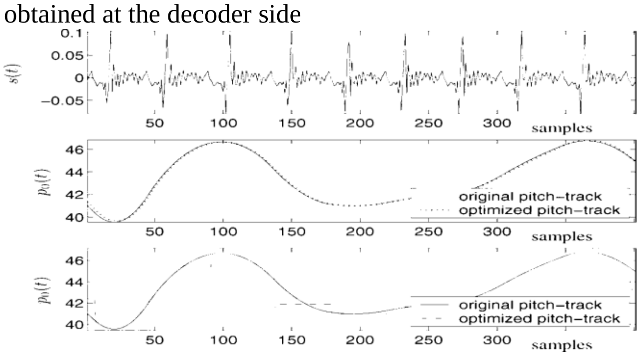


Fig2:Comparison of audio signals

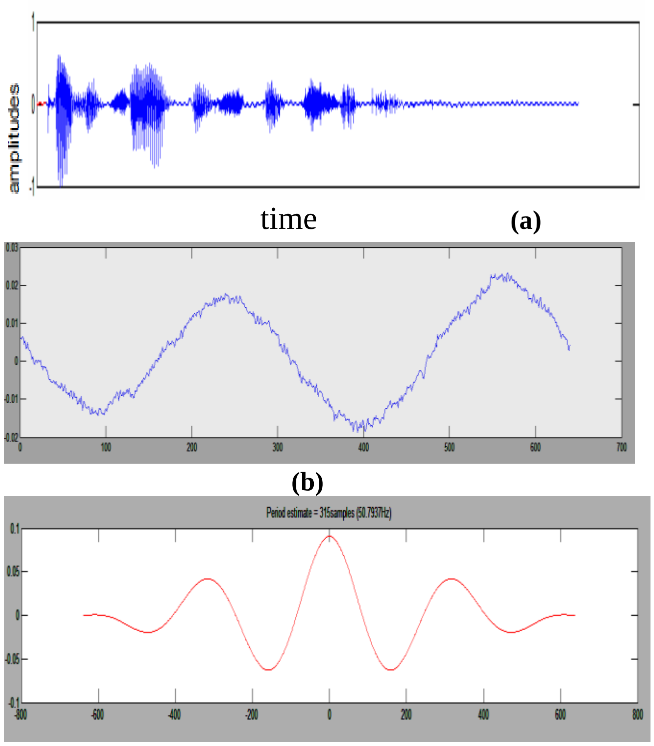


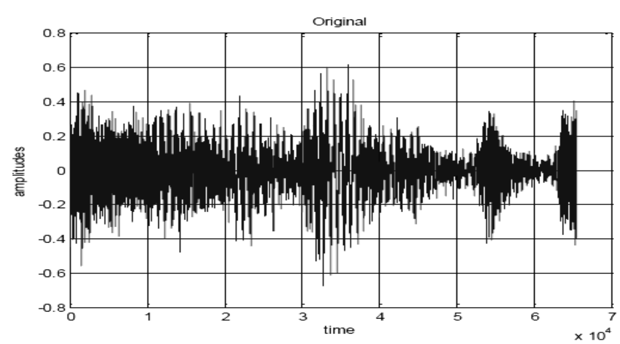
Fig 3.Audio signals(a)Original signal in database(b)One stepframe analysis (c) Estimated pitch by LMS filter

Audio recovery that must be analyzed both scientifically and also with software techniques, the comparison from Fig2,4 must revealed by the fact that anything in close similarity with the original pitch of sound signals either a voice signal or a sound of a instrument the identical pitch selection always be rate digitized. The graphical results obtained from GUI or database after applying least square minimization techniques also gives close comparison original signal. The signal path from Fig2 in comparison with original pitch in close function gives the same samples or frames between 40 and 46. The recovered signals must be obtained at the decoder side with uniformed data size.

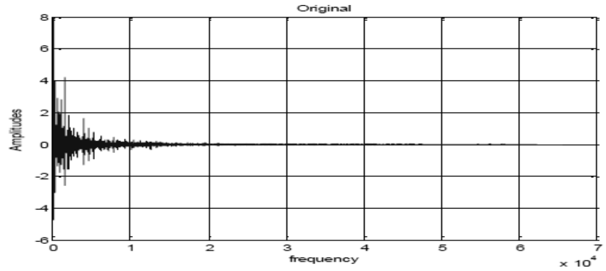
The above results pointed towards the appearance gain achieved the restoration of signals with frame by frame analysis**[11.** The above results can also be implemented for other types of signals such as image and speech. By taking recovered techniques for sparse signals, the compressed sensing provides a precious option for sparse signal recognition, especially if low complication is required at the acquirement**[12].**

Since the equivalent pitch of every sound from the database is easily locatable in the database system and Audio signal is also locatable in the final step when compressive sensing is applies to the audio signals for recovery and acknowledgment of audio signals**[9].** The filtering technique is based on the least mean square minimization (LMS) for filtering design which shows the estimated or original pitch result at each outline of the sound signal dropping the error upto least minimum value of roughly 50% and also reduces likely SNR ratio that gives the close resemblance for identifying a sound from the database which same pitch quality[15]. The signal is first sampled at the decoder side and then sampled at the uniformed rate to obtain better signal quality. The reconstruction rate depends upon the pitch estimation rate to be obtained at the decoder side with proper GUI system design. The sampled signals recovered with proper sampled size are obtained after applying compressed sensing and pitch estimation techniques. The advantage of using filtering techniques based on linear minimization is necessary for original pitch detection of the sampled sound signal which must be comparable to the input signal and input signal must received after filtering with compressed data.

The least square minimization technique are edge sensitive to signals whether it is audio or image signal processing. For preserving signal quality the sampling must be quantized as well as filtering.



(a)



(b)

Fig4: (a)Original audio signal from database
(b) Reconstructed signal

6. CONCLUSION:

In this paper, we present an optimized-based framework or algorithm for compression of audio signal and evaluation of the results with compressed sensing results to assess the exactness of samples for compression audio signals. Simulation based results conclude that filtering at each step of audio or speech signals based structure reforms the compressed sensing based advance in signal influence**[10].** The results obtained after the simulation or filtering designed by an LMS filter help the user to store various voices based on the pitch estimation based approach. The presentation grow of the conventional structure rationalize the require for a new class of sensing matrix for the reduced sensing, taking into account the built-in properties of the sparse signal like the perceptual effect of the signal content and a better estimate of the location of zeros in the sparse signal.

7. ACKNOWLEDGMENT:

This research work has been acknowledged by my respected advisor and Dean of University College of Engineering and Technology Prof. Dr. Ghulam Yasin. They are highly accredited for offering tools and equipments needed.

8. REFERENCES:

1) Salami, R.; Laflamme, C.; Adoul, J-p; Kataoka, A.; Hayashi,S.;Moriya,T.; Lamblin, C.; Massaloux, D.; Proust, S.; Kroon, P.; Shoham, Y., "Design and description of CS-ACELP: a toll quality 8 kb/s speech coder,"Speech and Audio Processing, *IEEE Transactions on* , vol.6, no.2, pp.116,130, Mar 1998
2) Ghido, F.; Tabus, I., "Sparse Modeling for Lossless Audio Compression," Audio, Speech, and Language Processing, *IEEE Transactions on* , vol.21, no.1, pp.14,28, Jan. 2013 doi: 10.1109/TASL.2012.2211014
3) Candes, E.J.; Romberg, J.; Tao, T., "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," Information Theory, *IEEE Transactions on* , vol.52, no.2, pp.489,509, Feb. 2006.
4) D. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006.
5) Michael Lustig, David Donoho, John M. Pauly, "Sparse MRI: The Application of Compressed *International Conference of the IEEE* , vol., no., pp.5734,5737, Aug. 30 2011-Sept. 3,2011
6) Donoho, D., "Compressed sensing," IEEE Trans. Inf. Th., Vol.52, No.4, pp 1289-1306, Apr 2013.
7) Huang Feng-Long, "An Effective approach for Chinese speech recognition on small size of vocabulary", Signal & Image Processing: An International Journal (SIPIJ) Vol.2, No.2, June 2011.
8) Candes, E.J. and Wakin, M.B., "An introduction to compressive sampling," IEEE Sig Proc. Mag., pp 21-30, Mar 2012.
9) Tropp, J.A. and Gilbert, A.C., "Signal recovery from random measurements via orthogonal matching pursuit," IEEE Trans. Inf. Th., Vol.53, No.12, pp 4655-4666, Dec 2014.
10) Rauhut, H., Schnass, K. and Vandergheynst, P., "Compressed sensing and redundant dictionaries," IEEE Trans. Inf. Th., Vol.54, No.5, May 2008.
11) Figueiredo, M.A.T., Nowak, R.D. and Wright, S.J., "Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems," IEEE J. Sel. Top. Signal Proc., 2007.
12) Aggarwal, R.K. and Dave, M., "Acoustic Modelling Problem for Automatic Speech Recognition System: Conventional Methods (Part I)", International Journal of Speech Technology (2011) 14:297–308.
13) Goodwin, M., "Adaptive Signal Models: Theory, Algorithms and Audio applications," Ph.D. Thesis, Univ California, Berkeley, USA, 1997.
14) Ghido, F.; Tabus, I., "Sparse Modeling for Lossless Audio Compression, "Audio, Speech, and Language Processing, *IEEE Transactions on* , vol.21, no.1, pp.14,28, Jan. 2013 doi: 10.1109/TASL.2012.2211014
15) N. Roman and D. Wang, "Pitch-based monaural segregation of reverberant speech," *J. Acoust. Soc. Amer*, vol. 120, no. 1, pp. 458–469, Jul.2006