International Journal of Pure and Applied Mathematics
Volume 119 No. 14 2018, 111-114
ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue
ijpam.eu

DCT Based Compressive Sensing For Audio Signal Acquisition

Electa Alice Jayarani.A[1] , Mahabaleswara Ram Bhatt[2], Geetha.D.D[3]

[1]*ECE, Reva University, Bangalore,India*
[2]*Dept of Medical Electronics, BMS College of Engineering, Bangalore,India.*
[3]*ECE, Reva University, Bangalore,India.*
**Corresponding author and email*
*[*1]electalice@gmail.com*
*[*2]mr.bhatt@rediffmail.com*
*[*3]dgeetha@reva.edu.in*

Abstract

The analog signal is converted into digital using the following steps, sampling, quantization, and encoding. In all analog to digital converter (ADC), to convert the signal to discrete signal Nyquist sampling is used widely. The N samples obtained from the above will get reduced to M samples (M≪N) when we perform compression for digital storage or for digital transmission. In this paper, we propose a non-uniform sampling called as compressive sensing on an audio signal in ADC instead of traditional sampling. The proposed algorithm is based on DCT based basis matrix for compressive sensing. The results proofs that the signal can be successfully recovered from fewer samples so that the function of ADC can be reduced.

Keywords— *Audio Signal, Compressive Sensing, Analog to Digital conversion.*

1. Introduction

In nature all audio signals are in analog in nature i.e. they are continuous in time. To convert to a digital signal, the analog signal undergoes sampling, quantization, and encoding in ADC. The audio signal reaches the human ear in the form of pressure varies continuously in time. Fig.1. shows the general block diagram of an audio signal hearing system.

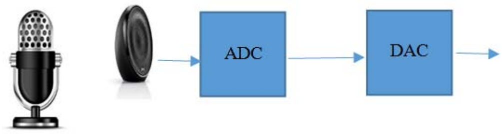

Fig. 1 Audio signal hearing system

In order to reduce the number of samples for digital storage or for digital transmission, we propose compressive sensing on audio rather than Nyquist sampling. Candes Emmanuel J and Romberg Justin [1, 2] proved that the stable signal recovery is possible from an incomplete and inaccurate measurement in 2006. This gives a gain in the compressive sensing technique and various research is going on this new algorithm. The compressive sensing is used in various multimedia files such as image and biomedical images [3, 4]. As the human Auditory System (HAS) is more sensitive than the Human Vision System (HVS), the small changes in the audio can be heard by the human [5]. Because of the challenge in the Human Auditory System (HAS), less research is undergone on audio using compressive sensing. In this article, we propose compressive sensing on an audio signal based on DCT. The original signal can be recovered from the compressed data using $l1$ minimization.

2. Compressive Sensing

Compressive sensing is a new sensing paradigm where sampling the signal with a much lower sampling rate than the Nyquist rate but guarantee the good reconstruction with the less samples. Most of the natural signals are not sparse but when expressed in proper basis coefficient the signal will be sparse. For example, consider the host signal as $a(t), a \in R^N$ is not sparse in nature but when expressed in frequency domain i.e.

$$A = \varphi a$$

Where $\varphi \in R^{NxN}$ is Discrete Fourier Transform (DFT) matrix and $A \in R^N$.

The widest tool used to analyze the frequency content of any signal is Fourier Transform in digital signal processing. The Discrete Fourier transform [6] is given as

$$X(K) = \sum_{n=0}^{N-1} x(n) e^{\frac{-j2\pi Kn}{N}} \cdots\cdots\cdots (1)$$

And inverse DFT is given by

$$x(n) = \frac{1}{N} \sum_{K=0}^{N-1} X(K) e^{\frac{j2\pi Kn}{N}} \cdots\cdots\cdots (2)$$

DFT represents a signal by a set of orthogonal sinusoidal fu nctions and DFT is a unitary transform.

A. Discrete Cosine Transform

Discrete cosine transform DCT represents a signal by a cosine function oscillating at different frequencies. DCT is given by

$$X(K) = \alpha(K) \sum_{n=0}^{N-1} x(n) \cos\left[K\left[\frac{\pi}{M}\left(n+\frac{1}{2}\right)K\right]\right] \cdots\cdots\cdots (3)$$

where $K = 0,1,\cdots,N-1$

$$\alpha(K) = \begin{cases} \sqrt{\frac{1}{N}} \; if\ K = 0 \\ \sqrt{\frac{2}{N}} \; if\ K \neq 0 \end{cases}$$

DCT is similar to DFT but representing only real numbers and it is widely used in JPEG compression.

B. Compressive Sensing on Audio

Considering the transformed signal A is K sparse and the compressive sensing is given by

$$Y = \Phi A \cdots\cdots\cdots\cdots (4)$$

Where $\Phi \in R^{MxN}$ is the measurement matrix. Here for our paper, we have considered the measurement or basis sensing matrix is based on the Discrete Cosine Transform (DCT).

And $Y \in R^{M}$ is the non-uniform sampled data with M samples which is very much less than N ($M \ll N$) samples. The original host audio signal can be reconstructed from the compressed data Y using $l1$ minimization if the following condition is satisfied

$$M \geq CKln\left(\frac{N}{K}\right) \cdots\cdots\cdots (5)$$

where $C > 0$ is a universal constant independent of K, N, M. In equation (5), M is directly proportional to K and hence if the sparsity is considered small then the measurement M can also be chosen small in comparison with N, so that the solution of an underdetermined system of linear equation is reasonable. Different sparse k signal and the corresponding M measurement are listed in Table. I.

Table. I. Different K and M

S.No	K	M
1.	3	≥ 14
2.	5	≥ 20
3.	7	≥ 25
4.	10	≥ 32
5.	13	≥ 39
6.	15	≥ 43
7.	20	≥ 51

The above table is calculated for the length of $N = 256$, It is noted that M can be selected such that $0 \leq \frac{M}{N} < 1$.

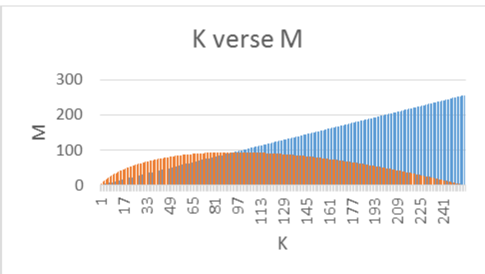

Fig.2. $K\ vs\ M$

Fig 2. Shows the analysis of different $K\ vs\ M$ for a length of $N = 256$. It can be observed, the M values increases and starts to decrease and hence we can choose the M as per the K sparse state. For our experiment, we have tried with different K sparse signal with different measurement matrix.

3. Proposed Algorithm

In order to reduce the size of the samples for digital storage or digital transmission, we propose non-uniform sampling during signal acquisition only rather than reducing the size of samples during compression. The proposed algorithm is as follows:
Consider a host signal $a(t), a \in R^{N}$ is continuous in nature. When the host signal is converted to the frequency domain the signal is sparse.

$$A = \varphi a \cdots\cdots\cdots (6)$$

Let's consider a signal is K sparse and we calculate the measurement parameter M for compressive sensing as

$$M \geq CKln\left(\frac{N}{K}\right)$$

Where $C = 1$ & $N = 256$

Perform the non-uniform sampling on the K sparse signal using

$$Y = \Phi A \cdots\cdots\cdots\cdots (7)$$

Where A $\in R^{MxN}$, $\Phi \in R^{NxN}$ is the DCT coefficient and $Y \in R^{M\,x\,N}$

At the receiver end to convert the digital data into an analog signal to be recognized by humans, the signal has to be reconstructed from the sampled data. We can successfully reconstruct the host signal from less sampled data using $l1$ minimization method using

min ($\|A\|$) subject to $Y = \Phi A$

The perceptual quality of the reconstructed audio is measured using a Signal to Noise ratio (SNR). The SNR is

K		7	10			15			20
M		26	36	78	104	36	78	104	52
SNR (dB)	Guitar	22.6723	28.3507	34.5665	33.9192	27.1575	34.2531	33.5009	25.3612
	Handel	10.2786	19.5650	20.8717	20.9864	20.6705	21.2048	21.5069	18.2151
	Piano	15.7357	23.2751	30.5590	32.0801	23.2499	30.1964	33.2670	30.1706
	440 Hz	27.1674	33.9697	37.1612	36.5344	39.0404	38.1143	37.0946	29.6276
	1K hz	25.0596	33.4631	24.4572	22.3478	32.4533	23.9153	22.5442	33.1304

Table 2. Experimental Results one of the objective measurement parameters of perceptual quality and is given by

$$SNR = 10\ \log_{10} \frac{\sum_{i=1}^{n}(Y)^2}{\sum_{i=1}^{n}(Y - Y')^2}\ dB \cdots\cdots\cdots (8)$$

Where Y is the original host audio signal
And Y' is the recovered audio signal

4. Simulation Results

The proposed algorithm is simulated using MATLAB2017 in Intel

Core i5 processor. For our experiments, we have considered various audio signals such as guitar, piano, bass, Handel and various frequency signals such as 440 Hz and 1 KHz. For better performance, all video is divided into frame size of 256. Various frame size is also tested but as there is no variation in performance, the results of different frame size are not discussed here. All the audio signal is tested with different sparsity K and different measurement M and the corresponding SNR is measured and listed in table 2.

It is observed from the table 2 that with less samples we can successfully reconstruct the original host signal without disturbing the perceptual quality of the audio signal. The comparison of the original host signal and recovered signal of two different audio is shown in fig 3

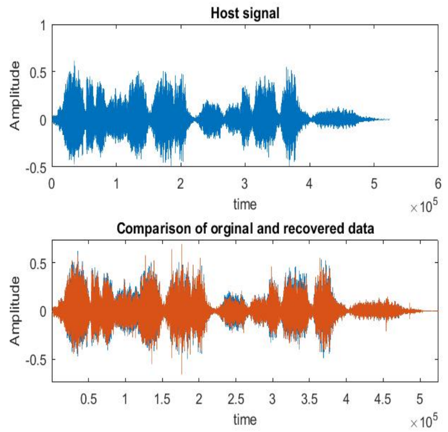

Fig. 3.a. Comparison of original and recovered bass audio signal with $M = 78$ for $N = 256$ therefore $\frac{M}{N} = 0.3$ with SNR of 21.58dB.

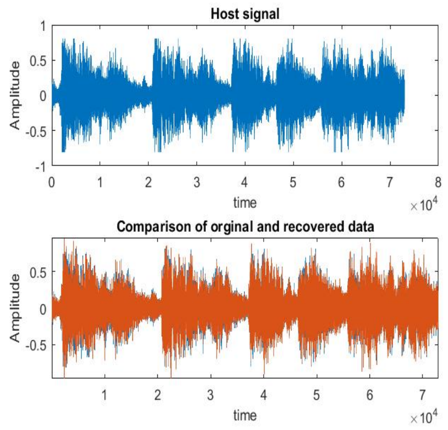

Fig. 3.b. Comparison of original and recovered Handel audio signal with $M = 44$ for $N = 256$ therefore $\frac{M}{N} = 0.17$ with SNR of 20.6705dB.

From the above, it is clearly observed that the host signal can be recovered with less samples and thus it can be used for audio signal acquisition. Thus the proposed algorithm can reduce the size at the time signal acquisition rather than in compression.

5. Conclusion

For a traditional digital storage or digital transmission, the size of the signal is reduced at the time of compression. But in our paper, we propose a non-uniform sampling to reduce the size of the samples during the signal acquisition only. By reducing the samples size the task of ADC in real time will get reduced and the successful recovery from the less samples adds advantage to our proposed method than the traditional method. In future, the signal acquisition based compressive sensing model can be simulated and can be implemented in hardware.

Reference

[1] Candes Emmanuel J, Romberg Justiin K, Tao Terence (2006), “Stable signal recovery from incomplete and inaccurate measurements”, *Communication on pure and Applied mathematics*,59(8):1207 – 1223.doi:10.1002/cpa.20124.

[2] Donoho D.L (2006), “Compresses sensing”, *IEEE Transaction on Information Theory*, 52(4): 1289-1306.doi:10.1109/TIT.2006.871582

[3] Pham Hong Ha, W. Lee and V. Patanavijit, "An introduction to compressive sensing for digital signal reconstruction and its implementation on digital image reconstruction," *2014 International Electrical Engineering Congress (iEECON)*, Chonburi, 2014, pp. 1-4. doi: 10.1109/iEECON.2014.6925959

[4] T. Mantecón, A. Mantecón, C. R. del-Blanco, F. Jaureguizar and N. García, "Hand-gesture-based human-machine interface system using Compressive Sensing," *2015 International Symposium on Consumer Electronics (ISCE)*, Madrid, 2015, pp. 1-2. doi: 10.1109/ISCE.2015.7177828

[5] Y.Lin and W.H.Abdulla, "*Audio Watermark: A Comprehensive Foundation Using MATLAB*,"Springer International Publishing Switzerland 2015.

[6] Steven K Thompson, “*Sampling*”, A John Wiley & Sons Inc Publication, 3rd edition, 446 pages, 2012.

[7] Simon Foucart, Holger Rauhut, “*A Mathematical Introduction to Compressive sensing”*, Springer New York Heidelberg Dordrecht London, 634pages, 2013.