Advances in Social Science, Education and Humanities Research (ASSEHR), volume 130
2nd International Forum on Management, Education and Information Technology Application (IFMEITA 2017)

Application of compressed sensing theory in the sampling and reconstruction of speech signals

Xiwen Tang[1], Shilong Wu[1] , Rui Dong[1] and Guang Xia[2]

[1] School of Radar Confrontation, National University of Defense, Hefei, China

[2] Institute of Automotive Engineering, Hefei University of Technology, Hefei, China

Keywords: Compressed Sensing; Speech Signals; DCT; Orthogonal Matching Pursuit Algorithm

Abstract: This paper studies the application of compressed sensing theory in speech signal sampling and reconstruction of speech signals. According to the sparsity of speech signals in the discrete cosine transform basis (DCT), we propose a speech compressed sensing (CS) system based on DCT domain which realizes sparse representation of speech signal in DCT domain. Utilizing Gauss random matrix as the measurement matrix and orthogonal matching pursuit algorithm (OMP), the performance of speech signal reconstruction is acquired. The simulation results show that the sparsity of the speech signal is higher in the DCT domain and the OMP algorithm can effectively improves the performance of reconstructed speech signals.

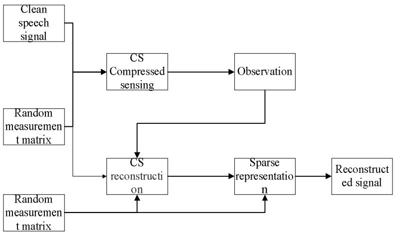
Introduction

Compressed sensing theory restores the original signal by reconstructing the observation sequence. If the signal is sparse in a transform domain, it can reconstruct the original signal approximately with lossless according to the observation set of the signal in a projection domain, that is, the observation can represent the signal in some sense[1][2].Compressed sensing (CS)is used as a sampling method for speech signals. Since the number of observations is less than the dimension of the signal, compression and sampling can be performed simultaneously, which breaks through the bottleneck of Nyquist sampling theorem. The speech compression and reconstruction algorithm based on compressed sensing theory has the following two advantages:1) The compression method of the inchoation is simple which only needs to be multiplied by the observation matrix as well as low computational complexity; 2) For the same compression scheme of the inchoation, the receiver can adopt different improved reconfiguration schemes to improve the reconfiguration performance. At home and abroad, the research of compressed sensing used in the field of speech signal processing is relatively less than that of the image field, and it is still in its infancy.

Speech CS system Based on Discrete Cosine Transform (DCT)

Framework of speech CS system based on DCT domain

Speech signal is a random process with time-varying and non-stationary characteristics. However, due to the speed of human muscle movement is slow, it can be approximated as a stationary process in 10~30ms, that is, the speech signal has short-time stability. According to the CS theory, firstly, the speech signal should be processed by frame. Then the signal is compressed and sampled frame by frame. Finally, the signal is reconstructed by frame by frame. The diagram of speech signal CS system is shown in figure 1. The key issues include random observation matrix selection, sparse signal representation, and reconstruction algorithm selection.



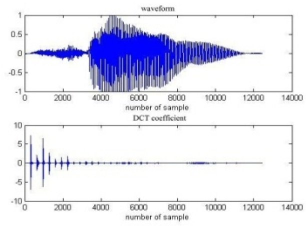


Figure 1 The Diagram of Speech Signal CS System   Figure 2 Speech Signal Waveform and DCT Coefficient

Sparse representation of speech signals in DCT domain

The sparse representation of the signal is that when the signal is projected to a domain, the absolute values of most transform coefficients are very small so as to ensure effectively extraction of the essential characteristics of signals, which simplifies the subsequent signal processing and reduces the cost of processing signals at the same time. The sparse bases mainly adopt the orthogonal basis. DCT bases, wavelet bases and KLT（Karhunen-Loeve Transform） bases are often used. Here we employ the DCT transform for the speech signal sparse representation.

(1) Discrete Cosine Transform (DCT)[3]

Discrete cosine transform (DCT) is divided into eight categories. In this paper, DCT- II is adopted for the speech signal sparse decomposition.

Given the sequence $x(n), n = 0, 1, \cdots, N-1$, the discrete cosine transform can be described as[7]

$$X_q(0) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x(n) \qquad (1)$$

$$X_q(\mathrm{k}) = \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} x(n) \cos\frac{(2n+1)k\pi}{2N}, k = 1, 2, \cdots, N-1 \qquad (2)$$

(2) Sparseness of speech signals in DCT basis

In order to ensure the speech signal compressible, a sparse basis must be chosen reasonably to guarantee the good sparseness of speech signal under the base. Here we employ the DCT transform for the sparse decomposition of speech signal "3".

The Matlab simulation results are shown in figure 2. The results show that the larger values in the coefficient of the speech signal "3" coefficients in the DCT domain account for only a small part, most of which are very small and can be approximately zero, which indicates that the speech signal in the DCT domain is approximately sparse. So the CS speech signal reconstruction in the DCT domain is feasible.

Observation matrix

The mainly observation matrices include random Gauss matrix, random Bernoulli distribution, Tope Liz and cyclic matrix, etc..[4][5]

The random Gauss matrix is the most widely used in compressed sensing. The matrix has the following characteristics: very strong randomness, can ensure that any sparse base A has strong correlation, to meet the RIP property. This kind of matrix not only has strong randomness, but also ensures strong irrelevance with any sparse base Ψ, so as to satisfy the RIP property. It can be proved that when $M \ge cKlog(N/K)$, there is a very large probability for observation matrix Φ to meet the RIP properties.

The specific design method is to build a matrix Φ of $M \times N$ size. The matrix follows the Gauss distribution with the mean value of 0 and variance of $1/M$:

$$\Phi_{i,j} \sim N\left(0, \frac{1}{M}\right).$$

The observation matrix plays a very crucial role in the restoration of speech signal. In this paper, random Gauss matrix is chosen as the observation matrix for speech compressed sensing system.

Reconstruction algorithm

Currently, the reconstruction algorithms based on compressed sensing are mainly divided into three categories, namely greedy algorithm, convex optimization algorithm and combinatorial algorithm. This paper mainly introduces two categories of greedy algorithm. Through analysis and comparison, we choose the orthogonal matching pursuit algorithm to reconstruct the speech signals which have been sparse decomposed

The basic algorithm of OMP is as follows:

1) Define X to represent the signal, and initialize the residual $e_0 = x$;
2) Select the atom which obtains the maximum absolute value of transvection with e_0, and express it as ϕ_1;
3) Constitute matrix Φ_t by confining the selected atom as column, and define orthogonal projection operator of matrix Φ_t column space as

$$P = \Phi_t \left(\Phi_t^T \Phi_t\right)^{-1} \Phi_t^T \quad (3)$$

The residuals e_1 are obtained through that e_0 subtracts its orthogonal projection on the space formed by Φ_t.

$$e_1 = e_0 - Pe_0 = (I - P)e_0 \quad (4)$$

4) Perform step (2) and step (3) iteratively for residuals.

$$e_{m+1} = e_0 - Pe_m = (I - P)e_m \quad (5)$$

Where I stands for unit matrix.

5) The algorithm does not stop until a specified stopping criterion is reached.

Speech Signal OMP Algorithm Recovery Simulation

As shown in Fig. 2.2, the DCT coefficient of the speech signal "3" is sparse. Its larger values are generated only in a few positions, and its sparsity K is about 7.

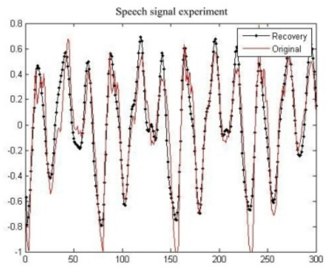
OMP algorithm reconstruction of speech signal "3" in DCT domain

(1) Parameter setting

The experimental signal is speech signal "3" and the signal length ranged from 4000 to 4299 with a total of 300 points. The length of observation signal Y is 200 and the signal sparsity is K=7.

(2) Simulation results and analysis

Recovery residuals: $ans = norm(x\_r - x) = 2.6856$



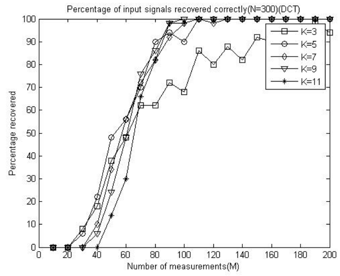


Figure 3 Reconstruction of Speech Signal Figure 4 Influence of M and K on Successful Recovery Rate

The simulation results show that the speech signal has high sparsity in the DCT domain, and the OMP algorithm can achieve effective reconstruction of speech signals.

Influence of parameter setting on success rate of recovery

(1) Parameter setting

The experimental signal is speech signal "3" and the signal length ranged from 4000 to 4299 with a total of 300 points. When K takes 3,5,7,9,11, the length of the observed signal Y increases from 10 to 200 at intervals of 10 points; The number of experimental repetitions under each observation signal length is 50; the successful recovery criterion is that when the recover residuals<3.5.

(2) Simulation results and analysis

The influence of the signal length M and the sparsity K on the percentage of input signals recovered correctly are shown in figure4.

Conclusion

This paper studies the application of compressed sensing theory in speech signal sampling and reconstruction of speech signals. According to the sparsity of speech signals in the discrete cosine transform basis (DCT), we propose a speech compressed sensing (CS) system based on DCT domain which realizes sparse representation of speech signal in DCT domain. Utilizing Gauss random matrix as the measurement matrix and orthogonal matching pursuit algorithm (OMP), the performance of speech signal reconstruction is acquired. The simulation results show that the sparsity of the speech signal is higher in the DCT domain and the OMP algorithm can effectively improves the performance of reconstructed speech signals.

References

[1] Sun Lin-hui. Research on the key issues of compressed speech sensing[D]. Nanjing University of Post and Telecommunications, 2012.

[2] DL. Donoho, Y. Tsaig. Extensions of compressed sensing[J]. Signal Processing,2006, 86(3): 533-548.

[3]Hu Guang-shu. Digital Signal Processing[M]. Bei Jing: Tsinghua University Press, 2003.

[4] Wang Qiang, Li Jia, Shen Yi. A Survey on Deterministic Measurement Matrix Construction Algorithms in Compressive Sensing[J]. Chinese Journal of Electronics, 2013, 41(10): 2041-2050.

[5] Shi Guang-ming, Liu Dan-hua, Gao Da-hua, et al. Advances in Theory and Application of Compressed Sensing[J]. Chinese Journal of Electronics, 2009, 37(5): 1070-1081.

[6] Joel A, Tropp, Anna C. Gilbert, et al. Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit[J]. IEEE Transactions on Information Theory, 2007, 53(12): 436-442.

[7] Wang Xia, Wang Kai, Wang Qing-yun, et al. Deterministic Random Measurement Matrices Construction for Compressed Sensing [J]. Journal of Signal Processing, 2014, 30(4): 436-442.