4th International Conference on Structural Health Monitoring on Intelligent Infrastructure (SHMII-4) 2009
22-24 July 2009, Zurich, Switzerland

Investigation of compressive sampling for structural vibration data

Yuequan Bao[1], Hui Li[1], James L. Beck[2], and Jinping Ou[1, 3]

[1] Harbin Institute of Technology, Harbin, China
[2] California Institute of Technology, Pasadena, USA
[3] Dalian Institute of Technology, Dalian, China

ABSTRACT: In structural health monitoring (SHM) of civil structures, data compression is often needed for saving the cost of data transfer and storage because of the large volumes of sensor data generated from the monitoring system. The traditional framework for data compression is to first sample the full signal, then to compress it. Recently, a new data compression method named *compressive sampling* (CS) has been presented, that can acquire the data directly in compressed form by using special sensors. In this paper, the potential of CS for data compression of vibration data is investigated using simulation of the CS sensor algorithm. The acceleration data collected from the SHM system of Shandong Binzhou Yellow River Highway Bridge and China National Aquatics Center are used to analyse the data compression ability of CS. For comparison, the wavelet transform based and Huffman coding methods are also employed to compress the data. The results show that CS is useful for compression of vibration data in SHM of civil structures and that CS works better for narrowband signals such as the Shandong Binzhou Yellow River Highway Bridge vibration signal than wideband signals such as the vibration signal from the National Aquatics Center. Finally, a design of analog-to-digital converter (ADC) based on CS technique (CSADC) is proposed in this paper and a simulation with analog signal is carried out to illustrate the ability of CSADC for acquiring data directly with compressed form.

1 INTRODUCTION

In recent years, many SHM systems have been installed on civil structures because they have the ability to ensure safety and to observe any damage evolution. Generally, a substantial number of sensors are required in SHM systems because of the large scale and complexity of civil structures. For example, the sensor system of the Wind and Structural Health Monitoring System (WASHMS) installed on Tsing Ma, Ting Kau and Kap Shui Mun bridges in Hong Kong, China, consists of approximately 800 sensors (Ko, 2003). The structural behavior of the bridges is measured 24 hours a day, seven days a week, so enormous amounts of sensor data are generated by the WASHMS. Data compression has, therefore, a potentially important role in SHM to save the cost of data transfer and storage.

For SHM of civil structures, data compression has attracted the interests of researchers in recent years. Zhang & Li (2006) presented a data compression method based on the lifting-scheme wavelet transforms, and this technique can achieve a high compression ratio. Wireless sensor

networks for SHM have been investigated because they not only can provide the same functionality at a much lower installation price but also can process data autonomously by embedded software (Lynch & Loh, 2005). Data compression techniques can be used to minimize bandwidth and power consumption during wireless communications, so that the lifetime of the wireless sensors can be maximized. Data compression is therefore a crucial aspect of wireless sensor networks. Xu et al. (2004) investigated the applicability of wavelet-based compression techniques to overcome the bandwidth limitations of wireless sensors. Lynch et al. (2003) used Huffman coding to reduce the size of the wirelessly transmitted data packets. However, all these methods belong to a traditional data compression framework.

Recently, a new data compression method called *compressive sampling* (CS) has been introduced (Candès 2006). CS is an alternative theory to Nyquist/Shannon sampling theory that can reconstruct images or signals accurately from far fewer data/measurements than what is usually considered necessary (Candès 2006). Candès (2006) suggests that this new technique may come to underlie procedures for sampling and compressing data simultaneously. In this paper, the application of CS for compression of structural vibration data is investigated.

2 COMPRESSIVE SAMPLING

For more detailed information on CS, the reader is referred to the references (Candès 2006, Candès & Tao 2005, Candès et al. 2006, Baraniuk 2007). Only a short introduction to CS is presented here.

Any discrete-time signal $x(t)$, $t = 1,\ldots,n$, in $\mathbb{R}$ can be represented in terms of an orthonormal basis of $n \times 1$ vectors $\{\psi_i\}_{i=1}^{n}$ as follows. Forming the $n \times n$ *basis matrix* $\mathbf{\Psi} = [\psi_1, \psi_2, \ldots, \psi_n]$ by taking the vectors ψ_i as columns, then $\mathbf{\Psi}^{-1} = \mathbf{\Psi}^T$ and the vector $x = [x(1), \cdots, x(n)]^T \in \mathbb{R}^n$ of n sampled values of $x(t)$ can be expressed as:

$$x = \sum_{j=1}^{n} \alpha_j \psi_j \text{, or } x = \mathbf{\Psi}\alpha \tag{1}$$

where the weighting or basis coefficient $\alpha_j = \langle x, \psi_j \rangle$ using the Euclidean scalar product, or $\alpha = \mathbf{\Psi}^T x$.

Suppose that we acquire an $m\,(< n)$ length data vector y by a linear projection of x:

$$y = \mathbf{\Phi}x = \mathbf{\Phi}\mathbf{\Psi}\alpha = \mathbf{\Theta}\alpha \tag{2}$$

where $\mathbf{\Theta} = \mathbf{\Phi}\mathbf{\Psi}$ is an $m \times n$ matrix and $\mathbf{\Phi}$ is called a *measurement matrix*. Since $\mathbf{\Phi}$ is an $m \times n$ matrix with $m < n$, the problem of recovering the signal x, which has n components, from the m measurements in y is ill-posed. But if x is known to be sparse in the basis $\mathbf{\Psi}$, i.e. many of the coefficients α_j are zero, then under some conditions α can be reconstructed exactly by solving the convex optimization program:

$$\text{(P1)} \quad \hat{\alpha} = \arg\min \|\tilde{\alpha}\|_1 \text{ such that } \mathbf{\Theta}\tilde{\alpha} = y \tag{3}$$

The solution $\hat{\alpha}$ can recover α exactly under the following conditions: (1) the signal x is *s-sparse* where a vector is defined to be s-sparse if it has at most s nonzero entries; (2) the matrix $\mathbf{\Theta}$ obeys a so-called *restricted isometry property* (RIP) (Candès 2006, Candès & Tao 2005), that is, there exists an isometry constant δ_s for the matrix $\mathbf{\Theta}$, where δ_s is defined as the smallest number such that

$$(1 - \delta_s)\|v_s\|_2^2 \le \|\mathbf{\Theta}v_s\|_2^2 \le (1 + \delta_s)\|v_s\|_2^2 \tag{4}$$

holds for all *s*-sparse vectors v_s.

Thus, if $\boldsymbol{\Theta}$ satisfies Eq. 4, then the signals that are *s*-sparse can be exactly recovered by the optimization problem (P1).

Suppose that $\boldsymbol{\Psi}$ is an orthonormal basis such as a wavelet basis or Fourier sinusoids basis, then we wish to construct a measurement matrix $\boldsymbol{\Phi}$ such that $\boldsymbol{\Theta} = \boldsymbol{\Phi}\boldsymbol{\Psi}$ has the RIP. In compressive sampling, $\boldsymbol{\Phi}$ is selected as a random matrix; for example, some matrices are given in Candès (2006) such as the $m \times n$ matrix $\boldsymbol{\Phi}$ formed by sampling independent and identically distributed entries $\phi_{i,j}$ from either the Normal distribution with mean zero and variance $1/m$ or the symmetric Bernoulli distribution. With overwhelming probability, these random Gaussian and Bernoulli matrices $\boldsymbol{\Phi}$ obey the RIP provided that

$$m \geq c \cdot s / \log(n/s) \tag{5}$$

where c is a specified constant (Candès et al. 2006). The matrix $\boldsymbol{\Theta} = \boldsymbol{\Phi}\boldsymbol{\Psi}$ then also obeys the RIP with $m \geq c \cdot s / \log(n/s)$ regardless of the orthonormal basis matrix $\boldsymbol{\Psi}$ (Candès 2006, Baraniuk 2007).

If the measurements are corrupted with noise (e.g. during transmission), consider the model

$$y = \boldsymbol{\Phi}x + e \tag{6}$$

where e is a bounded but uncertain error with $\|e\|_2 \leq \varepsilon$. The basis coefficients α can be reconstructed by solving the convex optimization problem:

$$\text{(P2)} \quad \hat{\alpha} = \arg\min \|\tilde{\alpha}\|_1 \text{ such that } \|\boldsymbol{\Theta}\tilde{\alpha} - y\|_2 \leq \varepsilon \tag{7}$$

where ε is the chosen bound on the size of the measurement error. As for problem (P1) in Eq. 3, problem (P2) also has a unique solution and the optimal basis coefficients $\hat{\alpha}$ may be used to represent the data; however, now exact recovery of x is not guaranteed.

3 COMPRESSIVE SAMPLING OF ACCELEROMETER SIGNALS FROM A BRIDGE SHM SYSTEM

Shandong Binzhou Yellow River Highway Bridge is a cable-stayed bridge with three towers located in Shandong, China, as shown in Fig. 1. It has a total length of 768m, consisting of two three-hundred meter main spans and two eighty-four meter side spans. The main tower is 125.28 meters high and the side tower is 75.778 meters high. The bridge was constructed in 2002 and opened to traffic in 2004. A SHM system has been installed on this bridge and operated since 2004. Fig. 2 shows the locations of the accelerometers placed on the bridge.

Fig. 1. Shandong Binzhou Yellow River Highway Bridge.

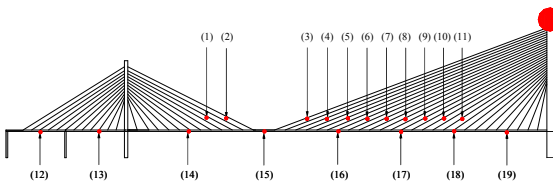


Fig. 2. Locations of the accelerometers.

Fig. 3(a) shows the acceleration response of the bridge deck collected from one accelerometer of the SHM system. The sampling frequency is 100Hz, and a total of 2,048 data samples are collected.

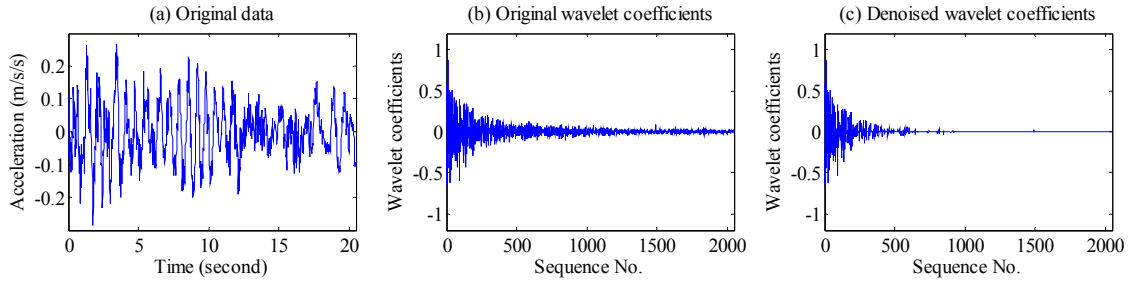


Fig. 3. (a) Acceleration response of bridge deck measured by accelerometer 13; (b) Original wavelet coefficients; (c) Denoised wavelet coefficients.

The wavelet coefficients of the acceleration data when decomposed by the Haar wavelet basis with five resolution levels are shown in Fig. 3(b). It can be seen that the wavelet coefficients are not sparse. After using threshold denoising, the wavelet coefficients shown in Fig. 3(c) are sparse in the Haar wavelet basis domain because 82.28% of the coefficients are zero. Denoting the denoised wavelet coefficients by α, the original acceleration can expressed as

$$x(t) = \mathbf{\Psi}\alpha + \tilde{e} \tag{8}$$

where $\mathbf{\Psi}$ is the Haar wavelet basis matrix and $\tilde{e}$ is the noise. The measured data can be expressed as:

$$y = \mathbf{\Phi}x = \mathbf{\Phi}x_d + \mathbf{\Phi}\tilde{e} = \mathbf{\Phi}\mathbf{\Psi}\boldsymbol{\alpha} + e \tag{9}$$

where x and $x_d = \mathbf{\Psi}\boldsymbol{\alpha}$ are the original acceleration data and denoised data, respectively.

We solve the (P2) problem of CS to determine the optimal basis coefficients $\hat{\alpha}$:

$$\hat{\boldsymbol{\alpha}} = \arg\min \left\|\tilde{\boldsymbol{\alpha}}\right\|_1 \text{ such that } \left\|\mathbf{\Phi}\mathbf{\Psi}\tilde{\alpha} - y\right\|_2 \le \varepsilon \tag{10}$$

where ε is an upper bound on the size of the noise contribution such that $\| e \|_2 \le \varepsilon$. Since $\| e \|_2 = \| \mathbf{\Phi}\tilde{e} \|_2 \le \| \mathbf{\Phi} \|_2 \| \tilde{e} \|_2 \le \varepsilon$, to set the threshold ε, the norm of measurement noise $\| \tilde{e} \|_2$ must be estimated, such as by prior information or by making measurements with a sensor without CS. In the examples below, we chose ε by minimizing the reconstruction error for a given number of measurements m (that is, a given compression ratio). In practice, one would not know the true signal x, so this optimum value of ε is not available; however, we have found that the results are not sensitive to the choice of ε.

3.1 The CS Results

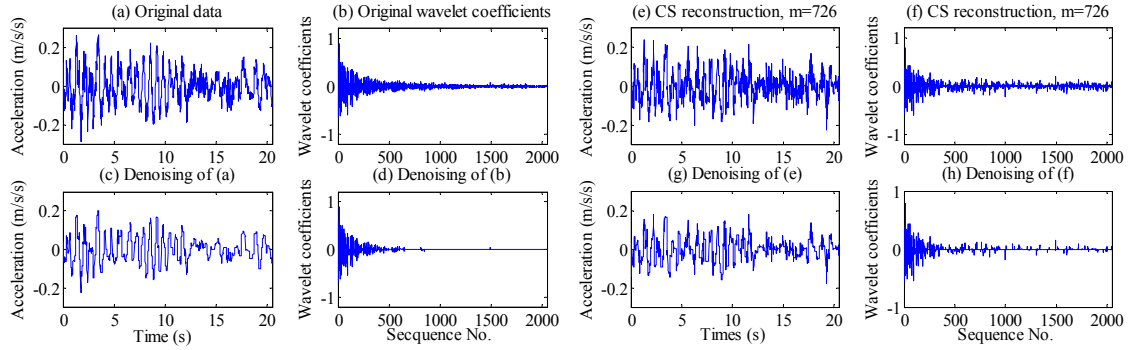


Fig. 4. CS results of acceleration data: (a) Original data, $n = 2048$; (b) Original wavelet coefficients; (c) Denoised data of (a), the reconstruction error is $\xi = 0.3525$; (d) Denoised wavelet coefficients used for (c), $s = 363$; (e) CS reconstruction data, $m = 2s = 726$, the reconstruction error is $\xi = 0.5507$; (f) Wavelet coefficients of CS reconstruction data; (g) Denoised data of (e), the reconstruction error is $\xi_d = 0.4280$; (h) Wavelet coefficients of (g).

The Haar wavelet basis is selected to form the orthogonal basis matrix $\mathbf{\Psi}$. In this example, $n = 2048$ and $s = 363$, where s is the nonzero wavelet coefficients of the denoised data. The CS results with $m = 2s$, $m = 3s$ and $m = 4s$ are shown in Fig. 4-6, respectively. The reconstruction error is calculated as

$$\xi = \left\| \hat{x} - x \right\|_2 / \left\| x \right\|_2 \tag{11}$$

where $\hat{x}$ and x are the reconstruction data and the original data, respectively. The reconstruction error of CS with denoising (CSDN) is calculated as

$$\xi_d = \left\| \hat{x}_d - x_d \right\|_2 / \left\| x_d \right\|_2 \tag{12}$$

where $\hat{x}_d$ and x_d are the denoised reconstruction data and the denoised original data, respectively.

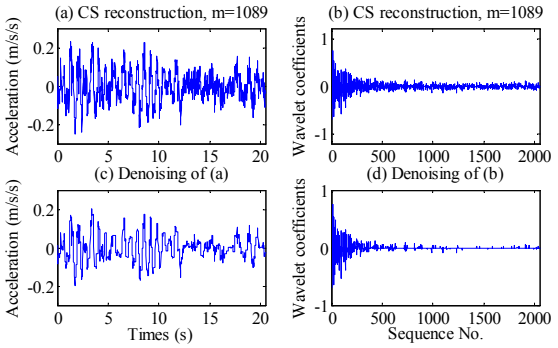


Fig. 5. CS results of acceleration data with $m = 3s = 1089$: (a) CS reconstruction data, the reconstruction error is $\xi = 0.3902$; (b) Wavelet coefficients of (a); (c) Denoised data of (a), the reconstruction error is $\xi_d = 0.2624$; (d) Denoised wavelet coefficients used for (c).

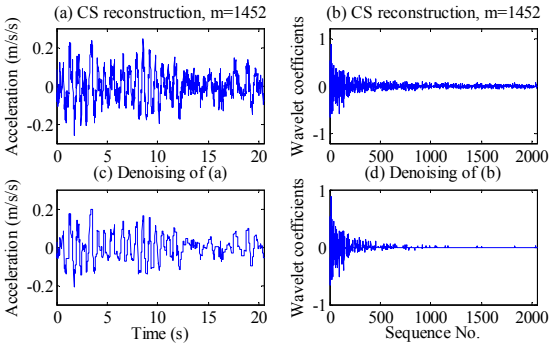


Fig. 6. CS results of acceleration data with $m = 4s = 1452$: (a) CS reconstruction data, the reconstruction error is $\xi = 0.2721$; (b) Wavelet coefficients of (a); (c) Denoised data of (a), the reconstruction error is $\xi_d = 0.1651$; (d) Denoised wavelet coefficients used for (c).

Fig. 4(e,f) compared with Fig. 4(a,b) show that CS can approximately recover the data when $m = 2s$ but it is somewhat noisy. After denoising the CS reconstruction data by a wavelet transform, the plots shown in Fig. 4(g,h) and Fig. 4(c,d) demonstrate that the CS denoised data results are much improved over the original and CS reconstruction data comparison. Fig. 5 shows the reconstruction results with $m = 3s$, which look much better than the results with $m = 2s$ in Fig. 4. Fig. 6 shows the reconstruction results with $m = 4s$, which are better than $m = 2s$ and $3s$. Clearly, the results are improved with increasing m, as expected. However, the goal of CS in this example is to compress the data, so, as expected, there is a trade-off between the accuracy of the reconstructed data and the amount of compression.

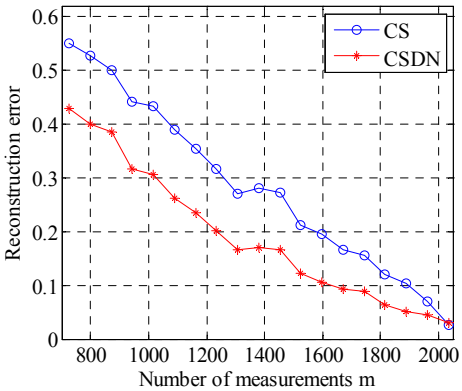


Fig. 7. Reconstruction error of CS for Binzhou Yellow River Highway Bridge acceleration data.

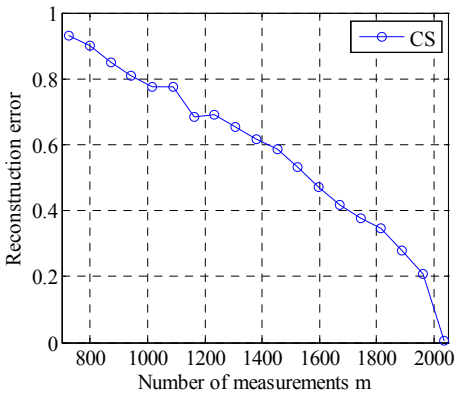
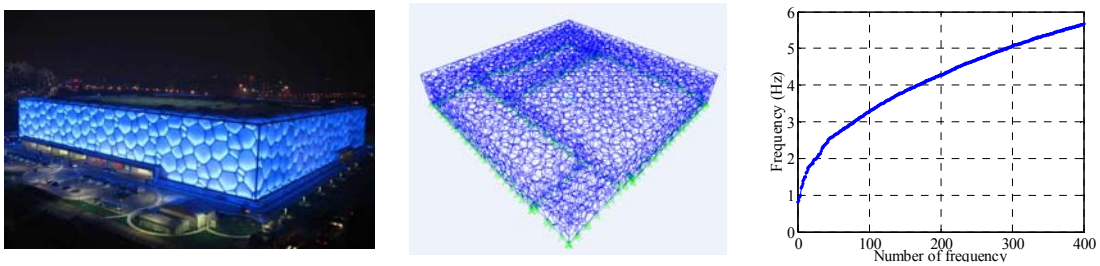


Fig. 8. Reconstruction error of CS for China National Aquatics Center acceleration data.

To investigate the relation between reconstruction error and the amount of measurements using CS, we choose a sequence of measurement numbers $m = (2s, 2.2s, \ldots, 5.6s)$, where $s = 363$. The results are shown in Fig. 7, which demonstrates how the reconstruction error decreases with an increase in the number of measurements.

4 COMPRESSIVE SAMPLING OF ACCELEROMETER SIGNALS FROM A SPACE STRUCTURE SHM SYSTEM

The China National Aquatics Center, also known as the 'Water Cube', is an aquatics center that was built for the swimming competitions of the 2008 Summer Olympics. The National Aquatics Center, shown in Fig. 9(a), has had an SHM system installed in it since 2007. Fig. 9(c) is the first 400 frequencies calculated from the finite element model shown in Fig. 9(b), which shows that the Water Cube is a structure with closely-spaced natural frequencies.



(a) (b) (c)

Fig. 9. (a) National Swimming Center; (b) Finite element model of National Aquatics Center; (c) First four hundred frequencies of National Swimming Center.

The acceleration data collected from the 'Water Cube' is shown in Fig. 10(a). The wavelet coefficients and Fourier spectrum are shown in Fig. 10(b) and Fig. 10(c), respectively, which show that the acceleration data is a wideband signal. The CS results are shown in Fig. 8, which demonstrates that the CS reconstruction error of National Aquatics Center acceleration data is more than the CS reconstruction error of the Shandong Binzhou Yellow Riever Highway Bridge acceleration data with the same number of measurements.

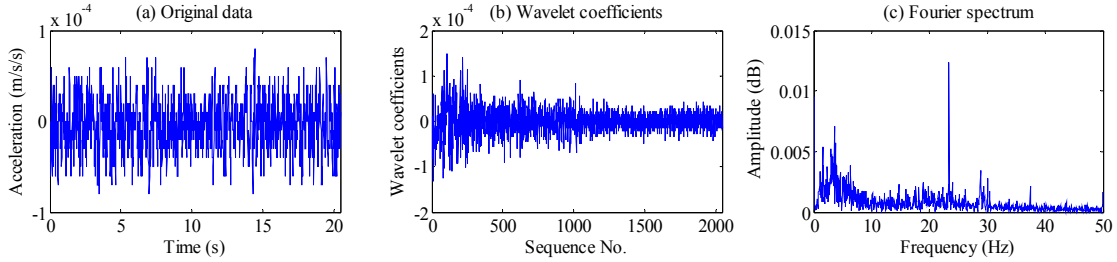


Fig. 10. (a) Acceleration data of National Aquatics Center; (b) Wavelet coefficients; (c) Fourier spectrum

5 COMPARISON WITH OTHER DATA COMPRESSION METHODS

For comparison, data compression using the wavelet transform and Huffman coding are also used. The acceleration data of Shandong Binzhou Yellow River Highway Bridge is used as a example for comparison in this section. The data compression performance is measured using the compression ratio

$$CR = N_a / N_c \tag{13}$$

where N_a and N_c represent the original data size and compressed data size, respectively. Huffman coding methods include two stages. The first stage is decomposing the acceleration data using the wavelet transform and quantifying the wavelet coefficients by threshold. In the

second stage, the wavelet coefficients are encoded with a binary code such as "100010010". The wavelet transform method, without also using another encoding approach, only recodes the nonzero wavelet coefficients and their locations. The original acceleration data and wavelet coefficients are all accurate to 4 decimal digits. The location information of the wavelet coefficients are recorded as integers. All the data sizes are measured in bytes, and the data compression results are shown in Fig. 11.

Fig. 11 shows the relation of compression ratio and reconstruction error of CS, wavelet data compression, static Huffman coding and adaptive Huffman coding methods. It can be seen from Fig. 8 that the compression ratio increases with an increase in reconstruction error for these four data compression methods. Clearly, with the same reconstruction error, the wavelet method has the best compression ratio, and CS with the wavelet basis is close to the Huffman coding methods for smaller reconstruction errors.

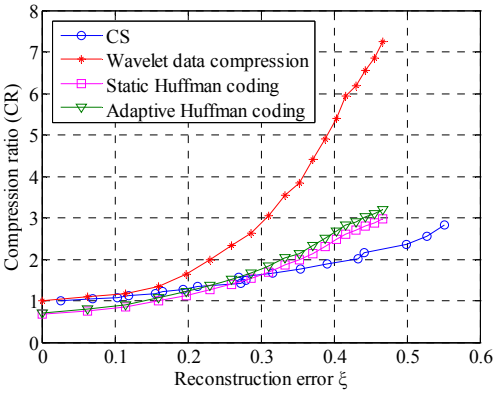


Fig. 11. Relation of compression ratio and reconstruction error.

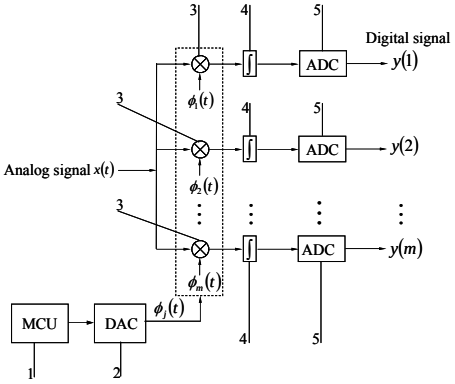


Fig. 12. Design of CSADC.

6 DESIGN OF ADC BASED ON COMPRESSIVE SAMPLING TECHNIQUE

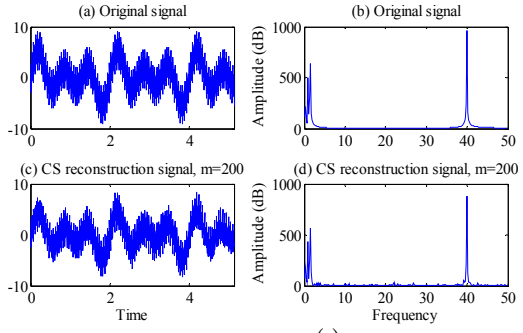


Fig. 13. CS results of signal $x_1(t)$: (a) Original data, n=512; (b) Fourier spectrum of (a); CS reconstruction signal with $m = 200$, $\xi = 0.1979$; (d) Fourier spectrum of (c).

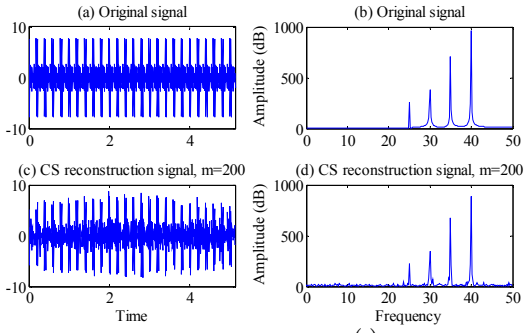


Fig. 14. CS results of signal $x_2(t)$: (a) Original data, n=512; (b) Fourier spectrum of (a); CS reconstruction signal with $m = 200$, $\xi = 0.2229$; (d) Fourier spectrum of (c).

The advantage of CS is that only the measurement matrix $\mathbf{\Phi}$ is needed when acquiring the data. The orthonormal basis $\mathbf{\Psi}$ is only used when recovering the data. Furthermore, the matrix $\mathbf{\Phi}$ does not depend on the signal $x(t)$, although the sparsity of the signal in some orthonormal basis must be estimated. Thus, it is possible to design a physical sampling device that directly records compressed discrete data from the analog signal. A new design of ADC based on CS (CSADC) is given in Fig. 12, which is composed by Micro Controller Unit (MCU) (1), traditional DAC (2), analog multiplier (3), integral circuit (4), and traditional ADC (5). Two simple analog signals $x_1(t) = \sin(\pi t) + 2\sin(2\pi t) + 3\sin(3\pi t) + 4\sin(80\pi t)$ and

$x_2(t) = \sin(50\pi t) + 2\sin(60\pi t) + 3\sin(70\pi t) + 4\sin(80\pi t)$ are used to illustrate the ability of CSADC and the results from simulating the CSADC design are shown in Fig. 13 and Fig. 14, respectively.

7 CONCLUSIONS

In this paper, the potential of CS for compressing vibration data for SHM is investigated for the first time by using real acceleration data collected from the SHM system of the Shandong Binzhou Yellow River Highway Bridge and China National Aquatics Center. Since CS sensors are not yet available, we digitally simulate their behavior. The results show that the achieved compression ratios by CS are not high because these data are not naturally sparse in the wavelet basis. The sparsity of the transformed signal for the chosen basis controls the applicability of CS, since greater sparsity means that CS can compress the data better. Also, CS works better for narrowband signals such as the Shandong Binzhou Bridge vibration signal than wideband signals such as the vibration signal from the National Aquatics Center. Although the compression ratio shown in this work is not better than the wavelet-based data compression method, CS can be designed into a physical device such as an analog-to-digital converter to acquire the data directly in a compressed form, in contrast to the traditional data compression methods. A design of CSADC is proposed in this paper and the simulation results using a simple analog signal show how the CSADC can directly record compressed discrete data from an analog signal, although the reconstruction errors are not be negligible. In future work, we will explore a Bayesian statistical approach that should allow better control of the releation between reconstruction error and the number of measurements.

ACKNOWLEDGEMENTS

This research is supported by the China Scholarship Council (CSC), which supported the first author while he was a visiting student researcher at the California Institute of Technology. This support is gratefully acknowledged. Also this research is supported by grants from China NSFC (Grant No. 50278029, and 50525823), which supported the first, second and fourth author.

REFERENCES

Ko, JM. 2003. Health Monitoring and Intelligent Control of Cable-Supported Bridges, Proceeding of the International Workshop on Advanced Sensors, Structural Health Monitoring and Smart Structures, Keio University, Japan, November 10-11.

Zhang, Y and Li, J. 2006. Wavelet-Based Vibration Sensor Data Compression Technique for Civil Infrastructure Condition Monitoring, *Journal of Computing in Civil Engineering*, 20(6): 390-399.

Lynch, JP and Loh, KJ. 2005. A Summary Review of Wireless Sensors and Sensor Networks for Structural Health Monitoring, *Shock and Vibration Digest*, 38(2): 91-128.

Xu, N., Rangwala, S., Chintalapudi, K., et al. 2004. A Wireless Sensor Network for Structural Monitoring, Proceedings of the ACM Conference on Embedded Networked Sensor Systems, Baltimore, MD, USA, November.

Lynch, JP, et al. 2003. Power-Efficient Data Management for a Wireless Structural Monitoring System, Proceedings of the 4th International Workshop on SHM, Stanford, CA, September 15-17: 1177-1184.

Candès, EJ. 2006. Compressive sampling, Proceedings of the International Congress of Mathematicians, Madrid, Spain: 1433-1452.

Candès, EJ and Tao, T. 2005. Decoding by Linear Programming, *IEEE Transactions on Information Theory*, 51: 4203-4215.

Candès, EJ, et al. 2006. Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information, *IEEE Transactions on Information Theory*, 52(2): 489–509.

Baraniuk, R. 2007. A Lecture on Compressive Sensing, *IEEE Signal Processing Magazine*, 24(4): 118-121.