Compressed Sensing Method for Human Activity Sensing using Mobile Phone Accelerometers

Daito Akimura*, Yoshihiro Kawahara*† and Tohru Asami*
{akimura, kawahara, asami}@akg.t.u-tokyo.ac.jp
*Graduate School of Information Science and Technology
The University of Tokyo
7-3-1, Hongo, Bunkyo-ku, Tokyo, Japan
†School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, Georgia 30308, USA

***Abstract*—This paper presents the first complete design to apply the compressed sensing (CS) theory to activity sensor data gathering for smart phones. Today, most of the mobile phones are equipped with multiple sensors, such as cameras, GPS, and accelerometers. By exploiting the sensing features, we capture many different events and share them over the mobile network. One of the most important challenges for such a participatory sensing system is to reduce the battery consumption of the mobile device. We overcome this challenge by reducing the communication data, without introducing intensive computation at mobile terminals. The CS technique consists of very simple matrix operations at the mobile side, and CPU-intensive reconstruction is performed on the resource-rich machine on the network side. Since CS is a lossy compression technique, the reconstructed signal contains errors depending on the degree of sparseness of the original signal. We evaluated the proposed method by using a large amount of real activity data consisting of six basic activities performed by 90 test subjects. We also implemented our method on the iPhone/iPod platform and showed that our method can reduce power consumption by approximately 16% as compared with ZIP compression, while maintaining the error below 10%.**

I. Introduction

Mobile devices such as smartphones have become a powerful sensing platform for capturing the user's daily activity data in the real world. A majority of people carry around mobile devices and charge them daily. They facilitates the deployment of sensors in wide areas and the installation of sensing applications[1]

Built-in sensors in mobile devices are not only typically used in personal applications such as health monitoring but can also be applied to urban area sensing. Examples include climate change data collection using barometers and ambient noise measures using microphones. By extracting the relevant data from the vast amount of context information, these applications can provide fine-tuned personalized services. In order to capture not only the environmental information but also the behavior of the users, accelerometers embedded in mobile devices are used to recognize human activities and transportation modes [2]. By capturing and collecting the characteristics and the dynamics of daily activities, we can implement social applications such as lifelogging and microblogging.

The main difference between mobile devices and dedicated sensing systems is that the former systems allow any user to participate in sensing campaigns[3]. By broadcasting such campaigns, service providers recruit participants and capture sensor data from their devices. These data are then sent to a central server on the network. This participatory sensing scheme has been applied in several studies. However, one of the most important considerations concerning the practical uses of the applications is the battery life of mobile devices. The mobile devices include phone calls, text messaging, and Web browsing. Hence, secondary sensing tasks should not consume considerable power.

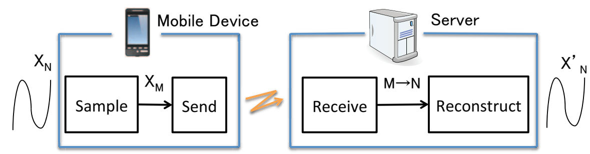

Fig. 1. Sensing System using a Mobile Device

CPU usage and wireless communication are the most salient factors for saving energy at mobile terminals. A band-limited analog signal that has been sampled can be perfectly reconstructed from an infinite sequence of samples if the sampling rate exceeds twice the highest frequency. However, in most applications, the Nyquist rate is so high that the signal contains too much information. This excessive amount of data results in the additional consumption of CPU and communication resources.

In this paper, we propose the use of a lightweight compression technology called compressed sensing (CS) to minimize the size of the data sent over a wireless link, with exceedingly simple calculations (Fig. 1). Although CS is a lossy compression that generates the reconstruction error, we can exploit the major components of the original signals by taking advantage of the sparseness of the acceleration signals.

Yang and Gerla have already proposed the use of the CS technology to body worn accelerometers[4]. However, the performance evaluation of this concept is carried out under limited conditions. Our objective is to analyze the performance of the CS method for participatory sensing applications. The relationship between compression error and activities and the impact of the battery lifetime using off-the-shelf mobile terminals have never been reported in the literature.

The contributions of this study are summarized as follows:

- Evaluation of the reconstruction error and the recognition accuracy of 6 basic human activities using the acceleration data corpus of 90 test subjects, as well as clarification of the variation in the error.

- Implementation of an iPhone/iPod application that continuously samples, compresses, and sends acceleration data to a server using CS, as well as evaluation of power consumption.

II. Theory of Compressed Sensing

When the original signals can be expressed as a large number of zero or nearly zero values and a small number of non-zero values, the signals are said to be "sparse." CS is a state-of-the-art data compression and reconstruction theory that exploits the fact that many natural signals are sparse or compressible in the sense that they have concise representations when expressed in the appropriate basis[5].

Let $\boldsymbol{\Psi}$ be an $N \times N$ basis matrix and $\boldsymbol{s} \in \mathbb{R}^n$ be a sparse expression of the original signal. Then, the original signal $\boldsymbol{x}$ can be expressed as follows:

$$\boldsymbol{x} = \boldsymbol{\Psi s} \tag{1}$$

Figs. 2 and 3 show an example acceleration signal of walking and the discrete cosine transform (DCT) coefficients of the original signal. These figures show that the acceleration signals can be expressed as sparse data with the basis transform.

In the CS theory, the original signal $\boldsymbol{x}$ is compressed by an $M \times N$ sampling matrix $\boldsymbol{\Phi}$, and the data $\boldsymbol{d} \in \mathbb{R}^m$ that are sent to a server are expressed using the following Eq. (3):

$$\begin{aligned} \boldsymbol{d} &= \boldsymbol{\Phi x} && (2) \\ &= \boldsymbol{\Phi \Psi s} && (3) \end{aligned}$$

where $\boldsymbol{\Phi}$ should be incoherent with $\boldsymbol{\Psi}$. Berinde et al. showed that the sparse random matrix introduced by them has this property with high probability[6]. Therefore, the N-dimensional original data $\boldsymbol{x}$ are compressed to the M-dimensional data $\boldsymbol{d}$ that is sent to the server (Fig. 1).

When $\boldsymbol{\Psi}$ and $\boldsymbol{\Phi}$ are known at the server, the original signal $\boldsymbol{x}$ can be reconstructed correctly only if it is possible to estimate $\boldsymbol{s}$ uniquely. Eq. (3) represents an ill-posed problem that generally cannot be solved since $M < N$. However, under the assumption that $\boldsymbol{x}$ can be expressed in the basis $\boldsymbol{\Psi}$, $\boldsymbol{s}$ can be estimated by solving the following L^1 norm minimization using linear programming:

$$\arg\min_{\boldsymbol{s}} ||\boldsymbol{s}||_1 \text{ subject to } \boldsymbol{\Phi \Psi s} = \boldsymbol{d} \tag{4}$$

Figs. 4 and 5 show the reconstructed walk signal and DCT coefficients, respectively. These figures show that L^1 norm minimization reconstructs the major components of the DCT coefficients of the original signal.

Now, using $\boldsymbol{\Phi}$, we can reduce the data traffic to M/N each time N data are sampled. Because the compression process does not require complex calculations other than one multiplication of the sparse matrix, it is not CPU-intensive. However, because CS is a lossy compression technique, the

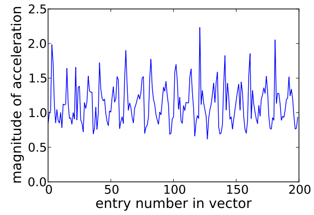

Fig. 2. Example of the Original Walking Signal

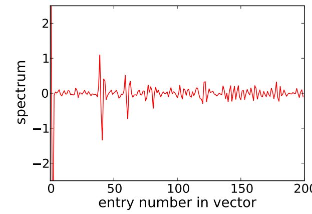

Fig. 3. DCT Coefficients of the Original Walking Signal

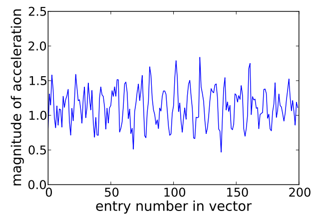

Fig. 4. Example of the Reconstructed Walking Signal (Compression Ratio = 0.50)

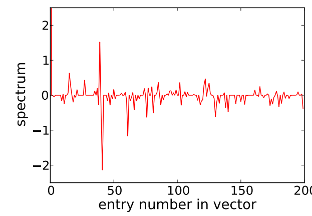

Fig. 5. DCT Coefficients of the Reconstructed Walking Signal (Compression Ratio = 0.50)

reconstructed acceleration signals contain errors. Therefore, it is important to adjust M/N, depending on the sparseness of the original signal.

III. Activity Recognition and the Effect of Data Compression

In this paper, we focus on user's activity monitoring based on acceleration sensor readings. Such activity monitoring information is useful for e-Health applications[7]. The objective of this section is to clarify how the reconstruction error impacts the activity recognition operation.

A. Compression and Reconstruction of Acceleration Data

The acceleration data obtained through mobile devices typically consist of 3 axes. This is mainly because the accelerometer is used to detect the mobile terminal's orientation as a human interface device. However, for the activity recognition application, the Euclidean norm is more convenient because it eliminates the effect of gravity resulting from the difference in the device orientation. Hence, the "original" signal $\boldsymbol{x}$ is calculated as $\boldsymbol{x} = \sqrt{Acc_x^2 + Acc_y^2 + Acc_z^2}$ using the readings from the accelerometer (Acc_x, Acc_y, Acc_z).

As mentioned in the previous section, signal $\boldsymbol{x}$ is compressed by multiplying it by the sampling matrix $\boldsymbol{\Phi}$. We chose the sparse random matrix introduced by Berinde et al. for $\boldsymbol{\Phi}$ [6]. Then $\boldsymbol{d} = \boldsymbol{\Phi x}$ is sent over the communication link and stored at the server. At the server side, a DCT is selected as the basis matrix $\boldsymbol{\Psi}$ as in Eq. (5).

B. Dataset

In the field of speech and image recognition, a large-scale corpus plays an important role. "HASC2010corpus" is the only and largest corpus of acceleration data attached to the human body[8] that aims at achieving the recognition and understanding of human activity through sensing. The corpus is composed of basic six basic activities (stay, walk, jog, skip, ascend stairs up, and descend stairs) performed by each user.

The corpus includes activity data of 540 subjects, and 96 of those have a full dataset. The total number of activity data files is 6791, with a total size of 966 Mbytes. Because unfortunately some of the datasets were incomplete, we obtained the data of 90 persons from the corpus and examined the individual differences depending on the CS compression ratio.

C. Activity Recognition

First, we investigated the recognition error and the individual variance of the activity data. We prepared 19 datasets that are reconstructed from different data compression ratios ranging from 5% to 95% and compared the recognition errors and the individual variances according to the type of the activity. The normalized root mean square error (NRMSE) was used to evaluate the errors between the original and the reconstructed signals.

Then, we conducted activity recognition using the reconstructed signals to check the effect of the recognition error. For activity recognition, we employed a method introduced by Kwapisz et al.[9]. On the basis of their work, we chose 10 seconds as the window size and the sampling rate of each acceleration signal was downsampled to 20Hz through a LPF. For constructing a classifier, 4 trials $\times$ 10 s samples of 6 activities for each of the 90 subjects are used: therefore, the total amount of data reaches 2160 $\times$ 10 s samples. We evaluated the relationship between the compression ratio and the recognition accuracy.

The following values are extracted from the norm of the acceleration data as features: average, variance, average absolute difference (AAD), zero-crossing rate (ZCR), binned distribution (10), energy ($0.5-1.0$Hz, $1.0-1.5$Hz, $1.5-2.0$Hz, $2.0-2.5$Hz, $2.5-3.0$Hz). As a classifier, we chose a J48 decision tree included in the Weka data-mining suite[10]. The accuracies were calculated by using a 10-fold cross validation. The same persons were not used for collecting data for learning and for testing the datasets. When we tested the reconstructed data, we also used the classifier learned with the original uncompressed data.

D. Results

Fig. 6 shows the average errors and the standard deviations of the reconstructed signals of the six basic activities. When the compression ratio is $M/N = 0.5$, it is possible to reconstruct the original signal with an accuracy of approximately 0.08 NRMSE. The error of the stay signals is greater than those of the other activities. However, the norms of the stay signals have small fluctuations; therefore, their NRMSE does not influence their general shape.

Table I shows the confusion matrix when the original (uncompressed) signals are classified in the way noted above. Overall, the recognition accuracy was 84%. In all the features, the energies heavily affected the accuracy. Most of the faults of the classification derive from failure to distinguish between the walk signals and the up/down stairs signals.

Fig. 7 shows the relationship between the compression ratio and the recognition accuracy of each activity. The recognition

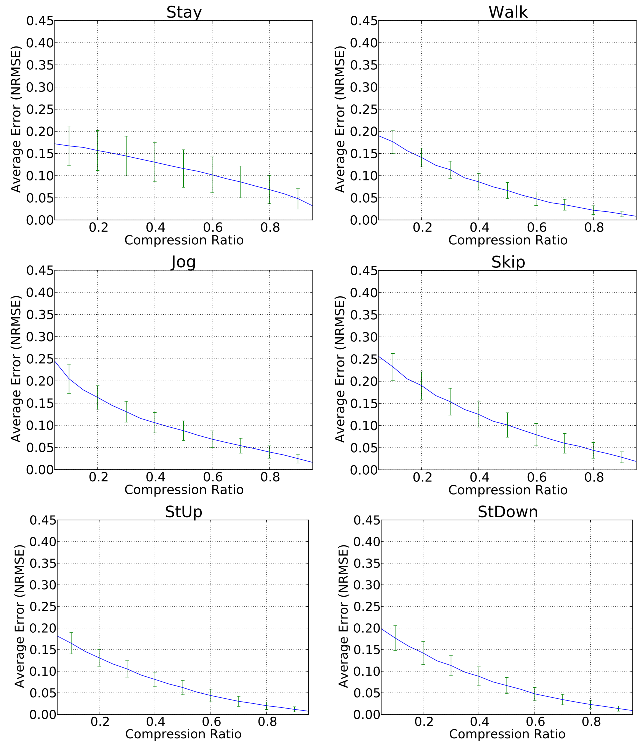

Fig. 6. Average Error and Variance of the Basic Activities

TABLE I
RECOGNITION ACCURACY OF THE ORIGINAL SIGNALS OF 90 PEOPLE (OVERALL: $\mathbf{84.0}\%$)

%		Predicted Class					
		Stay	Walk	Jog	Skip	StUp	StDown
Actual Class	Stay	**98.1**	1.4	0.0	0.0	0.3	0.3
	Walk	2.5	**77.2**	0.3	0.3	12.8	6.9
	Jog	0.3	1.7	**89.7**	6.4	0.0	1.9
	Skip	0.0	0.3	3.9	**93.3**	1.1	1.4
	StUp	0.0	11.7	0.0	0.3	**73.1**	15.0
	StDown	0.0	6.9	2.8	1.9	15.6	**72.8**

accuracy is greater than 70% even when the compression ratio is set to 50%. Overall, the accuracy decreases with an increase in the compression ratio. However, the degree of degradation differs depending on the activity. This is because each activity is recognized by a different set of features. For example, the recognition accuracy of the "stay" signal does not drop much. This is because the signals in the "stay" state originally show little change in the accelerations.

When the compression ratio is 50%, 70% of the all classification error arises because "walk" state is mistakenly classified as "up/down stairs", and vice versa. This tendency becomes more obvious when the compression ratio increases. More accurate classifiers to recognize "walk" and "up/down stairs" signals are requested when the data compression ratio needs to be small. Since we assume that the acceleration signals in DCT are sparse, the higher the compression ratio, the more the signals are sparsely reconstructed. Therefore, in future, we should also consider the effect of the reconstruction process in

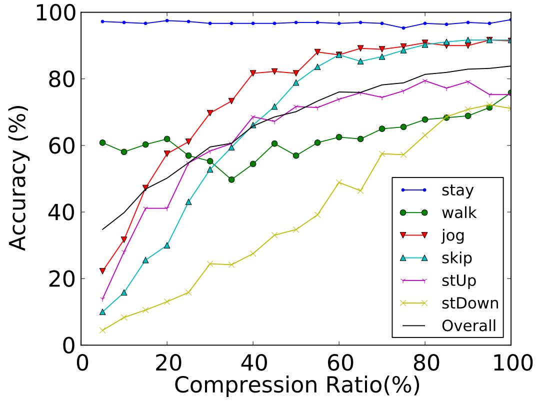

Fig. 7. Relationship between Compression Ratio and Recognition Accuracy of Basic Activities

order to recognize the activities from the reconstructed signals.

IV. Evaluation of the Power Consumption on the iPhone/iPod platform

To compare the power consumption of the proposed scheme, we have implemented an application in an iPod Touch. The iPod Touch is equipped with a built-in 3-axis accelerometer, CPU, and WiFi communication chip. The application continuously captures the acceleration data, compresses it, and uploads it to the server until the battery runs out. We chose the traditional ZIP compression algorithm, measured the battery duration, and compared it to the CS case.

Using the CS method, we determined $N = 400$ and chose a 400×200 sparse random matrix as the sampling matrix $\mathbf{\Phi}$. The acceleration data are sampled at 100Hz. At intervals of 4 s, the sequence of acceleration data is compressed by multiplying $\mathbf{\Phi}$ and stored in the local flash storage. The stored data are then sent to the server at intervals of 30 s by the HTTP POST method.

In the ZIP compression method, the acceleration data are compressed to a ZIP file by using a ziparchive library and uploaded by the same method. The compression ratio of ZIP changed according to the entropy of the signal. When the acceleration data of walking signals were compressed, the average compression ratio of ZIP was 0.43. Since the compression ratio of CS was set to $0.5 (= M/N)$, we set the sending interval to 35 s in the ZIP compression method so that the total amount of the uploaded data was same for both schemes.

A. Results

Fig. 8 shows the remaining battery power as time elapses. The battery duration was approximately 440 min and 370 min for the CS and ZIP systems, respectively. This implies that if CS is applied, power consumption is reduced by approximately 16% as compared to the ZIP compression scheme. Since the amount of uploaded data was the same in both cases, the result clearly indicates that the CS method is computationally lighter than the ZIP.

Owing to the fact that the CS compression ratio fully depends on M/N, if we can reconstruct the data that are sufficient to recognize human activities by using an M value smaller than that used in this experiment $(= N/2)$, power consumption will be further reduced.

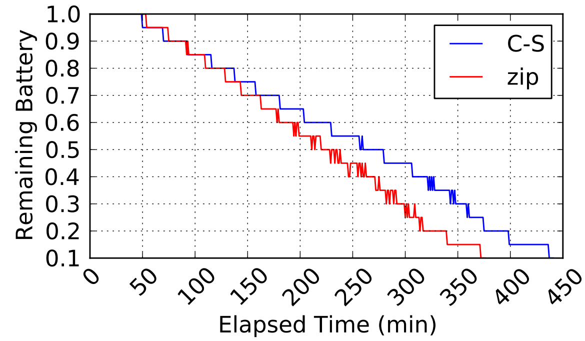

Fig. 8. Battery Duration for the Two Compression Method (CS and ZIP)

V. Conclusion

In this paper, a lightweight data compression system using CS is proposed for human activity sensing. This scheme is suitable for mobile-phone-based sensing systems because it can significantly reduce the acceleration data without heavy computation costs at the mobile terminals.

The evaluation results revealed that the proposed scheme could reduce power consumption by 16% as compared to the traditional ZIP compression method with no duty cycling, while keeping the reconstruction error to less than 10% and the recognition accuracy of the 6 basic activities over 70%.

Acknowledgment

This work was supported by a Grant-in-Aid for Scientific Research (A)(23240014).

References

[1] N.D. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury, and A.T. Campbell, “A survey of mobile phone sensing,” Communications Magazine, IEEE, vol. 48, no. 9, pp. 140–150, September 2010.

[2] S. Reddy, M. Mun, J. Burke, D. Estrin, M. Hansen, and M. Srivastava, “Using Mobile Phones to Determine Transportation Modes,” ACM Transactions on Sensor Networks (TOSN), Vol. 6, No. 2, pp. 1–27, February 2010.

[3] J. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy, and M. B. Srivastava, “Participatory sensing,” World Sensor Web Workshop, ACM Sensys’06, Boulder, Colorado, October 31, 2006.

[4] S. Yang and M. Gerla: “Energy-Efficient Accelerometer Data Transfer for Human Body Movement Studies,” IEEE SUTC, pp.304–311, California, USA, July 2010.

[5] E. J. Candés and M. B. Wakin, “An Introduction To Compressive Sampling,” IEEE, Signal Processing Magazine, Vol. 25, No.2, pp.21-30, March 2008.

[6] R. Berinde and P. Indyk, “Sparse recovery using sparse random matrices,h MIT-CSAIL TR-2008-001, 2008, Technical Report.

[7] Y. Kawahara, N. Ryu, T. Asami, “Monitoring Daily Energy Expenditure using a 3-Axis Accelerometer with a Low-Power Microprocessor,” e-Minds: International Journal on Human-Computer Interaction, Vol. 1, No. 5, pp.145–154 (2009)

[8] N. Kawaguchi *et al.*, “HASC Challenge: Gathering Large Scale Human Activity Corpus for the Real-World Activity Understandings,” Proceeding of AH’11, Article No.: 27, March 2011.

[9] J. R. Kwapisz, G. M Weiss, and S. A. Moore, “Activity Recognition using Cell Phone Accelerometers,” ACM SIGKDD Explor. Newsl. Vol. 12, No. 2, pp.74–82, March 2011.

[10] Available from http://www.cs.waikato.ac.nz/ml/weka/