

# A Novel Acceleration Data Compression Scheme for Wireless Sensor Network Application in Fresh Food Tracking System

Jun Chen Zhibo Pang Zhi Zhang Jie Gao Qiang Chen Lirong Zheng

iPack Vinn Excellence Center, School of Information and Communication Technology  
Royal Institute of Technology (KTH), Forum 120, 164 40 Kista-Stockholm, Sweden  
{chenj, qiangch, lirong}@kth.se

**Abstract** – An enhanced dual-layer wide area wireless sensor network (WSN) system was developed and applied in fresh food tracking application [1]. The sensor nodes of WSN provide a set of useful measurements about the environmental conditions during the fresh food transportation. Since the 3-axis acceleration data in such application account for more than 99% out of all collected source data, we propose a novel acceleration data compression scheme for this specific application, to overcome the resource and cost limitation imposed on embedded sensing nodes and wireless communication links in WSN. According to general fresh food transport behavior pattern, the acceleration data can be extracted into three components (D waveform, A waveform and P waveform). Each component could be compressed separately according to its distinct characteristic and then packetized as an encoded frame. In this paper we examine the experimental acceleration dataset from a field test (fresh melon fruit transport from Brazil to Sweden in 20 days) as a test case. Experimental results show that a high compression ratio (around 10:1) with the acceptable distortion could be achieved, confirming the effectiveness of the compression scheme proposed in this work.

**Keywords** – wireless sensor network, fresh food tracking, acceleration data, source data compression.

## I. INTRODUCTION

There are many feasible applications for wireless sensor network (WSN) [1],[2] such as industrial sensor networks, health monitoring, home automation, logistic data acquisition, etc. We developed an enhanced dual-layer dual-directional wide area wireless sensor network of a WAN-SAN coherent structure and applied it in fresh food tracking system [1]. According to a report [3], about 10% of the fresh fruits and vegetables coming from different parts of the world into the European market are damaged during the transportation process. The main causes of fresh food damage are abnormal temperature and humidity, microbial infections, composition of storage atmosphere and mechanical damage [4].

Mechanical damage of the fresh fruits and vegetables is one of the most common and severe defects. In the fresh food tracking system, a 3-axis accelerometer ( $\pm 8G$ , 10-bit output) [5] was used to measure the instantaneous acceleration (X-, Y-, and Z- axis) when the food product is experiencing mechanical impacts such as dropping, shaking, vibration, and sudden break. The acceleration data are measured with a sampling rate of 20Hz to monitor

most of the possible events that may cause mechanical damage to fresh fruits and vegetables. This makes 3-axis acceleration the densest data source with extreme amount of data in this application. It is highly desirable to have a highly efficient compression scheme for acceleration data for resource and cost saving.

It is also generally necessary to have data compression in embedded sensor nodes and networks because of limited cost budget and available resources, such as power source, memory space, processing capability, and communication bandwidth, etc, [6],[7],[8].

Commonly used data compression algorithms, such as bzip2 [9], compress [10], and gzip [11], were developed for desktops or servers. These algorithms mainly focus on the compression ratio with little attention paid to the resource availability. They are apparently not possible to use for embedded applications.

Some papers [7],[12] adopted two dimensional discrete cosine transform (2D-DCT) based method to compress the sensing data. The 2D-DCT would transform two dimensional (2D) data array from the spatial domain to the frequency domain and extract significant values. 2D-DCT based compression methods are widely used in video sensor networks [12], or in some kind of applications [7] the sensing data could be organized as multi-dimensional array to explore the frequency domain redundancy. But these methods are not suitable for our application. The events which may lead to mechanical damage are undetermined, so the frequency spectrum cannot be classified. In addition, 2D-DCT would also incur high computation cost.

Differential pulse code modulation (DPCM) based compression approach was also utilized in sensing data compression [13]. But in our case, the acceleration data didn't show significant temporal correlation. DPCM based compression would introduce extra accumulative error.

In this paper, we propose a novel lossy acceleration data compression scheme for the fresh food tracking system, to achieve the maximum utilization out of limited available resource. The acceleration data are extracted into three components (D waveform, A waveform and P waveform) rather than dealing with the source data directly. Then each component could be compressed separately according to its characteristic, and then packetized as an encoded frame.

The field test data (transportation of melon fruits from Brazil to Sweden in 20 days) are used as an experimental data source. Experimental results show that the proposed

scheme could achieve a high compression ratio (around 10:1) with the acceptable distortion.

In section II the WSN platform and the source data collection in the fresh food tracking application are described. Detailed design and analysis of the lossy acceleration data compression scheme are presented in section III. Experimental results from the field test data are shown in section IV. The conclusions are given in section V.

## II. WSN PLATFORM

### A. WAN-SAN architecture

The WSN system used in this paper is based on a dual-layer network topology: a Wide Area Network (WAN) layer comprising a Central Server and an amount of Master Sensor Nodes (MSN), and a Sensor Area Network (SAN) layer comprising a MSN and an amount of Slave Sensor Nodes (SSN) (Fig.1). The wireless WAN layer could use all kinds of commercially available infrastructure with wide coverage and data services. Local memory devices on MSN are also available to store the sensing data collected from SSNs.

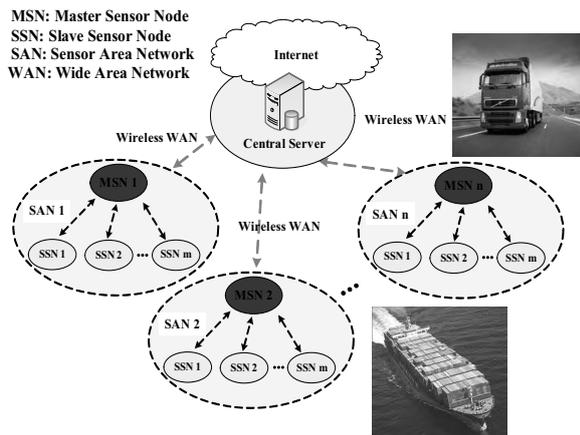


Fig.1. WAN-SAN architecture

The MSN acts as not only the coordinator in the SAN layer but also an independent sensor node in the WAN layer. In the SAN layer SSNs sense data and send data to MSN by RFID technology. In fresh food tracking application, SSNs measuring temperature, humidity and acceleration could be placed either on carton level or on pellet level.

Compact MSN hardware modules are manufactured, including an MSN sensor board which can measure temperature, relative humidity, different gases, and 3-axis acceleration data (Fig.2).

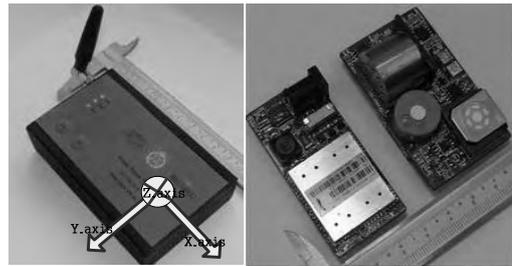


Fig. 2. Manufactured MSN hardware modules

### B. Source data collection in the fresh food tracking system

The MSN node of WSN would provide on container/cargo level an essential set of useful measurements, including temperature, relative humidity, concentrations of CO<sub>2</sub>, O<sub>2</sub> and ethylene, and 3-axis acceleration data, to monitor the transport conditions of fresh foods.

Table 1. Collected source data in one day.

	Temperature & Humidity	Gas Concentrations	Acceleration Data
Data Size	2.88KB	4.32KB	6.48MB
Percentage	0.044%	0.067%	99.89%

Among these source data, 3-axis acceleration data which reach up to 600bps or 6.48MB per day are the densest data (Table 1), where temperature/relative humidity and the gas concentrations are sampled every 2 minutes and acceleration data are sampled at 20Hz. The 3-axis acceleration data account for 99.89% out of all collected data. Obviously, the compression of 3-axis acceleration data is the crucial part in data reduction for this specific application.

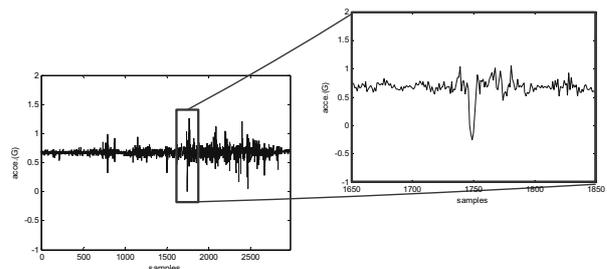


Fig. 3. Acceleration dataset (first 3000 samples of X-axis) from container truck with sampling rate of 20Hz

A typical acceleration dataset is used to analyze the proposed compression scheme in section III. The dataset was collected from a running container truck in 8 Hours, with totally 574,340 samples. Since the characteristics of the data on each axis (X-, Y-, and Z-axis) are quite similar, we pick the data of X-axis (with more vibrations) for analysis. The typical acceleration data are the superimposition of a minor vibration waveform and an occasional short-term high-amplitude pulse waveform (Fig.3).

### III. COMPRESSION SCHEME DESIGN AND ANALYSIS

#### A. Compression scheme design

Instead of dealing with the source acceleration data directly, D waveform, A waveform and P waveform are extracted from the original acceleration data as shown in Fig.4.

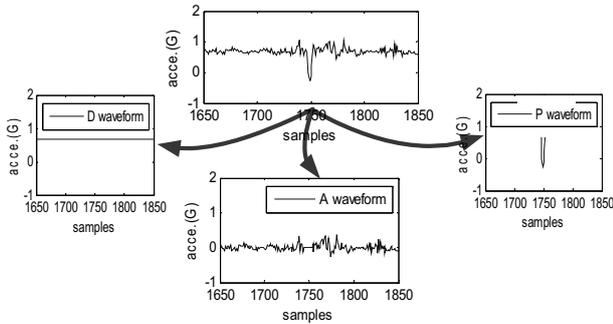


Fig.4. Decomposition of the source acceleration data

D waveform corresponds to the DC (direct current) component of the original signal, A waveform indicates the AC (alternating current) component of the original signal while P waveform represents the high-amplitude and short-term pulses of the original signal.

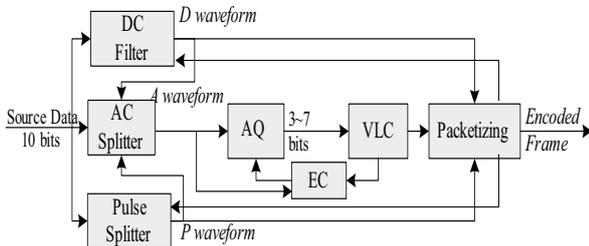


Fig.5. Encoder block schematic

The decomposition makes it possible that each component could be compressed separately according to its distinct characteristic and then packetized as an encoded frame (Fig.5). We could exploit each component's unique feature to make the compression more efficiently for a better overall performance.

The high-amplitude and short-term pulses (P waveform) rarely happen during fresh food transport (given a proper pulse level at  $0.75S_G$  in equation (1)). That's why we do not compress the P waveform, and instead keep its source data in the final encoded frames. And DC is the constant value in a frame which could be predicted from the previous frame. We could then only need to focus on the compression of A waveform.

Adaptive quantizer (AQ) is utilized to compress A waveform efficiently. The proposed scheme fixes

quantization step due to the accuracy of the accelerometer [5], and varies the quantizer bit size of the output sample between encoded frames.

Error Controller (EC) would inform AQ whether the distortion of the current encoded frame is beyond the preset level or not. According to this information, AQ decides the  $Q\_bits$  (quantizer bit size of an output sample) for the next frame. That is, AQ would increase  $Q\_bits$ , decrease  $Q\_bits$ , or leave  $Q\_bits$  unchanged to achieve the trade-off between compression ratio and distortion

A losses variable length coding (VLC) based compression method, is used to achieve further compression performance in the proposed scheme.

#### B. Compression performance criteria

In lossy data compression, the reconstructed data are not exactly the same as the original data. A distortion measure is a mathematical entity which specifies exactly how close the approximation is.

In this paper, mean magnitude of relative error (MMRE) is used as a criterion for distortion measurement which allows us to get an idea of the error introduced by the compression process, independently of the frame size (the amount of source data samples encoded in a single encoded frame).

$$MMRE = \frac{1}{N} \sum_{i=1}^N \frac{|e_i|}{|S_G|} \quad (1)$$

Where  $e_i$  corresponds to the error between source data and reconstructed data and  $S_G$  is the acceleration data measured by the accelerometer when a standard gravity applies.

Compression ratio (CR) is a commonly used criterion to quantify the reduction in data-representation size produced by compression process. The CR of one single frame is calculated by:

$$CR = \frac{S_{FH} + S_D + S_P + S_A}{B_S \times F_S} \quad (2)$$

Where  $S_{FH}, S_D, S_P, S_A$  correspond to the size of encoded frame header, the size of encoded D waveform, the size of encoded P waveform and the size of encoded A waveform (in bits), respectively. And  $B_S$  is the bit size of a source data sample while  $F_S$  is the frame size.

#### C. Compression performance analysis

We evaluate here the compression performance with different quantizer strategy: with fixed 3 bits, with fixed 5 bits, and with AQ of  $Q\_bits$  ranging from 3 bits to 7 bits (Fig.6). All the quantizer bit size mentioned above are including 1 sign bit. The average CR of AQ strategy is

around 13% higher than fixed 5 bits strategy while the average MMRE is slightly worse. Meanwhile, the distortion of fixed 3 bits strategy is unacceptable (MMRE>0.04), although the CR of fixed 3 bits strategy is higher. That is, our proposed AQ based compression scheme could achieve the optimal trade-off between compression ratio and distortion.

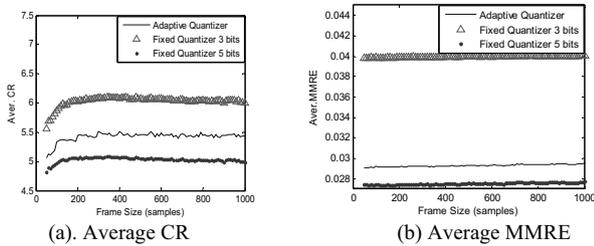


Fig.6. Compression performance on different quantizer strategy.

Another conclusion we can get from the evaluation is that the average CR is not correlated with the frame size when the frame size is greater than a proper value (around 200 samples) (Fig.6a). As aforementioned, the CR is mainly upon the performance of A waveform compression. When the amount of samples in an encoded frame is greater than some proper value, the average compression ratio of A waveform (AQ and VLC based approach) would not increase as the amount of the input source data samples increases.

#### IV. EXPERIMENTAL RESULTS

Field test was done for fresh food tracking system of WSN to monitor the environmental conditions inside containers during the melon transport from fruit farm in Brazil to Swedish supermarket (Fig.7).

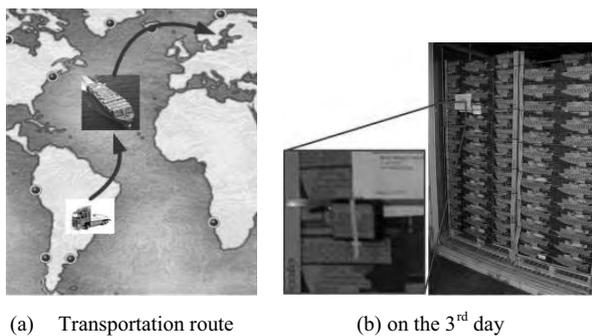


Fig.7. fresh melon fruit transport from Brazil to Sweden

The trip from Brazil to Sweden lasted for 20 days. There are approximately 33,082,000 acceleration data sampled per axis. We use this field test data to examine the effectiveness of the proposed compression scheme (Table.2).

Table 2 Compression performance of overall dataset

	X-axis	Y-axis	Z-axis	Overall
Aver. CR	10.26:1	11.53:1	10.44:1	10.74:1
Aver. MMRE	0.023	0.021	0.022	0.022

(with 20 days, totally 33,082,000 samples per axis and a frame size of 512)

As discussed above, the compression ratio is not correlated with the frame size when frame size is larger than 200 or so. This feature is important for our current system since the MSP430 micro controller [14] of the MSN node is of only 10 KB RAM. For the sake of the overall system performance, we fix the frame size as 512 samples.

We assess here the proposed compression scheme with more details using the dataset on the 3<sup>rd</sup> day (in container truck, MSN sensor node attached as shown in Fig.7b). As shown in Fig.8, the reconstructed signal is smoother than the original signal because lossy quantization smoothes out the small vibrations. And Q\_bits tends to increase when there are higher amplitude vibrations (Fig.9).

The CR of the whole dataset is higher than we explained in section III because there is less vibration in container ship. Our compression approach would achieve higher compression ratio for the data with less vibration (Table 3). And the sensor node is attached to the package with different angles for test during the whole trip, that's the reason why the compression performances of the 3 axes are almost the same (Table 2).

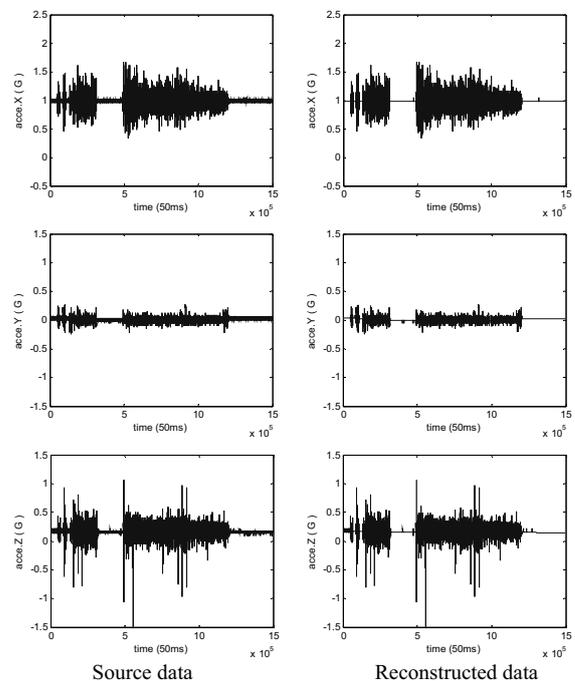


Fig.8. Source data and reconstructed data (X-axis, Y-axis and Z-axis) of the 3<sup>rd</sup> day dataset

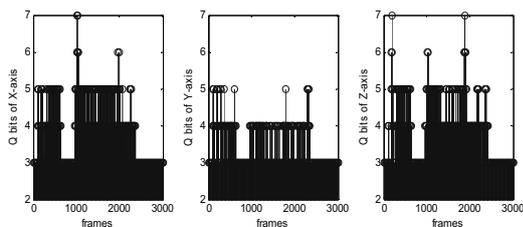


Fig.9. Q\_bits (X-axis, Y-axis and Z-axis) on the encoding frames of the 3<sup>rd</sup> day dataset

Table 3 Compression performance of the 3<sup>rd</sup> day dataset

	X-axis	Y-axis	Z-axis	Overall
Aver. CR	6.41:1	10.92:1	5.96:1	7.76:1
Aver. MMRE	0.029	0.024	0.030	0.028

(with 1 day, totally 1,636,000 samples per axis and a frame size of 512)

The CR of the 3<sup>rd</sup> day’s dataset is higher in Y-axis since it got less vibration than the other two axes.

V. CONCLUSIONS

With the emergence of various WSN applications, the application-specific source data compression approaches are often needed for resource and cost saving. An enhanced dual-layer dual-directional wide area wireless sensor network system is applied in fresh food tracking application. An essential set of useful measurements is provided to monitor the environmental conditions during fresh food (such as fruits and vegetables) transport. Since 3-axis acceleration data account for 99.89% data amount out of all collected data, we propose in this work a novel lossy acceleration compression scheme. The source acceleration data are extracted into three components (D waveform, A waveform and P waveform), then each component is compressed separately according to its distinct characteristic and then packetized as an encoded frame.

Field test data (fruit transport from Brazil to Sweden) are used for acceleration data compression experiment. Experimental results confirm that the compression scheme proposed in this work could achieve a high compression ratio (around 10:1) with the acceptable distortion.

ACKNOWLEDGMENT

This work was financially supported by Vinnova (The Swedish Governmental Agency for Innovation Systems) through the Vinn Excellence centers program at the Royal Institute of Technology (KTH).

REFERENCES

- [1] Zhibo Pang , Jun Chen, David Sarmiento M. , etc. “Mobile Wide Area Deployable Sensor System for Networked Services” IEEE Sensors 2009 Conference, October 2009, New Zealand
- [2] Gajbhiye. P. and Mahajan. A. “A Survey of Architecture and Node deployment in Wireless Sensor Network” Int’l. Conf. the Applications of Digital Information and Web Technologies 2008 (ICADIWT 2008), pp.426-430, Aug. 2008
- [3] Billerud AB. The Billerud book of maths. <http://www.billerund.com/>
- [4] Domingo Martinez-Romero, Maria Serrano, Angel Carbonell, etc. “Mechanical Damage During Fruit Post-Harvest Handling: Technical and Physiological Implications” Production Practices and Quality Assessment of Food Crops, Vol. 3, pp. 233 - 252. Springer Netherlands Publisher 2004
- [5] Freescale Semiconductor, ±2g/±4g/±8g Three Axis Low-g Digital Output Accelerometer , <http://www.freescale.com/>
- [6] Kimura. N. and Latifi. S. “A survey on data compression in wireless sensor networks” Proc. Int’l. Conf. Information Technology: Coding and Computing 2005 (ITCC2005) vol 2, pp.8 - 13 Vol. 2. Apr. 2005
- [7] You-Chiun Wang; Yao-Yu Hsieh; Yu-Chee Tseng; “Multiresolution Spatial and Temporal Coding in a Wireless Sensor Network for Long-Term Monitoring Applications” IEEE Trans. Computers, vol. 58, no. 6, pp. 827 - 838, Jun. 2009,
- [8] J. Haupt, W.U. Bajwa, M. Rabbat, and R. Nowak, “Compressed Sensing for Networked Data,” IEEE Signal Processing Magazine, vol. 25, no. 2, pp. 92-101, Mar. 2008.
- [9] <http://www.bzip.org/>
- [10] <http://ncompress.sourceforge.net/>
- [11] <http://www.gzip.org/>
- [12] Wu-chi Feng, Brian Code, Ed Kaiser, etc. “Panoptes: scalable low-power video sensor networking technologies” ACM Trans. Multimedia Computing and Applications, vol. 1, no. 2, pp. 151-167, May 2005.
- [13] Ben Greenstein, Christopher Mar, Alex Pesterev, etc. “Capturing high-frequency phenomena using a bandwidth-limited sensor network” Proc. ACM Int’l Conf. Embedded Networked Sensor Systems (SenSys’06) 2006
- [14] Texas Instruments. MSP430x161x Mixed Signal Microcontroller Datasheet. <http://www.ti.com/>