



# Muscles data compression in body sensor network using the principal component analysis in wavelet domain

Elmira Yekani Khoei<sup>1</sup>, Reza Hassannejad<sup>2\*</sup>, Behzad Mozaffari Tazehkand<sup>3</sup>

<sup>1</sup> Faculty of Computer, College of Engineering, East Azerbaijan Science and Research Branch, Islamic Azad University, Tabriz, Iran

<sup>2</sup> Faculty of Mechanical Engineering, University of Tabriz, Tabriz, Iran

<sup>3</sup> Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran

## Article Info



**Article Type:**  
Research Article

## Article History:

Received: 02 June 2014  
 Revised: 01 Nov. 2014  
 Accepted: 05 Nov. 2014  
 ePublished: 20 Feb. 2015

## Keywords:

Body sensor network  
 Data compression  
 Prediction  
 Principal component analysis  
 Wavelet transform

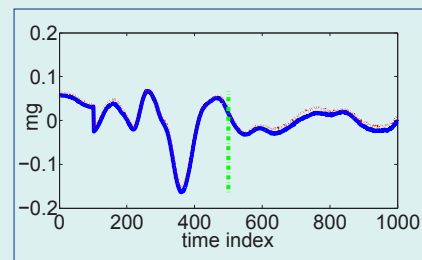
## Abstract

**Introduction:** Body sensor network is a key technology that is used for supervising the physiological information from a long distance that enables physicians to predict and diagnose effectively the different conditions. These networks include small sensors with the ability of sensing where there are some limitations in calculating and energy.

**Methods:** In the present research, a new compression method based on the analysis of principal components and wavelet transform is used to increase the coherence. In the present method, the first analysis of the main principles is to find the principal components of the data in order to increase the coherence for increasing the similarity between the data and compression rate. Then, according to the ability of wavelet transform, data are decomposed to different scales. In restoration process of data only special parts are restored and some parts of the data that include noise are omitted. By noise omission, the quality of the sent data increases and good compression could be obtained.

**Results:** Pilates practices were executed among twelve patients with various dysfunctions. The results showed 0.7210, 0.8898, 0.6548, 0.6765, 0.6009, 0.7435, 0.7651, 0.7623, 0.7736, 0.8596, 0.8856 and 0.7102 compression ratios in proposed method and 0.8256, 0.9315, 0.9340, 0.9509, 0.8998, 0.9556, 0.9732, 0.9580, 0.8046, 0.9448, 0.9573 and 0.9440 compression ratios in previous method (Tseng algorithm).

**Conclusion:** Comparing compression rates and prediction errors with the available results show the exactness of the proposed method.



## Introduction

The networks consist of individual sensors that are able to interact with human body. These sensors could operate through exchanging information, which is impossible by a single sensor. The application systems receive the results and send the necessary feedbacks to physicians. In most cases, application systems need real-time and high-resolution data. These demand shorter transference delay and higher sampling rate. The main problem is the limitation of band width of wireless channel and contention and collision among the data sent from sensors to the base station. In order to solve these problems, data compression technique is used to reduce the transfer rate. A QoS framework similar to IEEE 802.16 has been

suggested by Zhou et al for body sensor networks through which band width could be allocated to the sensor nodes.<sup>1</sup> Jea et al prepared a method in which sensors with higher priority are allowed to transfer their data with high exactness.<sup>2</sup> Baheti et al analyzed the scattered data resources and declared that if data sources are scattered, compressed sensing technique would be used to reduce the necessary sampling rate.<sup>3</sup> Chang et al studied the function of some compression lossy and lossless algorithm on motions that exhibit repetitive patterns such as running.<sup>4</sup> In other research, Zarei et al presented the compression method using principal component analysis for optimization of energy consumption in body sensor networks in which compression is carried



\*Corresponding authors: Reza Hassannejad, Email: [hassannejad@tabrizu.ac.ir](mailto:hassannejad@tabrizu.ac.ir)



© 2015 The Author(s). This work is published by BioImpacts as an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by-nc/4.0/>). Non-commercial uses of the work are permitted, provided the original work is properly cited.

out in two stages.<sup>5</sup> Compression method based on the theory of multiple principal component analysis theory for compression of data in wireless sensor network has been presented by Chen et al that principal component analysis theory has been used in different layers to omit the coherence between the measured raw data.<sup>6</sup> For slow changing environments, model-driven data acquisition models have been suggested by Deshpande et al and Silberstein et al.<sup>7,8</sup> However, the mentioned compression methods only have considered data compression by one sensor. By considering this correlation among the data received by different sensors, compression method based on overhearing and using temporal-spatial correlation has been suggested by Li et al and Tseng et al for wireless sensor networks and body sensor network.<sup>9,10</sup>

In this research, Tseng algorithm<sup>10</sup> is developed through principal component analysis and wavelet transform. The principal component analysis theory is used for increasing the correlation among the received data from different sensor nodes and reduction of sent data to base station and optimum use of the band width. Wavelet transform is used for the reduction of noise and collision among received raw data from sensors. Here, the sensors are used to collect the motional data to better diagnosis and treatment.

## Materials and methods

### Tseng algorithm

Tseng et al<sup>10</sup> represented a method for data compression using temporal and spatial correlation in 2011 in which temporal and spatial correlation was used to optimum use of band width of connective wireless channel restrictions and reduction of conflict among the sensors transfer.<sup>10</sup> In the present research, a network of body wireless sensor network with n sensor nodes has been considered in which every sensor node is equipped with m axis. Considering that in the present research, tri-axial accelerometers have been used, m is regarded 3. The mentioned algorithm has two offline and online phases. In offline phase, received data from the sensors in body sensor network are collected and sent for processing to a fusion center analyser. The main aim of the present phase is to achieve suitable order of the sensor node prediction. Block diagram of Fig. 1 shows the details of offline phase.

Online phase: the main aim of this phase is data compression according to produced compression tree in offline stage. There are two cases for each sensor. 1.  $V_i$  node which is the first level sensor node. 2.  $V_i$  node which is not first level sensor node. The stages shown in (Fig. 2) and (Fig. 3) are performed for case 1 and 2.

Since physiological signals studied in physical engineering

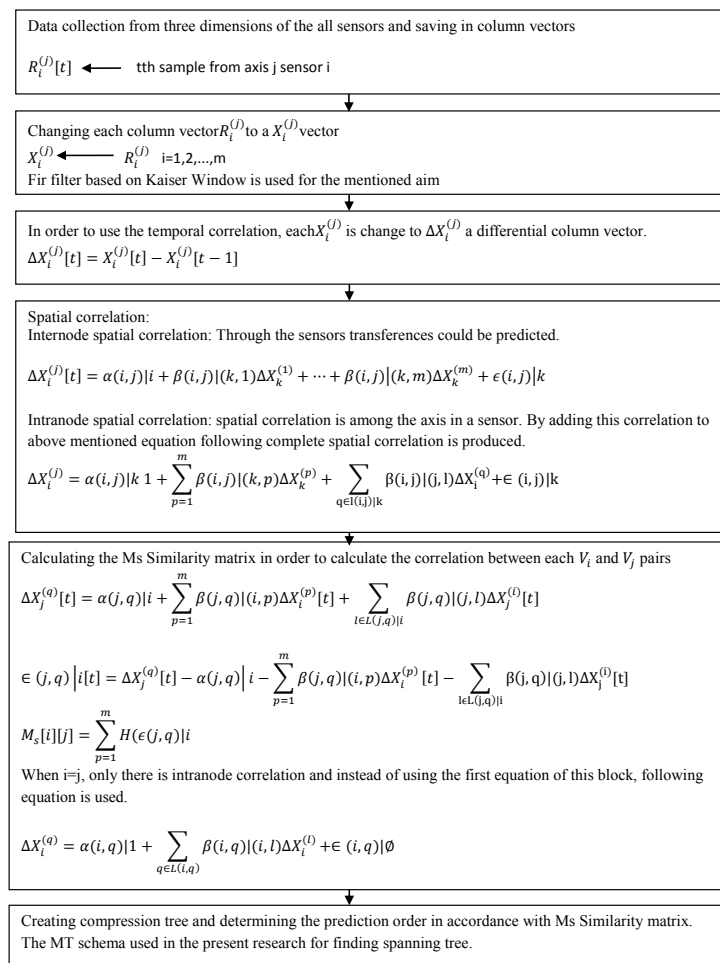


Fig. 1. Offline phase of Tseng algorithm.

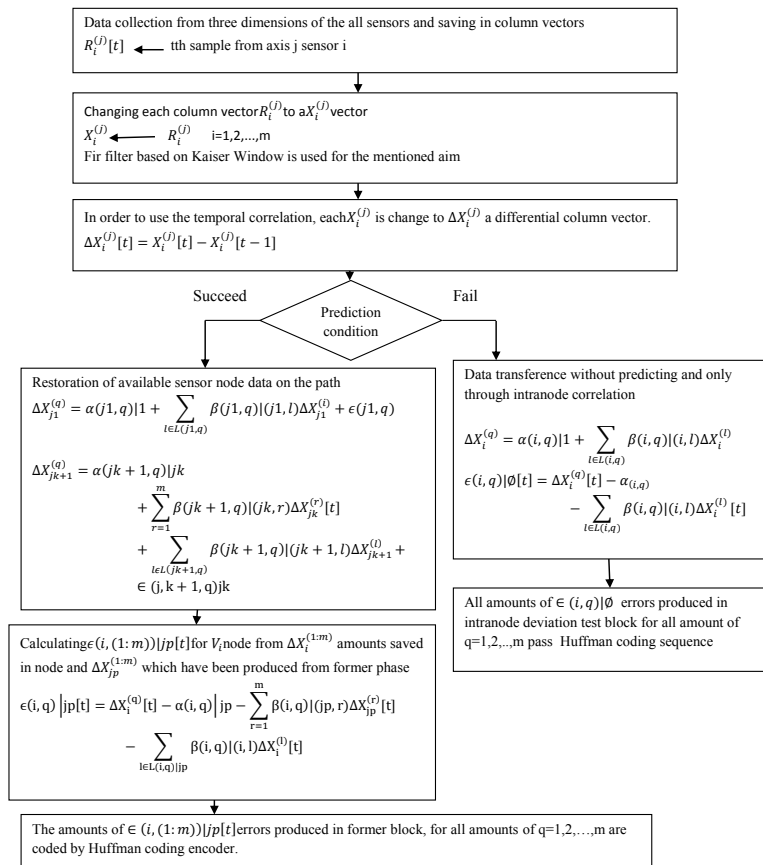


Fig. 2. Online phase of Tseng algorithm:  $V_i$  is a level-1 node.

are often weak, they are always exposed to noise and interference. In this paper, in order to increase compression rate, two methods of signal processing as principal component analysis and wavelet transform have been used. Principal component analysis is used for separating blind signals and increasing correlation among the received data from sensors. In addition, wavelet transform is also used to reduce noise and interference, increasing the compression rate and reducing the error between the real and predicted amounts. As a result, prediction error between the real amount and prediction amount decreases and compression rate increases in comparison with Tseng method.

**Proposed algorithm**

Proposed algorithm diagram block is depicted in Fig. 4 and Fig. 5. As it is shown in Fig. 4 and Fig. 5, after data collection, principal component analysis is performed on the data. Then, the obtained result of this stage is sent to online and offline stages with applying wavelet transform (as Figs 1, 2, and 3).

**Principal components analysis**

In principal component analysis stage shown in Fig. 4 and 5, the raw data that have been collected and processed in data collection stage, would be saved in  $R_i^{(1:m)}$  column vectors for each  $i=1,2,\dots,n$ . In this analysis, processed data

should be normalized and their average should be zero. Thus, according to equation I the average of all numbers of column vector  $R_i^{(q)}$  would be calculated for each  $i=1,2,\dots,n$

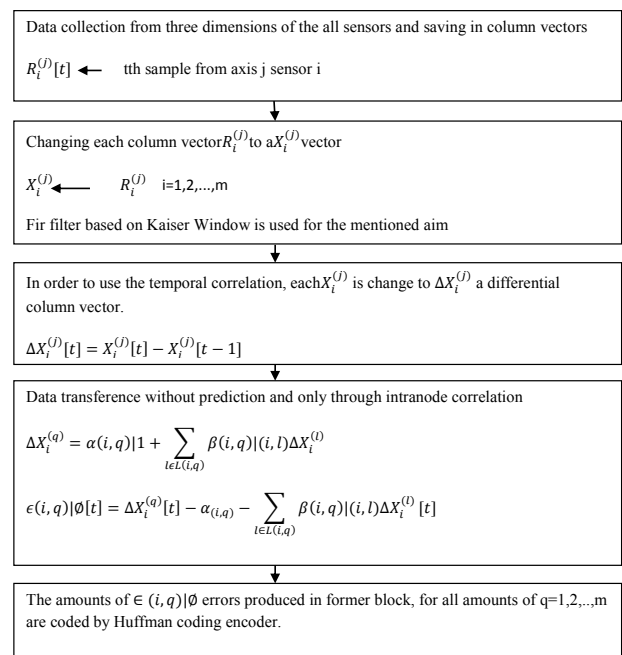


Fig. 3. Online phase of Tseng algorithm:  $V_i$  is non-level-1 node.

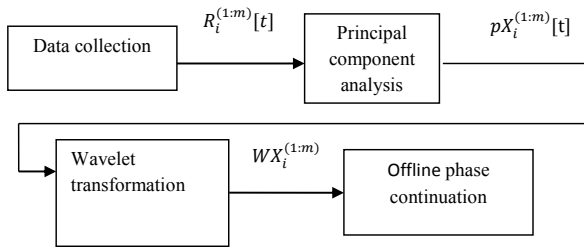


Fig. 4. Offline phase of the proposed algorithm.

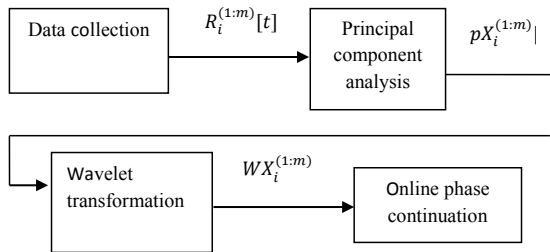


Fig. 5. Online phase of the proposed algorithm.

and  $q=1,2,\dots,m$ . This obtained average is subtracted from all numbers on the vector ( $R_i^{(q)}$ ) and it is saved in  $MR_i^{(q)}$  column vector, as follows:

$$mean = \frac{1}{n} \sum_{i=1}^n R_i^{(q)} \tag{Eq. (I)}$$

$$MR_i^{(q)}[t] = R_i^{(q)}[t] - mean$$

In this equation,  $n$  displays the number of digits on column vector,  $i$  displays the number of considered node and  $q$  displays the studied node's axis.

All the amounts (in all three dimensions of the node) in this analysis are placed in  $X$  group. Covariance matrix is calculated through equation II and through singular value decomposition would be factorized to eigenvalues and eigenvectors:

$$\begin{aligned} X &= [MR_i^1, MR_i^2, \dots, MR_i^q]^T \\ C_{xx} &= E\{XX^T\} = EDE^T \in R^{m \times m} \\ D &= diag[\lambda_1, \lambda_2, \dots, \lambda_m] \in R^{m \times m} \\ E &= [e_1, e_2, \dots, e_m] \in R^{m \times m} \end{aligned} \tag{Eq. (II)}$$

In equation II, matrix  $D$  is a diagonal matrix and  $\lambda_i$  is the eigenvalue of covariance matrix  $C_{xx}$  and  $e_i$  is the eigenvector corresponding to eigenvalue  $\lambda_i$ . Columns of matrix  $E$  is formed by eigenvectors.  $E$  is an orthogonal matrix.  $k_i$  Vectors are orthogonal, thus:

$$\begin{aligned} e_i \cdot e_j^T &= \delta_{ij} \\ \delta_{ij} &= \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \end{aligned} \tag{Eq. (III)}$$

Where,  $\delta_{ij}$  is unit impulse response.

According to equation IV, the matrix  $E$  of eigenvectors are multiplied to  $R_i^{(q)}$  and the data is saved in  $PX_i^{(q)}$ :

$$PX_i^{(q)} = E * R_i^{(q)} \tag{Eq. (IV)}$$

The theory of principal component analysis increases the correlation between the sensors. Therefore, it causes an increase in compression rate. According to equation IV, eigenvector ( $e_i$ ) forms a space with  $m$  dimensions. Therefore, a second space is produced that is dimensionally smaller. Then, matrix of  $E_s$  is defined as follows. The principal component analysis is defined as equation V:

$$\begin{aligned} E_s &= [e_1, e_2, \dots, e_n]^T \in R^{m \times n} \\ Y_s &= E_s^T X \in R^{m \times n} \end{aligned} \tag{Eq. (V)}$$

**Wavelet transform**

Fourier transform is extracting oscillatory components of frequency signals in time domain. The aim of wavelet transform is calculating components of signal that there are locally, like sudden changes of a noisy signal in a limited time frame. Wavelet transform could be stated as a number of mother wavelet transform functions. Using acts of contraction, expansion and translation in time domain, a family of wavelets based on mother wavelet could be created. Wavelet transform of one signal such as  $x(t)$  could be written as equation VI. In the present equation, function  $\Psi(t)$  is the mother wavelet and plays the role of  $e^{j\omega t}$  function in Fourier transform:

$$T_x(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t) \psi^* \left( \frac{t-b}{a} \right) dt \tag{Eq. (VI)}$$

In order to reduce the complexity, the mentioned transform is calculated in discrete form for each amount of  $a$  and  $b$ . Discrete wavelet transform is written according to equation VII:

$$\begin{aligned} \psi_{j,k}(t) &= 2^{\frac{j}{2}} \psi(2^j t - k) \\ X(t) &= \sum_{k=-\infty}^{+\infty} C_k \varphi(t-k) + \sum_{k=-\infty}^{+\infty} \sum_{j=0}^{+\infty} d_{j,k} \psi_{j,k}(t) \end{aligned} \tag{Eq. (VII)}$$

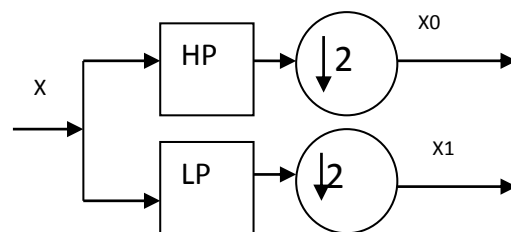


Fig. 6. Decomposition of wavelet signal in one level.

Where,  $\varphi(t)$  is called scale function or father wavelet and  $C_k$  and  $d_{i,k}$  parameters are considered as approximation coefficient and detail respectively that could be calculated as a set of low-pass and high-pass filters. These filters could be calculated according to functions of mother and father wavelet like equation VIII:

$$\begin{aligned}\varphi(t) &= \sum_k h(k) \sqrt{2} \varphi(2t-k) \\ \psi(t) &= \sum_k h(k) \sqrt{2} \psi(2t-k)\end{aligned}\quad \text{Eq. (VIII)}$$

Where  $h_0(k)$  and  $h_1(k)$  are the impulse responses of high-pass and low-pass filter (Fig. 6). depicts decomposition in one level.

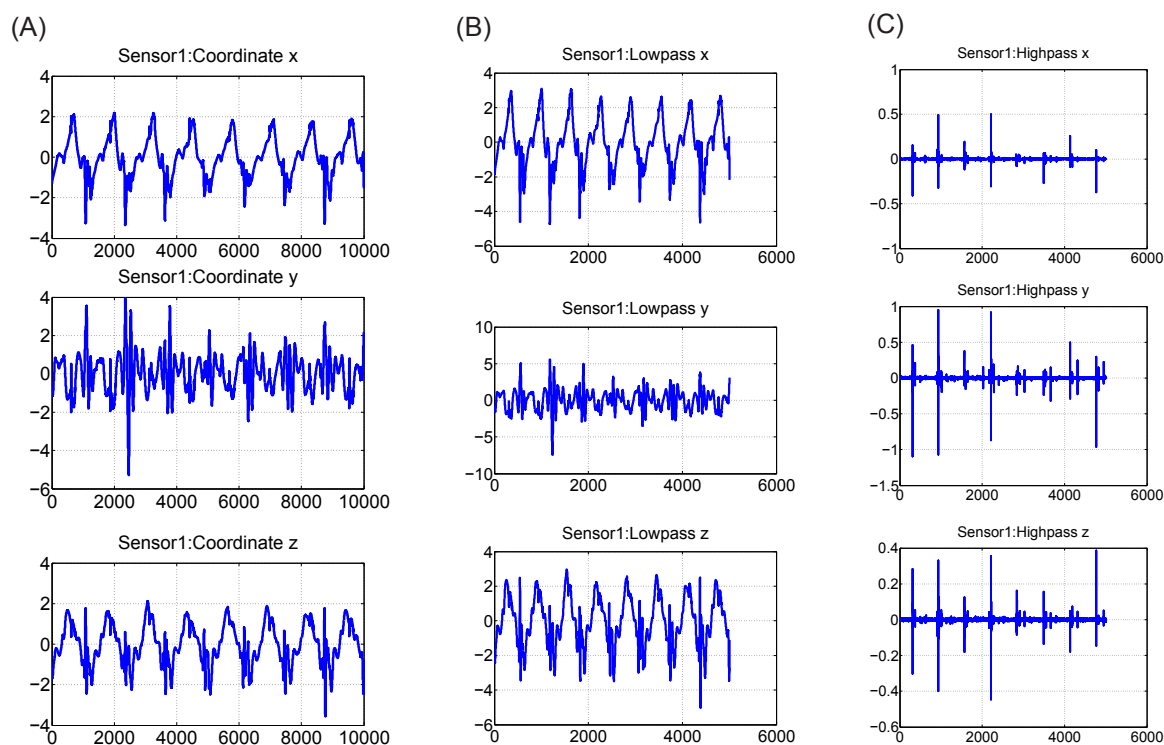
In the wavelet transform block (Figs. 4 and 5) removing the noise from received data of the sensors are eliminated. Since these data have noise and interference, in the proposed method Dabuchi 4 wavelet transform is used for removing the noise, signal compression and reduction of the interferences. Dabuchi 4 wavelet is suitable for multi resolution analysis of wavelet; moreover orthogonality is the other advantage of this wavelet. Some wavelets are without orthogonal property and there are problems of signal energy reduction and frequency leakage. Transformed  $PX_i^{(q)}$  data in previous block is given as entering signals to Dabuchi 4 wavelet transform, then wavelet transform decomposes base signal to two signal levels according to high-pass and low-pass filters like

Fig. 7. As, it is shown in Fig. 7, high-pass signal is the noisy signal that is put aside. Low-pass signals in every dimension of each sensor are saved in  $WX_i^{(l,m)}$  for each  $i=1,2,\dots,n$ . The remaining process continues over the low-pass signal whose noises are reduced.

## Results

### Simulation process

A wireless body sensor network of two nodes, in which, each node is equipped with three axes, is considered. Twelve different test conditions for patients with low back pain (LBP, test.a), weak muscles (test.b), calf strain (test.c), chondromalacia of patella (test.d), disc herniation (test.e), spinal canal stenosis (test.f), DJD of spine (test.g), hamstring shortness (test.h), medial meniscal tearing (test.i), DJD of knee (test.j), DJD of hip (test.k) and anterior collateral ligament tearing (test.m) are considered. As an example, [Supplementary 1](#) shows seven models of these tests. Both of the sensor nodes are installed on the body of the patients in order to receive and monitor the movements. Each sensor is equipped with a LPC1768 Cortex-M3 microprocessor, one RFM70 module and an ADXL335 tri-axial accelerometer. In order to make a comparison between Tseng algorithm<sup>10</sup> and proposed method, the data are collected from sensors. Then, collected data is processed in base station and two considered compression methods are applied on the processed data in the base station. Every test is carried out on the considered patient and the saved results in the nodes are transferred to base



**Fig. 7.** A sample of decomposition of the collected data: According to discrete wavelet transform: (A) displays the collected raw data; (B) displays the low-pass part of the decomposed signal; and (C) displays the high-pass part of the signal.

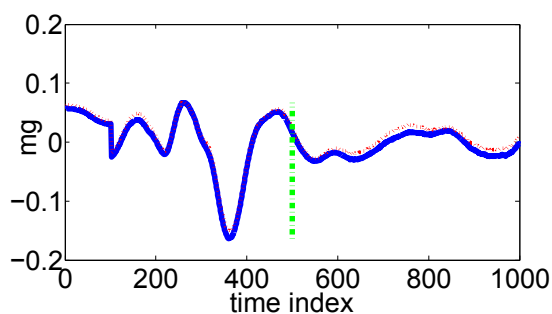


Fig. 8. A sample of real and predicted data.

station through a trustable mechanism. Some parts of the processed data are considered as a set of learning sequence and remaining part is considered as test data.

For every test case, two methods of data compression are compared:

1. Data compression method using the temporal-spatial correlation (Tseng algorithm<sup>10</sup>): in this method only temporal-spatial correlation is used to compress received data.
2. Compression method using wavelet transform, principal component analysis and temporal-spatial correlation: in the proposed method, in addition to temporal-spatial correlation, wavelet transform and principal component analysis are also used to increase the compression rate and reduce the error between the real amount and the prediction amount.

**Discussion**

**Evaluation of results**

For each compression method, a compression model would be created using the learning sequence in offline phase and then online compression rate would be measured using the data test set. For each method, the obtained results from offline phase include transference order tree,  $\alpha$ s and  $\beta$ s correlation coefficient and Huffman code word for each sensor. Fig. 8 shows a sample of prediction result. Where the black signal represents the raw collected data and the

signals with dotted line displays the predicted data. As, it is shown in the figure, the presented prediction method, creates a prediction data which is much closer to the raw data with a low error. This has led to the reduction of prediction error; therefore the proposed method increases the compression rate and optimized bandwidth usage.

Fig. 9 shows the compression rate for both methods (proposed and Tseng methods). Compression rate is defined as the ratio of the size of compressed data to that of the non-compressed data, and it is used as main criterion of the evaluation. In fact, size of the compressed data is always smaller than size of the non-compressed data. Therefore, by decreasing compression rate, data would be more compressed. In fact, non-compressed data is the same with obtained data in diff-coding stage. Compressed data is amount of the obtained error in online stage. As it was mentioned in the online section, the stages of obtaining prediction error depend on the level of considering node. In Fig. 9, it is observed that compression rate in the proposed method (gray column) has been increased and it has led to optimum use of the band width and reduction of the conflict between the sent data by the sensors.

Fig. 10 shows the examples of error between the real and predicted amounts. The error between the real and predicted amount is achieved through the subtraction of real data from the predicted one. In order to increase the correlation among the received data from the nodes and reduce the noise, the theory of principal component analysis and wavelet transform is used. As it was observed in Fig. 10, error in the proposed method in each axis node is close to zero and shows more reduction because of using the analysis of principal components and wavelet transform (part B, Fig. 10) relative to Tseng algorithm (part A, Fig. 10).

Since produced error is a datum that should be sent to the base station instead of raw data, reducing the error improves the band width that is considered as the main aim of the present research. Fig. 11 shows the restored data that is calculated from the sum of predicted amount

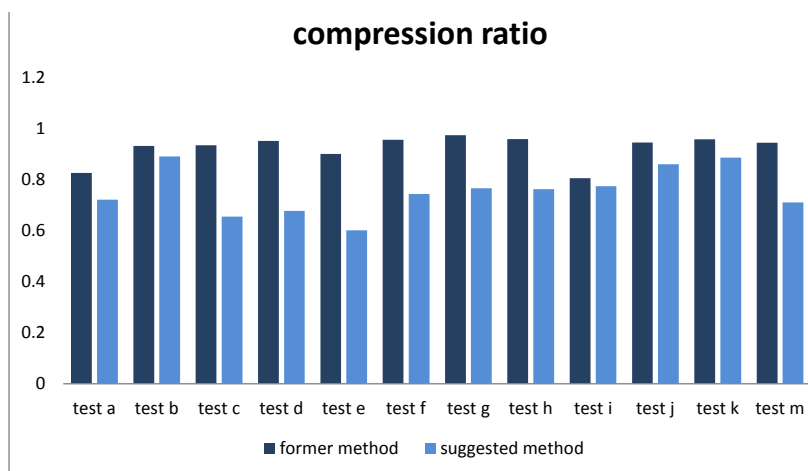
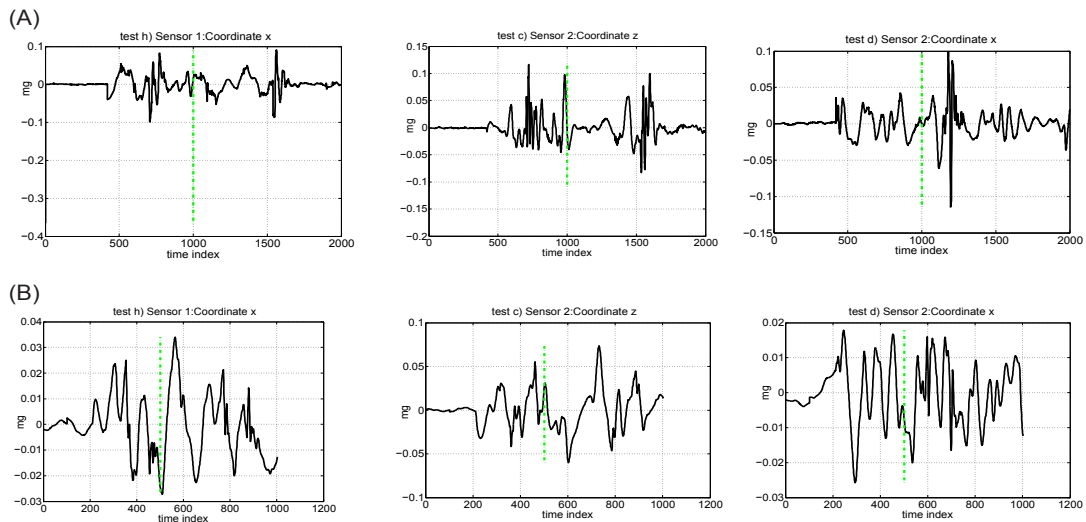
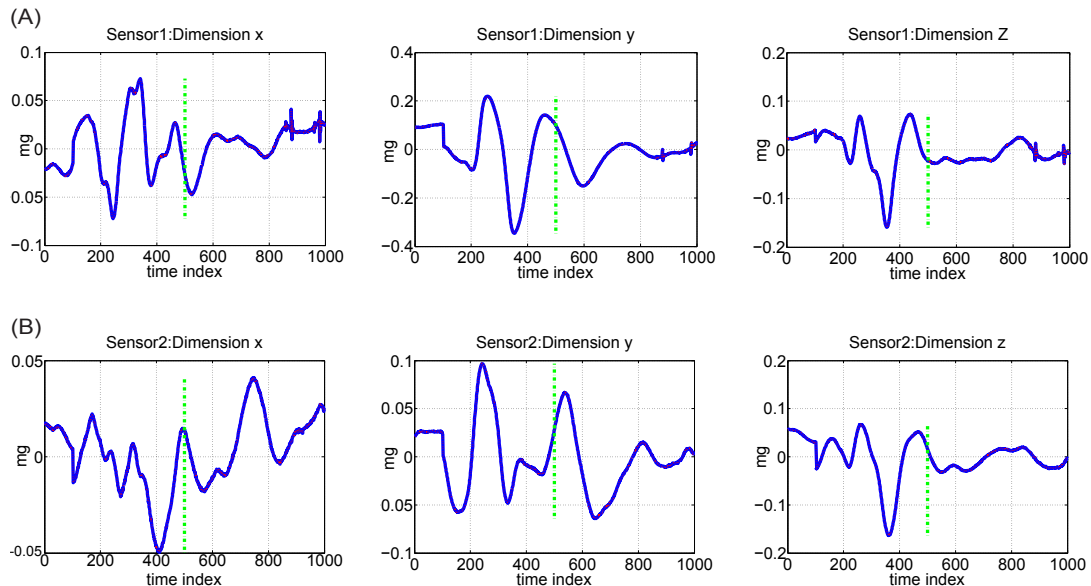


Fig. 9. The diagram of comparing compression rate between two proposed method and Tseng algorithm in all twelve tests.



**Fig. 10.** The examples of error between the real and predicted amounts in testes g, h, i, j, k and m in both compression methods: part (A) displays the error in Tseng method and part; and part (B) displays the error in proposed method.



**Fig. 11.** An example of the restored data that displays a good coincidence between the raw data and predicted one.

with the produced error according to equation IX.

$$\Delta X_i^{(j)} = \alpha(i, j) |k| + \sum_{p=1}^m \beta(i, j) |(k, p) \Delta X_k^{(p)} + \sum_{q \in L(i, j) |k} \beta(i, j) |(j, l) \Delta X_i^{(q)} + \varepsilon(i, j) |k \quad \text{Eq. (IX)}$$

In this figure signals are displayed with black colour which shows the principle data, and signals displayed with dotted line shows restored data obtained by the proposed method. From Fig. 11 it could be clearly observed that the restored data coincides with raw data. It demonstrates the correctness of data prediction method. Vertical lines in Figs. 8, 10, and 11 separate offline and online stages. The signal in the right side of the dotted line shows the

produced signal in offline stage and right side represents the produced signal in online stage.

**Conclusion**

A compression method based on the theory of principal component analysis, wavelet transform and temporal-spatial correlation is presented. The theory of principal component analysis has been used to increase compression rate and the correlation among the received data from different nodes. Wavelet transform is also used for better reduction of the noise and interferences in raw data received from sensors. The proposed method reduces: 1- the data volume sent by the sensors to the base station 2- interference and conflict among the sent

data and 3- error between the real amount and predicted amount in comparison with the method available in the literature. Furthermore, it leads to the optimum use of wireless channel's band width. Since transference of one bit data is equal to executing 1000 code in CPU order, thus this reduction in volume of sent data through data compression causes the optimum use of energy in sensor batteries and increase battery life. The importance of this problem could be observed clearly where body wireless sensor network especially planted ones, are needed to be used for a long time like diabetic patients. The amount of compression rate in tests a to m for 30000 data using Tseng method is 0.8256, 0.9315, 0.9340, 0.9509, 0.8998, 0.9556, 0.9732, 0.9580, 0.8046, 0.9448, 0.9573, 0.9440 and for proposed method is 0.7210, 0.8898, 0.6548, 0.6765, 0.6009, 0.7435, 0.7651, 0.7623, 0.7736, 0.8596, 0.8856 and 0.7102 which shows an increase in compression rate in proportion to Tseng method (see Fig. 9). The results display the usefulness of proposed algorithm in comparison to Tseng algorithm. The proposed method makes it possible to receive correct data by physicians and physiotherapists from the body sensor networks for better treatment and diagnosis of illness.

#### Acknowledgments

We thank Abdolvahab Moghadamm, specialist in physical medicine and rehabilitation, because of patient referral for experimental exercises, and vibration and modal analysis, laboratory of Mechanical Engineering Department, University of Tabriz for their sincere collaboration.

#### Ethical issues

There is none to be declared.

#### Competing interests

The authors declare no conflict of interests.

#### Supplementary materials

Supplementary file contains the supplementary 1.

#### References

1. Zhou G, Li Q, Li J, Wu Y, Lin S, Lu J, et al. Adaptive and radio-agnostic qos for body sensor networks. *ACM Transactions on Embedded Computing Systems (TECS)* **2011**; 10: 48. doi: 10.1145/2043662.2043672
2. Jea D, Wu W, Kaiser WJ, Srivastava MB, editors. Approximate data collection using resolution control based on context. *4th International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2007)*; **2007**: Springer.
3. Baheti PK, Garudadri H, editors. An ultra low power pulse oximeter sensor based on compressed sensing. *Wearable and Implantable Body Sensor Networks, 2009 BSN 2009 Sixth International Workshop on*; **2009**: IEEE.
4. Cheng L, Hailes S, Cheng Z, Fan FY, Hang D, Yang Y, editors. Compressing inertial motion data in wireless sensing systems—an initial experiment. *Medical Devices and Biosensors, 2008 ISSS-MDBS 2008 5th International*

#### Research Highlights

##### What is current knowledge?

- ✓ The rapid growth in physiological sensors, low-power integrated circuits, and wireless communication has enabled a new generation of wireless sensor networks, now used for purposes such as monitoring health.
- ✓ Such applications may require real-time and high-resolution sensing data for various purposes, such as visualizing motions, recognizing body conditions, and diagnosing diseases. These requirements imply shorter transmission delays and higher sampling rates.
- ✓ The main problem in this regard is the limitation of band width of wireless channel and connection and collision between the data sent from sensors to the base station.
- ✓ In order to solve these problems data compression technique is used to reduce the transference rate.

##### What is new here?

- ✓ In this research a model of compression based on the theory of principal component analysis and wavelet transform is presented.
- ✓ The proposed method reduces: 1- the data volume sent by the sensors to the base station 2- interference and conflict among the sent data 3- error between the real amount and predicted amount in comparison with the method available in the literature.
- ✓ This method leads to optimum use of wireless channel's band width.

5. Zarei S, Farokhi F, editors. Data Reduction in Body Sensor Networks Using Wavelet Principal Components Analysis. *Intelligent Computation and Bio-Medical Instrumentation (ICBIMI), 2011 International Conference on*; **2011**: IEEE.
6. Chen F, Wen F, Jia H, editors. Algorithm of data compression based on multiple principal component analysis over the WSN. *Wireless Communications Networking and Mobile Computing (WiCOM), 2010 6th International Conference on*; **2010**: IEEE.
7. Deshpande A, Guestrin C, Madden SR, Hellerstein JM, Hong W, editors. Model-driven data acquisition in sensor networks. *Proceedings of the Thirtieth international conference on Very large data bases-Volume 30*; **2004**: VLDB Endowment.
8. Silberstein A, Puggioni G, Gelfand A, Munagala K, Yang J, editors. Suppression and failures in sensor networks: A Bayesian approach. *Proceedings of the 33rd international conference on Very large data bases*; **2007**: VLDB Endowment.
9. Li J, Deshpande A, Khuller S, editors. On computing compression trees for data collection in wireless sensor networks. *INFOCOM, 2010 Proceedings IEEE*; **2010**: IEEE.
10. Wu CH, Tseng YC. Data compression by temporal and spatial correlations in a body-area sensor network: A case study in pilates motion recognition. *Mobile Computing, IEEE Transactions on* **2011**; 10: 1459-72. doi: 10.1109/tmc.2010.264