



## Feature Dimension Reduction of Multisensor Data Fusion using Principal Component Fuzzy Analysis

H. Kashanian, E. Dabaghi\*

Department of Computer Science & Software Engineering, Ferdows Islamic Azad University, Ferdows, Iran

### PAPER INFO

#### Paper history:

Received 21 October 2016

Received in revised form 17 January 2017

Accepted 09 February 2017

#### Keywords:

Principle Components Fuzzy Analysis

Multisensor Data Fusion

Ubiquitous Activities Monitoring

Intelligent Devices

Health Care

### ABSTRACT

These days, the most important areas of research in many different applications, with different tools, are focused on how to get awareness. One of the serious applications is the awareness of the behavior and activities of patients. The importance is due to the need of ubiquitous medical care for individuals. That the doctor knows the patient's physical condition, sometimes is very important. Of course, there are other important applications for this information. There are a variety of methods and tools for measurement, gathering, and analysis of the physical behaviors and activities' information. One of the most successful tools for this aim are ubiquitous intelligent electronic devices, specifically smartphones, and smart watches. There are many sensors in these devices, some of which can be used to understand the activities of daily living. As an output result, these sensors produce many raw data. Thus, it is needed to process these information and recognize the individual behavior of the output of this processing. In this paper, the basic components of the analysis phase for this process have been proposed. Simulations validate the benefits and superiority of this method.

doi: 10.5829/idosi.ije.2017.30.04a.06

## 1. INTRODUCTION

In recent decades, the ubiquitous devices have made a considerable growth in technology and applications. As an instance, the use of such devices has entered into medical care. Different types of sensors such as accelerometers, magnetometer and gyroscope, are generally found in every device. The ubiquitous nature of these devices is discovered especially by considering the global sales of such devices. It is estimated that only smartphone sales has surpassed 837 million units in 2013 [1]. In addition, the rate of worldwide growth of 26% for both tablets and smart phones, between 2012 and 2016, was predicted. These devices are portable and thus can be remained close to the user in a long time [2]. In addition, these devices are becoming usual, not only in the developed economies, but worldwide [3]. Mobile devices are solutions for healthcare remote monitoring of those who are mostly at risk. It is predicted that the prevalence of cardiovascular disease and stroke among

Americans increases on an average of 20.75% by 2030. Therefore, there are persuasive reasons to do studies on effectiveness of such devices in the applications related to mobile healthcare. There are multiple applications for these devices, including detection of falls [4-7] and preventing falls [8, 9]. A number of studies investigate the potential role of individual sensors including accelerometers [10-13] and GPS [14] in the activity detection. Recently, multisensory fusions attract more attention in this area [15-17] or in some other applications [17-19]. In recent developments, a smart tool, such as a mobile phone can be used as a gateway for one or more embedded sensors [20]. A number of articles have tried to study the detection and understanding of the physical activities using dedicated sensors, which are often tied to the user by straps or tapes [21-23]. However, most recent researches have focused on mobile devices embedded sensors. In some works [24], Android-based smartphone built-in accelerometer sensor was used for data collection of everyday activities in 29 patients. Activities included walking, jumping, going up and down the stairs, sitting and standing. However, the authors only examined a

\*Corresponding Author's Email: [e.dabaghi.ac@gmail.com](mailto:e.dabaghi.ac@gmail.com) (E. Dabaghi)

mobile phone's accelerometer, and did not use the other sensors data. In the literature [25], a dual-axis accelerometer beside a sensitive optical sensor in eWatches was used to recognize six activities: standing, sitting, running, walking, going up and down the stairs. The authors gained 92 percent precision. Ganti et al. [16] recorded the data from four sensors of a Nokia N95; including a microphone, accelerometer, GPS and GSM. Eight separate activities that have been recorded include aerobic exercise, cooking, desk, driving, eating, hygiene, meeting, and watching television. Selected features included estimating energy expenditure, skewness of acceleration size, and microphone signal. The authors decided to use a three hidden Markov model (HMM) and the average results were 66%. The potential role of cell phones and smart watches in everyday human behavior elicitation, using a single sensor or multisensor, was also studied [26], for example, patient indoor/outdoor situation detection using GPS and light sensors. A feasibility study was done which included 10 participants to evaluate the devices ability to distinguish between nine daily activities. Activities included walking, running, biking, standing, sitting, elevator ascent/descent, going up and down the stairs. The authors also investigated the indoor/outdoor distinguish ability of this device. This data was used to train five known test machine-learning algorithms: C4.5, CART, Naïve Bayes, MLP and support vector machine. The authors claimed that some correct models can reach up to 100% recognition of all cases [26]. However, they did not properly prove the validity of their claims, and did not describe the preprocessing and post-processing algorithms in details.

One impressive problem in all these studies is the big size and dimensions of information, which prolong the process and in fact cause all the methods unusable. Therefore, as a basic solution, the majority of articles summarize the information, for example using Principle Component Analysis (PCA) approach. PCA is very useful in feature dimension reduction, and has lots of applications [27-31].

Here, a novel Principle Component Fuzzy Analysis approach is introduced to solve the described problem using the advantages of combination of fuzzy logic into the traditional PCA. The main reason of this combination is the nature of human behaviors and activities that are not precise and distinctly partitioned from each other. Sometimes running in a person is very similar to walking in another person or also can be very similar to a new person periodically falling down! Thus, the fuzzy logic view can be more successful to crisp approaches. In the other words, for coping with this type of human based uncertain information, one should treat very similar to the main character. In this type of problems, certain, crisp and determined boundary

solutions lead to several self-constructed mistakes in information processing, detection and inference. Fuzzy logic was first raised in 1964 by Dr. Zadeh to deal with this kind of uncertain problems.

The second problem is the presence of noise in the received information from the sensors. Information received from almost all sensors are noisy, but mobile sensors are inherently more noisy. Scientists omit this noise by a variety of filters such as low pass filters, integrator, received interval data average, etc. Using fuzzy logic in the proposed Principle Components Fuzzy Analysis (PCFA), also helps to noise filtering and effectively removing small noises.

According to above-mentioned problems, in this paper, traditional PCA is combined with fuzzy logic (FL) and a novel method called PCFA is proposed for noise cancelation and summarize the sensors information for use in health care. Problems included the need to compress due to the large volume of data, and the need to remove noise that is observed in the sensor data.

Section 2 describes the traditional principle component analysis; in section 3 the novel fuzzy based version is explained. In addition, some discussion is detailed in section 4. To illustrate the validity of our novelty, several simulations are done and reported in section 5.

## 2. PRINCIPLE COMPONENT ANALYSIS

The PCA is based on the analysis of covariance or correlation matrices. If we consider the data set as follows:

$$X = \{x^1, \dots, x^n\} \subset \mathbb{R}^p \quad (1)$$

and covariance matrix as Equation (2):

$$M_{ij} = \frac{1}{p-1} \sum_{k=1}^p (x_i^k - \bar{x}_i)(x_j^k - \bar{x}_j), \quad i, j = 1, \dots, n \quad (2)$$

where  $n$  is the number of samples and  $p$  is the number of the original variables. Let us name the orthonormal eigenvalues of matrix  $M$  by  $e^i$  and corresponding eigenvalues by  $\lambda_i$ . The principle components of data  $X$ , will appear as a linear combination of main variables as follows:

$$PC_i = e_1^i y^1 + e_2^i y^2 + \dots + e_n^i y^n \quad (3)$$

In which  $y^i$  is  $i^{\text{th}}$  main variable ( $y_j^i = x_j^i$ ),  $e_j^i$  is  $j^{\text{th}}$  element of  $i^{\text{th}}$  eigenvector of matrix  $M$  and  $PC_i$  is the  $i^{\text{th}}$  principle component. The constraint  $(e_1^i)^2 + (e_2^i)^2 + \dots + (e_n^i)^2 = 1$  is true for all the components, and the following constraints, too.

$$\begin{aligned}
 e^{iT} e^i &= 1, & \text{for any } i &\in \{1, \dots, n\} \\
 e^{iT} e^j &= 0, & \text{for any } i, j &\in \{1, \dots, n\}, i \neq j \\
 e^{iT} M e^i &= 1, & \text{for any } i &\in \{1, \dots, n\} \\
 e^{iT} M e^j &= 0, & \text{for any } i, j &\in \{1, \dots, n\}, i \neq j \\
 M &= \lambda_1 e^1 e^{1T} + \lambda_2 e^2 e^{2T} + \dots + \lambda_n e^n e^{nT}
 \end{aligned}
 \tag{4}$$

The main property of the new variable is uncorrelation.

### 3. TOWARD PRINCIPLE COMPONENTS FUZZY ANALYSIS

Here, we propose to utilize sugeno fuzzy system to improve PCA. Consider the dataset  $X$ , assuming that the matrix  $L$  to be defined as  $L = (L^1, L^2, \dots, L^s)$ , and each one determines one of  $s$  clusters. Mapping  $X$  into  $s$  fuzzy clusters is achieved by minimizing the following objective function:

$$J(P, L) = \sum_{i=1}^s \sum_{j=1}^n (A_i(x^j))^m d^2(x^j, L^i)
 \tag{5}$$

In which  $P = \{A_1, A_2, \dots, A_s\}$  is fuzzy partitions and  $A_i(x^j) \in [0, 1]$  is the fuzzy membership function of data  $x^j$  in fuzzy cluster  $A_i$ .

Covariance matrix is calculated as follows:

$$C = \frac{\sum_{k=1}^p A(x^k)^m \cdot (x_i^k - \bar{x}_i) \cdot (x_j^k - \bar{x}_j)}{\sum_{k=1}^p A(x^k)^m}, i, j = 1, \dots, n
 \tag{6}$$

Fuzzy membership function is calculated as:

$$A(x^j) = \frac{\alpha}{1 - \alpha} \frac{1}{\frac{\alpha}{1 - \alpha} + (d^2(x^j, L))^{\frac{1}{m-1}}}
 \tag{7}$$

where  $d$  is the Euclidian distance between input data and center of cluster,  $m$  and  $\alpha \in (0, 1)$  are design parameters. The effects of change in these two design parameters are shown in Figures 1 and 2.

The number of fuzzy membership functions is directly equal to the number of fuzzy clusters (principle components).

Thus, the proposed method determines  $A(x^j)$ , the best values that describe the best fuzzy sets, and the relation with its linear prototype (the first principal component), using Equation (5).

### 4. DISCUSSION

Utilizing PCFA instead of using PCA can solve the two problems. The first problem is large volume of data beside uncertainties in the data, and the second problem is presence of noise in sensor data. This claim is proven

by simulations conducted and the results have been announced in the next section.

The presented method using the advantage of fuzzy approach, maps the noisy data into the fuzzy memberships and even in this first step, the effect of noise is paled significantly. Then, by eigenvector and covariance analysis, decrease the number of features and only send the principle features to clustering methods. Thus, the resulted system is more efficient and robust to environmental noises.

The proposed method has many applications, i.e. wherever the PCA is used the proposed PCFA might work better, and overall in problems where we need to decrease the number of features into only principle components, we can use the PCFA; an experimental application of this method is studied in the next section.

### 5. SIMULATION

In this paper, the database taken from mobile sensors by Mike Stanley<sup>2</sup> is used and the discriminant ability for three actions including fast picking up the mobile phone, picking up the phone, and slow picking up the mobile phone are studied. Sensors used in this study are gyroscope, accelerometer, and magnetometer, which each has three degrees of freedom.

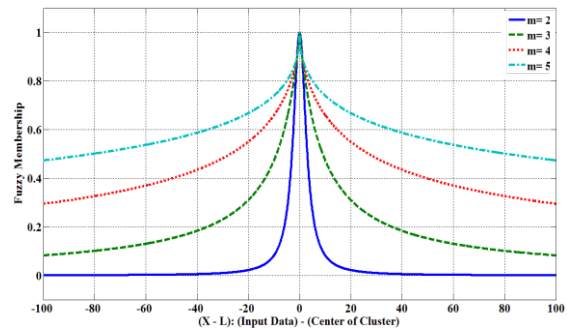


Figure 1. Fuzzy membership function, the effect of altering the design parameter,  $m$ .

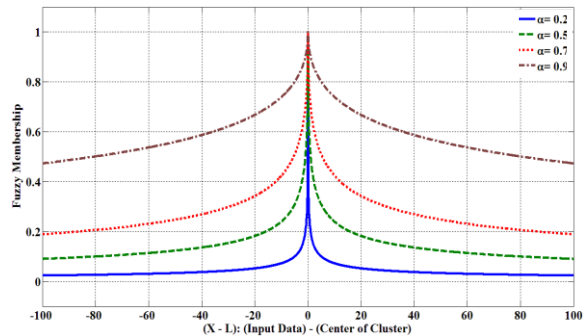


Figure 2. Fuzzy membership function, the effect of altering the design parameter,  $\alpha$

<sup>2</sup> <http://www.nxp.com/products/sensors/sample-data-sets-for-inertial-and-magnetic-sensors:sensordata>

In the literature [26] both Samsung Galaxy Nexus smartphones and Motorola MotoActv smartwatch were used in order to collect data from all sensors. Data are collected from sensors in Nexus from, data 3-axial accelerometer sensor, 3- axial magnetometer, 3- axial gyroscope, GPS, light and pressure, and from the 3-axial accelerometer of smartwatch. The preprocessing reshapes the raw data into the interval. On both devices, the accelerometer raw data can be somewhat sporadic (intermittent), about 90 Hz and 15 Hz, respectively, for phone and watch.

With full interpolation, the authors focused on signal filtering. The use of such a filter is not particularly insignificant. Two accelerometers aim to separate the dynamic components according to human movement, and static component due to gravity, filtered with lowpass filters. Preprocessing was applied also to pressure and other sensors, although the gravity is not an issue for the magnetometer or gyroscope. Similarly, pressure sensor accepts all frequencies lower than 0.1 Hz.

PCA method is used in most of the algorithms, in order to reduce the dimensions of the feature vector. The main component follows the data direction with the biggest power or changes. PCA algorithm is fully described elsewhere [32]. PCA involves the calculation of covariance matrix of eigenvalues and eigenvectors. PCA is used for both data of smartphone and smartwatch. It is shown that 29 principle features from 53 main features, have 95% of the whole information, and the important features of all sensors (accelerometer, magnetometer, gyroscope and etc.) are seen in the principle features.

In this paper, using the novel PCFA instead of PCA is proposed, that helps dramatically to reduce the number of features on systems with a big data size and eliminates the noise, and as a result, the system is robust against environmental disturbances.

Here, a data set is utilized that is related to data taken from the three sensors, gyroscope, accelerometer and magnetometer, and each of these sensors sends information in three degrees of freedom. Sensors data can be seen in Figures 3-11; these data have been received for three activities, fast picking up the mobile phone, picking up the phone, and slow picking up the mobile phone. The collected data are noisy and have other uncertainties of environmental.

Before we can learn how to use such huge amount of data in the system, we must reduce the number of features, i.e., the basic features are found and the rest are deleted.

The result of using PCFA instead of PCA is summarized in Table 1.

As seen in the table, discrimination is done far better by the proposed PCFA. It reduced the number of features and accuracy is higher than PCA.

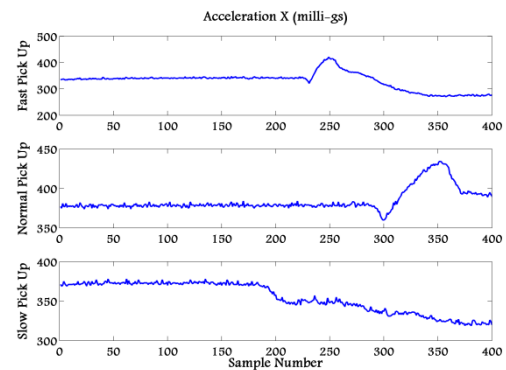


Figure 3. Accelerometer data of axes X

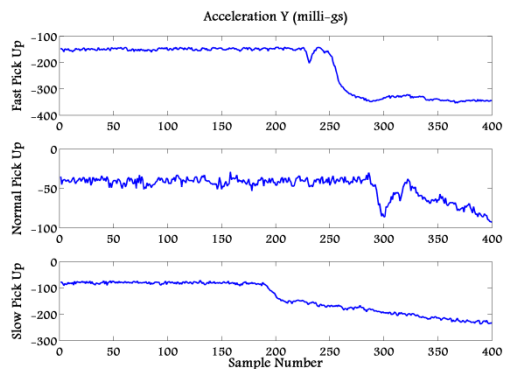


Figure 4. Accelerometer data of axes Y

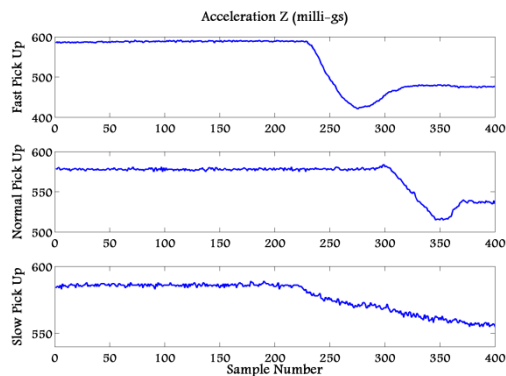


Figure 5. Accelerometer data of axes Z

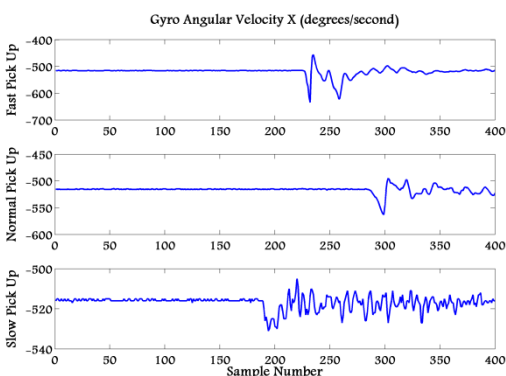


Figure 6. Gyroscope data of axes X

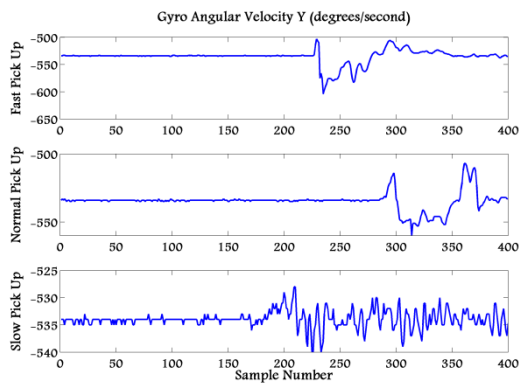


Figure 7. Gyroscope data of axes Y

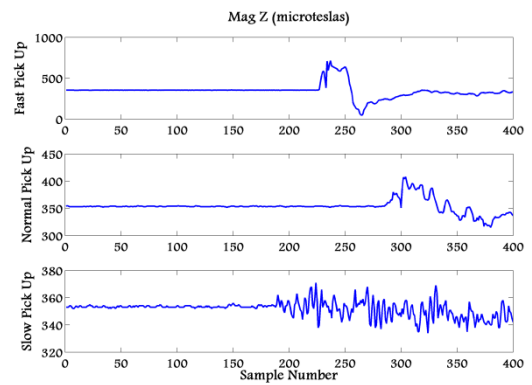


Figure 11. Magnetometer data of axes Z

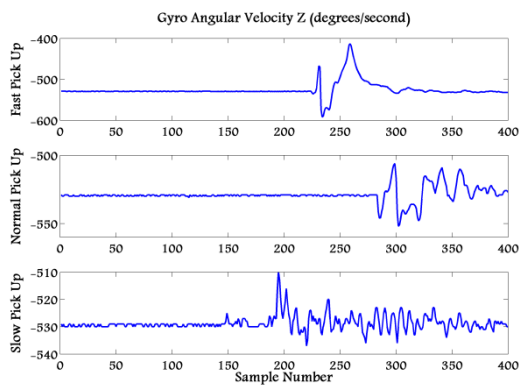


Figure 8. Gyroscope data of axes Z

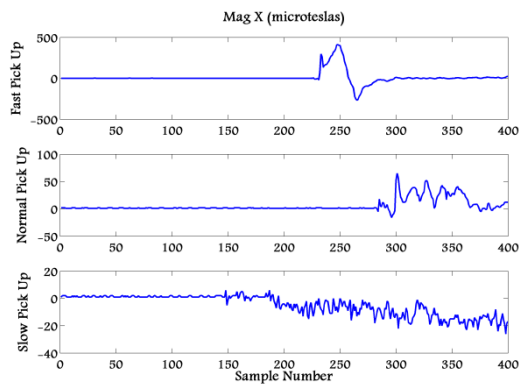


Figure 9. Magnetometer data of axes X

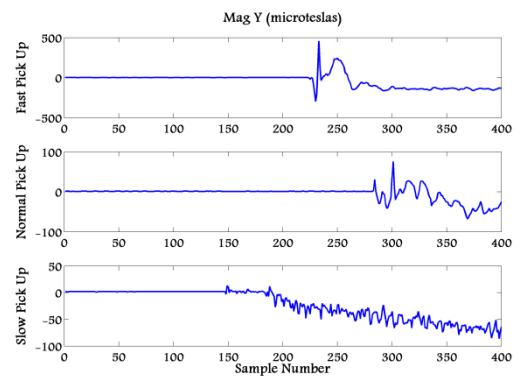


Figure 10. Magnetometer data of axes Y

TABLE 1. A comparison between PCFA and PCA

Method	Main feature num	Principle components num	Compression accuracy
PCA	444	29	3e-25
PCFA	444	2	9e-27
PCA	835	50	4e-24
PCFA	835	2	2e-24
PCA	444	25	8e-23
PCFA	444	2	2e-28

6. CONCLUSION

In this paper, Principal Components Fuzzy Analysis (PCFA) is presented. The proposed algorithm is utilized to reduce the size of data and eliminates the noise of received data from sensors for remote healthcare patient monitoring. Use of the proposed PCFA instead of traditional PCA, leads to accuracy improvement, increases the speed of computations, better noise cancelation, and results in making a robust system in patient behavior recognition according to multisensor fusion of mobile devices. Simulations validated the advantages and benefits of this approach. The resulting benefits include fewer principle features and higher accuracy in the presence of noise.

7. REFERENCES

1. Favell, A., "Global mobile statistics 2014 part a: Mobile subscribers; handset market share; mobile operators", *mobiThinking*, (2014).
2. Kearney, A., "The mobile economy 2013", *London: GSMA*, (2013).
3. Sanou, B., "The world in 2013: Ict facts and figures", *International Telecommunications Union*, (2013).
4. Bianchi, F., Redmond, S.J., Narayanan, M.R., Cerutti, S. and Lovell, N.H., "Barometric pressure and triaxial accelerometry-based falls event detection", *IEEE Transactions on Neural*

- Systems and Rehabilitation Engineering*, Vol. 18, No. 6, (2010), 619-627.
5. Silva, M., Teixeira, P.M., Abrantes, F. and Sousa, F., "Design and evaluation of a fall detection algorithm on mobile phone platform", in International Conference on Ambient Media and Systems, Springer., (2011), 28-35.
  6. Zhang, T., Wang, J., Liu, P. and Hou, J., "Fall detection by embedding an accelerometer in cellphone and using kfd algorithm", *International Journal of Computer Science and Network Security*, Vol. 6, No. 10, (2006), 277-284.
  7. Sposaro, F. and Tyson, G., "Ifall: An android application for fall monitoring and response", in Engineering in Medicine and Biology Society., EMBC. Annual International Conference of the, IEEE., (2009), 6119-6122.
  8. Hausdorff, J.M., Rios, D.A. and Edelberg, H.K., "Gait variability and fall risk in community-living older adults: A 1-year prospective study", *Archives of Physical Medicine and Rehabilitation*, Vol. 82, No. 8, (2001), 1050-1056.
  9. Menz, H.B., Lord, S.R. and Fitzpatrick, R.C., "Acceleration patterns of the head and pelvis when walking are associated with risk of falling in community-dwelling older people", *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, Vol. 58, No. 5, (2003), M446-M452.
  10. Bieber, G., Koldrack, P., Sablowski, C., Peter, C. and Urban, B., "Mobile physical activity recognition of stand-up and sit-down transitions for user behavior analysis", in Proceedings of the 3rd International Conference on Pervasive Technologies Related to Assistive Environments, ACM., (2010), 50-59.
  11. Bieber, G., Voskamp, J. and Urban, B., "Activity recognition for everyday life on mobile phones", in International Conference on Universal Access in Human-Computer Interaction, Springer., (2009), 289-296.
  12. Yang, J., "Toward physical activity diary: Motion recognition using simple acceleration features with mobile phones", in Proceedings of the 1st international workshop on Interactive multimedia for consumer electronics, ACM., (2009), 1-10.
  13. Sun, L., Zhang, D. and Li, N., "Physical activity monitoring with mobile phones", in International Conference on Smart Homes and Health Telematics, Springer., (2011), 104-111.
  14. Terrier, P. and Schutz, Y., "How useful is satellite positioning system (GPS) to track gait parameters? A review", *Journal of Neuroengineering and Rehabilitation*, Vol. 2, No. 1, (2005), 28.
  15. Parkka, J., Ermes, M., Korpipaa, P., Mantjarvi, J., Peltola, J. and Korhonen, I., "Activity classification using realistic data from wearable sensors", *IEEE Transactions on Information Technology in Biomedicine*, Vol. 10, No. 1, (2006), 119-128.
  16. Ganti, R.K., Srinivasan, S. and Gacic, A., "Multisensor fusion in smartphones for lifestyle monitoring", in Body Sensor Networks (BSN), International Conference on, IEEE., (2010), 36-43.
  17. Morón, M.J., Luque, R. and Casilari, E., "On the capability of smartphones to perform as communication gateways in medical wireless personal area networks", *Sensors*, Vol. 14, No. 1, (2014), 575-594.
  18. Moshiri, B., Asharif, M.R. and Hoseinnezhad, R., "A new approach to self-localization for mobile robots using sensor data fusion", *International Journal of Engineering Transactions B*, Vol. 15, (2002), 145-156.
  19. Moussavi Khalkhali, A., Moshiri, B. and Momeni, H., "Designing a home security system using sensor data fusion with dst and dsmt methods", *International Journal of Engineering-Transactions A: Basics*, Vol. 22, No. 1, (2008), 13.
  20. Han, M., Lee, Y.-K. and Lee, S., "Comprehensive context recognizer based on multimodal sensors in a smartphone", *Sensors*, Vol. 12, No. 9, (2012), 12588-12605.
  21. Godfrey, A., Conway, R., Meagher, D. and ÓLaighin, G., "Direct measurement of human movement by accelerometry", *Medical Engineering & Physics*, Vol. 30, No. 10, (2008), 1364-1386.
  22. Moe-Nilssen, R., "A new method for evaluating motor control in gait under real-life environmental conditions. Part 1: The instrument", *Clinical Biomechanics*, Vol. 13, No. 4-5, (1998), 320-327.
  23. Staudenmayer, J., Pober, D., Crouter, S., Bassett, D. and Freedson, P., "An artificial neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer", *Journal of Applied Physiology*, Vol. 107, No. 4, (2009), 1300-1307.
  24. Kwapisz, J.R., Weiss, G.M. and Moore, S.A., "Activity recognition using cell phone accelerometers", *ACM SigKDD Explorations Newsletter*, Vol. 12, No. 2, (2011), 74-82.
  25. Maurer, U., Smailagic, A., Siewiorek, D.P. and Deisher, M., "Activity recognition and monitoring using multiple sensors on different body positions", in Wearable and Implantable Body Sensor Networks., BSN. International Workshop on, IEEE., (2006), 4 pp.-116.
  26. Guiry, J.J., van de Ven, P. and Nelson, J., "Multi-sensor fusion for enhanced contextual awareness of everyday activities with ubiquitous devices", *Sensors*, Vol. 14, No. 3, (2014), 5687-5701.
  27. Esmaileyan, Z. and Marvi, H., "A database for automatic persian speech emotion recognition: Collection, processing and evaluation", *International Journal of Engineering*, Vol. 27, No., (2013), 79-90.
  28. Srikanth, S., Sudha, K. and Raju, Y.B., "Fuzzy load frequency controller in deregulated power environment by principal component analysis".
  29. Dehghan, H., Pouyan, A.A. and Hassanpour, H., "Detection of alzheimer's disease using multitracer positron emission tomography imaging", *International Journal of Engineering, Transactions A: Basics*, Vol. 27, No. 1, (2014), 51-56.
  30. AhilaPriyadharshini, R. and Arivazhagan, S., "Object recognition based on local steering kernel and svm", *International Journal of Engineering-Transactions B: Applications*, Vol. 26, No. 11, (2013), 1281-1288.
  31. Hamidi, H. and Daraee, A., "Analysis of pre-processing and post-processing methods and using data mining to diagnose heart diseases", *International Journal of Engineering-Transactions A: Basics*, Vol. 29, No. 7, (2016), 921.
  32. Marsland, S., "Machine learning: An algorithmic perspective, CRC press, (2015).

# Feature Dimension Reduction of Multisensor Data Fusion using Principal Component Fuzzy Analysis

H. Kashanian, E. Dabaghi

Department of Computer Science & Software Engineering, Ferdows Islamic Azad University, Ferdows, Iran

---

## PAPER INFO

چکیده

---

### Paper history:

Received 21 October 2016

Received in revised form 17 January 2017

Accepted 09 February 2017

---

### Keywords:

Principle Components Fuzzy Analysis

Multisensor Data Fusion

Ubiquitous Activities Monitoring

Intelligent Devices

Health Carer

امروزه، مهمترین حوزه های تحقیقاتی در کاربردهای متفاوت، با ابزارهای متفاوت، روی کسب آگاهی تمرکز کرده اند. یکی از جدی ترین کاربردها، آگاهی از رفتارها و فعالیت های بیماران است. اهمیت این موضوع به خاطر این است که افراد احتیاج به مراقبت های پزشکی همیشه همراه دارند. اینکه پزشک وضعیت جسمی بیمار را در هر لحظه ای بداند، گاهی بسیار پراهمیت می شود. البته، برای این حوزه از تحقیقات کاربردهای مهم دیگری نیز متصور است. تاکنون برای اندازه گیری، جمع آوری و تحلیل داده های رفتار و فعالیت های فیزیکی بیمار، ابزارها و روش های متفاوت و متنوعی، در نظر گرفته شده است. که یکی از موفقترین ابزارها برای این مهم، دستگاه های الکترونیکی هوشمند همیشه همراه، و به طور خاص تلفن های هوشمند همراه، و ساعت های هوشمند است. حسگرهای زیادی در این دستگاه ها وجود دارند، برخی از آنها میتوانند برای فهمیدن فعالیت های زندگی روزمره به کار گرفته شوند. به عنوان نتیجه خروجی، این حسگرها داده های خام زیادی تولید میکنند. بنابراین، احتیاج است که این اطلاعات پردازش شده، و از خروجی این پردازش ها رفتار فرد بازشناسی شود. در این مقاله، روش تحلیل فازی اجزاء اصلی برای فشرده سازی این اطلاعات، معرفی شده است. شبیه سازی ها، مزایا و برتری روش پیشنهادی را تایید می کنند.

doi: 10.5829/idosi.ije.2017.30.04a.06

---