Activity Recognition with Wristband Based on Histogram and Bayesian Classifiers

Yi-Cong Huang, Wing-Kuen Ling
College of Information Engineering,
Guangdong University of Technology.
Guangzhou, China.
thisisecong@outlook.com, yongquanling@gdut.edu.cn.

Chi-Wa Cheng, Chun-Hung Li, Chong-Yan Chen
Add Care Limited.
Hong Kong, China.
victor.gmbox@gmail.com, chli@add-care.net,
chongyanchen_hci@utexas.edu.

***Abstract*—With the increasing popularity of wristbands and their applications on the health monitoring, the recognition of various activities has become a significant feature which helps the users to assess and review their daily practices. As a result, it can improve their health via the recorded information. The core task of the activity recognition is to automatically classify a large amount of motion data recorded by the accelerometers to one of the predefined activity classes such as eating, walking, etc. In contrast to the conventional recognition approaches, the algorithms are required to be simple, efficient and accurate as the battery powers and the allowable computational powers of the wristbands are very limited. Some approaches directly transfer the motion data to the mobile phones by the bluetooth so that more sophisticated computations can be performed. Nevertheless, the transmission of the raw digital motion data can consume a lot of power and occupy a considerable bandwidth. This paper proposes an algorithm that highly reduces the communications between the wristbands and the mobile phones by firstly preprocessing and labeling the raw motion data based on their characteristics. Only the labels and some encoding values are transferred to the mobile phones for performing the classification. In order to achieve the high accuracy, the classification is performed by combining different approaches such as the histogram approach and the Bayesian classifier approach. The obtained results show that the proposed approach achieves a good accuracy, which just uses some labels and the encoding value information.**

Keywords-activity recognition; machine learning; histogram; Bayesian classifier; model combination; wristband.

I. Introduction

For the wristband application, the priority is to reduce both the transmission bandwidth and the use of the floating computation. Therefore, this paper proposes a method based on the cluster labeling of the motion data and the relationships among the encoding values of the motion data to reduce the use of the floating computation and the transmission bandwidth. Then, the integer motion feature is transmitted to the cloud server via the bluetooth and internet. The classification algorithm deployed on the cloud processes the motion feature and returns the result of the activity recognition to the wristband.

II. Related Work

The performances of different algorithms on the activity recognition have been evaluated via the hidden Markov model method, the random forest approach and the deep learning algorithm. For the hidden Markov model, it is a stochastic finite state machine. It models an activity pattern by learning the transition probabilities among its non-observable states such that the likelihood of the observation of a temporal sequence of the symbols representing the activity is maximized [1]. For the random forest, it is an ensemble learning algorithm for performing the classification, the regression and other tasks. It operates by constructing a number of decision trees at the training phase as well as outputting the class that is the mode of the classes of the individual trees by performing the classification or the mean prediction of the individual trees by performing the regression [2]. For the deep learning algorithm, it uses the recurrent neural networks. Here, the long short term memory cells do not require or almost do not require the features [3]. The data can be fed directly into the neural network who acts as a black box for performing the modeling.

However, the hidden Markov model method and the random forest approach require a lot of floating computations for performing the feature extraction. Hence, it is not suitable for the wristband application because of its limited computational resource. Besides, although the deep learning algorithm can achieve a good accuracy, sending the wristband data to the mobile phone via the bluetooth and to the cloud computing server via the internet require a considerable bandwidth.

Therefore, reducing the use of the floating computation and the transmission bandwidth are important for the wristband application. This paper is to address these issues.

III. RAW MOTION DATA LABELING AND ENCODING

As the wristbands have a very limited computational resource and the battery power, the obtained motion data has better to be handled with either the mobile phone or the cloud computing server. Nevertheless, the direct transmission of the raw motion data by the bluetooth requires a considerable amount of bandwidth and the battery power. We first cluster and encode the data according to their motion characteristics. Then, only the cluster label and the encoding value are transmitted to the mobile phones. Hence, the communications between the devices can be reduced significantly. For example, if the x, y and z axes of the accelerometers have the sampling rate of 25Hz and each acceleration value is represented by 16 bits, then the data rate transmitted to the mobile phone is 9 Kbytes per minute. Just storing and transmitting 1 minute data for a wristband is challenging. On the other hand, if the data is segmented into a 4 seconds overlapping block with overlapping a 2 seconds data, then there are only 30 blocks for one minute data with each of them being represented by a 1 byte cluster label. Thus, it only required to transmit these 30 bytes which is acceptable for the wristband application.

A. Raw Data Preprocessing

Figure 3.1 shows a typical raw motion data collected by the accelerometers. From Figure 3.1, it is clear to see that the raw accelerometer data contains some high frequency noises and the gravity components (the average DC values). Therefore, the first step of the preprocessing should remove the high frequency noise.

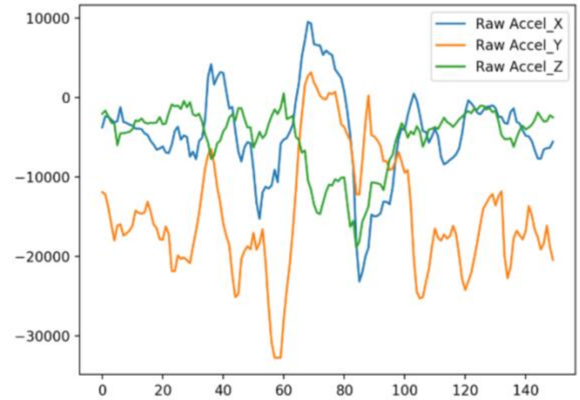

Figure 3.1. The raw accelerometer data.

For our proposed algorithm, a denoising filter is applied for performing the moving average operation. That is, the transfer function of the filter is $H(z)=\frac{1-w}{1-wz^{-1}}$. This operator is chosen because it is easy to implement. In particular, $w=0.7$ is chosen. Hence, it can reduce the required computational power. Figure 3.2 shows the result processed by the moving average filter. It can be seen that it can achieve a considerably good result.

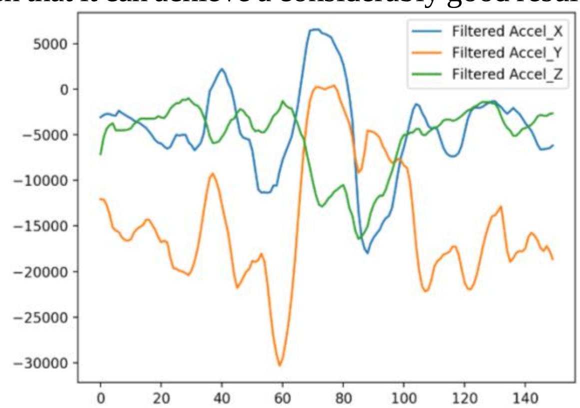

Figure 3.2. The filtered accelerometer data.

For the activity recognition, the angular information is very useful for performing the classification. However, the attitude algorithm may require the floating computations. Thus, after performing the denoising, the relationships among the encoding values of the motion data are used instead. In particular, the encoding values of the data between the x axis and the y axis is defined as

$$E_{xy}=\begin{cases}0 & x-y\geq 500\\ 1 & |x-y|<500\\ 2 & y-x\geq 500\end{cases},$$

that between the y axis and the z axis is defined as

$$E_{yz}=\begin{cases}0 & y-z\geq 500\\ 1 & |y-z|<500\\ 2 & z-y\geq 500\end{cases}$$

and that between the x axis and the z axis is defined as

$$E_{xz}=\begin{cases}0 & x-z\geq 500\\ 1 & |x-z|<500\\ 2 & z-x\geq 500\end{cases}.$$

From Figure 3.2, it is obvious to see that the gravity components will affect the dynamic range of the motion data. Therefore, it is required to remove the gravity components. To address this issue, define the local gravity components of the x axis, y axis and z axis as $g_x(k)$, $g_y(k)$ and $g_z(k)$, respectively. Define the filtered outputs corresponding to the x axis, y axis and z axis as $f_x(k)$, $f_y(k)$ and $f_z(k)$, respectively. First, $g_x(0)$, $g_y(0)$ and $g_z(0)$ are initialized as the means of the first two points of the corresponding filtered outputs, respectively. That is,

$$g_x(0)=\frac{f_x(0)+f_x(1)}{2},$$

$$g_y(0)=\frac{f_y(0)+f_y(1)}{2}$$

and

$$g_z(0)=\frac{f_z(0)+f_z(1)}{2}.$$

Then, the gravity components are updated based on the convex weighting of the next filtered outputs and the current gravity components. That is,

$$g_x(k+1)=\frac{f_x(k+1)+(c-1)g_x(k)}{c},$$

$$g_y(k+1)=\frac{f_y(k+1)+(c-1)g_y(k)}{c}$$

and

$$g_z(k+1)=\frac{f_z(k+1)+(c-1)g_z(k)}{c}.$$

Here, $c=20$ is chosen as the convex weighting coefficient. Finally, define $\tilde{f}_x(k+1)$, $\tilde{f}_y(k+1)$ and $\tilde{f}_z(k+1)$ as the corresponding gravity removed outputs, respectively. Here, the gravity removed components are defined as the filtered outputs minus the updated gravity components. That is,

$$\tilde{f}_x(k+1)=f_x(k+1)-g_x(k+1),$$

$$\tilde{f}_y(k+1)=f_y(k+1)-g_y(k+1)$$

and

$$\tilde{f}_z(k+1)=f_z(k+1)-g_z(k+1).$$

Figure 3.3 shows the results of the gravity removed components. It can be seen that the proposed method can significantly remove the gravity components.

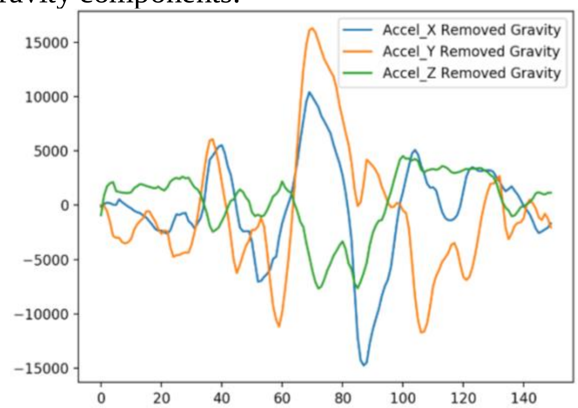

Figure 3.3. The gravity removed accelerometer data.

B. Motion Data Clustering and Encoding

For our experiment, every 20 seconds motion data of a 1 minute motion data are collected continuously by the wristband. To perform the clustering, the 20 seconds data is divided into the blocks of data with each block of data containing a 4 seconds data with overlapping a 2 seconds data. Some useful motion features of these 4 seconds data blocks such as the energies, the time indices corresponding to the zero crossing points as well as the time indices of the crossing points of three different threshold values are extracted. Based on these motion feature vectors, a k means clustering algorithm [4] with the one norm operator as the criterion is performed for assigning the feature vectors. Here, the feature vectors are classified into 9 clusters. Since the features are related to the energies and the time indices of the crossing points of different threshold values, the obtained cluster labels are related to the magnitudes of the motions.

In our experimental setup, only a 20 seconds motion data is taken out from a 1 minute motion data. As each a 4 seconds data block consists of an overlapping 2 seconds data block, there are totally 9 data blocks for each 1 minute motion data. For these 9 data blocks, both the cluster labels and the encoding values discussed in Section 3.1 are used to form the feature vector of each data block. Here, the integer labels are used for the representation of both the cluster labels and the encoding values of each data block. Since each data block only requires 1 byte for the representation of the cluster labels and 3 bytes for the representation of the encoding values, the new feature vector representing the 20 seconds motion data only requires 9 bytes for representing the cluster labels and 3 bytes for representing the encoding values to transmit to the cloud server showed at figure 3.4. The main advantage of the proposed algorithm is to reduce the required memory for the processing and the bandwidth for the transmission via the bluetooth technique.

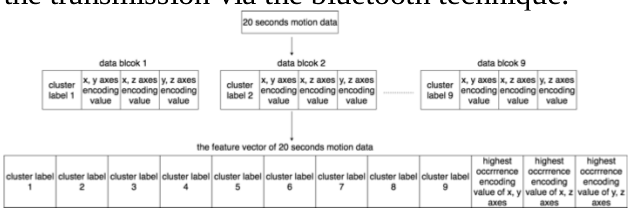

Figure 3.4. Feature extraction of 20 seconds motion data.

IV. HISTOGRAM AND BAYESIAN COMBINED APPROACH

First, the histograms of these 9 cluster labels of these 20 seconds motion data from the 1 minute motion data are computed. Since different histograms represent different classes of the motion activities [5], the histogram classifier is used to perform the recognition of the motion activities. For improving the robustness of the recognition of the motion activities, the histograms are transformed to the cumulative histograms. When a new feature vector is obtained, the cosine similarities between the cumulative histogram of the new feature vector and the cumulative histograms of the new feature vectors of different classes of the motion activities are computed. The class of the motion activities corresponding to the highest similarity is the estimated class of the motion activities.

Second, the Bayesian classifier is used to calculate the conditional probabilities of the encoding values for the given classes of the motion activities as well as the prior probabilities of the classes of the motion activities. When a new feature vector is obtained, the probabilities of the new feature vector corresponding to different classes of the motion activities are computed [6]. The class corresponding to the maximum probability of the new feature vector is the estimated class of the motion activities.

It is worth noting that both the histogram classifier and the Bayesian classifier return the probability vectors of different classes of the motion activities. To combine these two probability vectors, different significant scores are multiplied to these probability vectors. Here, the scores are dependent on the classes of the motion activities and the estimation accuracies of these individual classifiers. The combination result is obtained by applying the softmax function to the weighted probability vectors. Figure 4.1 shows the flowchart of combining the results of the individual classifiers.

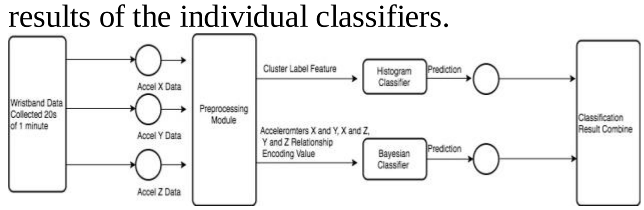

Figure 4.1. The flowchart of the combined classifier.

V. MOTION PERIOD PREDICTION VIA THE MULTIRESOLUTION VOTING

Besides, the motion period is also meaningful for the long time daily activity recognition. To perform the motion period estimation, the multiresolution voting method is used to improve the accuracy [7]. In particular, the original 1 minute motion data is expanded to a 5 or 7 minutes motion data via the multiresolution voting method. Then, the class of the motion activities with the highest occurrence is chosen. On the other hand, the multiresolution voting method can also reduce the recognition error estimated using different time frames of the motion data. Figure 5.1 shows the flowchart of the multiresolution voting method.

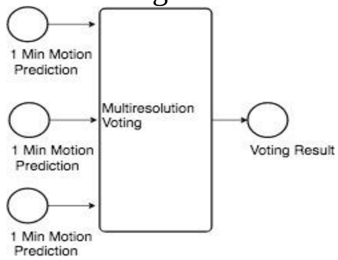

Figure 5.1. The flowchart of multiresolution voting method.

VI. EXPERIMENTAL RESULT AND DISCUSSION

In this paper, 8 people wear the wristbands and each person conducts 5 different motion activities such as walking, office working, eating, sporting and resting. Table I shows the estimated average accuracies of these motion activities based on our proposed method without applying the multiresolution voting method on the 1 minute motion data. Table II and Table III show the estimated average accuracy of these motion activities based on our proposed method with applying the multiresolution voting method on the 5 minutes and 7 minutes motion data, respectively. It is worth noting that our proposed method can achieve the acceptable results even though the multiresolution voting method has not been applied. When the multiresolution voting method is applied, the average accuracies are improved. This is because the multiresolution voting method could perform the motion period prediction and reduce the recognition error estimated using different time frames of the motion data. This is important for the estimation of the long time daily activity and reviewing the human daily practices for the further health monitoring.

TABLE I. THE ESTIMATED ACCURACIES OF THE MOTION ACTIVITIES BASED ON THE 1 MINUTE MOTION DATA.

Motion activities	***Eat***	***Walk***	***Rest***	***Office Work***	***Sport***	
Average accuracies	44%	51%	78%	69%	43%	63% (Total average accuracy)

TABLE II. THE ESTIMATED ACCURACIES OF THE MOTION ACTIVITIES BASED ON THE 5 MINUTE MOTION DATA.

Motion activities	***Eat***	***Walk***	***Rest***	***Office Work***	***Sport***	
Average accuracies	67%	60%	87%	87%	55%	78% (Total average accuracy)

TABLE III. THE ESTIMATED ACCURACIES OF THE MOTION ACTIVITIES BASED ON THE 7 MINUTE MOTION DATA.

Motion activities	***Eat***	***Walk***	***Rest***	***Office Work***	***Sport***	
Average accuracies	72%	62%	92%	95%	61%	83% (Total average accuracy)

VII. CONCLUSION

For the wristband application, the transmission bandwidth is very limited. To address this difficulty, this paper proposes to combine both the histogram classifier and the Bayesian classifier to perform the recognition of the motion activities. In particular, the histogram classifier employs the cluster label feature and the Bayesian classifier employs the encoding value features for performing the recognition. Besides, the multiresolution voting method is employed to estimate the motion period prediction and to reduce the recognition error estimated using different time frames of the motion data. Here, the integer valued features instead of the floating point features are employed. This could improve the efficiency for the data transmission. This is useful for the estimation of the long time daily activity and reviewing the human daily practices for the further health monitoring.

ACKNOWLEDGMENT

We would like to thank the Shutang Information Technology(Shenzhen) Limited Company for providing us the experimental device.

REFERENCES

[1] Raul Ramos Garcia. Using Hidden Markov Models to Segment and Classify Wrist Motions Related to Eating Activities. Clemson University.

[2] Zengtao Feng, Lingfei Mo and Mengli. A Random Forest-Based Ensemble Method for Activity Recognition. 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).

[3] Guillaume Chevalier. LSTMs for Human Activity Recognition, 2016. https://github.com/guillaume-chevalier/LSTM-Human-Activity-Recognition.

[4] Shizeng Lu , Hongliang Yu, Xiaohong Wang, Qiang Zhang, Fanjun Li, Zhao Liu and Fangqian Ning. Clustering Method of Raw Meal Composition Based on PCA and Kmeans. 2018 37th Chinese Control Conference (CCC).

[5] Fei Xue, Yu-Jin Zhang. Image Class Segmentation via Conditional Random Field over Weighted Histogram Classifier. 2011 Sixth International Conference on Image and Graphics.

[6] A. Madabhushi and J.K. Aggarwal. A Bayesian approach to human activity recognition. Proceedings Second IEEE Workshop on Visual Su rveillance (VS'99) (Cat. No.98-89223).

[7] U. Bhattacharya and B. B. Chaudhuri. A Majority Voting Scheme for Multiresolution Recognition of HandprintedNumerals. Seventh International Conference on Document Analysis and Recognition, 2003.Proceedings.

AUTHORS' BACKGROUND

Your Name	Title*	Research Field	Personal website
Yi-Cong Huang	Master student	Machine Learning, Signal Processing, Embedded System	None
Wing-Kuen Ling	Full professor	Time-Frequency Analysis, Optimization Theory, Nonlinear Signal Processing System, Control Theory	http://wkling.gdut.edu.cn
Chi-Wa Cheng	Doctor	Machine Learning, Data Ming, Natural Language Processing	None
Chun-Hung Li	Assistant professor	Machine Learning, Data Mining, Network Intelligence	https://www.comp.hkbu.edu.hk/~chli/index.html
Chong-Yan Chen	Master student	Machine Learning, Data Mining, Human Computer Interaction	None

*This form helps us to understand your paper better, the form itself will not be published.

*Title can be chosen from: master student, Phd candidate, assistant professor, lecture, senior lecture, associate professor, full professor