
IEICE

TRANSACTIONS

on Information and Systems

VOL.E94-D
NO.6
JUNE 2011



A PUBLICATION OF THE INFORMATION AND SYSTEMS SOCIETY

The Institute of Electronics, Information and Communication Engineers

Kikai-Shinko-Kaikan Bldg., 5-8, Shibakoen 3chome, Minato-ku, TOKYO, 105-0011 JAPAN

Discrimination between Upstairs and Downstairs Based on Accelerometer

Yang XUE[†], Nonmember and Linwen JIN^{†*}, Member

SUMMARY An algorithm for the discrimination between human upstairs and downstairs using a tri-axial accelerometer is presented in this paper, which consists of vertical acceleration calibration, extraction of two kinds of features (Interquartile Range and Wavelet Energy), effective feature subset selection with the wrapper approach, and SVM classification. The proposed algorithm can recognize upstairs and downstairs with 95.64% average accuracy for different sensor locations, i.e., located on the subject's waist belt, in the trousers pocket, and in the shirt pocket. Even for the mixed data from all sensor locations, the average recognition accuracy can reach 94.84%. Experimental results have successfully validated the effectiveness of the proposed method.

Key words: acceleration data, vertical acceleration calibration, feature selection, activity recognition, tri-axial accelerometer

1. Introduction

Accelerometers are currently widely studied for activity recognition which has been applied in many applications including human computer interaction [1], [2], elderly care [3], [4], context-awareness [5], [6], and preventive healthcare [7], [8], thanks to their accuracy in the detection of human body movements, small size, and reasonable power consumption [7]. Although the most prevalent everyday activities (sitting, walking, running, vacuuming) have been successfully recognized [2]–[7], climbing upstairs and downstairs are still hard to distinguish in a few studies [6], [9], [10], [18]–[20]. Some past works have demonstrated 24.07% to 84% recognition rates for upstairs and downstairs using acceleration data, which are summarized in Table 1. Most of these works compute recognition results with data collected from a very small number of subjects under artificially constrained laboratory settings. Some also evaluate recognition performance on data collected from a large number of subjects, but recognition rates dropped a lot. Lastly, most researchers focus on time-domain features, such as mean [6], [9], [10], standard deviation or variance [6], [9], [10], energy, and correlation [9], [10], and extract these features usually from each of the three axes of the accelerometer.

Especially, the cause resulting in hard recognition of downstairs and upstairs is due to the confusion between them. In [9], all 19 downstairs activities are recognized as upstairs, and 12 upstairs are recognized as downstairs in 21

Table 1 Summary of some typical methods on the classification of upstairs and downstairs.

Ref.	Extracted features	Classifier	No. Subjects	No. Classes	Recognition accuracy
[9]	Mean, variance	Bayes	1	2	about 30% upstairs, about 31% downstairs
[18]	IGC, AC	1, distance	24	2	74.00% upstairs, 51.67% downstairs
[9]	Mean, Standard Deviation, Energy, Correlation	C4.5, KNN, SVM, NaïveBayes	2	2	0% upstairs, 43.68% downstairs
[19]	Mean, Standard Deviation, Energy, Correlation	C-NN	6	2	43.33% upstairs, 68.33% downstairs
[10]	Mean, Standard Deviation, Energy, Correlation	148	8	2	78.50% upstairs, 53.67% downstairs
[20]	Raw data, PCA, PCA-90	MLP	6	4	84% upstairs, 83% downstairs
[18]	FPL, DCT	GMM	98	2	63.99% upstairs, 75.25% downstairs

upstairs activities. In [10], 4 downstairs are recognized as upstairs in 11 downstairs activities when J48 is used as a classifier, and 2 upstairs are recognized as downstairs in 13 upstairs activities. It is worthwhile to note that these are only 2 subjects in [9] and 6 subjects in [10], which shows that, even if a very small number of subjects, downstairs is often confused with upstairs, not to mention the case of a large number of subjects. So, we can see that discrimination between downstairs and upstairs remains a research challenge.

In view of this, in this paper, an algorithm for the discrimination between upstairs and downstairs is presented. Activity recognition results are based on acceleration data collected from a number of subjects (32 males and 10 females) using only one tri-axial accelerometer located alternatively on the waist belt, in the trousers pocket, and in the shirt pocket under naturalistic conditions. Both time-domain and time-frequency features are extracted only from the vertical axis of the accelerometer that has the largest contribution on classification of upstairs and downstairs.

2. Classification Algorithm

Our classification algorithm to distinguish upstairs and downstairs activities is shown in Fig. 1. The detailed procedure of the proposed algorithm is introduced as follows.

2.1 Calibration of the Vertical Acceleration Signal

In order to achieve robustness with regard to sensor location, subjects put the accelerometer on their waist belt, in

Manuscript received November 9, 2010.

Manuscript revised February 21, 2011.

*The authors are with South China University of Technology, Guangzhou, 510640, China.

E-mail: celwjin@scut.edu.cn

DOI: 10.1109/tnsis.2011.1173

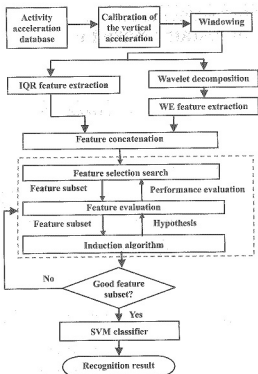


Fig. 1 The block diagram of the classification algorithm.

their trousers pocket and shirt pocket respectively. In addition to fixing it on the waist belt, the accelerometer located in the trousers pocket or shirt pocket is not fixed. As the sensor is not fixed to the body, it may move randomly in the pocket (e.g., rotation) which can continuously change the orientation between accelerometer and subject's body. In order to extract the actual vertical acceleration signal (acceleration signal of gravitational direction), a calibration algorithm [11] is used. The actual vertical acceleration signal is calibrated as:

$$a'_z = (a'_x(i), a'_y(i), a'_z(i)) \cdot D_\perp \quad (1)$$

where $(a'_x(i), a'_y(i), a'_z(i))$ is the dynamic acceleration signal; $D_\perp = \frac{(a'_{x_mean}, a'_{y_mean}, a'_{z_mean})}{\|(a'_{x_mean}, a'_{y_mean}, a'_{z_mean})\|}$ is the actual gravitational direction; $(a'_{x_mean}, a'_{y_mean}, a'_{z_mean})$ is an average acceleration vector calculated from tri-axial acceleration signal when the sensor is static.

2.2 Feature Extraction

The following two features are extracted from the calibrated vertical acceleration signal to identify upstairs and downstairs. (1) Interquartile Range (IQR) and (2) Wavelet Energy (WE). Figure 2 shows differences in the two feature values, which are effective to discriminate between upstairs and downstairs.

IQR is equal to the difference between the third and first quartiles in descriptive statistics. Just like variance and

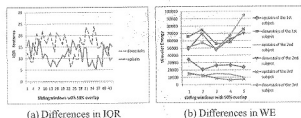


Fig. 2 Differences in feature values between upstairs and downstairs.

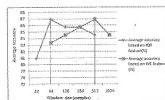


Fig. 3 Performance comparison for the different window size.

standard deviation, IQR is a measure of statistical dispersion, but not affected by outliers or extreme values. This feature can support discrimination between activities with similar mean values and avoid the effect on range caused by extreme values in the acceleration data. From Fig. 2(a), it can be seen that the distribution values of IQR for upstairs are very low when the distribution values of IQR for downstairs are high, and vice versa.

Wavelet energy (WE) is calculated as the sum of the squared decomposed wavelet coefficients of the vertical acceleration signal. Because the low-frequency components in the vertical direction correspond to the gravity, the high-frequency components are calculated as the energy of the vertical direction. Daubechies wavelet of order 5 is used to decompose the calibrated vertical acceleration signal five levels. Wavelet coefficients of high-frequency components in the four and five level are extracted. Fig. 2(b) shows that the wavelet energy of downstairs is far greater than that of upstairs. The form of the WE can be given by

$$WE = \sum_{i=4}^5 |cD_i|^2 \quad (2)$$

where cD_i are coefficients of details in the i level.

The IQR feature and WE feature are both extracted on sliding windows with 50% overlap which has been demonstrated success [12], [13]. Figure 3 illustrates the performance comparison for the different window size based on IQR feature and WE feature, respectively.

According to Fig. 3, the IQR feature is extracted using a window size of 64 samples and the WE feature is extracted using a window size of 512 samples, respectively. Then these two features are concatenated as a feature set.

2.3 Feature Subset Selection (FSS) with the Wrapper

For the above extracted feature set, some features may be

irrelevant or redundant and not contribute to improve the recognition accuracy. Furthermore, the computational speed may be slow because of the high dimension of the feature set. Thus, feature subset selection (FSS) is taken into account in the classification algorithm. The wrapper approach is one of the well-known approaches for FSS in machine learning [14] and can select effective feature subsets without ignoring the induction algorithm.

The FSS wrapper algorithm conducts a search for a good subset using the induction algorithm itself as part of the evaluation function. The accuracy of the induced classifiers is estimated using accuracy estimation techniques [14]. The wrapper approach to FSS is described as follows:

2.3.1 Feature Selection Search

A feature selection search requires a state space, an initial state, a termination condition, and a search engine [14]. The search space organization that we chose is such that each state represents a feature subset. For n features, there are n bits in each state, and each bit indicates whether a feature is present (1) or absent (0). Adding or deleting a single feature from a state is chosen to use as operators, which determine the connectivity between the states. The size of the search space for n features is $O(2^n)$, so it is impractical to search the whole space exhaustively, unless n is small. We chose a best-first search engine that starts with the empty set of features and searches forward. Termination condition is 5 backtracking, which depends on the search engine.

2.3.2 Feature Evaluation

Since we do not know the actual accuracy of the induced classifier, we use accuracy estimation [15] as both the heuristic function and the evaluation function. The accuracy estimation method used is five-fold cross-validation [15], repeated multiple times with a small penalty (0.1%) for every feature. The number of repetitions is determined by the standard deviation of the accuracy estimate.

2.3.3 Induction Algorithm

The induction algorithm used is a Support Vector Machine (SVM) [16] with One-versus-One strategy (OVO), in which a set of binary classifiers are constructed using corresponding data from two different classes. In classification we use the voting strategy of "Max-Wins" to produce the output. In case that two classes have identical votes, though it may not be a good strategy, now we simply select the one with the smallest index.

The induction algorithm is run on the dataset, usually partitioned into internal training and holdout sets, with different sets of features removed from the data. The feature subset with the highest evaluation is chosen as the optimal set on which to run the induction algorithm. The resulting classifier is then evaluated on an independent test set that was not used during the search. In this paper, our dataset is

tested using leave-one-subject-out cross-validation.

3. Experimental Design and Results

3.1 Experimental Data

The researchers usually collected data from a very small number of subjects, and each activity is often performed more than twice by the same subject [6], [9], [10], [20]. In this study, however, we use a more challenge dataset, SCUT-NAA [17], which contains 1278 samples of ten activities using only one tri-axial accelerometer in naturalistic settings. 42 different subjects (32 males and 10 females) placed the accelerometer alternatively on their waist belt or in their trousers pocket or shirt pocket as they performed each activity. The data generated by the tri-axial accelerometer was transmitted to a PDA wirelessly over Bluetooth. The only two activities, namely, downstairs and upstairs, are used in this paper. Figure 4 shows a subject's example of the acceleration signal along the x -, y -, and z -axis respectively.

3.2 Performance Comparison and Analysis

To validate the effectiveness of the proposed classification algorithm, we carry out a leave-one-subject-out cross-validation method. To justify the necessity and effectiveness of the proposed vertical acceleration calibration and FSS wrapper, the following recognition performance are studied. 1) The recognition results both with calibration and without calibration are compared in Table 2. These results show that the total average accuracy with calibration is 95.64%, increasing by 3.18% compared with the results without calibration. 2) Table 3 shows the recognition results both using the FSS wrapper and not using the FSS wrapper. The total average accuracy using the FSS wrapper is 95.64%, increasing by 13.90% compared with the results without the FSS wrapper step. This indicates that the FSS wrapper approach can significantly improve the recognition performance.

The performance comparison of the proposed features (IQR and WE) against the widely used time-domain features

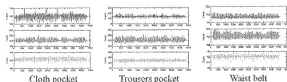


Fig. 4 The acceleration signal of downstairs for different sensor locations.

Table 2 Recognition Results with or without calibration (%).

Sensor location	Waist belt		Shirt pocket		Trousers pocket	
Calibration or Not	No	Yes	No	Yes	No	Yes
downstairs	97.62	100	90.48	92.86	97.62	100
upstairs	90.48	95.24	88.10	90.48	90.48	95.24
average	94.05	97.62	89.29	91.67	94.05	97.62

Table 3 Recognition results with or without FSS wrapper (%).

Sensor location FSS or Not	Waist belt		Shirt pocket		Trousers pocket	
	No FSS	FSS wrapper	No FSS	FSS wrapper	No FSS	FSS wrapper
downstairs	90.48	100	76.19	92.86	83.33	100
upstairs	83.33	95.24	80.95	90.48	76.19	95.24
average	86.90	97.62	78.57	91.67	79.76	97.62

Table 4 Accuracy based on three features for different sensor locations.

Sensor location Features	Waist belt			Shirt pocket			Trousers pocket		
	TDF	FFT	IQR+WE	TDF	FFT	IQR+WE	TDF	FFT	IQR+WE
downstairs	90.48	90.48	100	76.19	83.33	92.86	88.10	97.62	100
upstairs	92.86	92.86	95.24	85.71	92.86	90.48	90.48	90.48	95.24
average	91.67	91.67	97.62	80.95	88.10	91.67	89.29	94.05	97.62

Table 5 Accuracy based on three features for the mixed data.

Features	TDF	FFT	IQR+WE
downstairs	90.48	91.27	92.06
upstairs	90.48	90.48	97.62
average	90.48	90.87	94.84

(TDF) [9], [10], [17] and FFT features [17] is summarized in Table 4. The Four traditional time-domain features (mean, standard deviation, energy, and correlation) and the first 32 FFT coefficients are extracted from each axis of acceleration data, respectively, as time-domain feature set and FFT feature set. Then, the FSS wrapper is used for time-domain feature set and FFT feature set. Finally, SVM is also used as a classifier for these two features.

Overall, the recognition rates based on the proposed features are highest for every different sensor location, and the average of which outperforms time-domain features with a 8.34% accuracy improvement and FFT features with a 4.37% accuracy improvement, respectively. Although the sensor is located in different position, our proposed features perform better, the algorithm based on which recognizes upstairs and downstairs with 95.64% average accuracy.

Table 5 shows the recognition results based on TDF, FFT, and IQR+WE for the mixed data from all sensor locations, with the wrapper feature selection. From Tables 4 and 5, it is worthwhile to note that the average accuracy (94.84%) for the mixed data is only 0.80% lower than the average accuracy of three different sensor locations, which seems acceptable because the mixed data increases the complexity of the acceleration signal and further increases the difficulty of the classification. Even so, the recognition rate based on the proposed features is the highest for the mixed data. Obviously, all these results show the superiority of the proposed features, comparing with previous widely used features.

To demonstrate the proposed algorithm can effectively recognized the difference between downstairs and upstairs, we analyze the confusion matrices. Table 6 shows the confusion matrix for different sensor settings. It can be seen

Table 6 Confusion matrix for different sensor settings down = downstairs; up = upstairs.

		down = downstairs; up = upstairs							
		Recognized as							
		Waist belt		Shirt pocket		Trousers pocket		Mixed data	
		down	up	down	up	down	up	down	up
		42	0	39	3	42	0	116	10
down	up	2	40	4	38	2	40	3	123

that, in 42 subjects, no one's downstairs activity is recognized as upstairs and only 2 subjects' upstairs activities are recognized as downstairs when the sensor is located on the waist belt or in the trousers pocket. Even when the sensor is located in the shirt pocket, only 3 of 42 subjects' downstairs activities are recognized as upstairs and 4 subjects' upstairs are recognized as downstairs. Finally, for the mixed data from all sensor locations, only 3 of 126 upstairs activities are recognized as downstairs and 10 of 126 downstairs are recognized as upstairs. Overall, the misidentification between downstairs and upstairs has been effectively reduced.

4. Conclusions

This paper presents an algorithm for the classification of human upstairs and downstairs using a tri-axial accelerometer. First, the vertical acceleration signal is calibrated through estimating the gravitational direction. Then, two features (IQR and WE) are extracted from the calibrated acceleration signal. Thereafter, FSS wrapper approach is used to select an efficient feature set. Finally, SVM is adopted as a classifier. The average accuracy using the proposed algorithm for the different sensor locations is 95.64%, which outperforms time-domain features with a 8.34% accuracy improvement and FFT features with a 4.37% accuracy improvement, respectively. Although the sensor is located in different position, our proposed features perform better. The algorithm recognizes upstairs and downstairs with 94.84% accuracy even for the mixed data from all sensor locations. The experimental results have confirmed the effectiveness of the proposed algorithm.

Acknowledgements

This work is supported in part by NSFC (Grant No. U0735004 and 61075021) and Fundamental Research Funds for the Central Universities, SCUT (No. 2009ZZ0014). We sincerely thank the helpful comments and suggestions given by the anonymous reviewers.

References

- [1] E.S. Choi, W.C. Bang, S.J. Cho, J. Yang, D.Y. Kim, and S.R. Kim, "Beatbox music phone: Gesture-based interactive mobile phone using a tri-axis accelerometer," IEEE Int. Conf. Industrial Technology, pp.97-102, Hong Kong, Dec. 2005.
- [2] J.Y. Yang, J.S. Wang, and Y.P. Chen, "Using acceleration measurements for activity recognition: An effective learning algorithm for constructing neural classifiers," Pattern Recognit. Lett., vol.29,

- pp.2213-2220, 2008.
- [3] E. Tapia, S. Intille, and K. Larson, "Activity recognition in the home using simple and ubiquitous sensors," *Proc. 2nd Int. Conf. Pervasive Computing*, pp.158-175, Vienna, Austria, April 2004.
- [4] S.K. Song, J. Jang, and S. Park, "An efficient method for activity recognition of the elderly using tilt signals of tri-axial acceleration sensor," *Proc. 6th Int. Conf. Smart Homes and Health Telematics*, pp.99-104, IA, USA, June 2008.
- [5] C. Randell and H. Muller, "Context awareness by analyzing accelerometer data," *4th Int. Symposium on Wearable Computers*, pp.175-176, Atlanta, Georgia, USA, Oct. 2000.
- [6] N. Kern, B. Schiele, and A. Schmidt, "Multi-sensor activity context detection for wearable computing," *European Symposium on Ambient Intelligence*, pp.220-232, Veldhoven, the Netherlands, 2003.
- [7] M. Ermes, J. Pärkkä, J. Mäntylä, and I. Korhonen, "Detection of daily activities and sports with wearable sensors in controlled and uncontrolled Conditions," *IEEE Trans. Inf. Technol. Biomed.*, vol.12, no.1, pp.20-26, 2008.
- [8] A. Godfrey, R. Conway, D. Meagher, and G. Ólaighin, "Direct measurement of human movement by accelerometry," *J. Medical Engineering & Physics*, vol.30, pp.1364-1386, 2008.
- [9] N. Ravi, N. Dandekar, P. Mysore, and M.L. Littman, "Activity recognition from accelerometer data," *American Association for Artificial Intelligence*, pp.1541-1546, 2005.
- [10] D. Maguire and R. Frisby, "Comparison of feature classification algorithms for activity recognition based on accelerometer and heart rate data," *9th IT & T Conference*, pp.1-10, France, Oct. 2009.
- [11] Z. He, Z. Liu, L. Jin, L.X. Zhen, and J.C. Huang, "Weightlessness feature — A novel feature for single tri-axial accelerometer based activity recognition," *19th Int. Conf. Pattern Recognition*, pp.1-4, Florida, USA, Dec. 2008.
- [12] K.-V. Laerhoven and O. Cakmakci, "What shall we teach our pants?" *4th Int. Symposium on Wearable Computers*, pp.77-83, Oct. 2000.
- [13] R.W. Deval and S. Dunn, "Real-time motion classification for wearable computing applications," *Technical Report of MIT Media Laboratory*, pp.1-14, Dec. 2001.
- [14] R. Kohavi and G.H. John, "Wrappers for feature subset selection," *Artif. Intell.*, vol.97, pp.273-324, 1997.
- [15] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," *Int. Joint Conf. on AI*, pp.1137-1143, 1995.
- [16] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer Press, New York, 1999.
- [17] Y. Xue and L. Jin, "A naturalistic 3D acceleration-based activity dataset & benchmark evaluations," *IEEE Int. Conf. Syst., Man, Cybern.*, pp.4081-4085, Istanbul, Turkey, 2010.
- [18] R.K. Ibrahim, E. Ambikairajah, B. Celler, N.H. Lovell, and L. Kilmartin, "Gait patterns classification using spectrum features," *IET Irish Signals and Systems Conference*, pp.98-102, Galway, Ireland, June 2008.
- [19] F. Foerster, M. Sineja, and J. Fahrenberg, "Detection of posture and motion by accelerometry: A validation study in ambulatory monitoring," *Computers in Human Behavior*, vol.15, pp.571-583, Sept. 1999.
- [20] J. Mäntylä, J. Himberg, and T. Seppänen, "Recognizing human motion with multiple acceleration sensors," *IEEE Int. Conf. Syst., Man, Cybern.*, pp.747-752, AZ, USA, Oct. 2001.