Real Time Accelerometer-Based Gait Recognition Using Adaptive Windowed Wavelet Transforms

Jian-Hua Wang
Department of Electrical Engineering,
National Taiwan University,
Taipei, Taiwan, R.O.C.
bawan_eawan@yahoo.com.tw

Jian-Jiun Ding
Department of Electrical Engineering,
National Taiwan University,
Taipei, Taiwan, R.O.C.
djj@cc.ee.ntu.edu.tw

Yu Chen
Department of Electrical Engineering,
National Taiwan University,
Taipei, Taiwan, R.O.C.
hck13kimo@gmail.com

Hsin-Hui Chen
Department of Electrical Engineering,
National Taiwan University,
Taipei, Taiwan, R.O.C.
d98942018@ntu.edu.tw

Abstract—**This paper presents a real time gait recognition system using the wavelet transform. The activity signal is acquired from three-axis accelerometers on mobile phones. It is first decomposed into wavelet coefficients with eight levels. Several statistical measures, such as power, mean, variance, energy, and the energy of neighbor difference, are calculated from these coefficients. Furthermore, the adaptive window size is adopted to well fit the footstep of each person. The selected features are also adjusted adaptively to improve the accuracy. The simulation results show that the proposed method has reliable recognition accuracy both in the real-time and the long-term cases.**

I. INTRODUCTION

In recent years, with the fast development of Micro Electro Mechanical Systems (MEMS), the 3-axis accelerometer have become a powerful tool for acquiring activity signals and embedded in novel devices, such as the Wii and the smart phone, for several applications. For example, it is combined with the Hidden Markov Model (HMM) for gesture recognition in Wii-controllers [1] and mobile games [2]. These systems receive the quantized activity signals from the 3-axis accelerometer and classify them into different activities by using the HMM. This process has been widely used in acceleration signal processing. However, it only considers the effects of acceleration in the time domain.

To extract more discriminative features for activity recognition, both the time domain and the frequency domain information should be considered. In [3]-[5], the fast Fourier transform (FFT) was applied for frequency-domain analysis, and the features from both the two domains are combined to recognize walking or other activities. The FFT, however, cannot tell the time when each frequency component occurs. In [6][7], the wavelet Transform (WT) was adopted in their architectures. The advantage of the wavelet transform is that the wavelet coefficients can describe the intensity of different basebands at any time. The characteristic of gait pattern was exactly depicted from the detail coefficients of the WT. The result made a mechanism to distinguish whether the activity was walking or not. But this simple algorithm was not sufficient to analyze complicated or multiple activities. The fractal dimension was a meticulous analytical method in the WT [8][9]. In their work they found the trend which can present the differences between level walking and stairway walking. But gait pattern is a relative result due to subject's physical condition. Therefore, this method cannot precisely identify the activity of a person. In [10], they combined the WT with machine learning algorithms and proposed a complete recognition system. Although the accuracy of the system is high, it must utilize two accelerometers because their method did not take the advantages of the characteristics of wavelet coefficients, and leads to unnecessary devices.

Our goal is to develop a reliable real-time gait recognition system using a single smart phone. The system adopts the WT and classifies the activities into walking, ascending stairs, descending stairs and jogging from acceleration signals. In addition to that only one smart phone is required, the proposed system has the advantages of "real time", "high accuracy", "stability", and "adapting the window size with the change of the footstep", compared with other systems.

This paper is organized as follows: In Section II, our proposed method is described. The simulation results are presented in Section III. The conclusion is given in Section IV.

II. METHOD

A. Subjects and Experimental Set-Up

Our study focuses on young subjects: There are 8 male participants (age 25 ± 2, height 174.5 ± 3, weight 71.5 ± 5) in the experiments, and all of them were asked to perform four gait activities (walking, ascending stairs, descending stairs, and jogging). The experimental data was collected from the smart phone HTC Aria, which contains a 3-axis accelerometer inside. All subjects placed the device at a pocket on the hip

978-1-4577-1729-1/12/$26.00 ©2012 IEEE.

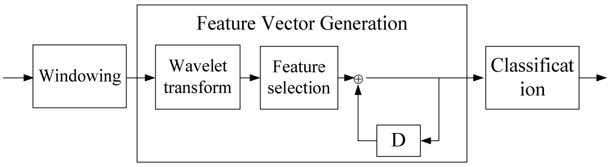


Figure 1. System flowchart.

because, as mentioned in [3], this place has higher recognition accuracy. The sampling rate is around 50 Hz. The window size is 25-point width and will adaptively change in our experiments. The windows did not overlap with each other and signals were obtained directly from these consecutive windows.

B. Wavelet Transform

The discrete wavelet transform [11] was adopted in our system. It decomposes a signal $x[n]$ into an approximation signal $x_{2^{\alpha},L}$ and a detail signal $x_{2^{\alpha},H}$ by using the mother wavelet function $g[k]$ and the scaling function $h[k]$, respectively. The definition of the WT is as follows:

$$x_{2^a,L}[n] = \sum_k x_{2^a-1,L}[k]g[k-2^a n], \tag{1}$$

$$x_{2^a,H}[n] = \sum_k x_{2^a-1,L}[k]h[k-2^a n] \tag{2}$$

where α means the stage level in wavelet decomposition. It is related to frequency resolution. The parameter k is the scaled shifting of the basis function in the time domain. These two functions emphasized the low-frequency and high-frequency bands of the signal $x[n]$ which have the similar behaviors with those of the low-pass and-high pass filters, respectively.

On the other hand, the original signal $x[n]$ can also be expanded by the mother wavelet function and the scaling function. It means that the signal $x[n]$ can be constructed from these two functions. Its definition is as follows:

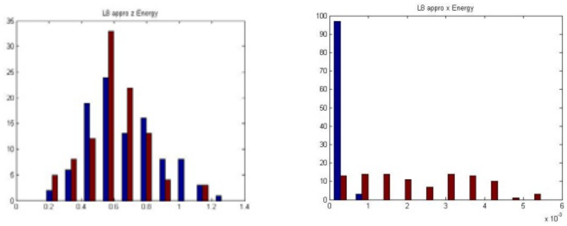
$$x[n] = \sum_{a=1}^{J}\left\{\sum_k x_{2^a-1,L}[k]h[n-2^a k] + \sum_k x_{2^a-1,L}[k]g[n-2^a k]\right\} \tag{3}$$

where J means the number of decomposition levels.

C. Feature Set Generation

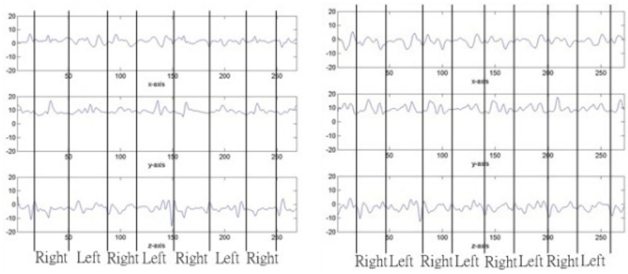
The flowchart of the proposed system is plotted in Fig. 1. There are totally 3 stages in the system. In the first stage, the signals obtained from the 3-axis accelerometer are separated window by window. In the second stage, the major features of the personal gait pattern are extracted with the aid of the WT.

Finally, machine learning algorithms are applied to classify the gait pattern into one of the four possible activities. We will discuss each stage in detail in Sections D, E, and F, respectively.



(a) z-axis (b) x-axis

Figure 2. The statistical diagrams of the energies of the wavelet coefficients at the 8th level. The blue bars mean the number of energy of the walking pattern and the red bars mean that of the upstairs pattern.



(a) Walking (b) Upstairs

Figure 3. The accelerations in the 3 axes measured by the accelerometer for the actions of "Walking" and "Upstairs".

D. Windowing

The acceleration signals are accessed in real time in the system. Therefore, the system must cut a sequence of data into consecutive windows before data analysis. There are two kinds of windows: the fixed-size one and the adaptive-size one. The fixed size window is 25 sampling point width, since the sampling rate of the device is about 50 Hz and the speed of walking is about 1.5 ~ 2 footsteps per second. Considering the different walking habits for different people, we proposed an adaptive size windowing method, which is composed of several steps as follows:

1) The initial window size is set as 40-point width.

2) The local minimum position of the acceleration signal (P_{min}) is searched from the initial 40 sampling points.

3) The distance between P_{min} and the first sample in the corresponding window must be longer than 15 sampling points. If not, find another P_{min}.

4) The duration of the window is set from the first data position to P_{min}.

5) The first data position of next window is P_{min} + 1. Return to *3)* and find the duration of the next window.

E. Feature Selection

Feature selection is an important process in the recognition system. The advantage of the WT is that the wavelet coefficients imply the details in different bands. The system can find out the most suitable features of the anterior-posterior, vertical, medio-lateral and signal vector magnitude (SVM) signals on these coefficients by the following equations:

- Power of maximum signal:

$$P_I = MAX\{I[n]\}^2 . \quad (4)$$

- Mean:

$$m_I = \frac{1}{N+1}\sum_{n=0}^{N} I[n] . \quad (5)$$

- Variance:

$$v_I = \frac{1}{N+1}\sum_{n=0}^{N}\{I[n]-m_I\}^2 . \quad (6)$$

- Energy:

$$E_I = \frac{1}{N+1}\sum_{n=0}^{N}|I[n]|^2 . \quad (7)$$

- The energy of neighbor difference:

$$NDE_I = \frac{1}{N}\sum_{n=1}^{N}|I[n]-I[n-1]|^2 \quad (8)$$

where I can be x, y, z, or SVM, and x, y, and z are the anterior-posterior, vertical, and medio-lateral signals, respectively. The last parameter SVM is

$$SVM = \sqrt{x^2+y^2+z^2} . \quad (9)$$

In sum, there are 5 × 4 = 20 features at each of the 8 levels in the system. Not every feature will be a distinctive feature. For example, from Figs. 2 and 3, it can be seen that the features acquired from x-axis and y-axis are useful for distinguishing the actions of walking and upstairs but those acquired from z-axis are not useful for distinguishing. Therefore, the proposed system only picks 40 features from wavelet coefficients.

The proposed system then put the feature set of the previous window with that of the current sample window together. This will generate a new feature set, which is helpful for analyzing the relationship between two adjective footsteps.

F. Classification

The final stage in our system is to determine what the action is from the WT feature set. The proposed system obtained 40 features from the block of feature vector generation in Fig. 1. There are several machine learning algorithms that can be used for classification, such as the Gaussian mixture model (GMM) [13], the test decision tree (J48) [15], and logistic regression [15]. The GMM can be performed by the Bayes net toolbox (BNT) [14].

III. SIMULATIONS

Two cross validation methods were adopted in simulations. There were 8 persons participating in the simulations. We asked them to perform the specific activities, such as walking, ascending stairs, descending stairs, and jogging.

A. The Accuracy of Each Procedure in the System

First, the confuse matrix of the proposed recognition system using different types of windows are compared. In Tables I-III, two fixed-size sampling window, one adaptive-size sampling window, and two adaptive-size sampling window are adopted, respectively. Ten-fold cross validation [15] was introduced for these experiments. Compared Table

TABLE I. CONFUSE MATRIX FOR THE FEATURE SET OF TWO FIXED-SIZE SAMPLING WINDOWS

Activities	***Walking***	***Upstairs***	***Downstairs***	***Jogging***	***Accuracy***
Walking	*307*	*34*	*47*	*6*	*77.92%*
Upstairs	*85*	*221*	*60*	*0*	*60.38%*
Downstairs	*63*	*12*	*362*	*12*	*80.62%*
Jogging	*5*	*3*	*13*	*435*	*95.39%*

TABLE II. CONFUSE MATRIX FOR THE FEATURE SET OF ONE ADAPTIVE-SIZE SAMPLING WINDOW

Activities	***Walking***	***Upstairs***	***Downstairs***	***Jogging***	***Accuracy***
Walking	332	0	0	0	100%
Upstairs	0	250	81	0	75.53%
Downstairs	0	37	299	36	80.38%
Jogging	1	5	19	350	93.33%

TABLE III. CONFUSE MATRIX FOR THE FEATURE SET OF TWO ADAPTIVE-SIZE SAMPLING WINDOWS

Activities	***Walking***	***Upstairs***	***Downstairs***	***Jogging***	***Accuracy***
Walking	332	0	0	0	100%
Upstairs	0	275	54	2	83.08%
Downstairs	0	19	323	30	86.83%
Jogging	1	4	14	356	94.93%

TABLE IV. CONFUSE MATRIX FOR THE PROPOSED SYSTEM WITH DECISION TREE.

Activities	***Walking***	***Upstairs***	***Downstairs***	***Jogging***	***Accuracy***
Walking	332	0	0	0	100%
Upstairs	0	284	39	8	85.8%
Downstairs	3	29	283	57	76.08%
Jogging	1	4	19	351	93.6%

TABLE V. CONFUSE MATRIX FOR THE PROPOSED SYSTEM WITH LOGISTIC REGRESSION.

Activities	***Walking***	***Upstairs***	***Downstairs***	***Jogging***	***Accuracy***
Walking	331	0	0	1	99.7%
Upstairs	0	318	11	2	96.07%
Downstairs	0	44	324	4	87.1%
Jogging	0	13	23	339	90.4%

III with Table I, it can be seen that the concept of the adaptive-size window, which is proposed in this paper, is indeed helpful for improving the accuracy of gait recognition. Moreover, when compared Table III with Table II, one can see that using the feature set acquired from two sampling windows is also helpful for improving the accuracy, since the features acquired from two sampling windows can reveal the relationship between two adjective footsteps. This relationship is hard to observe when extracting the features only from the current window.

B. Accuracies of Different Classification Algorithm

We then apply different machine learning algorithms in this section. In Tables III, IV, and V, the feature sets are all acquired from two adaptive-size sampling windows, but the used classification algorithms are the GMM, the decision tree, and logistic regression. All the results show that the proposed system has very high accuracy. Especially, when logistic regression is applied, the average accuracy of the proposed

TABLE VI. CONFUSE MATRIX FOR KWAPISZ'S SYSTEM [12] WITH LOGISTIC REGRESSIOIN.

Activities	*Walking*	*Upstairs*	*Downstairs*	*Jogging*	*Accuracy*
Walking	16	5	2	1	66.67%
Upstairs	4	19	1	0	79.17%
Downstairs	6	1	16	1	66.67%
Jogging	2	0	1	21	87.5%

TABLE VII. CONFUSE MATRIX FOR THE PROPOSED SYSTEM WITH LOGISTIC SYSTEM.

Activities	*Walking*	*Upstairs*	*Downstairs*	*Jogging*	*Accuracy*
Walking	272	35	6	19	81.93%
Upstairs	22	293	16	0	88.52%
Downstairs	19	36	317	0	85.22%
Jogging	18	8	16	332	88.77%

algorithm is 93.32%. These results prove that the proposed system works well no matter what machine learning algorithms is applied.

C. Compared with Other Systems

The final simulation compares our proposed system with Kwapisz's system [12]. In this simulation, we introduce eight-fold cross validation because their algorithm acquires feature set from the 10 second data. The number of features of our system and Kwapisz's system are 40 and 43, respectively. The experiment results are shown in Tables VI and VII. The simulation results in Tables VI and VII showed that proposed system has higher accuracy than Kwapisz's. The simulation results indicated that the proposed method has the **advantages** as follows:

- **Real-time**: Because the length of window period of the proposed system is adaptive and close to a footstep period, our proposed system can process the 3-axis acceleration signals immediately and classifies these signals into a specific result in real time. By contrast, in Kwapisz's system, the feature set is produced directly from 10 sec data, which is a regular long period. However, in real cases, people merely perform the same activity in 10 sec. It means that their system may not have good performance when a person changes his activities at an uncertain period.
- **Stability**: The simulation results showed that the proposed system has a stable result, and the results maintained in high accuracies (always above 81%). In comparison, when using Kwapisz's system, although in some case the accuracy is high, sometimes the accuracy drops to 66%. The simulation results show that the feature set produced by the WT in the proposed system has higher reliability for gait recognition.
- **Adaptation**: In the proposed system, the adaptive window is applied, and the selected wavelet features are adjusted after each time of gait recognition.

IV. CONCLUSION

In our work, the WT, the adaptive window, and two-window features are utilized to design an accurate algorithm for gait recognition. Because the WT can analyze the 3-axis acceleration signal effectively and emphasize the characteristics of each activity in a short period, it provides more effective features for gait recognition. Moreover, the adaptive window can well fit the step size of different persons and the two-window features can explore the relationship between two adjacent footsteps, they can improve the accuracy of gait recognition. Furthermore, with a little modification, our proposed system can also be applied for other gesture recognition and activity classification problems.

REFERENCES

[1] T. Schlömer, B. Poppinga, N. Henze, and S. Boll, "Gesture recognition with a Wii controller," *TEI '08: Proceedings of the 2nd International Conference on Tangible and Embedded Interaction, ACM*, New York, pp. 11-14, 2008.

[2] M. Joselli and E. Clua, "GRmobile: A framework for touch and accelerometer gesture recognition for mobile games," *Brazilian Symposium on Games and Digital Entertainment*, pp. 141-150, 2009.

[3] L. Bao and S. Intille, "Activity recognition from user-annotated acceleration data," *Proc. 2nd Int. Conf. Pervasive Comput.*, pp. 1-17, 2004.

[4] M. Ermes, J. Parkka, J. Mantyjarvi, 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, Jan. 2008.

[5] D. M. Karantonis, M. R. Narayanan, M. Mathie, N. H. Lovell, and B. G. Celler, "Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring," *IEEE Trans. Inf. Technol. Biomed.*, vol. 10, no. 1, pp. 156-167, Jan. 2006.

[6] P. Barralon, N. Vuillerme, and N. Noury, "Walking detection with a Kinematic sensor: Frequency and wavelet comparison," *Proceedings of the 28th IEEE EMBS Annual International Conference*, New York, USA, 2006.

[7] N. Bidargaddi, S. Antti, L. Klingbeil, and M. Karunanithi, "Detecting walking activity in cardiac rehabilitation by using accelerometer," *Proceedings of Third International Conference Intelligent Sensors, Sensor Networks and Information Processing* Melbourne, Australia, pp. 1-7, 2007.

[8] M. Sekine, T. Tamura, M. Akay, T. Togawa, and Y. Fukui, "Analysis of acceleration signals using wavelet transform," *Methods of Information in Medicine*, pp. 183-185, 2000.

[9] M. Sekine, T. Tamura, M. Akay, T. Togawa, and Y. Fukui, "Discrimination of walking pattern using wavelet-based fractal analysis," *IEEE Trans. Neural. Syst. Rehabil. Eng.*, vol. 10, pp. 188-196, 2002.

[10] J. Mantyjarvi, J. Himberg, and T. Seppanen, "Recognizing human motion with multiple acceleration sensors," *Proceeding of the IEEE International Conference on Systems, Man, and Cybernetics*, pp. 747-752, 2001.

[11] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 1998.

[12] J. R. Kwapisz, G. M. Weiss, and S. A. Moore. "Activity recognition using cell phone accelerometers," *ACM SIGKDD Explorations Newsletter*, vol. 12, Dec. 2010.

[13] D. A. Reynolds, "Gaussian mixture models," in *Encyclopedia of Biometric Recognition*, Springer, Feb. 2008.

[14] K.P. Murphy "The Bayes net toolbox for MATLAB," *Technical Report* 94720-1776, Dept. Comp. Sci., University of California at Berkeley, CA, 2001.

[15] E. Alpaydin, *Introduction to Machine Learning*, 2nd Ed., MIT Press, 2010.