Cost-Effective Activity Recognition on Mobile Devices

Jian Cui
Tsinghua University
Beijing, China
cuijian.hh@gmail.com

Bin Xu
Tsinghua University
Beijing, China
xubin@tsinghua.edu.cn

ABSTRACT

Activity recognition using motion sensors, which denotes a person's posture, has become one of the most important research topics in body sensor network. With the rapid development of monitoring and sensing applications, activity recognition on mobile devices or portable platforms has drawn lots of attentions. Constrained by computing capability and energy budget, activity recognition on mobile devices faces the challenge of hungry energy consumption. Most of the existing work focus on modeling and recognizing activities accurately,however, without computational cost consideration. In this paper, we present WCF: Wavelet Coefficients based Features for activity recognition, which are cost-effective on feature extraction. WCF are extracted from wavelet domain of the sensory raw data and describes the *timbre* and *rhythm* properties of activities. Feature space in WCF is hierarchical. Compared with other features, WCF are extracted with less computational cost and redundancy. And WCF have better classification accuracy on activity recognition task. Experiments are conducted on the large public data set USC-HAD to recognize 11 kinds of activities, and our approach outperforms others by reducing 55% $\sim$ 75% computational cost as well as achieving 96.23% classification accuracy.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous; D.2.8 [**Software Engineering**]: Metrics—*complexity measures, performance measures*

General Terms

Theory

Keywords

activity recognition, feature extraction, cost reduction

BODYNETS 2013, September 30-October 02, Boston, United States

DOI 10.4108/icst.bodynets.2013.253656

1. INTRODUCTION

With the development of sensor technologies, representing and understanding human states with sensory data have become an important application in body sensor network. Among these states, recognizing human activities and behaviors have been a key topic, especially in human interaction, human sports analyzing, and healthcare assistance. Activity Recognition(AR), which identifies the activity that a user performs, is attracting tremendous amount of attentions. The advance of MEMS technologies make sensors smaller, richer in categories, and more powerful. Sensors such as acceleration sensors, gyroscope sensors, magnetometer sensors, and audio sensors have been widely embedded in personal equipments to provide rich physical readings. With these sensory data, we have a better way to model and understand human activities.

Activity recognition applications fall across a broad range of disciplines. Clinicians and medical researchers can monitor and diagnose patients using the continuous activity data. AR can be helpful to doctors diagnosing such conditions as they monitor daily activities in order to detect deviations from a typical routine or deterioration of a patient's current physical status [1]. Similar work have been done in the field of smart homes and workplaces automation using stationary sensors to detect the locations and activities of people in the room [2]. Fall detection, especially for the elder, can help prevent many serious injury. Mobile activity recognition systems can autonomously send an alert for help if the elder have fallen down [3]. Higher level activity recognition (such as 'having lunch','working') may be used to customize mobile phone to benefit the user by disabling incoming calls or setting the device to silent mode [4]. These applications are almost done through continuous sensing on mobile devices, such as smart phones(equipped with motion sensors) or portable equipments with sensor nodes.

Recent work on AR include: prototyping wearable sensor systems [2, 3, 1], developing pattern recognition and machine learning algorithms to model and recognize activities [5], and building human activity data sets [6]. Most of these work focus on recognizing human activity more correctly or analyzing them more deeply. But few research focus on recognizing activities efficiently with less energy consumption to apply it widely in human's normal life. However, many problems exits in these work. As discussed previously, providing status of people's activity continuously on portable devices is very meaningful and practical. Activity recognition task on these mobile and portable devices face one common problem: energy limitation.

In this paper, we propose a novel method of feature extraction called Wavelet Coefficients based Features(WCF). Our main contributions of WCF include: 1)WCF are extracted with less computational cost compared with traditional features. 2)WCF have less redundancy among features in theoretically. 3)WCF describe timbre and rhythm characters of activity signals, and improves the classification accuracy of activity recognition task. As far as we know, we are the first to apply wavelets in feature construction of activity recognition.

The paper is organized as follows. Section 2 mainly presents the preliminary and related work of feature extraction on activity recognition task. Section 3 describes our approach of WCF in detail. Section 4 conducts experiments on public dataset to evaluate our design. Finally, section 5 concludes our paper.

2. PRELIMINARY AND RELATED WORK

2.1 Processing Steps of Activity Recognition Task

Figure 1 shows the main steps of activity recognition task. Sensory raw data are collected from embedded sensors in first step. In this step, sensory raw data are segmented into separate instances for all channels, as shown in figure 2. Instance segmentation are based on the sliding window with overlap strategy. One instance is obtained in a window. This overlap strategy can help avoid losing information between successive instances. Through feature extraction on windows, raw data are transformed into feature candidates at second step. In the third step(optional), we select a subset of the feature candidates as the final feature vector. At last, we use the classifier trained from machine learning algorithm to recognize human activities. The classifier is trained from labeled instances off-line through the same processing steps.

Two main parts in the activity recognition task are: 1)transforming sensory raw data into feature vector; 2)learning the classifier from labeled instances. More importantly, these two parts are also the main energy consumers in activity recognition task. Although the cost of learning algorithms is very high, it's often done off-line. This cost has little affects on energy efficiency. However, the cost of the feature extraction can not easily be ignored. This step needs to scan the whole window more than one time and conduct complex numeric calculations, which must been done online. From the point of application, cost budget from upper application provided for AR is limited. This means that feature extraction should have less computational complexity. In this paper, we focus on the second part to propose a better approach to reduce the computational complexity of feature extraction as well as reaching higher classification accuracy.

2.2 Traditional Feature Extraction

Researchers have developed many methods to extract features from activity signals. Two kinds of features are proved to be effective in activity recognition: statistical features and physical features [7]. These two kinds of features are designed from different points of view and describe different characters of the original sensory data.

2.2.1 *Statistical Features*

Statistical features(SF) are used to extract mathematical statistical characters of the sensory data. SF are regarded as the general features which describe the overall information of raw data. The cost of each feature in SF is low, however, the task of activity recognition always needs a large set of SF to get a satisfactory classification accuracy. These features are computed from each axis of both accelerate and gyroscope sensors. Table 1 shows the statistical features used in this paper. Values of the cost column in table 1 is the computational cost of each feature. Some of them have been intensively investigated in previous studies and in proven to be effective for activity recognition. The exact meaning of each feature can be found from [8, 9]. We don't explain in detail because of space limitation.

We use the CPU running time on the same PC to measure the computational cost. The values of computational cost vary among different computers. However, we only care about the relative values among features for comparison and evaluation.

2.2.2 *Physical Feature*

The second set of features are called physical features(PF), which are derived based on our physical interpretations of human motion. Physical features are calculated based on the physical meanings of human activity. On the other hand, PF have different practical meanings in different kinds of data. In comparison, most of the physical features are calculated based on data from multiple sensor channels. In other words, sensor fusion is performed at feature level for physical features. Table 2 shows the physical features used in this paper and their computational cost. We can find from the table, most of PF cost more than SF. Energy features in PF, such as 'Average Acceleration(Rotation) Energy', are calculated based on FFT making the cost higher than others.

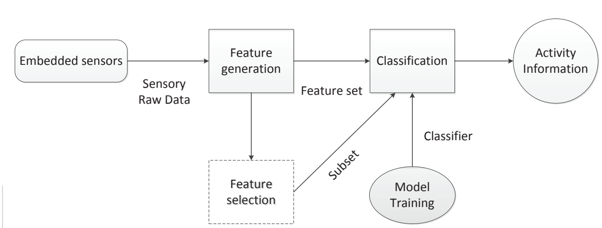


Figure 1: Processing steps

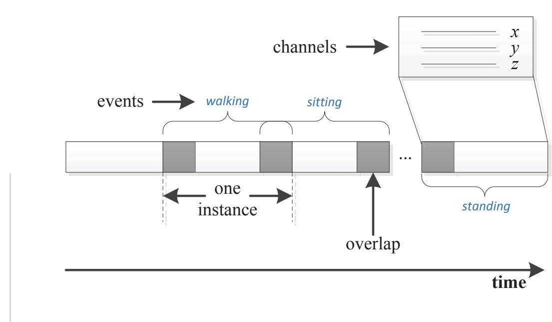


Figure 2: Segmentation strategy

Table 1: Statistical features

Statistical feature	Cost(ms)
Mean	0.0363
Median	0.0776
Standard Deviation	0.1146
Variance	0.0042
Averaged derivatives	0.0350
Root Mean Square	0.0302
Skewness	0.3215
Kurtosis	0.3070
Interquartile Range	0.4488
Zero Crossing Rate	0.0208
Mean Crossing Rate	0.0254

Table 2: Physical features

Physical feature	Cost(ms)
Movement Intensity mean	0.1896
Movement Intensity variance	0.5242
Normalized Signal Magnitude Area	0.0732
Eigenvalues of Dominant Directions	0.4101
Averaged Velocity along Heading	0.0588
Averaged Velocity along Gravity	0.0288
Averaged Rotation Angles on Gravity	0.0260
Dominant Frequency	0.9060
Energy	1.0670
Averaged Acceleration Energy	1.1659
Averaged Rotation Energy	1.1677

2.2.3 Disadvantages of SF and PF

In the previous section, we have presented the definitions of statistical and physical features. However, many problems exist in these features.

First, the computational cost of features extraction is very high, especially for physical feature calculation in table 1 and 2. PF achieve better accuracy than SF through more expensive calculations. However, PF are designed for specific sensor types. General features can be applied more widely and have better scalability. Second, there are bulk redundancy in both kinds of features, especially for statistical features. There are no strategies to guarantee redundancy generation. Redundant features do not provide information to improve the classification accuracy. Some even confuse the classifier rather than help discriminate these activities. If too much redundancy exists among features, the stability of the feature selection will be very poor. Finally, for statistical features usage, the high vector dimensionality makes the performance degrade sharply when there are not enough training data to learn all the parameters of activity models reliably.

3. OUR APPROACH

To address the problems above, we propose a wavelet coefficients based features extraction method, we call it WCF for short in following sections. WCF are designed for general purposes so that when new sensor types are involved, feature extraction can still be efficient. WCF provide a hierarchical feature space for sensory data. And we can select preferred subset on demand. WCF reduce the computational cost of feature extraction and increase the classification accuracy.

3.1 Wavelet Theory and Application

Wavelet transformation is a tool that divides data, functions, or operators into different frequency components. Each component has a resolution that matches its scale. In wavelet theory, $L^2(R)$ are used to stand for the wavelet function space [10]. The fact that $L^2(R)$ is decomposed into an infinite wavelet subspace is equivalent to the statement that $\psi_{j,k}$,$j,k \in Z$ span an orthonormal basis of $L^2(R)$. An arbitrary function $f \in L^2(R)$ is expressed as:

$$f(x) = \sum_{j,k \in Z} d_{j,k}\psi_{j,k}(x), \tag{1}$$

where $\psi_{j,k}(x)$ is called the *wavelet basis function* and $d_{j,k} = \langle f, \psi_{j,k} \rangle$ is called *wavelet coefficients* [11].

Note that j controls the resolution(frequency domain) and k controls the location(time domain). The *wavelet coefficients* $d_{j,k}$ is a combination of time and frequency information. If data in some location are relatively smooth (it can be represented by low-degree polynomials), then its corresponding wavelet coefficients should be fairly small by the vanishing moment property of wavelets.

Large coefficient mean important information at this moment on this frequency band. Wavelet transformation also has a hierarchical and multi-resolution decomposition structure property. Taking the popular Discrete Wavelet Transformation(DWT) [11] as an example, a n-level DWT transforms the original raw data into wavelet domain on different subbands. The coefficients can be grouped into two parts, one subband of low-frequency coefficients and n subbands of high-frequency coefficients. High-frequency coefficients are the detail information of original data in different scales, which are described as D_n at level n. Low-frequency coefficients are the approximation of the original data, which are described as A. These subbands construct the hierarchical structure of wavelet coefficients. Because of the orthogonality of the wavelet basis $\psi_{j,k}$ [12], features extracted from different subbands have less redundancy and relevance. Additionally, we can also follow [13] using *symmetrical uncertainty(SU)* as the correlation measure. $SU_{x,y}$ donates redundancy among features extracted from $d_{x,k}$ and $d_{y,k}$. $SU_{x,c}$ and $SU_{y,c}$ can describe the relevance between class concept with features extracted from $d_{x,k}$ and $d_{y,k}$, respectively. We use SU to describe the relevance in equation 2.

$$SU_{x,y} < \min\{SU_{x,c}, SU_{y,c}\} \tag{2}$$

where x, y is decomposition level in DWT.

3.2 WCF Design

Based on the analysis in previous section, we use DWT as a preprocessing step. First, we transform the sensory data into wavelet domains through a n-level DWT composition. Then, we calculate features from each subband of the wavelet domains. The detail of WCF extraction is described in algorithm 1.

WCF extraction algorithm takes sensory data window, decomposition level and window size as the input. Output are the features of instances. The symbols in algorithm 1 are defined as follow in table 3.

3.2.1 Sensory Raw Data Transformation

Activity raw data are often collected through multiple directions based on the devices or sensor nodes. For motion sensors, accelerate and gyroscope sensors collect data on x,

Table 3: Definition of the symbols in algorithm 1

symbol	definition
$Activity_{instance}$	store one instance of the sensory data as $[s_c^1, s_c^2, \ldots, s_c^k]$
wcf	feature set that consisted of $wcf_{timebre}$ and wcf_{ryhthm}
C	the channel set of the sensory data
dwt	Discrete Wavelet Transformation
A	store the low-level coefficients of dwt
D_j	store the high-level coefficients of dwt
$mean$	calculate the mean of the instance
$variance$	calculate the variance of the instance
$skewness$	calculate the skewness of the instance
$energy$	calculate the energy of the instance
$normalize$	normalize the features across all the instances

Algorithm 1 WCF extraction

Require:
1: The sensory data window $Activity_{instance}$;
2: The decomposition level n;
3: The window size k;
Ensure:
4: WCF feature set $wcf = \{wcf_{timbre}, wcf_{rhythm}\}$;
5: $wcf \leftarrow \emptyset$;
6: **for** each $c \in C$ **do**
7: $\quad window_c \leftarrow [s_c^1, s_c^2, \ldots, s_c^k]$;
8: $\quad A \leftarrow dwt(window_c, `a`)$;
9: $\quad$ **for** $j = 0$ to n **do**
10: $\quad\quad D_j \leftarrow dwt(window_c, `d`, j)$;
11: $\quad$ **end for**
12: $\quad$ **for** each $subband$ in $\{\biguplus D_j \cup A\}$ **do**
13: $\quad\quad wcf_{timbre} \leftarrow wcf_{timbre} \cup mean(subband)$;
14: $\quad\quad wcf_{timbre} \leftarrow wcf_{timbre} \cup variance(subband)$;
15: $\quad\quad wcf_{timbre} \leftarrow wcf_{timbre} \cup skewness(subband)$;
16: $\quad\quad wcf_{rhythm} \leftarrow wcf_{rhythm} \cup energy(subband)$;
17: $\quad$ **end for**
18: $\quad wcf \leftarrow wcf \cup normalize(wcf_{timbre})$;
19: $\quad wcf \leftarrow wcf \cup normalize(wcf_{rhythm})$;
20: **end for**
21: **return** wcf;

y, and z axis respectively. In this paper, we call these directions as *channels*(at line 6 algorithm 1). The original activity instances(at line 1 algorithm 1) can be formatted as in equation 3.

$$Activity_{instance} = \{{s_c}^t | c \in C, t \in [T_1, T_m]\} \tag{3}$$

where C is the channel set of the instance, and $T_m - T_1$ is the length of instance window k(at line 3 algorithm 1). The first step of WCF extraction transforms $Activity_{instance}$ into wavelet domains through n level DWT(from line 7 to line 10 algorithm 1). Parameters 'a' and 'd' stands for getting low-frequency and high-frequency coefficients respectively.

3.2.2 Definition of Timbre and Rhythm

We consume that different activities have their own timbre and rhythm characters just like music. For example, dancing is just like music and is made up of different low-level activities. Different kinds of music have different characters in timbre and rhythm. WCF are designed to extract the timbre and rhythm of human activities. After transformation, sensory raw data are divided into different subbands. We regard the histogram of each subband as a probability distribution. From probability theory, a probability distribution can be characterized by its moments [14]. In this paper, we define timbre and rhythm based on music genres classification problem [15]. **Timbre features** are defined as the *mean*, *variance*, and *skewness* of each subband. These calculation are the same as the ones in SF (table 1) extraction. **Rhythm features** are defined as *energy* feature [15]. Energy features are calculated as the percentage of current subband energy in sum of all subband energy, as expressed in equation 4. These features are all calculated on the wavelet domain(line 12 to line 17 algorithm 1).

$$energy(subband) = \sum |d_{subband}| / \sum_{j=1}^{n} \sum |d_j|. \tag{4}$$

3.2.3 Feature Normalization

For a feature set, features are calculated based on different methods so that the unit of features are different. Without a normalization among features, most learning algorithms will have poor performance. In this paper, we normalize all the features into a uniform scale using equation 5[7].

$$f_{normalization} = (f_{raw} - \mu)/\sigma \tag{5}$$

where μ and σ are the empirical mean and standard deviation of a particular feature across all activity classes.

4. EXPERIMENTAL RESULTS AND EVALUATION

In this work, we choose Support Vector Machines(SVM) as our main learning algorithm to evaluate the efficiency. Because the dimensionality of feature set ranges from 10 to 100. SVM has been proven effectively in handling this kind of feature set [16]. And we also use k-Nearest Neighbor algorithm(kNN) and J48 decision tree to help evaluate our WCF in activity classification. Without loss of generality, we use precision, recall and F1-score to evaluate feature sets. Window size in this paper is set as 500 with 100 overlap. Wavelet basis function used in WCF extraction is 'db5' which has best effects in this work [12].

4.1 Dataset

In our work, we use USC-HAD dataset [6] which are collected on the sensing platform called MotionNode. This dataset is collected from 13 subjects with divergence in gender, age, height, and weight. The type of sensors are accelerometer and gyroscope on 3-axis respectively, and six channels in total. 11 activities are used in this paper: walking forward, walking left, walking right, walking upstairs, walking downstairs, running forward, jumping, sitting, standing, sleeping, and elevator. We divide the original data stream into fixed length window with 20% overlap just as discussed in section 2.

4.2 Evaluation of SF and PF

In this section, we first evaluate the computational cost and classification accuracy of SF and PF. We use the classification accuracy of SVM classifier on different feature sets.

Figure 3 shows the accuracy-cost with the SF adding into feature vector. The left y axis stands for classification accuracy. The right one stands for total computational cost of current feature set. Heuristically, we use a simple feature selection strategy in our evaluation. We add a group of SF each time into the feature vector and evaluate the classification accuracy. One group is one set of features extracted from the six channels. There are totaly 66 SF from six channels(11 kinds of SF in each channel). We can get that the accuracy increases slowly as features are added into the feature vector. However, the computational cost increases sharply after the sixth group of features added.

Figure 3: Accuracy-cost of SF

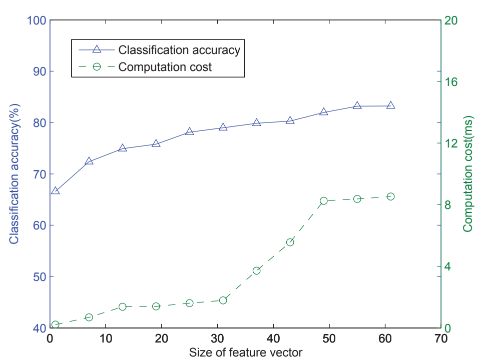


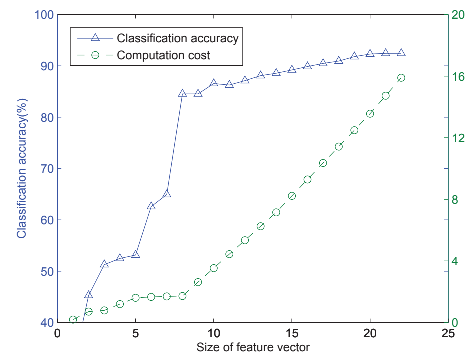
Figure 4 shows the accuracy-cost of PF similar with SF. Almost every PF is extracted from multiple sensor channels. And there are 22 features in total. So the average cost of PF may be higher than SF. For PF, the evaluation strategy is more simple. We add one PF each time into feature vector. We can see that accuracy increases sharply when the size ranges from 5 to 10, however, the cost also increases suddenly.

Generally speaking, PF have better performance than SF in accuracy but has higher computational cost.

4.3 Evaluation of WCF

In this section, we evaluate WCF from two aspects: accuracy improving and cost reducing. WCF combine the advantages of SF and PF. The specific features(such as PF) with practical meanings for activity can improve the accuracy. And the general features(in SF) save computational cost. WCF approach costless general calculation to extract specific features.

Figure 4: Accuracy-cost of PF



WCF have many feature candidates in a hierarchical space. In practical applications, a subset of WCF is selected based on demand. Heuristically, we choose timbre and rhythm features from different subbands as a subset to best approximate the WCF full set.

First, we should find the most valuable subband to calculate timbre features. We take decomposition level of DWT $n = 7$. Table 4 shows the classification accuracy of using each subband's timbre features independently. D_k stands for the timbre features on $k-th$ high-frequency subbad across all channels. And A stands for the timbre features on the low-frequency subband. We can see that A is the most important features of WCF, and using only A can archieve 92.2% accuracy.

Table 4: Accuracy of each subband's timbre features

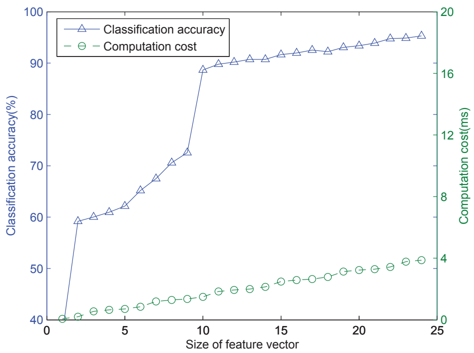
Timbre	Accuracy(%)
D_1	42.6
D_2	48.5
D_3	53.6
D_4	55.1
D_5	55.0
D_6	56.5
D_7	51.5
A	**92.2**

Second, we add A from all channels into feature vector. Then we add rhythm features of one subband D_k into feature vector each time. Figure 5 shows the classification accuracy when each subband's rhythm feature D_k is added into feature vector.

Table 5: A_α timbre features + subband-n rhythm features

Feature set	Accuracy(%)
$A + D_1$	94.30
$A + D_2$	94.16
$A + D_3$	94.73
$A + D_4$	93.73
$A + D_5$	94.20
$A + D_6$	**95.29**
$A + D_7$	93.63

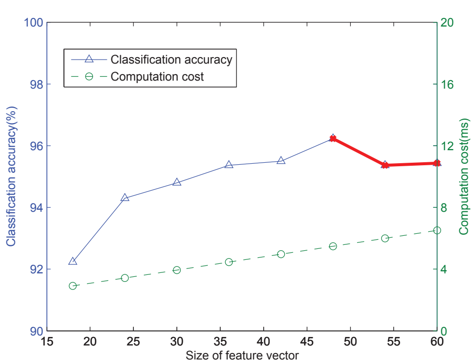
Figure 5: Accuracy-cost of WCF'



We can see that timbre features A with level-6 rhythm features D_6 of each channel can reach 95.29% accuracy. In following parts, we call the subset composed of A timbre and D_6 rhythm features as WCF' in short. We have proved the accuracy performance of WCF'. But we care more about the computational cost of WCF'. The timbre features in WCF' are same as in SF in calculating $mean, variance, skewness$. Feature calculation has nothing different but on wavelet coefficients. We have mentioned that the computational cost of wavelet transformation is not high. The computational complexity is $O(N)$. Here, we quantify it as the same way with the other cost measurements, CPU running time. The cost of 1-level decomposition on a window is about 1×10^{-1} ms, and 7-level is about 3.8×10^{-1} ms. We share the decomposition cost to every feature equally. Figure 5 shows the accuracy-cost of WCF'. We can find that WCF' reduce much more cost than SF and increase much more accuracy than PF. Similarly with PF, when the size over 10, the accuracy increase slowly. However, cost of PF increase sharply but WCF' remains steady.

We can conclude that WCF' are cost-effective and have better accuracy. In addition, we use different classification algorithms to evaluate WCF' from a general view. In this part, we use kNN and decision tree(J48) with SVM to evaluate feature set together. Table 6 shows best accuracy-cost performance for SF, PF, and WCF' using different classification algorithms respectively. We can see that WCF' outperforms SF and PF both in accuracy and cost.

Figure 6: Best-effort of WCF



So far, we have proven the effectiveness of WCF' on computational cost and classification accuracy compared with SF and PF. However, we have still not made best-effort for WCF to get the the max classification accuracy. WCF' are a subset selected from WCF heuristically. In detail, we only involve one rhythm($energy$) feature for each instance. With the decomposition level increasing and more rhythm features being added to feature vector, the accuracy would make a further improvement. We add one rhythm feature of six channels into feature set with A_α each time, and measure accuracy and cost. After adding all rhythm features, we get figure 6. We can find much information in this figure. First, we can get higher accuracy as cost increase, and we get new best accuracy 96.23%. Second, we find that adding level-6 and 7 rhythm features will decrease the accuracy. This means that these two level don't make any contribution to accuracy when cooperating with the other five levels' rhythm features. This is a general problem in feature selection. We will do future research on feature selection based on WCF.

Table 6: Performance of SF, PF, and WCF' on SVM, KNN, and J48

	Accuracy(%)			Cost(ms)	Size
	SVM	KNN	J48		
SF	83.23	81.83	76.88	8.5	66
PF	92.43	90.62	89.52	15.9	22
WCF'	**95.29**	**93.15**	**93.55**	**3.9**	24

The amount of instances ranges variously among activities. The overall accuracy is a mixed criterion for evaluation. Additionally, we use precision, recall, and F-measure to evaluate different feature sets together. Table 7 shows the classification results for each activity from different criterions. We can see from the table that the accuracy performance of WCF also has a big improvement.

Figure 7: Accuracy among classifying different amounts of activities

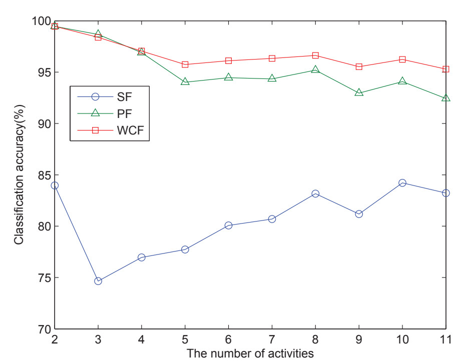


Figure 7 shows the performance of WCF in classifying different amounts of activities. For different kinds of requirements, the activities set we aim to recognize are different. A better approach should have steady performance across different tasks. In figure 7, PF and WCF always outperform SF. And PF have no obvious differences with WCF when classifying few activities. However, WCF outperform PF when the amount of activities increases.

Table 7: Accuracies of Activity Recognition

	% of Activities Correctly Classified									instances
	WCF			SF			PF			
	precision	recall	F1-score	precision	recall	F1-score	precision	recall	F1-score	
walkingF	97.99	94.82	96.38	89.75	88.62	89.18	88.92	90.48	89.69	896
walkingL	97.28	95.97	96.62	88.35	89.14	88.74	87.60	86.36	86.98	588
walkingR	95.87	97.42	96.64	87.92	85.37	86.62	86.59	89.02	87.79	630
walkingU	94.40	96.05	95.22	86.33	87.43	86.88	85.11	84.21	84.66	465
walkingD	89.77	95.07	92.34	92.18	93.23	92.70	91.68	90.73	91.20	430
runningF	100	99.48	99.74	93.50	92.41	92.95	92.45	92.22	92.33	384
jumping	100	100	100	83.68	84.14	83.91	83.99	84.13	84.06	212
sitting	89.04	93.71	91.31	87.25	88.42	87.83	88.39	86.54	87.46	603
standing	94.38	82.89	88.27	90.26	91.14	90.70	87.23	82.25	84.67	535
sleeping	100	100	100	92.05	91.65	91.85	93.23	90.15	91.66	899
elevator	82.78	93.71	87.91	80.17	82.17	81.16	81.25	83.69	82.45	360
overall	96.23			89.97			91.08			6002

5. CONCLUSION

In this work, we propose a cost-effective feature extraction method for activity recognition. We call it Wavelet coefficients based Feature extraction(WCF). WCF can be widely used on energy limitation devices to perform activity recognition task in an efficient way. WCF reduce computation cost and improve classification accuracy in feature extraction step. Besides computational cost reduction, WCF provide a hierarchical feature space in a general way. It's a feature specific design for activity recognition but can still be easily applied to other sensor data type. Through sufficient experiments on USC-HAD dataset, WCF outperform SF and PF both on computational cost reduction and accuracy improvement.

In our future work, we will focus on improving WCF in order to select feature adaptively based on different datasets and requirements. WCF shall also provide user-defined tradeoff between accuracy and cost.

Acknowledgment

This work is partly supported by NSF of China, with grant No.61170212.

6. REFERENCES

[1] Lau S.L. Supporting patient monitoring using activity recognition with a smartphone. *Proceedings of International Symposium on Wireless Communication Systems*, 2010.

[2] Tapia E.M., Intille S.S., and Larson K. Activity recognition in the home using simple and ubiquitous sensors. *IEEE Pervasive Computing*, pages 158–175, 2004.

[3] He Y. Fall detection by built-in tri-accelerometer of smartphone. *Proceedings of International Confrence on Biomedical and Health Informatics,Hong Kong, China*, pages 184–187, Jan 2012.

[4] Loseu V., Ghasemzadeh H., Ostadabbas S., Raveendranathan N., Malan J., and Jafari R. Applications of sensing platforms with wearable computers. *Proceedings of International Conference on Pervasive Technologies Related to Assistive Environments*, 2010.

[5] Wenyao Xu, Mi Zhang, Alexander A. Sawchuk, and Majid Sarrafzadeh. Co-recognition of human activity and sensor location via compressed sensing in wearable body sensor networks. *International Conference on Wearable and Implantable Body Sensor Networks*, pages 124–129, 2012.

[6] Mi Zhang and Alexander A. Sawchuk. Usc-had: A daily activity dataset for ubiquitous activity recognition using wearable sensors. *Proceedings of International Conference on Ubiquitous Computing*, Sep. 2012.

[7] Mi Zhang and Alexander A. Sawchuk. A feature selection-based framework for human activity recognition using wearable multimodal sensors. *International Conference on Body Area Networks*, pages 92–98, 2011.

[8] L. Bao and S. Intille. Activity recognition from user-annotated acceleration data. *IEEE Pervasive Computing*, pages 1–17, 2004.

[9] T. Huynh. Analyzing features for activity recognition. *Joint sOc-EUSAI conference*, 2005.

[10] Mojsilovic A, Popovic M.V., and Rackov D.M. On the selection of an optimal wavelet basis for texture characterization. *IEEE Transactions on Image Processing*, 9:2043–2050, 2000.

[11] S. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674–693, 1989.

[12] I. Daubechies. Ten lectures on wavelets. *Capital City Press Montpelier Vermont*, 1992.

[13] Lei Yu and Huan Liu. Efficient feature selection via analysis of relevance and redundancy. *Journal of Machine Learning Research*, 5, 2004.

[14] A. David and S. Panchanathan. Wavelet-histogram method for face recognition. *Journal of Electronic Imaging*, 9(2):217–225, 2000.

[15] Tao Li, Mitsunori Ogihara, and Qi Li. A comparative study on content-based music genre classification. *SIGIR*, pages 282–289, July 28 2003.

[16] C. Hsu. A practical guide to support vector classification. *Bioinformatics*, 1(1):1–16, 2010.