Human Activity Recognition for a Content Search System
Considering Situations of Smartphone Users

Tomohiro Mashita
Cybermedia Center, Osaka Univ.

Kentaro Shimatani
Grad. Sch. of Information Science and Tech., Osaka Univ.

Mayu Iwata
Grad. Sch. of Information Science and Tech., Osaka Univ.

Hiroki Miyamoto
Grad. Sch. of Information Science and Tech., Osaka Univ.

Daijiro Komaki
Grad. Sch. of Information Science and Tech., Osaka Univ.

Takahiro Hara
Grad. Sch. of Information Science and Tech., Osaka Univ.

Kiyoshi Kiyokawa
Cybermedia Center, Osaka Univ.

Haruo Takemura
Cybermedia Center, Osaka Univ.

Shojiro Nishio
Grad. Sch. of Information Science and Tech., Osaka Univ.

Abstract—Smart-phone users can search for information about surrounding facilities or a route to their destination. However, it is difficult to get or search for information while walking because of low legibility. To address this problem, users have to stop walking or enlarge the screen. Our previously proposed system for smart-phone switches the information presentation policies in response to the user’s context. In this paper we describe our context recognition mechanism for this system. This mechanism estimates user context from sensors embodied in a smart-phone. We use a Support Vector Machine for the context classification and compare four types of feature values consisting of FFT and 3 types of Wavelet Transforms. Experimental results show that recognition rates are 87.2 % with FFT, 90.9 % with Gabor Wavelet, 91.8 % with Haar Wavelet, and 92.1 % with MexicanHat Wavelet.

Keywords—Context aware system, Context recognition,

I. INTRODUCTION

Smart-phones have enabled us to provide information in various situations. Some users use smart-phones to get or search for information while walking. However, it is difficult and dangerous to read detailed information while walking. In this case, users have to stop and enlarge the content. To address this problem, a system should recognize the user context and switch the content, the interface, or the methods of display.

We have proposed a system for smart-phone that switches the information presentation policies in response to the user’s context[1]. This paper introduces the details of the context recognition mechanism for our context aware system.

II. ACTIVITY RECOGNITION

Our system recognizes whether a user is walking, standing, or running. This is achieved by using an accelerometer equipped smart-phone. Smart-phones are used or held in various ways and we should not limit the way of holding the accelerometer values should be transformed to values reflecting the users’ activities and should be independent from the way of holding or using the mobile device.

A. Recognition Mechanism

The human body is bouncing in a cycle when a person is walking or running. We applied frequency analysis to obtain the feature values for the activity recognition. Fig. 1 shows the flow of the activity recognition method.

1) Invariant Detection: In this system, the values from the accelerometer should be transformed to the invariants that are independent from the variation of the posture of the smart-phone. To detect invariants, the system applies Principal Component Analysis (PCA) to the accelerometer data.

Figure 1. Flow of the activity recognition

Accelerometer
1-D Time-Series
Projection
1st Principal Axis
PCA
Frequency Analysis
FFT
Gabor
Haar
MexicanHat
Feature Vector
Activity
Standing
Walking
Running
Frequency Component
Maximum
Minimum
SVM
values up to a particular time $\gamma$ in the past. The system
takes the one-dimensional time-series data $f(t)$ which is
independent from the variation of the posture of the smart-
phone by projecting the accelerometer values to the first
principal axis.

2) Feature Values Extraction: We adopt Wavelet Transform
and Fourier Transform as the frequency analysis for the
feature values extraction. Wavelet Transform is expressed as

$$ T_0^a(f) = \frac{1}{\sqrt{a}} \int f(t) \psi \left( \frac{t-b}{a} \right) dt, \quad (1) $$

where $\psi$ is the mother wavelet, $a$ is the scale parameter,
and $b$ is the shift parameter. The transformed value $T_0^a(f)$ has
the phase originating from the walking/running cycle and
this phase should be canceled. The maximum and minimum
values of $\{T_0^1(f), \ldots, T_0^n(f)\}$ are used as the feature values.
These feature values are independent from the phase when
the $\tau$ is longer than the walking/running cycle. Our system
applies $N_a$ scale parameters and obtains $2N_a$ dimensional
feature vectors for activity classification.

3) Classification: The classification is achieved using a
Support Vector Machine (SVM) with a RBF kernel. In our
system, we applied one-against-one strategy to the multiclass
classification because Kreibel et al. [2] reported that the
one-against-one strategy is better than the one-against-all
strategy. The parameters of the classifier are detected using
the supervised dataset and cross-validation.

III. IMPLEMENTATION

We use an iPhone 4S (Apple inc.) as a smart-phone
and our system is implemented into it. The accelerometer
equipped in an iPhone 4S can sample at 100 Hz. However,
the sampling interval can become unstable at the high
sampling rate. Therefore, we adopt 10 Hz sampling.

We take 4 samples (0.4 sec.) as the walking cycle $\tau$, 30
samples (3 sec.) as the PCA interval $\gamma$, and $a \in 1, \ldots, 32$
as the scale parameter of the mother wavelets. Therefore,
our system obtains 64 dimensional feature vectors every 0.1
second. We implemented three types of the mother wavelet,
Gabor (Eq. 2), MexicanHat (Eq. 3), Haar (Eq. 4).

$$ \psi_{\text{Gabor}}(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{t^2}{2\sigma^2}} e^{i\omega t}, \quad (2) $$

$$ \psi_{\text{Mexican}}(t) = \frac{2}{\sqrt{2\sqrt{\pi}\sigma}} \left( 1 - \frac{t^2}{\sigma^2} \right) e^{-\frac{t^2}{2\sigma^2}}, \quad (3) $$

$$ \psi_{\text{Haar}}(t) = \begin{cases} 1 & (0 \leq t < 1/2) \\ -1 & (1/2 \leq t < 1) \end{cases}, \quad (4) $$

where $\sigma = 1$, $\omega = 1$.

IV. EXPERIMENT

We conducted experiments to evaluate the activity recogni-
tion mechanism. We obtained 180 datasets from 15 sub-
jects for learning and testing. The datasets consist of three
types of the activities –standing, walking, and running–and
three types of the holding situations of the smart-phone –in
a hand, in a bag, and in a trousers pocket. Each dataset is
obtained by a subject that continues one kind of activity and
one holding situation for about 30 seconds.

We evaluated the performance of the feature values with
FFT or each Wavelet Transform. Table I shows the recog-
nition rate of each feature value. The feature values with
Wavelet Transform are better than that with FFT. The recog-
nition rate by Wavelet Transform is higher than 90%. We
think that our activity recognition mechanism is unaffected
by the holding situations of the smart-phone.

V. CONCLUSION

We proposed an activity recognition method for a content
search system considering situations of mobile users. Our
method classifies users’ activity into standing, walking, and
running using the sensors equipped in their smart-phone.
Our system extracts the feature values using PCA and fre-
quency analysis to avoid the variation of the device holding
situations. The classification is achieved using SVM with
RBF kernel. We adopted FFT and three types of Wavelet
Transform as the frequency analysis and compared those.
Experimental results show that recognition rates are 86.6 %
with FFT, around 92 % with Wavelet Transforms.

Future works includes experimentation and analysis of the
difference among individuals and more varied situations of
holding the device. Furthermore, our system should recog-
nize more detailed contexts including walking and running
speed, relationships with date and location, and so on.

ACKNOWLEDGMENTS

This research was partially supported by The Global
Center of Excellence Program of the Ministry of Education,
Culture, Sports, Science and Technology, Japan.

REFERENCES

Devices based on User Context Recognition,” The 1st In-
ternational Workshop on Ambient Information Technologies,
2012. (to appear)

machines,” pp.255–268, MIT Press, Cambridge, MA, USA,
1999.

<table>
<thead>
<tr>
<th>Table I</th>
<th>Recognition rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFT</td>
<td>87.2%</td>
</tr>
<tr>
<td>Gabor</td>
<td>90.9%</td>
</tr>
<tr>
<td>Haar</td>
<td>91.8%</td>
</tr>
<tr>
<td>MexicanHat</td>
<td>92.1%</td>
</tr>
</tbody>
</table>