Kobe University Repository : Kernel

PDF issue: 2026-11-30

Study on human behavior recognition using discrete HMM and continuous HMM

Su, Yancong

Murao, Hajime

(Citation)
ICIC Express Letters,6(3):791-796

(Issue Date)
2012-03

(Resource Type)
journal article

(Version)
Accepted Manuscript

(URL)
https://hdl.handle.net/20.500.14094/90001843

KOBE UNIVERSITY

ICIC Express Letters
Volume **4**, Number **5**, October **2010**

ICIC International ©2010 ISSN 1881-803X
pp. 1–ISII2010-000

A STUDY ON HUMAN BEHAVIOR RECOGNITION USING DISCRETE HMM AND CONTINUOUS HMM

Y. Su[1] and H. Murao[1]

[1]Graduate School of Intercultural Studies
Kobe University
1-2-1 Tsurukabuto, Nada, Kobe, 657-8501, Japan
{ gensou.so; hajime.murao }@mulabo.org

Received February 2010; accepted April 2010

Abstract. *In this paper, we propose a method to cluster the time-series of sensory data from mobile device automatically, and apply it to human behavior recognition. Both of a discrete HMM and a continuous HMM are applied to this purpose for comparing. In the experiments, we used an iPhone to collect 3-axis acceleration data for 4 types of user's behavior, including "Running", "Walking", "Climbing up stairs", and "Going down stairs". As a result, 4 classes were generated automatically. 80% of the sensory data were classified to these classes correctly by using discrete HMM, and 90% by using continuous HMM.*
Keywords: Behavior Recognition, HMM, Time-series Data, Acceleration

1. **Introduction.** Recently, by the development of sensing technology, various sensors have been produced. Such as acceleration sensor, gyro sensor, magnetic sensor and air-pressure sensor. By using these sensors, it is able to measure a lot of spatiotemporal data of human behavior much more easily and richly than before. There are many researches in order to understand the human behaviors based on these data. And, because these sensors become more and more cheaper and small-sized, the living commodity, which is embedded some sensor, is on the increase. Especially, the mobile phone, which is one of the most fixated goods in our daily life, includes acceleration sensor and gravity sensor usually. With the intuitive interface, so many never achieved features have been realized, that let it has a lot of attentions. The sensor, which can be mounted into human's body or something always cooperating with human behavior, is called "wearable sensors". Some wearable sensors adjust various purposes, and have already spread over many different fields in our daily life. Against this background, the goal of this research is not to develop a new special device. It is to propose a method for human behavior recognition by using some already implemented sensors, which are available to collect time-series data easily and comfortably.

Once the user's behavior is recognized correctly, it can be expected to change the service contents adaptively according the reorganization result. The service provided in this way is named "context-aware services". Some works aim to improve the quality of life in various fields, and propose their methods for providing context-aware service. For example, Doungs' ADL (Activity of Daily Living) monitoring in remote nursing system[1], Motomuras' early childhood injury prevention monitoring[2], Ryus' energy expenditure estimation method for NEAT (Non-Exercise Activity Thermogenesis)[3] and Ohnos' analysis method for customer patterns from the trajectory in supermarket[4]. In addition to this kind of laboratorial scale works, there are also a lot of new services for practical use in commercial providing, such as automatic controlled home appliances (air conditioner, light, etc.), restaurant and transfer information recommendation service. In the future,

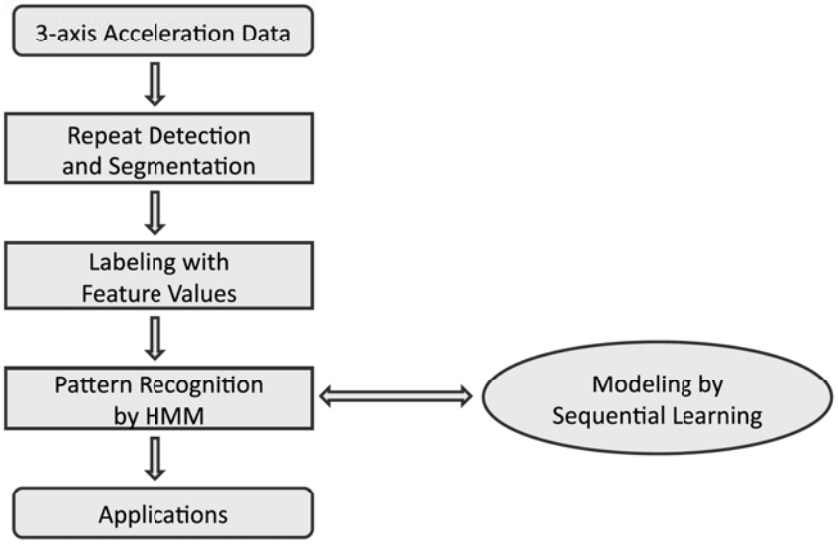


FIGURE 1. The flow of process in the proposed method

with the appearing of more and more context-aware services, it could be expected to become a new industry.

In order to achieve context-aware service described above, whatever its purpose, the sensor data is necessary. As an element of context, we could mention time, location, behavior etc. But due to the diversity and complexity in human activities, there are a considerable number of measurement difficulties. Previously, several modeling methods for human behavior have been proposed[5][6][7]. However, it is necessary to specify the types of behavior and classify the training data beforehand by handwork. Therefore, preparing the training data and deciding the number of classifications are heavy burdens for the system designer and the user. In this study, we propose a method to solve this problem that how to collect the activity information easily, and build a usefully model to recognize the user's behavior context accurately.

2. System Description.

2.1. **Summary.** Figure 1 shows the processing flow of the proposed system. First, as a pretreatment section, it detects the repeat points of the time-series of 3-axis acceleration data obtained from mobile device, and segments these data by the repeat points. Then in the next, these segmented data will be labeled as an observation symbol sequence of HMM. The labeling operation is based on the feature values such as movement direction and so on. In the next, the pattern recognition section calculates the probability of these symbol sequences drawn from every HMM. Depending on the result, it will decide the data should be used as a training data for the already existing model or the new additional one.

2.2. **Data Collection.** We use the mobile phone that has a 3-axis accelerometer for easing the data collection. Nowadays almost all the people using mobile phone, carry and use the device everywhere and always for daily life. A lot of mobile phones have more and more advanced functions, with some integrated sensors, such as GPS, accelerometer and gyro sensor. It allows saving many times and troubles from wearing some special sensor devices. Hence, some researches consider using the device for monitoring the user's context, and evaluate its effectiveness[8][9][10]. In this paper, with supposing the application in daily life, the basic behaviors such as "walking" and "running" are the objects to be modeled and recognized.

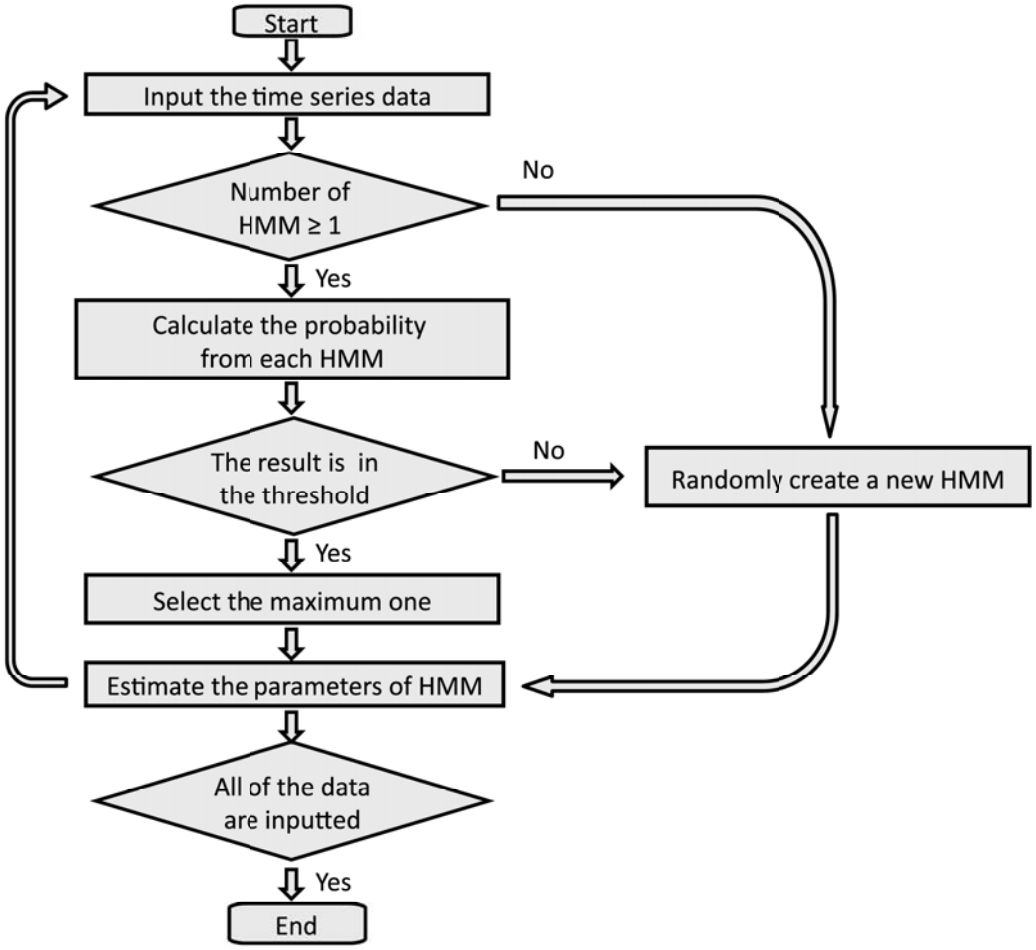


FIGURE 2. The flow of process in the proposed method

2.3. **Modeling and Recognition.** We apply Hidden Markov Model (HMM) for modeling and recognizing the user's behavior contexts in this paper. The models are constructed by the sequential learning that repeats recognizing and training processes based on HMM. It is able to save a lot of time and work for deciding the number of classifications, and increase new models easily and automatically[11][12]. Figure 2 shows the processing flow of the proposed system.

3. **Experiments.** "Walking", "Running", "Climbing up the stairs" and "Going down the stairs" are the 4 objects of human behavior in the experiments. We used iPhone and Kawaguchi's application "HASC Logger[13]" to collect the 3-axis acceleration data. With attending iPhone to a place on the trousers where the experimental subject usually put his/her mobile in, 20-seconds acceleration data were collected at the sampling rate 100Hz. They were transferred to a computer by "HASC Tool[14]" via wireless network as the input of the labeling process.

In the labeling process, we focused on the movement direction, or the direction of acceleration as feature value. The time-series sequence of these labels was used as an observed symbol sequence for HMM. Figure 3 shows the thresholds in a cube for labeling. The acceleration data is a directional and continuous. $A(X)$ is in the lateral direction that means "left" or "right", $A(Y)$ is in the longitudinal direction that means "up" or "down", and $A(Z)$ is in the depth direction that means "forward" or "backward". $max(X)$, $max(Y)$ and $max(Z)$ mean the average maximum value of $A(X)$, $A(Y)$ and $A(Z)$ in 10 times of "Resting standing" context. $min(X)$, $min(Y)$ and $min(Z)$ are in the opposite. Table 1 shows some examples. By comparing with the thresholds, the raw data were labeled with "Left" or "Right", "Up" or "Down", and "Forward" or "Backward" in the case of

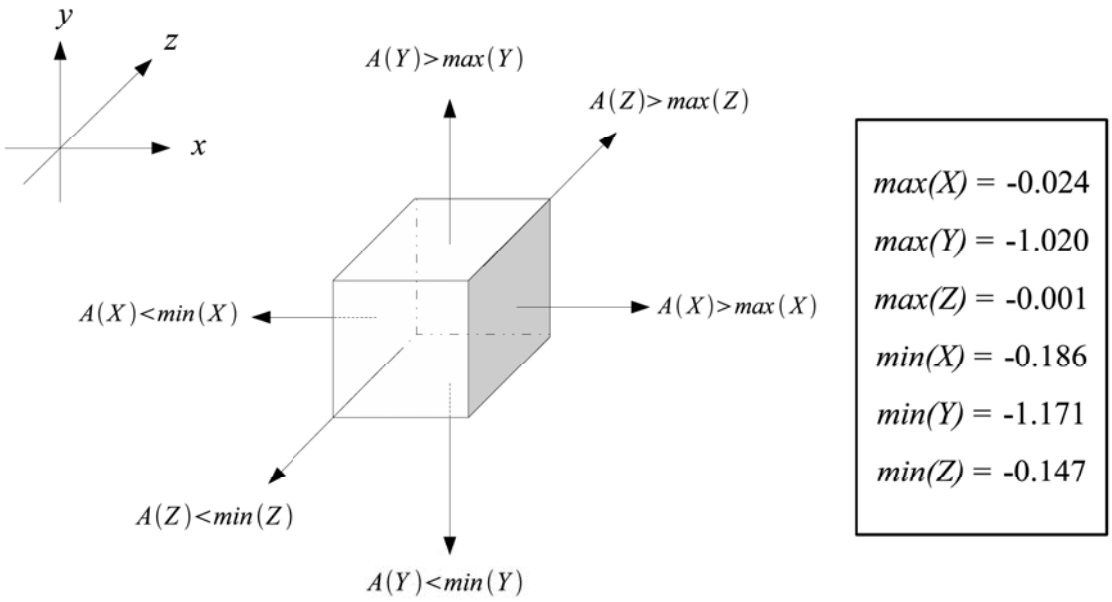


FIGURE 3. The thresholds in a cube

TABLE 1. Some examples of labeling

X	Y	Z	Symbol
$A > max$	$A > max$	$A > max$	RUF
$A < min$	$A < min$	$A < min$	LDB
$min < A < max$	$A > max$	$A < min$	CUB
$A > max$	$min < A < max$	$A < min$	RCB
$A < min$	$A > max$	$min < A < max$	LUC

exceeding the corresponding one, or "Center" in the case of no exceeding. As a result, a total of 27 different combinations were used for the generation of symbol sequence. Table 2 shows the subjective classification of the sequences. There are 5 sequences in each behavior context, and all of these 20 sequences were inputted to the system randomly.

In the modeling and recognition process, we used discrete HMM as the following for model construction and pattern recognition. A new HMM $\lambda = (\pi, a, b)$ will be generated randomly with all of its parameters (π: Initial state prior probabilities, a: State transition probabilities and b: Observation emission probabilities).

- Training algorithm: Baum-Welch Algorithm
- Number of hidden states: 4
- Number of observable symbols: 27 (equals to the number of label combinations)

We also used continuous HMM to compare the classification performance. The different with discrete HMM is that the observation emission probabilities b is following a Gaussian distribution in continuous HMM. Because of this, the observation symbols of continuous HMM are continuous values. The labeling process becomes not necessary, and we used the raw acceleration data as the input.

4. **Main Results.** Table 3 shows the clustering results when the threshold $w$was 500. By collating with Table 2, almost all of the sequences in the HMM_1, which was created by discrete HMM, are "Running" and "Walking". But it divided "Climbing up the stairs" and "Going down the stairs" into HMM_2 and HMM_3 accurately. Similarly, continuous HMM clustered "Running" and "Walking" correctly, but its HMM_3 includes all of the sequences of "Climbing up the stairs" and "Going down the stairs". It indicates that

TABLE 2. The list of the observation sequences

Behavior Context	Sequence #
Running	0, 1, 2, 3, 4
Walking	5, 6, 7, 8, 9
Climbing up stairs	10, 11, 12, 13, 14
Going down stairs	15, 16, 17, 18, 19

TABLE 3. The clustering result of $w = 500$

	by discrete HMM	by continuous HMM
HMM_1	0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 18	0, 1, 2, 3, 4
HMM_2	10, 11, 12, 13, 14	7, 8, 9
HMM_3	15, 16, 17, 19	5, 6, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19

TABLE 4. The clustering result of $w = 250$

	by discrete HMM	by continuous HMM
HMM_1	2, 3, 4, 5, 6	0, 1, 2, 3, 4, 15
HMM_2	0, 1, 7, 8, 9	6, 7, 8, 9
HMM_3	10, 11, 12, 13, 14	5, 10, 11, 12, 13, 14
HMM_4	15, 16, 17, 18, 19	16, 17, 18, 19

the discrete HMM, which has a labeling process using the movement direction as feature value, is effective against the behaviors on the stairs, or the activities that have violent direction conversion. In the opposite, the continuous HMM that uses the raw data in this paper performed better for the classification of the activities on a flat land. Both of them have succeeded the clustering to a limited extent. However, just only 3 classes were outputted for 4 types of behavior context. It is a different with the subjectivity of human. We tried using a lot of different thresholds to solve this problem.

Table 4 is the clustering results when the threshold was 250, and shows clearly that the same numbers of classification with the subjectivity of human were obtained by both of the two modeling process. Probably, HMM_1 is the model of "Running", HMM_2 is "Walking", HMM_3 is "Climbing up the stairs" and HMM_4 is "Going down the stairs". 80% of the data were recognized as these models correctly by the discrete HMM, and 90% by the continuous HMM. It confirmed that the sequential learning, which repeats recognizing and training based on HMM, is useful to categorize the human behaviors. It confirmed that the sequential learning, which repeats recognizing and training based on HMM, is useful to categorize the human behaviors. And by comparing the results, we can observe that the classification accuracy depends on the threshold. Consequently, we need to study how to find the optimum solution in the future work.

5. **Conclusion.** In this paper, we applied both of a discrete HMM and a continuous HMM to cluster the time-series of sensory data from mobile device automatically. In the experiments, we used an iPhone to collect 3-axis acceleration data, and labeling them with the direction of acceleration. As a result, depending on a good threshold setting, 80% of the sensory data were categorized correctly by using discrete HMM, and 90% in the case of continuous HMM. Moreover, we could observe the usefulness of our proposed method. In the future, we will include much more behavior contexts, and examine the optimal parameters.

Acknowledgment. This work is supported by Grant-in-Aid for Scientific Research of the Japanese Ministry of Education, Culture, Sports, Science and Technology (assignment number: 21500140). Our heartfelt appreciation goes to the other members of Murao Laboratory, whose comments and suggestions were inestimable value for our study. We would also like to express our gratitude for their moral support and warm encouragement.

REFERENCES

[1] T. V. Duong, H. H. Bui, D. Q. Phung and S. Venkatesh, Activity Recognition and Abnormality Detection with the Switching Hidden Semi-Markov Model, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol.1, pp.838–845, 2005.

[2] Y. Motomura, Y. Nishida and T. Yamanaka, Probabilistic Modeling for Preventing Children's Accidents, *IEICE technical report. Neurocomputing*, vol.104, no.759, pp.73–76, 2005

[3] N. Ryu, Y. Kawahara, A. Kobayashi and T. Aasami, A energy expenditure estimation method for non-exercise activity thermogenesis using accelerometer, *The Special Interest Group Technical Reports of IPSJ*, vol. 2008, no.40, pp. 67-74, 2008.

[4] A. Ohno, T. Inamoto and H. Murao, Fundamental Study of Clustering Images Generated from Customer Trajectory by Using Self-Organizing Maps, *ICIC Express Letters*, vol.4, no.5(B), pp.1979–1984, 2010.

[5] A. K. Dey, G. D. Abowd and D. Salber, A Conceptual Framework and a Toolkit for Supporting the Rapid Prototyping of Context-aware Applications, *Human-Computer Interaction*, vol.16, no.2, pp.97–166, 2001.

[6] C. Breazeal and B. Scassellati, Robots that Imitate Humans, *Trends in Cognitive Sciences*, vol.6, no.11, pp.481–487, 2002.

[7] A. Billard, S. Calinon, and F. Guenter, Discriminative and Adaptive Imitation in uni-manual and bi-manual Tasks, *Robotics and Autonomous Systems*, vol.54, no.5, pp.370–384, 2006.

[8] T. Nakamura, T. Ogawa, K. Kiyokawa and H. Takemura, Implementation and Evaluation of a User Context Recognition Mechanism for Continuous Wearable Learning Systems, *Technical report of IEICE. Multimedia and virtual environment*, vol.106, no.396, pp.17–22, 2006.

[9] K. Tabuchi, F. Naya, R. Ohmura, H. Noma, K. Kogure and F. Kishino, Evaluation of data collection parameters on the daily activity classification with accelerometers, *Technical report of IEICE. PRMU*, vol.106, no.73, pp.43–48, 2006.

[10] N. Hashida, R. Ohmura and M. Imai, Contribution of Personalization to Activity Recognition Accuracy Using Accelerometers in Daily Life, *The Special Interest Group Technical Reports of IPSJ*, vol.108, no.138, pp.69–74, 2008.

[11] W. Takano and Y. Nakamura, Realtime Unsupervised Selftuning Segmentation of Behavioral Motion Patterns Based on Probabilistic Correlation and Its Application to Automatic Acquisition of Proto-Symbols, *Journal of the Robotics Society of Japan*, vol.27, no.9, pp.1046–1057, 2009.

[12] D. Kulic, W. Takano and Y. Nakamura, Incremental Learning, Clustering and Hierarchy Formation of Whole Body Motion Patterns using Adaptive Hidden Markov Chains, *International Journal of Robotics Research*, vol.27, no.7, pp.761–784, 2008.

[13] Nobuo Kawaguchi, HASC Logger: http://itunes.apple.com/jp/app/id384897105

[14] HASC Tool: http://hasc.jp/hc2010/hasctool.html