33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

A Semi-supervised Hidden Markov Model-based Activity Monitoring System

Min Xu, Long Zuo, Satish Iyengar, Albert Goldfain and Jim DelloStritto
Blue Highway LLC.
2-212 Center for Science and Technology, Syracuse, NY, 13244
{mxu, agoldfain, jdellostritto}@blue-highway.com

***Abstract*— Most existing human activity classification systems require a large training dataset to construct statistical models for each activity of interest. This may be impractical in many cases. In this paper, we proposed a semi-supervised HMM based activity monitoring system, that adapts the HMM for a specific subject from a general model in order to alleviate the requirement of a large training data set. In addition, using two triaxial accelerometers, our system not only identifies simple events such as sitting, standing and walking, but also recognizes the behavior or a more complex activity by temporally linking the events together. Experimental results demonstrate the feasibility of our proposed system.**

I. INTRODUCTION

The continuing objective of wireless body sensor networks is to create a non-intrusive reliable instrument for the collection of (inertial and physiological) data streams from clinically relevant positions on the body. The collected data could be analyzed for a number of end applications ranging from clinical studies (activity monitoring [1], disease progression, drug reaction etc) to entertainment and lifestyle applications(athletic training, robotics). Recent advances in wearable sensing technologies enable us to collect longitudinal data from real subjects performing their normal activities without interfering with their daily routines. Automatic capture and analysis of human daily activities has been an active research area due to a number of potential applications. In particular, such activity monitoring system can assist in care giving for elderly group's daily lives, help to identify and track their regular performance of daily activities and find any abnormal behavior pattern due to disease, environmental effects and seasonal change. For example, the activity pattern in the long-term data is very useful for a study of a mental/-physical disorder due to seasonal change known as "seasonal affective disorder". Also it can be used in clinical studies for prediction, diagnosis, and treatment of several diseases and disorders, such as: Parkinson's disease [2], multiple sclerosis [3], severe fatigue or pain, sleep disorders [4], and stroke [5].

Design of such systems has been considered in the past (e.g. [6]–[8]). Most systems tend to identify simple short duration events, such as walking, sitting, running, standing and climbing stairs. Few systems aim to recognize complex human activities such as working in the office, shopping and dining, which are longer in duration. The main challenge is to find realistic yet adaptable statistical models for complex human activities. Our goal is to identify both simple events and complex activities. We decompose a complex activity into a sequence of simple events. The structure of data flow is shown in Fig. 1.

In a preliminary work presented in [9], a rule based hierarchical classification strategy has been proposed to classify human activities without training data. Empirically derived rules based on the inputs of extracted features and temporal patterns of event sequence (formulated as a Markov chain) are used to identify simple events and complex activities, respectively. However, it is difficult to develop a general rule across subjects and for different events independent of activities. For example, someone's sitting posture when he works in office could be different from his sitting posture when he watches TV. This implies that the same event in the context of different activities could vary from person to person. Thus, in this paper, we extend our work in [9] and propose an approach that is more data-driven.

One of the most popular approaches for activity classification is to use HMM. Specifically, two approaches can be found in the literature. In [8], [10], the authors define events/activities such as walking, sitting and climbing stairs, as states, which compose a Markov chain to represent the observation data. Viterbi decoding algorithm is then used to estimate the state sequence (event/activity sequence) to exploit the temporal dependency among the events/activities. A second approach as adopted by [7], [11], [12] is to model each event/activity as an HMM, and then calculate the likelihood of each model to classify events/activities. As a Markov model needs to be trained for each activity, the computational cost increases. Motivated by the fact that a complex activity can be decomposed into a sequence of simple events and by the fact that such simple events (primitive motor movements) are detectable via wearable sensors, our approach is to combine the two approaches together. A complex activity is modeled by a Markov model and the short-term events are modeled as the states in each Markov model. The system outputs both the short-term simple events and the long-term complex activities.

In addition, good HMMs require a large training dataset, which makes it inconvenient in practical application. In order to overcome this, we propose the use of Bayesian adaptation techniques, which adapt a universal model to produce the model for specific subject in an unsupervised manner. Two adaptation algorithms are employed for such purpose. This framework can thus be considered as a semi-supervised learning technique.

The details of this system is presented in Section 2. Section 3 describes the experimental results that demonstrate the classification accuracy of our proposed approach. Finally, a concluding remark is given in Section 4.

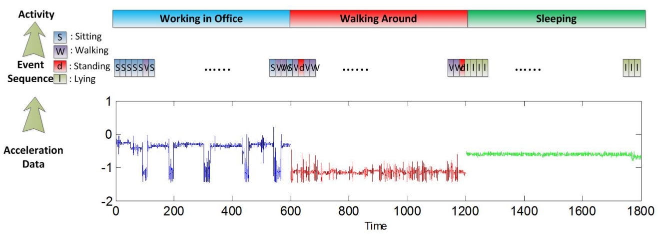

Fig. 1. Data flow sequence

II. System Description

A. Data acquisition

To recognize the activity of a person's daily life, two Freescale triaxial accelerometers are required, with one attached to the waist, and the other one attached to the left thigh. As we want to recognize the posture of the subject, we positioned the accelerometer such that Y axis of the accelerometer is aligned with the head and the Z axis is perpendicular to the torso and leg. Also the Y axes of the accelerometers are aligned to the gravity line when the subject stands upright. To avoid the artifact due to motion of accelerometers relative to the body, accelerometers are fixed by iPod bands and then are firmly attached to the body. The sampling rate of the accelerometer is 30HZ.

B. Feature extraction

Feature extraction is an important step to reduce the dimensionality of the data space as the raw acceleration data sequence is too large to be processed by the classifier. Physical activity is a complex phenomenon generally characterized by its intensity, duration, frequency, environment and the type of activity (rest, walk, run, etc.) [13]. Features selected should be able to represent relevant kinetic information from acceleration data and characterize different motion patterns. Thus, what type of features to be selected highly depends on the type of events/activities to recognize.

Different types of features calculated in time domain, frequency domain and wavelet domain, such as mean, energy, entropy, standard deviation, wavelet coefficients calculated from a small window have been investigated in the literature. However, postural features are important but often ignored by the community. As the three activities we are considering are mostly composed by different postures (sitting, standing, lying), we use two features from each sensor to represent the data. The first feature metric is a binary variable, with 1 indicating moving status and 0 indicating static status. It is equal to 1 when the standard deviation of acceleration data within a moving window is greater than a threshold and equal to 0 when the standard deviation is less than a threshold. The threshold is determined based on our experiments and is set to $0.3g$ in our current system. Note that the threshold could be adjusted or learned by supervised learning for different subject groups. The second feature is the inclination angle of the Y axis of the sensor, which indicates the orientation of the torso and leg. This feature can be used to distinguish three main postural conditions, standing, sitting, and lying.

C. HMM modeling

The HMM based classifier has been widely applied to recognize different human activities. Most HMM based algorithms model each activity using a HMM and experimentally or arbitrarily choose the number of hidden states without attaching physical significance to the states. Our work is also based on Markov modeling and each feature is modeled as a Gaussian mixture model, albeit with the key difference. We define each state as one simple event and it is achieved by interpreting Gaussian mixture model for each state. Fig. 2 illustrates how the Markov model is used to represent the complex activity. Assume that a Markov model is denoted by $\lambda = (A, B, \pi)$, where A denotes the state transition probability distribution, B denotes the observation symbol probability distribution in each state, and π denotes the initial state distribution. Given the values of these probability distributions, a HMM is built for a given observation sequence. A typical training process of a HMM is to adjust the model parameters such that *a posteriori* probability distribution of observation sequence given the model λ is maximized. And this training process includes two steps:

1) Initialize the model distribution (A, B, π)
2) Use the Baum-Welch method [14] to estimate the model parameters

To initialize the observation probability distribution in each state, we first performed the clustering analysis to find the initial model parameters. And then a rule based algorithm is applied to interpret the state by attaching an event to the mean value of the features of the state. For example, at the working scenario, the clustering analysis returns two clusters corresponding to the sitting event and walking event, respectively.

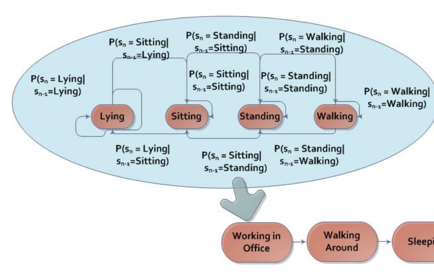

Fig. 2. Markov chain of events

After the HMMs are built for all the activities, we calculate the model likelihood values for all possible models using Viterbi algorithm, and select the activity whose likelihood is highest. At the same time, the corresponding state sequence generated by the selected model will be outputted. The structure of such system is shown in Fig. 3.

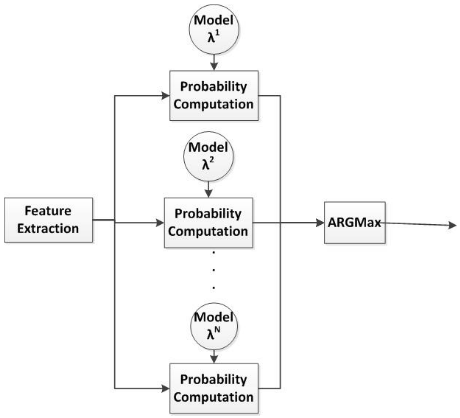

Fig. 3. System structure

D. Bayesian adaptation algorithms

In practice, it is often not convenient to collect a large amount of training data for each activity. Thus, it is necessary to build a universal model that allows user to directly use the model and recognize the activity without training data. On the other hand, due to the the large variety of motion patterns for different subjects, a single universal model does not generalize well to individual motion pattern. For these reasons, we used Bayesian adaptation algorithms [15] to adapt a universal model to a specific subject.

The universal model is typically trained using a large training database involving different subjects. In the Bayesian adaptation approach, prior knowledge of the distribution of HMM parameters from the universal model is incorporated into the modeling process and adapted to the data from specific subject. Even if some areas of the feature space are less represented in the training data, the prior information about the parameters can help to overcome the problem [15]. Two adaptation algorithms are employed here depending on whether the training data is available or not. We denote the model parameters in the universal HMM as w^{univ} and the model parameters after adaptation as w^{adapt}. In our experiment, we use the Gaussian mixture model to represent the probability of observations conditioned on each state. Thus, the model parameters to be adapted are the mean value, standard deviation and the weight of each component in the model.

The first adaptation algorithm requires some training data and consists two steps in adaptation. First, the model parameters of the new training data are calculated, denoted as w^{new}. Then, the new model parameters are adapted using $w^{adapt} = \alpha * w^{univ} + (1 - \alpha) * w^{new}$ [15], where α is the scale weighting factor that controls the balance between the universal model and new estimates. The smaller the value of α, the more contribution the new training data gives to the adapted model.

The second adaptation algorithm does not require any training data with the cost of possible higher classification error. When the new data is received, the built HMM is applied to estimate the class. And at the same time, the model parameters corresponding to this class are updated. This algorithm assumes that the classification result of the first step is accurate, and therefore, it could lead to higher classification error when the new data has very different motion pattern.

III. Experimental Results

In our experiments, we designed three activities, working in office, sleeping and walking around. For the activity of working in office, the subject sits in front of desk most of time, and stands up and then walks sometimes. For the activity of sleeping, the subject keeps lying posture most of time. For the walking around activity, the subject spends most of the time walking. In total, six subjects aging from 22 to 30 involving both females and males participated this study. Each subject was asked to perform the three activities with each for 20 minutes. In the experiment, we have used 5-second window for feature extraction and estimation of simple event, and 80-second window for estimation of complex activity. The Leave-One-Out technique is used to evaluate the performance. Four tests are performed:

1) For each subject, use 10-minute data to train the model and then the remaining data to test the model.
2) For each subject, three universal HMMs corresponding to three activities using 5-minute data from the other five subjects are trained. Then, 5-minute data from this subject is used to adapt the universal model to a personal model and the remaining data is used for testing. This test uses the first adaptation algorithm. In this case, the weighting factor α is given 0.5.
3) For each subject, three universal HMMs corresponding to three activities using 5-minute data from the other five subjects are trained. Then, 5-minute data from this subject is classified and used to adapt the universal model to a personal model. The remaining data is used for testing. This test uses the second adaptation algorithm. In this case, the weighting factor α is given 0.5.
4) For each subject, three universal HMMs corresponding to three activities using 5-minute data from the other five subjects are trained. Then, the universal models are used to test the subject's data for testing.

The results of four tests are displayed in Table 1-4. Test 1 does not employ the universal model and only relies on the person's training data to build a model. Test 2 and 3 employ the universal model and combine it with a small personal training data set to build an adapted model. Test 4 uses the universal model to directly classify the person's activities without generating an adapted model. From the data we collected from six subjects, Subject 2's motion pattern is not consistent, and therefore, when his training data size is not large enough, the classification performance is very bad, as we observed from Table 1. Comparing Table 2 and 3 with

Table 1, it is observed that incorporating the universal model to generate the adapted model improves the classification accuracy. In Table 4, the classification accuracy of Subject 2 is higher than that of the other three tests, which implies that the universal model is robust and characterizes good activity patterns as it is built from a larger training data set. However, without adapting the universal model to personal data, the classification performance from Subject 4 and Subject 6 degrades in Table 4, which implies that Subject 4 and Subject 6 have more personal motion pattern and therefore, the use of adapted model improves the classification performance.

TABLE I
CLASSIFICATION ACCURACY OF TEST 1 (%)

	Walking Around	Sleeping	Working
Subject 1	100	100	100
Subject 2	25	75	25
Subject 3	100	100	100
Subject 4	100	100	100
Subject 5	100	100	100
Subject 6	100	100	100

TABLE II
CLASSIFICATION ACCURACY OF TEST 2 (%)

	Walking Around	Sleeping	Working
Subject 1	100	100	100
Subject 2	50	75	37.5
Subject 3	100	100	100
Subject 4	100	100	100
Subject 5	100	100	100
Subject 6	100	100	100

TABLE III
CLASSIFICATION ACCURACY OF TEST 3 (%)

	Walking Around	Sleeping	Working
Subject 1	100	100	100
Subject 2	0	100	75
Subject 3	100	100	100
Subject 4	100	100	100
Subject 5	100	100	100
Subject 6	100	100	100

TABLE IV
CLASSIFICATION ACCURACY OF TEST 4 (%)

	Walking Around	Sleeping	Working
Subject 1	100	100	100
Subject 2	0	100	100
Subject 3	100	100	100
Subject 4	100	100	90
Subject 5	100	100	100
Subject 6	100	100	73

IV. CONCLUSION

In this paper a framework of HMM based real-time activity classification strategy is proposed. By defining the level of activity as a Markov chain of different events, our system uses a semi-supervised HMM based classification process to recognize different level of activity. The output of the system provides both the event status of the subject and level of the activity, providing valuable insight towards the subject's health condition and enabling us to find some interesting and useful behavior patterns in the long term data. Some experimental results demonstrate that our system can achieve good classification performance of daily activities without a large training data.

REFERENCES

[1] C.-F. Juang and C.-M. Chang, "Human body posture classification by a neural fuzzy network and home care system application," *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, vol. 37, no. 6, pp. 984 –994, Nov. 2007.

[2] M. Mancini, C. Zampieri, F. Horak, and L. Chiari, "Accelerometry-based longitudinal biomarkers of stance posture in early parkinson's disease," *Gait and Posture*, vol. 29, no. 1, p. e25, 2009.

[3] K. Zabjek, S. Hill, W. Gage, C. Danells, V. Closson, B. Maki, and W. McIlroy, "Gait and standing posture in patients with multiple sclerosis," *Gait and Posture*, vol. 21, p. S136, 2005.

[4] H. Pigot, B. Lefebvre, J. Meunier, B. Kerherv, A. Mayers, and S. Giroux, "The role of intelligent habitats in upholding elders in residence," in *5Th international Conference on Simulations in Biomedicine*, 2003, pp. 497–506.

[5] T. Batisheva, L. Rusina, and D. Skvortsov, "Biomechanical symptoms from gait and posture in patients after stroke," *Gait and Posture*, vol. 21, p. S105, 2005.

[6] 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 Recogn. Lett.*, vol. 29, no. 16, pp. 2213–2220, 2008.

[7] R. K. Ganti, P. Jayachandran, T. F. Abdelzaher, and J. A. Stankovic, "Satire: a software architecture for smart attire," in *Proceedings of the 4th international conference on Mobile systems, applications and services*. New York, NY, USA: ACM, 2006, pp. 110–123.

[8] A. Mannini and A. M. Sabatini, "Machine learning methods for classifying human physical activity from on-body accelerometers," *Sensors*, vol. 10, pp. 1154–1175, 2010.

[9] M. Xu, S. Iyengar, A. Goldfain, A. RoyChowdhury, and J. DelloStritto, "A two-stage real-time activity monitoring system," in *8th International Conference on Body Sensor Networks (BSN)*, 2011.

[10] J. He, H. Li, and J. Tan, "Real time daily activity classification with wireless sensor networks using hidden markov model," in *29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2007, pp. 3192–3195.

[11] D. O. Olgun and A. Pentland, "Human activity recognition: Accuracy across common locations for wearable sensors," in *IEEE 10th International Symposium on Wearable Computers*, 2006, pp. 11–13.

[12] J. Boyd and H. Sundaram, "A framework to detect and classify activity transitions in low-power applications," in *Proceedings of the 2009 IEEE international conference on Multimedia and Expo*, ser. ICME'09. Piscataway, NJ, USA: IEEE Press, 2009, pp. 1712–1715. [Online]. Available: http://portal.acm.org/citation.cfm?id=1698924.1699355

[13] "Medicine and science," *Sports and exercise*, vol. 32, no. 8, pp. S439–S516, 2000.

[14] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," in *Proceedings of the IEEE*, 1989, pp. 257–286.

[15] D. Zhang, D. Gatica-perez, S. Bengio, and I. Mccowan, "Semi-supervised adapted hmms for unusual event detection," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR*, 2005, pp. 611–618.