Sensor-Based Human Activity Recognition in a Multi-user Scenario

Conference Paper · November 2009
DOI: 10.1007/978-3-642-05408-2_10 · Source: DBLP

Liang Wang
Nanjing University

Tao Gu
RMIT University

Xianping Tao
Nanjing University

Sensor-Based Human Activity Recognition in a Multi-user Scenario

Liang Wang[1], Tao Gu[2], Xianping Tao[1], and Jian Lu[1]

[1] State Key Laboratory for Novel Software Technology, Nanjing University
[2] Department of Mathematics and Computer Science, University of Southern Denmark
wangliang@ics.nju.edu.cn, gu@imada.sdu.dk, txp@nju.edu.cn, lj@nju.edu.cn

Abstract. Existing work on sensor-based activity recognition focuses mainly on single-user activities. However, in real life, activities are often performed by multiple users involving interactions between them. In this paper, we propose Coupled Hidden Markov Models (CHMMs) to recognize multi-user activities from sensor readings in a smart home environment. We develop a multimodal sensing platform and present a theoretical framework to recognize both single-user and multi-user activities. We conduct our trace collection done in a smart home, and evaluate our framework through experimental studies. Our experimental result shows that we achieve an average accuracy of 85.46% with CHMMs.

Keywords: Multi-user activity recognition, probabilistic model.

1 Introduction

The problem of recognizing human actions and activities based on video camera has been studied in computer vision since a decade ago [18, 7]. With the availability of low-cost sensors and the advancement of wireless sensor networks, researchers in ubiquitous computing are recently interested in deploying various sensors to collect observations, and recognizing activities based on these observations. This in turn supports many potential applications such as monitoring activities of daily living (ADLs) [8] for the elderly.

Recognizing human activities based on sensor readings is interesting since sensors can capture many useful low-level features of human users, their living environments and human-to-environment interactions. It is also a challenging task because sensor data are noisy and human activities are complex in nature. Existing work on sensor-based activity recognition mainly focuses on recognizing activities of a single user [10, 9, 13, 15]. However, activities involving multiple users collaboratively or concurrently are common in our daily lives, especially in a home setting. For example, family members always watch TV together in a living room, and prepare meals together in a kitchen.

From social psychology point of view, people often form groups to perform certain activity collectively not only because they share socially relevant features but also because they interact and rely on each other to achieve specific goals. Among others, two distinctive features – social interdependence and task interdependence – are salient [14]. Foremost, they are socially interdependent because they rely on one another for feelings of connectedness and positive emotional outcomes. This is especially demonstrated in a

M. Tscheligi et al. (Eds.): AmI 2009, LNCS 5859, pp. 78–87, 2009.

home environment among family members whether affiliated by consanguinity, affinity, or co-residence.

Understanding activities of multiple users is important for not only psychologists, but also for computer scientists. In the vision of ambient intelligence, to a great extent, people who share socially relevant characteristics or features like gender, age, or an interest in the environment are brought together to form a group. Then, the focus and interaction again are often on multiple human users within the group instead of single ones.

Recognizing activities of multiple users is more challenging than that of a single user. The challenge is to find a suitable model to capture the interactions between users and perform inference using these observations. In this paper, we propose a temporal probabilistic models – Coupled Hidden Markov Models (CHMMs) – to model user interactions and recognize multi-user activities. CHMMs are a multi-chained variant of the basic Hidden Markov Models (HMMs); it couples HMMs with temporal, asymmetric influences. We design a wearable sensor platform capable of capturing observations of human users and their interactions. Based on this platform, we conduct real-world trace collection, and evaluate our model through comprehensive experiments to demonstrate the effectiveness.

In summary, the paper makes the following contributions.

- We investigate the problem of sensor-based, multi-user activity recognition in a smart home setting, and propose a temporal probabilistic model to recognize activities in a multi-user scenario.
- We develop a multimodal, wearable sensor platform to capture observations of both users and their interactions, and conduct trace collection involving two users in a real smart home.
- We conduct experimental studies to evaluate our proposed model for multi-user activity recognition.

The rest of the paper is organized as follows. Section 2 discusses the related work. In Section 3, we present the design of our wearable sensor platform. Section 4 describes our proposed activity model, and Section 5 reports our empirical studies. Finally, Section 6 concludes the paper.

2 Related Work

In ubiquitous computing, researchers are recently interested in recognizing activities based on sensor readings. Recognition models are typically probabilistic based, and they can be categorized into static classification or temporal classification. Typical static classifiers include naïve Bayes used in [10], decision trees used in [10], k-nearest neighbor (k-NN) used in [6], and Support Vector Machine used in [6]. In temporal classification, state-space models are typically used to enable the inference of hidden states (i.e., activity labels) given the observations. We name a few examples here: HMMs used in [9, 16], Dynamic Bayesian Network (DBN) used in [13] and Conditional Random Fields (CRFs) used in [15].

Some work has been done in modeling interaction process in computer vision. Oliver et al. [11] proposed and compared HMMs and CHMMs for modeling interactions between people and classifying the type of interaction based on observations collected from video camera. Gong et al. [4] developed a dynamically multilinked HMMs model to interpret group activities based on video camera. Park et al. [12] presented a synergistic track- and body-level analysis framework for multi-person interaction and activity analysis in the context of video surveillance. An integrated visual interface for gestures was designed in [2] as a platform for investigating visually mediated interaction with video camera. However, their system only tackled simple gestures like waving and pointing.

Most of the existing work on multi-user activity recognition used video data only. One possible reason is that it is rather hard to determine hidden parameters of HMMs in the case of multimodal group action or activity recognition, where features from each modal are concatenated to define the observation model [19]. Wyatt et al. [17] presented a privacy-sensitive DBN-based unsupervised approach to separating speakers and their turns in a multi-person conversation. They addressed the problem of recognizing sequences of human interaction patterns in meetings with two-layer HMMs using both audio and video data. Different from these work, we deploy a multimodal sensor platform, and show that how sensor modality plays a role in multi-user activity recognition.

3 Multimodal Wearable Sensor Platform

We built our wearable sensor platform as shown in both Figure 1 and Figure 2. This platform measures user movement (i.e., both hands), user location, human-object interaction (i.e., objects touched and sound), human-to-human interaction (i.e., voice) and environmental information (i.e., temperature, humidity and light). To capture acceleration data, we used a Crossbow iMote2 IPR2400 processor and radio board with the ITS400 sensor board, as shown in Figure 1d. The ITS400 sensor board also captures environmental information such as temperature, humidity and light. To capture object use,

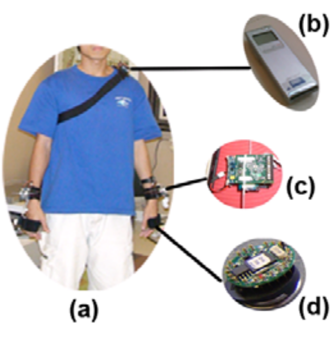

Fig. 1. (a) Wearable sensor set, (b) Audio recorder, (c) iMote2 with ITS400, (d) RFID wristband reader

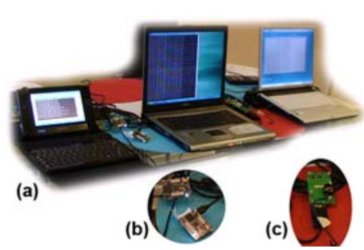

Fig. 2. (a) Servers for logging sersor readings, (b) iMote2 receiver, (c) Mica2Dot receiver

we built a customized RFID wristband reader which incorporates a Crossbow Mica2Dot wireless mote, a Skyetek M1-mini RFID reader and a Li-Polymer rechargeable battery. The wristband is able to detect the presence of a tagged object within the range of 6 to 8 cm. To capture the vocal interaction between users, we use a commercial audio recorder to record audio data as shown in Figure 1b. To determine user identity, the device IDs of each iMote2 set and RFID reader are logged and bound to a specific user.

The sampling rate of the RFID readers is set to 2 Hz and the sampling rate of the 3-axis accelerometer in each iMote2 is set to 128 Hz, and the sampling rate of audio recorder is set to 16 KHz. When a user performs activities, the acceleration readings from each iMote2 set are transmitted wirelessly to a local server (shown in Figure 2a, left or middle) which runs on a laptop PC with an iMote2 IPR2400 board connected through its USB port. When a user handles a tagged object, the RFID wristband reader scans the tag ID and sends it wirelessly to another server (shown in Fig. 2a, right) that can map the ID to an object name. In addition, human voice and environmental sound are recorded by the audio recorder. All the sensor data are logged separately with timestamps, and will be merged into a single text file as the activity trace for each user.

4 Multi-chained Temporal Probabilistic Model

In this section, we first describe our problem statement, then present a multi-chained temporal probabilistic activity model.

4.1 Problem Statement

We formulate our multi-user activity recognition problem as follows. We assume that there are a number of training datasets, where each training dataset corresponds to each user. Each training dataset O consists of T observations $O = \{o_1, o_2, \dots, o_T\}$ associated with activity labels $\{A_1, A_2, \dots, A_m\}$, where there are m activities and activities can be single-user ADLs or multi-user ADLs. For a new sequence of observations corresponding to a user, our objective is to train an appropriate activity model that can assign each new observation with the correct activity label.

4.2 Feature Extraction

After obtaining sensor readings, we first need to extract appropriate sensor features. We convert all the sensor readings to a series of *observation vectors* by concatenating all of the data observed in a fixed time interval which is set to one second in our experiments. Our feature extraction process generates a 47-dimensional *observation vector* every second. Different types of sensors require different processing to compute various features.

For acceleration data, we compute the features including mean, variance, energy, frequency-domain entropy, and correlation. The DC feature is the mean acceleration value over the window. Variance is used to characterize the stability of the signal. The energy feature captures the data periodicity, and is calculated as the sum of the squared discrete FFT component magnitudes of the signal. Frequency-domain entropy helps to discriminate activities with similar energy values, and is calculated as the normalized information entropy of the discrete FFT component magnitudes of the signal. Correlation is calculated between every two axes of each accelerometer and between all pairwise combinations of axes on different accelerometers. This feature aims to find out the correlation among the different axes of the two accelerometers.

For audio data, we compute both time-domain and frequency-domain features. The time-domain features measure the temporal variation of audio signal, and consist of three features. The first one is the standard deviation of the reading over the window, normalized by the maximum reading in the window. The second one is the dynamic range defined as (max - min) / max, where min and max represent the minimum and maximum readings in the window. The third is Zero-Crossing Rate (ZCR), which measures the frequency content of the signal and is defined as the number of time-domain zero crossings in the window. In the frequency-domain, we compute two features – centroid (the midpoint of the spectral power distribution) and bandwidth (the width of the range of frequencies that the signal occupies).

For RFID reading or location information, we use object name or location name directly as features. For each RFID wristband reader, we choose the first object in a one-second window since a user is unlikely to touch two or more objects in such a short interval. If no RFID reading is observed or in the presence of a corrupted tag ID, the value will be set to NULL.

We then transform these observation vectors into feature vectors. A feature vector consists of many feature items, where a feature item refers to a feature name-value pair in which a feature can be numeric or nominal. We denote a numeric feature as $numfeature_i$. Suppose its range is $[x, y]$ and an interval $[a, b]$ (or in other forms, $(a, b]$, $[a, b)$, or (a, b)) is contained in $[x, y]$. We call $numfeature_i@[a, b]$ a numeric feature item, meaning that the value of $numfeature_i$ is limited inclusively between a and b. We denote a nominal attribute as $nomfeature_j$. Suppose its range is $\{v_1, v_2, ..., v_n\}$, we call $nomfeature_j@v_k$ a nominal feature item, meaning the value of $nomfeature_j$ is v_k.

The key step of transformation is to discretize numeric features. We follow the entropy-based discretization method [3], which partitions a range of continuous values into a number of disjoint intervals such that the entropy of the partition is minimal.

Then we can directly combine the feature name and its interval into a numeric feature item. For the nominal feature, the feature name and its value are combined as a nominal feature item. For the $LEFTOBJ$ and $RIGHTOBJ$ features, we merge them into one feature by computing $LEFTOBJ \cup RIGHTOBJ$ without losing any essential objects during the user-object interaction due to user's handedness. All the feature items will be indexed by a simple encoding scheme and used as inputs to the probabilistic model described in the next section.

4.3 Coupled Hidden Markov Models

After feature extraction, we obtain a sequence of feature vectors for each user, where a feature vector $f = \{f_1, f_2, \ldots, f_T\}$ is associated with activity labels $\{A_1, A_2, \ldots, A_m\}$. A typical temporal probabilistic framework to model each user's sequence is HMMs which consist of a hidden variable and an observable variable at each time step. In this case, the hidden variable is an activity label, and the observable variable is a feature

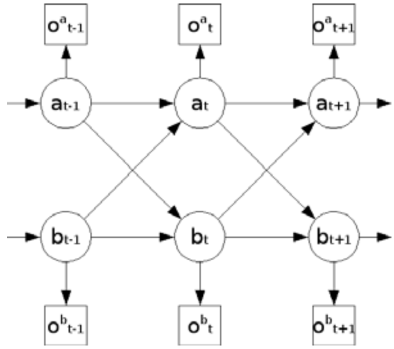

Fig. 3. Structure of CHMM

vector. For multiple sequences of observations corresponding to multiple users, we can factorize the basic HMMs into multiple channels to model interaction processes. We exploit CHMMs to model multi-user sequences involving user interactions. The CHMMs was originally introduced in [1], and it couples HMMs to capture inter-user influences across time.

To illustrate, as shown in Figure 3, there are two sequences of states A and B with observation O^a and O^b, respectively, at each time slice t. A two-chain CHMMs can be constructed by bridging hidden states of its two component HMMs at each time slice with the crosswork of conditional probabilities $P_{a_t|b_{t-1}}$ and $P_{b_t|a_{t-1}}$.

Inference for CHMMs. The posterior of a state sequence through fully coupled two-chain CHMMs is defines as follows.

$$P(S|O) = \frac{\pi_{a_1} P(o_1^a|a_1)\pi_{b_1} P(o_1^b|b_1)}{P(O)} \prod_{t=2}^{T} [P_{a_t|a_{t-1}} P_{b_t|b_{t-1}} P_{a_t|b_{t-1}} P_{b_t|a_{t-1}} P(o_t^a|a_t) P(o_t^b|b_t)] \quad (1)$$

where π_{a_1} and π_{b_1} are the initial probabilities of states, $P_{a_t|a_{t-1}}$ and $P_{b_t|b_{t-1}}$ are the inner-chain state transition probabilities, $P_{a_t|b_{t-1}}$ and $P_{a_t|b_{t-1}}$ are the inter-chain state transition probabilities modeling the interactions, $P(o_t^a|a_t)$ and $P(o_t^b|b_t)$ are the output probabilities of the states, we employ the Gaussian distribution in this case.

The CHMMs inference problem is formulated as follows. Given an observation sequence O, we need to find a state sequence S which maximizes $P(S|O)$. The inference algorithm – Viterbi – for HMMs could be applied to CHMMs as well with some modifications. The key point is, for each step, we need to compute both the inner-chain and inter-chain state transition probability, i.e., $P_{a_t|a_{t-1}} P_{b_t|b_{t-1}}$ and $P_{a_t|b_{t-1}} P_{b_t|a_{t-1}}$. The algorithm outputs the best state sequence S which involves two state sequences S_a and S_b corresponding to the recognized activity sequences for the two users.

Parameter Estimation in CHMMs. There are many existing algorithms for training HMMs such as Baum-Welch. Since a two-chain CHMM C can be constructed by joining two component HMMs A and B and taking the Cartesian product of their states, we define our training method as follows. We first train A and B following the maximum likelihood method, and then, we couple A and B with inter-chain transition probabilities which can be learnt from training datasets. This method is efficient since we do not need to re-train the CHMMs.

4.4 Activity Recognition Using CHMMs

For each user, the wearable sensors produce an observation sequence. The observation sequence will be divided into a training sequence and a testing sequence. We first compute a feature vector extracted from the observation sequence in every one second interval, which each feature vector contains many feature values as described in Section 4.2. We then obtain a sequence of feature vectors for each user. These sequences, corresponding to multiple users, will be the inputs to a CHMM.

To apply CHMMs for recognizing activities of multiple users, the model will be first trained from multiple training sequences corresponding to multiple users, and the trained models is then used to infer activities given multiple testing sequences. For example, to apply CHMMs, we build a CHMM for multiple users where each single-chain HMM is used for each user and each hidden state in the HMM represents one activity for the user. Given multiple training sequences, we train the CHMM using the parameter estimation method described in Section 4.3. When testing, the multiple testing sequences are fed into the trained CHMM, the inference algorithm then outputs multiple state sequences for each single-chain HMM as the labeled sequences for each of the users.

5 Experimental Studies

We now move to evaluate our proposed activity recognition models. We use the dataset collected in [5]. The dataset contains the observations of two subjects performing a total number of 21 activities (shown in Table 1) over a period of two weeks.

5.1 Evaluation Methodology

We use ten-fold cross-validation to generate our training and test datasets. We evaluate the performance of our methods using the time-slice accuracy which is a typical technique in time-series analysis. The time-slice accuracy represents the percentage of correctly labeled time slices. The length of time slice Δt is set to 1 second as our experiment shows different Δt doest not affect accuracy much. This time-slice duration is short enough to provide precise measurements for most of activity recognition applications. The metric of the time-slice accuracy is defined as follows.

$$Accuracy = \frac{\sum_{n=1}^{N} [predicted(n) == ground\_truth(n)]}{N} \tag{2}$$

where $N = \frac{T}{\Delta t}$.

Table 1. ADLs Performed in Our Trace Collection

Single-user ADLs				*Multi-user ADLs*	
0	brushing teeth	8	vacuuming	15	making pasta
1	washing face	9	using phone	16	cleaning a dining table
2	brushing hair	10	using computer	17	making coffee
3	making pasta	11	reading book/magazine	18	toileting (with conflict)
4	making coffee	12	watching TV	19	watching TV
5	toileting	13	eating meal	20	using computer
6	ironing	14	drinking		
7	making tea				

Table 2. Accuracy for CHMMs

User	**CHMMs**		
	Single-user ADLs	Multi-user ADLs	Overall
user 1	74.79%	96.91%	82.22%
user 2	85.11%	95.91%	88.71%
overall	79.95%	96.41%	**85.46%**

5.2 Results

In this experiment, we evaluate the accuracy of our proposed model. Table 2 shows the result. We achieve an accuracy of 85.46% for CHMMs. CHMMs perform better in multi-user ADLs than single-user ADLs. This probably can be explained as follows. For multi-user ADLs, the observations from both users are useful in the model since both users are performing the same activity. Hence, coupling two single-chain HMMs is effective to recognize multi-user activities. However, when recognizing single-user ADLs for one user, the observations from other users will become background noise, and hence they will have negative effects in model prediction. In this case, coupling two single-chain HMMs may weaken the posterior probability we compute for single-user activities.

Table 3 presents the confusion matrix for CHMMs. The activity serial numbers in both tables are identical to the numbers in Table 1. The values in each table are percentages, indicating that the percentage of the entire observation sequence for each activity is predicted correctly, and the percentages are predicted as other labels. For CHMMs, the activities such as *brushing hair* (single-user ADL), *toileting* (single-user ADL), *using computer* (single-user ADL), *toileting* (multi-user ADL) and *watching TV* (multi-user ADL) give the highest accuracies, and the activities *brushing teeth* (single-user ADL), *reading book/magazine* (single-user ADL), *watching TV* (single-user ADL) perform the worst. Most confusion takes place in the following cases:

Case 1: A single-user ADL is predicted as another single-user ADL. For example, the result shows that, for *making coffee*, while 74.7% of its entire observations are predicted correctly, 25.3% of them are predicted as *making pasta*.

Case 2: A single-user ADL is predicted as a multi-user ADL. For example, the result shows that, for *reading book/magazine* (single-user ADL), 24.4% of its observations are predicted as *watching TV* (multi-user ADL).

Case 3: A multi-user ADL is predicted as another multi-user ADL. For example, for *making pasta* (multi-user ADL), the result shows that 10.0% of its observations are predicted as *cleaning a dining table* (multi-user ADL).

Table 3. Confusion Matrix of CHMMs

		Ground Truth Activities																				
		0	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*
Predicted Activities	*0*	27.8	2.4	0	0	0	0	0	0	0	0	0	0	0	0.1	0	0	0	0	0	0	0
	1	72.2	97.6	0	0	0	0	0	0	0	0	0	0	0	5.0	0	0	0	0	0	0	0
	2	0	0	100.0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	3	0	0	0	92.9	25.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	4	0	0	0	7.1	74.7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	5	0	0	0	0	0	100.0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	6	0	0	0	0	0	0	93.7	1.1	0	0	0	0	0	0	0	0	0	0	0	0	0
	7	0	0	0	0	0	0	1.7	93.9	3.5	0	0	0	0	0	0	0	0	0	0	0	0
	8	0	0	0	0	0	0	0	0	95.6	0	0	0	0	0	0	0	0	0	0	0	0
	9	0	0	0	0	0	0	0	0	0	100.0	0	0	0	0	0	0	0	0	0	0	0
	10	0	0	0	0	0	0	0	0	0	0	100.0	0	0	0	0	0	0	0	0	0	0
	11	0	0	0	0	0	0	0	0	0	0	0	68.0	1.7	0	0	0	0	0	0	0	0
	12	0	0	0	0	0	0	0.4	0	0	0	0	4.8	43.3	0	0	0	0	0	0	0	0
	13	0	0	0	0	0	0	0.4	0	0	0	0	0.4	0	84.9	0	0	0	0	0	0	0
	14	0	0	0	0	0	0	0.3	0	0	0	0	0.4	0.1	0	90.2	0	0	0	0	0	0
	15	0	0	0	0	0	0	2.5	0	0	0	0	0.6	0.1	0	0	90.0	4.9	0	0	0	0
	16	0	0	0	0	0	0	0.3	0	0	0	0	0.4	0	5.5	0	10.0	95.1	8.8	0	0	0
	17	0	0	0	0	0	0	0.3	5.0	0	0	0	0.6	0.1	0.1	0	0	0	91.2	0	0	0
	18	0	0	0	0	0	0	0.2	0	0	0	0	0.4	0.1	4.4	0	0	0	0	100.0	0	0
	19	0	0	0	0	0	0	0	0	0.1	0	0	24.4	54.4	0	0.2	0	0	0	0	100.0	0
	20	0	0	0	0	0	0	0	0	0.8	0	0	0	0.1	0	9.7	0	0	0	0	0	100.0

6 Conclusions and Future Work

In this paper, we study the fundamental problem of recognizing activities of multiple users from sensor readings in ubiquitous computing, and propose a multi-chained temporal probabilistic model. Our evaluation results demonstrate the effectiveness of our model.

For our future work, we will investigate a possible future research direction – to recognize activities in a more complex scenario where single-user ADLs and multi-user ADLs are mixed with *interleaved* (i.e., switching between the steps of two or more activities) or *concurrent* (i.e., performing two or more activities simultaneously) activities. Recognizing activities in such a complex situation can be very challenging while we consider both single-user and multi-user activities at the same time, and hence, an in-depth study is required.

Acknowledgments. This work is supported by the National 973 project of China under Grant 2009CB320702, the National Natural Science Foundation of China under Grants NSFC 60721002 and NSFC 60736015, and the Jiangsu Natural Science Foundation under Grant BK2008017.

References

1. Brand, M.: Coupled hidden markov models for modeling interacting processes. Technical Report (November 1997)
2. Du, Y., Chen, F., Xu, W., Li, Y.: Recognizing interaction activities using dynamic bayesian network. In: Proc. of ICPR 2006, Hong Kong, China (August 2006)

3. Fayyad, U., Irani, K.: Multi-interval discretization of continuous-valued attributes for classification learning. In: Proc. Int'l. Joint Conf. on Artificial Intelligence, San Francisco (1993)
4. Gong, S., Xiang, T.: Recognition of group activities using dynamic probabilistic networks. In: Proc. of ICCV 2003, Nice, France, October 2003, pp. 742–749 (2003)
5. Gu, T., Wu, Z., Wang, L., Tao, X., Lu, J.: Mining Emerging Patterns for Recognizing Activities of Multiple Users in Pervasive Computing. In: Proc. of the 6th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous 2009), Toronto, Canada (July 2009)
6. Huynh, T., Blanke, U., Schiele, B.: Scalable recognition of daily activities from wearable sensors. In: Hightower, J., Schiele, B., Strang, T. (eds.) LoCA 2007. LNCS, vol. 4718, pp. 50–67. Springer, Heidelberg (2007)
7. Ivanov, Y.A., Bobick, A.F.: Recognition of visual activities and interactions by stochastic parsing. IEEE Trans. Pattern Recognition and Machine Intelligence 22(8), 852–872 (2000)
8. Katz, S., Ford, A.B., Moskowitz, R.W., Jackson, B.A., Jaffe, M.W.: Studies of illness in the aged. the index of adl: A standardized measure of biological and psychological function. Journal of the American Medical Association 185, 914–919 (1963)
9. Lester, J., Choudhury, T., Borriello, G.: A practical approach to recognizing physical activities. In: Fishkin, K.P., Schiele, B., Nixon, P., Quigley, A. (eds.) PERVASIVE 2006. LNCS, vol. 3968, pp. 1–16. Springer, Heidelberg (2006)
10. Logan, B., Healey, J., Philipose, M., Munguia-Tapia, E., Intille, S.: A long-term evaluation of sensing modalities for activity recognition. In: Krumm, J., Abowd, G.D., Seneviratne, A., Strang, T. (eds.) UbiComp 2007. LNCS, vol. 4717, pp. 483–500. Springer, Heidelberg (2007)
11. Oliver, N., Rosario, B., Pentland, A.: A Bayesian computer vision system for modeling human interactions. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(8), 831–843 (2000)
12. Park, S., Trivedi, M.M.: Multi-person interaction and activity analysis: A synergistic track- and body- level analysis framework. Machine Vision and Applications: Special Issue on Novel Concepts and Challenges for the Generation of Video Surveillance Systems 18(3), 151–166 (2007)
13. Philipose, M., Fishkin, K.P., Perkowitz, M., Patterson, D.J., Fox, D., Kautz, H., Hähnel, D.: Inferring activities from interactions with objects. IEEE Pervasive Computing 3, 50–57 (2004)
14. Smith, E.R., Mackie, D.M.: Social Psychology. Routledge, London (1999)
15. van Kasteren, T.L.M., Noulas, A.K., Englebienne, G., Kröse, B.: Accurate activity recognition in a home setting. In: Proc. of International Conference on Ubiquitous Computing (Ubicomp 2008), Korea (September 2008)
16. Wang, S., Pentney, W., Popescu, A.M., Choudhury, T., Philipose, M.: Common sense based joint training of human activity recognizers. In: Proc. Int'l. Joint Conf. on Artificial Intelligence, Hyderabad (January 2007)
17. Wyatt, D., Choudhury, T., Bilmes, J., Kautz, H.: A privacy sensitive approach to modeling multi-person conversations. In: Proc. IJCAI, India (January 2007)
18. Yacoob, Y., Black, M.J.: Parameterized modeling and recognition of activities. In: Proc. of International Conference on Computer Vision, ICCV 1998 (1998)
19. Zhang, D., Gatica-Perez, D., Bengio, S., McCowan, I., Lathoud, G.: Modeling individual and group actions in meetings: A two-layer hmm framework. In: Proc. of CVPRW 2004, Washington, DC, USA (2004)