Evaluating a new classification method using PCA to human activity recognition

M'Hamed Bilal Abidine, Belkacem Fergani
Faculty of Electronics and Computer Sciences
USTHB
Algiers, Algeria
abidineb@hotmail.com

***Abstract*—The ability to recognize human activities from sensed information is very important for ubiquitous computing applications using smart identification technologies. This paper addresses a new discriminative supervised method combining the classical Principal Components Analysis with the correlation criterion for performing activity recognition in a smart home. We conduct several experiments, demonstrate the achievement of our method and show its promising results using real world dataset.**

Keywords-activity recognition; ubiquitous computing; smart home; sensors network; machine learning

I. INTRODUCTION

In 2030, nearly one out of two households will include someone who needs help performing basic activities of daily living (ADL) such as eating, bathing, dressing and toileting [1]. A sensor-based technology in the home is the key of this problem. Particularly in elderly care, ADL are used to assess the cognitive and physical capabilities of an elderly person [2].

The ability to identify the behaviour of people in a smart home is at the core of ubiquitous computing applications. Smart systems are equipped with sensor networks able to automatically recognize activities about the occupants and assist humans. They must be able to recognize the ongoing activities of the users in order to suggest or take actions in an intelligent manner [3]. In this approach, sensors can be attached to either an actor under observation or objects that constitute the environment.

The sensor data collected in smart home often needs to be analysed using data mining and machine learning techniques to build activity models and perform further means of pattern recognition [4]. Recognizing a predefined set of activities is a classification task: features are extracted from the space and time information collected in the sensor data and then used for classification. Feature representations are used to map the data to another representation space with the intention of making the classification problem easier to solve. In most cases, a model of classification is used that relates the activity to sensor patterns. The learning of such models is usually done in a supervised manner and requires a large annotated datasets recorded in different situations.

State of the Art methods used for recognizing activities can be divided in two categories: generative models and discriminative models [5]. The generative methods perform well but require data modeling and are generally time consuming. Discriminative ones give good performance with a fast prediction speed but require generally also a large training data for supervised algorithms which in turn are also time consuming especially for large datasets such as Wireless Sensor Networks data.

Previous works on activity recognition has used a wide variety of generative models (e.g., Hidden Markov Models (HMM) [3]). Few works utilizing discriminative models have been developed in the activity recognition field. Motivated by the needs of activity recognition problems, and point out the advantage of the discriminative ones, we have developed in this paper, a new supervised classification method combining Principal Component Analysis [6] and the correlation criterion for performing activity recognition in a smart home. We have conducted several experiments using real ADL dataset in order to achieve the effectiveness of our discriminative method.

The rest of the paper is organized as follows. Section II presents related works in activity recognition. Section III describes our approach used in the activity recognition field. Section IV describes the dataset used in this paper and discusses the experimental results. Finally, Section V concludes by summarizing our findings.

II. RELATED WORKS

Activity recognition has been performed using different types of data. We can cite for example, state change sensors [7], motion detectors [8], cameras [9], accelerometers [10], RFID tags and sensors [3][11], electrical signatures [12], GPS [13] and various types of sensors [14]. These technologies include different levels of complexity and technological challenges in terms of price, intrusiveness, installation and the type of data they output [3][15].

The ubiquitous computing community has many interesting and creative ideas of how activity recognition can be applied. Activity recognition supervised models are based on large annotated datasets. Annotation has been performed in many different ways. The least interfering method is to use cameras [8]. Other examples are self-reporting methods, keeping an activity diary on paper or using a PDA [16].

978-1-4673-5214-7/13/$31.00 ©2013 IEEE

Methods used for recognizing activities can be divided in two categories: generative models (e.g., Hidden Markov Models (HMM) [3], Naive Bayes Classifier (NBC) [7]) and discriminative models (e.g., Support Vector Machines (SVM) [14], Conditional Random Fields (CRF) [3] and *k*-Nearest Neighbor (*k*-NN) [17]). In this study, our objective is to develop a new discriminative method, named Principal Components Classification (PCC) to perform automatic recognition of activities from binary sensors patterns in a smart home. It is based on Principal Components Analysis method which permits to find an optimal hyperplan to best discriminate binary classes according to a correlation criterion.

III. *Principal Components Classification*

The method developed in this paper is based on the principle of the PCA [5][6] that is a standard linear technique for dimensionality reduction. The goal of PCA is to find an orthonormal subspace whose basis vectors correspond to the directions with maximal variances, i.e., the main components that best describe the scatter of all the projected samples, see Figure 1.

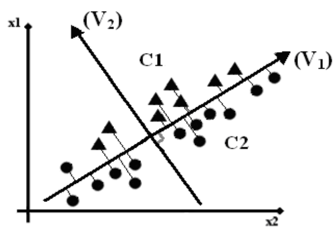

Figure 1. Principal components of two-variable dataset.

Consider a random vector $x \in R^n$ with the observations x_i, $i \in [1,...,m]$. In PCA, data matrix $X \in R^{m*n}$ are first centered $x \Leftarrow x - \mu(X)$ with $\mu(X) = \{\mu_i = E\{X_i\} \mid i = 1,...,n\}$. Then PCA diagonalizes the covariance matrix $Cov_{(X)}$

$$Cov_{(X)} = \frac{1}{m-1} X^T X. \quad (1)$$

This problem leads to solve the eigenvalue equation

$$\lambda V = Cov_{(X)} V. \quad (2)$$
$$\|V\| = 1.$$

Thus, it was thought to exploit this interesting property that is principal components for data classification in the supervised binary case: $y_i \in \{-1,+1\}$ where -1 is class label C_1 and +1 is class label C_2. In PCA, the first main component retains the most information about data, but it is not necessarily the optimal component V_{opt} that provides the optimal hyperplane (D_{opt}) for data discrimination. Hence, our method named Principal Components Classification (PCC) consists in searching a vector V_{opt} in an eigenvector basis for optimally classifying data. Our classification method requires a learning phase with training dataset $\{(x_1, y_1),(x_2, y_2)...(x_m, y_m)\}$. This phase is achieved as follows: we construct a component vector basis $V = \{V_i \mid i = 1,...,n\}$. Then, we search the decision function given in **(3)** of separation hyperplane (D) for each V_i by projecting the data on each component (xV_i) and taking the sign such that

$$f_i(x) = Sgn(x.V_i). \quad (3)$$

where Sgn is the sign function $(Sgn(z) = +1 \ \textit{if} \ z > 0, Sgn(z) = -1 \ \textit{otherwise})$.

Then, we search the optimal separation hyperplane (D_{opt}) which correspond the optimal main component $V_{opt} \in R^n$ that gives the best discrimination between two classes C_1 and C_2. In Figure 2, the first main component is V_1 but the optimal main component is V_2.

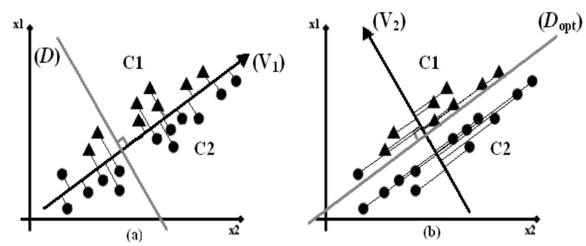

Figure 2. Classification PCC between two classes using optimal separating hyperplane (D_{opt}) with V_2 is V_{opt}.

The search of the optimal main component V_{opt} is determined by the maximum of correlation (*Corr*) between vector $f_i(x) = \{f_{1,i}(x), f_{2,i}(x),..., f_{m,i}(x) \mid i = 1,...,n\}$ and the original vector classes $y = \{y_1, y_2, \dots, y_m\}$, such that $Max\,(abs\,(Corr\,(f_i(x), y)))$. The *Pearson's correlation coefficient* is used

$$Corr\,(X = f_i(x), Y = y) = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}. \quad (4)$$

where σ_{XY} is the covariance between X and Y and σ_X, σ_Y are the standard deviation of X, Y, respectively.

The correction factor sign (S) in (5) is introduced to respect the choice of the label vector y with -1 is class label C_1 and +1 is class label C_2. The factor S is obtained by the sign of the correlation between the optimal function $f_{(V_{opt})/s=1}(x)$ given in (5) and y. This factor becomes negative (*resp*. positive) if the two vectors $f_{(V_{opt})/s=1}(x)$ and y are in the opposite direction (*resp*. collinear), see Figure 3. However, the decision function of optimal hyperplan (D_{opt}) is

$$f_{(V_{opt})}(x) = S * Sgn(x.V_{opt}) \text{ with } S \in \{-1,+1\}. \quad (5)$$

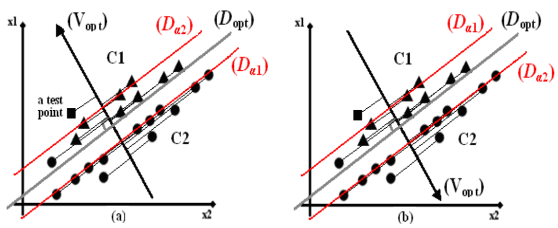

Figure 3. Classification PCC between two classes using optimal separating hyperplane (D_{opt}). (a) show the case when correlation is negative between $f_{(V_{opt})/s=1}(x)$ and y, (b) correlation is positive.

After choosing the optimal major component V_{opt}, we search the optimal factor translation $\alpha_{opt} \in R$ of our separation hyperplane that provides a good separation between classes. The choice of α_{opt} parameter is determined by taking the maximum of correlation between $f_{(V_{opt},\alpha)}(x)$ given in (6) and y, i.e., $Max(abs(Corr(f_{(V_{opt},\alpha)}(x), y)))$, such that

$$f_{(V_{opt},\alpha)}(x) = S * Sgn(x.V_{opt} + \alpha). \quad (6)$$

where the optimal translation factor value is set in the range $\alpha \in [\alpha_1,...,\alpha_2]$ with $\alpha_2 = -\alpha_1 = abs(\max(xV_{opt}))/2$.

The square test sample (see Figure 3) should be classified with the decision function

$$f_{(V_{opt},\alpha_{opt})}(x) = S * Sgn(x.V_{opt} + \alpha_{opt}). \quad (7)$$

The entire PCC training algorithm for each step is summarized below

TABLE I. PCC TRAINING ALGORITHM

Step1. Compute the covariance matrix $Cov_{(X)}$ from training data using (1).
Step2. Evaluate eigenvectors $V_i, i \in \{1...n\}$ using (2).
Step3. Compute the decision function for each eigenvector V_i using (3).
Step4. Determine the optimal eigenvector V_{opt} using $Max(abs(Corr(f_i(x), y)))$.
Step5. Find correction factor sign S of the function in (5). • If $Corr(f_{(Vopt)/s=1}(x), y) < 0$, S = -1. • Otherwise, S = +1.
Step6. Find the parameter $\alpha \in [\alpha_1,...,\alpha_2]$ using $Max(abs(Corr(f_{(V_{opt},\alpha)}(x), y)))$.

For Multiclass discrimination, the one-versus-one strategy [4] is used to classify between each pair with the binary decision function given in (7). This method consists in constructing $N(N-1)/2$ classifiers and each one is trained on data from two classes, a voting strategy is used for testing.

IV. EXPERIMENTAL RESULTS

In this section, we first give the details of our experimental setup, and then describe the dataset utilized in them. Finally, we present the acquired results.

A. Setup and Performance Measures

We separate our data into a test and a training set using a "leave one day out cross validation" approach [3]. In this approach, a classifier is designed using (l - 1) days and evaluated on the one remaining day; this is repeated l times, with different training sets of size (l - 1) and report the average performance measure. In this way, we get inferred labels for the whole dataset by concatenating the results acquired for each day.

We evaluate the performance of our model by measuring the accuracy. This measure is defined as follows

$$Accuracy \ : \frac{\sum_{i=1}^{m}[\inf erred(i)=true(i)]}{m}. \quad (8)$$

in which [a = b] is a binary indicator giving 1 when true and 0 when false. m is the total number of samples.

Sensors output are binary and represented in a feature space which is used by the model to recognize the activities performed. The "raw" sensor representation gives a 1 when the sensor is firing and a 0 otherwise. According to [3], the "change point" and "last" feature representation give better results than using "raw" sensor data directly. The "change point" representation gives a 1 when the sensor reading changes and the "last" representation gives a 1 for the sensor that changed state last and 0 for all other sensors. These experiments were run using our method PCC with the Last feature representation which gives better results of activity classification.

B. Database

We used Kasteren's realworld dataset [3] in order to recognize ADL. This dataset was recorded using a wireless sensor network in home with a single occupant of a 26-year-old man. He lives alone in a three-room apartment where 14 digital state change sensors were installed. Examples of sensors used include reed switches to measure open-close states of doors and cupboards; pressure mats to measure sitting on a couch or lying in bed; float sensors to measure the toilet being flushed.

The dataset consists of 245 actions for 7 different activities over l = 28 days, sensed using RFID technology. A list of activities that were annotated with information about class distribution can be found in Table II. Annotation was done by the subject himself at the same time the sensor data was recorded. A Bluetooth headset combined with speech recognition software was used for annotation.

TABLE II. NUMBER (NB) OF ACTIONS, NUMBER (NB) OF OBSERVATIONS AND PERCENTAGE OF TIME ACTIVITIES OCCUR IN THE HOUSE DATASET

ADL	Nb of actions	Nb of observations	Percentage of time (%)
Leaving	34	22617	63.9
Toileting	114	380	1.1
Showering	23	265	0.7
Sleeping	24	11601	32.8
Breakfast	20	109	0.3
Dinner	10	348	1.0
Drink	20	59	0.2

C. Results

We tested our method PCC under Matlab environment. The experiments were run using the last feature representation. Our proposed method PCC has good results with 96.9% for accuracy. The accuracy shows the percentage of correctly classified samples. In order to find out which activities are relatively harder to be recognized, we analyzed the confusion matrice of our method in Table III, which gives information about the actual and predicted classification results given by the classifier. It mainly gives better results for the 'Leaving', 'Showering' and 'Sleeping' activities but it has worse results for the others activities. Especially, the dominant activities with many samples 'Leaving' and 'Sleeping' are better recognized. Most confusion occurs between the three kitchen activities 'Breakfast', 'Dinner' and 'Drink'. It can be seen that the kitchen activities were performed in the same location using the same set of sensors are in general hard to recognize. For example, 'Toileting' and 'Showering' are more separable because they are in two different rooms, which makes the information from the door sensors enough to separate the two activities. Therefore, the sensors location used in smart home strongly influence recognition performance.

TABLE III. THE CONFUSION MATRIX FOR OUR PROPOSED METHOD. THE VALUES ARE PERCENTAGES.

Actvity	**Le**	**To**	**Sh**	**Sl**	**Br**	**Di**	**Dr**
Leaving	**98.7**	0.8	0.3	0.0	0.1	0.0	0.0
Toileting	46.3	**42.6**	1.3	3.1	2.4	4.0	0.2
Showering	10.6	5.6	**82.6**	0.0	1.1	0.0	0.0
Sleeping	0.2	0.2	0.2	**99.4**	0.0	0.0	0.0
Breakfast	25.7	28.4	4.6	1.0	**31.2**	7.3	1.8
Dinner	24.4	37.3	25.3	0.0	6.9	**5.2**	0.9
Drink	18.6	18.6	5.1	0.0	10.2	22.0	**25.4**

V. CONCLUSION AND FUTURE WORK

This paper introduces a new discriminative multi-class classification technique for activity recognition. This method based on PCA and the correlation criterion performs well. This method is sensitive to a dominant class 'Leaving' and 'Sleeping', but in general, without considering the context of class, the global recognition rate obtained with this method is good and results are promising. In the future, it would be interesting to use other datasets for performing our classification method. We aim to improve activities recognition of our proposed method PCC by exploiting the non linear relationships among the data using kernel-PCA method.

REFERENCES

[1] M.Wallace "Best practices in nursing care to older adults" in *Try This,* issue no. 2, Hartford Institute for Geriatric Nursing, 2007.

[2] S. Katz, T. D. Down, H. R. Cash, Progress in the development of the index of ADL. *Gerontologist*, 10:20–30, 1970.

[3] T. V. Kasteren, A. Noulas, G. Englebienne, B. Krose, "Accurate activity recognition in a home setting," In *Proc. the 10th Int. Conf. Ubiquitous Computing*, Tokyo, Japan, Sept. 11- 14, pp.1-8, 2008.

[4] C. Bishop, Pattern Recognition and Machine Learning. Springer. New York, ISBN: 978-0-387-31073-2, 2006.

[5] R. O. Duda, P. E. Hart, and D.G. Stork. Pattern Classification *(2nd edition) John Wiley and Sons*, 2000.

[6] I. Jolliffe, Principal component analysis, Springer Verlag, New York, 1986.

[7] E. M. Tapia, S. S. Intille, and K. Larson, "Activity recognition in the home using simple and ubiquitous sensors," in *Proc. Pervasive Computing*, Vienna, Austria, pp. 158 – 175, April 18-23, 2004.

[8] B. Logan, J. Healey, M. Philipose, E. M. Tapia and S. S. Intille, "A long-term evaluation of sensing modalities for activity recognition," In *Ubicomp* '07, 483–500, 2007.

[9] W. Lao, J. Han, P. H. N. de With, "Automative video-based human motion analysis for consumer surveillance system," *IEEE Trans. Consumer Electronics*, 55(2): 591-598, 2009.

[10] J. Lester, T. Choudhury, N. Kern, G. Borriello, B. Hannaford. "A hybrid discriminative/generative approach for modeling human activities," In *Proc. Int. Joint Conf. Artificial Intelligence*, Edinberg, UK, pp.766-772, Jul. 30-Aug. 5, 2005.

[11] M. Philipose, K. P. Fishkin, M. Perkowitz, D. J. Patterson, D. Hahnel, D. Fox & H. Kautz, Inferring Activities from Interactions with Objects, *IEEE Pervasive Computing,* Vol.3, No.4, pp. 50-57, 2004.

[12] S. Tsukamoto, H. Hoshino, and T. Tamura, "Study on indoor activity monitoring by using electric field sensor," in *International Conference of the International Society for Gerontechnology*, Pisa, Tuscany, Italy, June 4-7 2008.

[13] L. Lio, D. Fox, H. Kautz, "Extracting places and activities from GPS traces using hierarchical conditional random fields," *Int. J. Robotics Research*, 26(1): 119-134, 2007.

[14] A. Fleury, M. Vacher, N. Noury, " SVM-Based Multi-Modal Classification of Activities of Daily Living in Health Smart Homes : Sensors, Algorithms and First Experimental Results," *IEEE Transactions on Information Technology in Biomedicine*, Vol. 14(2), pp. 274-283, March 2010.

[15] A. Schmidt, *Ubiquitous Computing - Computing in Context*. Ph.D. Dissertation, Lancaster University, 2002.

[16] S. S. Intille, E. M. Tapia, J. Rondoni, and al., "Tools for studying behavior and technology in natural settings," *In Ubicomp*, pages 157–174, 2003.

[17] S. W. Lee and K. Mase, "Activity and location recognition using wearable sensors," *IEEE Pervasive Computing*, Vol.1, No.3, pp.24-32, 2002.