Accelerometer and GPS Sensor Combination Based System for Human Activity Recognition

Sahak Kaghyan
Armenian-Russian (Slavonic) University
Yerevan, Armenia
sahak.kaghyan@gmail.com

ABSTRACT

Cell phones and other mobile devices become a part of human culture and change activity and lifestyle patterns. Mobile phone technology continuously evolves and incorporates more and more sensors for enabling advanced applications. Latest generations of smart phones incorporate GPS and WLAN location finding modules, vision cameras, microphones, accelerometers, temperature sensors, etc. The availability of these sensors in mass-market communication devices creates exciting new opportunities for data mining applications. Particularly healthcare applications exploiting build-in sensors are very promising. These devices open a wide range of opportunities of using their potential in different branches like healthcare, financing and so on.

Current paper introduces an approach which allows recognizing activity, performed by human, using a smartphone, equipped with acceleration and positioning sensors. The classification stage was based on "learning with teacher" method. Incoming signal sequences collected from sensors of mobile device were analyzed using support vector machines (SVM) learning method.

Keywords

Human activity recognition, SVM, mobile devices, signal processing.

1. INTRODUCTION

Nowadays the Internet and personal computer are most common ways to connect people, allowing them to exchange information between each other. On the other side, none of these is able to reach each person anywhere and anytime like the cell phone does. Moreover, when the Internet became widely accessible on smartphones, cell phones, tablets and other mobile devices, achieve new wave of interest to them and made improved positions of mobile device manufacturers. So, what concerns to mobile technologies in general, now they are becoming ubiquitous all over the world, changing the way we communicate, conduct commerce, and provide care and services. Certainly some of the most compelling benefits of mobile technologies are in the areas of disease prevention, chronic disease management and improving healthcare delivery. For all the advances that are occurring in mobile health, or mHealth, its full potential for one very large group of beneficiaries – older adults and the persons who support them – is only starting to emerge. One of ways to help persons with health diseases is to give doctors an opportunity to monitor remotely their patients' life activity via cellular phones and smartphones they care. Thus, it is not a surprise that activity classification problems are widely discussed in this branch. Articles [1], [5], [7] present results and methods of activity detection via attaching sensors to different parts of body. The data to recognize human's movement activity is from the physical hardware sensors, and the combination of the accelerometer, the compass sensors and GPS are the most commonly used sensor devices. Existing activity recognition systems are constrained by practical limitations such as the number, location and nature of used sensors. Other issues include ease of deployment, maintenance, costs, and the ability to perform daily activities unimpeded. Sensors' outputs might vary for the same activity across different subjects and even for the same individual. Errors can also arise due to variability in sensor signals caused by differences in sensor orientation, placement, and from environmental factors such as temperature sensitivity.

1.1. Activity recognition approaches and machine learning

Human activity recognition (HAR) matured in recent years which will enable many health promotions and intervention applications. There are no standardized performance evaluation strategies. Recent efforts on designing public datasets might be one of the approaches to address this problem. Generally, activity recognition (AR) aims to identify the actions carried out by a person given a set of observations of itself and the surrounding environment. Three main classes of activity recognition are considered including coarse location tracking, video stream analysis and inertial navigation systems (INS) such as accelerometers. Sensor data are typically communicated from sensors to servers for data processing. Alternatively signal processing can be performed in mobile devices such as smart-phones. Many authors usually don't use standard tests for accuracy rate checks and validity of most reported results depends on testing specifics. There is no consensus even on a standard list of activities, but most of the reports include "walking", "sitting", "jogging" and "standing" patterns. Recognition can be accomplished, for example, by exploiting the information retrieved from inertial sensors such as accelerometers [4]. In some smartphones these sensors are embedded by default and we benefit from this to classify a set of physical activities (standing, walking, laying, walking upstairs and walking downstairs) by processing inertial body signals through a supervised Machine Learning (ML) algorithm for hardware with limited resources. So, in general, activity recognition algorithms can be divided into two major categories. The first one is based on supervised and unsupervised machine learning methods. Supervised learning requires the use of labeled data upon which an algorithm is trained.

Following training the algorithm is then able to classify unknown data. The steps are the following: (1) acquire sensor data representative of activities, including labeled annotations of what an actor does and when, (2) determine the input data features and its representation, (3) aggregate data from multiple data sources and transform them into the application dependent features, e.g., through data fusion, noise elimination, dimension reduction and data normalization, (4) divide the data into a training set and a test set, (5) train the recognition algorithm on the training set, (6) test the classification approach on the test set, and finally (7) to apply the algorithm in the context of activity recognition. Steps (4) to (7) can be repeated with different partitioning of the training and test sets in order to achieve better performance. The algorithms and models for supervised learning and activity recognition include Hidden Markov Models (HMM), dynamic and naive Bayes networks [9]-[11], decision trees [12], nearest neighbor [4], [13] and SVM [18] approaches. HMMs and Bayes networks are currently the most commonly used methods in activity recognition even though they require extensive computational resources. Multicore computers and clusters are typically used for these types of classifications. The number of machine learning models that have been used for activity recognition varies almost as greatly as the types of activities that have been recognized and types of sensor data that have been used. Solutions range from naive Bayes classifiers to support vector machines [1]-[6].

On the other hand, all approaches have limitations and efficiency strengths depending on sensor “hardware” that was used during information retrieving and used algorithms. A special focus of the paper is on mobile devices that are inherently wearable and equipped with GPS, accelerometers and so on that can be used to assess activity. Unsupervised learning is based on unlabeled data and applies the following steps: (1) acquire unlabeled sensor data, (2) aggregate and transform them into features; (3) model data by e.g. clustering techniques. The second broad category exploits logical modeling and reasoning. The steps are the following: (1) use a logical formalism to explicitly define and describe a library of activity models, (2) aggregate and transform sensor data into logical terms, and (3) perform logical reasoning based on observed actions, which could explain the observations.

2. RELATED WORKS

Recognizing a predefined set of activities is a recognition (classification) task: features are extracted from the space-time information collected by sensors 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 (human labeling) and requires a large annotated datasets recorded in different settings. Smart phones include various sensors such as gyroscopes, accelerometers, proximity sensors and have become affordable and ubiquitous. Convenient user interfaces make them attractive for all population groups. Oner et al [1] presented early work on a pedometer mobile application that was coupled with e-mail to notify medical assistants or family members. Their purpose was to use mobile smart phone to detect a fall event regardless of the phone position or orientation. Algorithm that was introduced in article was based on acceleration peak detection and was tested for different conditions. Das et al [3] introduced an attempt to recognize activity using Motorola Droid smartphone. Activity classification was done through several stages: data acquisition, signal processing, feature extraction and classification. Using the nearest neighbor classifier the program could predict patterns or activities with 93% accuracy after it had been calibrated for a particular user. Individual gestures were recognized with as much accuracy as activities but once again machine learning had to come first. A better classifier would work for more people without having to go through the individual training process.

3. SVM METHOD IN SIGNAL PROCESSING

Many activity recognition systems use one or several wearable sensors attached to different parts of human body to collect data and transfer them to a nearby server station. There exists a vast literature on wearable sensors, mostly accelerometers. Massive sensor deployments and related studies are constrained in number. A large class of activity recognition methods exploits sensors embedded in mobile devices, which potentially overcomes deployment constraints due to broad availability. In terms of wide usability, smartphones that are equipped with various sensors (audio, video or motion detection) can be considered as a perfect tool for short-term physical activity recognition. Broader list of useful mobile device sensors includes imaging camera, microphones, accelerometers, gyros and compasses, ambient light detectors, proximity sensors, location sensors (combination of GPS, WLAN and network), WLAN and other wireless network signal readings.

Our current research does recognition process based on machine learning method. SVM classifier works on training sets. Implemented mobile application based on Android operating system, allows user of smartphone to create and calibrate training sets in order to improve recognition accuracy. For each of primitive activities (walking, sitting, jogging, etc.), user can input number of patterns matching to his actions. These patterns asynchronously save in SQLite portable mobile database of smartphone. For each activity signals were collected for 10 seconds (Δt time interval). Having a set of predefined patterns allowed us to create hyper planes in order to distinguish different activities from each other (Fig.1).

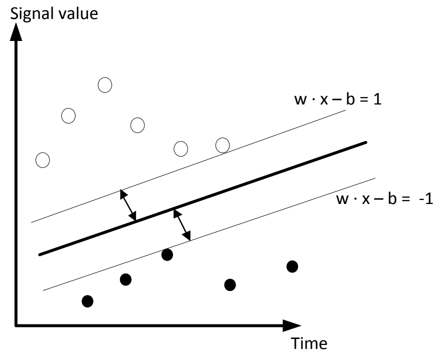

Fig. 1. Finding the optimal hyper plane to divide signal sets.

Consider that we have a finite set of labels representing each activity – A = {A_1… A_N}. Each training set for each activity performed Δt seconds. Let us denote as Δ the set of feasible distances between two neighbor points of two different instances of single activity A_i as follows: Δ = {d_1, …, d_N}. Each activity has an own set of trained sequences, necessary for calibration. Let us denote them as follows:

$$TS\_A_i = \{TS\_A_i^1, \dots, TS\_A_i^{s_i}\}, s_i \geq 0, 1 \leq \mathrm{i} \leq \mathrm{N}$$

Each $TS\_A_i^{s_p}$ element (1≤i≤N, 1≤p≤N) is a sequence of two dimensional or three dimensional vectors (depending on sensor, from which data is acquired). Currently we shall be using three dimensional vector notations. Generally, each activity can hold different number of training sets, but for ease of marking, let us consider that $s_i = s_j$, 1 ≤ i, j ≤ N. As a result, we can represent the whole training set of activities with the following matrix:

$$TS\_A = \begin{pmatrix} TS\_A_1^1 & \cdots & TS\_A_1^{s_1} \\ \vdots & \ddots & \vdots \\ TS\_A_N^1 & \cdots & TS\_A_N^{s_N} \end{pmatrix} \qquad (1)$$

where $TS\_A_i^j$ has the following structure:

$$\begin{cases} TS\_A_i^j = \left(p_1^{A_i^j}, \quad \dots, \quad p_t^{A_i^j}\right) \\ p_k^{A_i^j} = \left(x_k^{A_i^j}, \quad y_k^{A_i^j}, \quad z_k^{A_i^j}\right) \\ 1 \leq \mathrm{i} \leq \mathrm{N}, \ 1 \leq \mathrm{k} \leq \ \mathrm{t}, \ 1 \leq \mathrm{j} \leq s_i \end{cases} \qquad (2)$$

Let us also denote as *r* number of rows of current matrix and as *h* – the number of columns.

Unlabeled pattern must be processed through the following stages 1) noise reduction, 2) feature extraction, 3) learning and inference and 4) activity recognition. We used a median filter with 5 sequential neighbor points comparing to avoid noisy points. Having, the list of sequences representing the given activity we constructed an average weighted sequence for each labeled activity using Δ set of feasible bounds between neighbor points as limitation tool. As a result we gain one sequence of average signal which represents an activity and each point in sequence, which also has an own weight. So, when the unlabeled/ test activity is compared with every instance of average activities set, along with distance of points will be calculated importance of comparing points depending on weight of average point. Here we denote as L_{AVG} the set of average sequences representing each activity:

$$\begin{cases} L_{avg} = \{L_{avg}^1, \quad \dots, \quad L_{avg}^N\}, \ 1 \leq \mathrm{i} \leq \mathrm{N} \\ L_{avg}^i = \left(P_{avg}^{i\,1}, \quad \dots, \quad P_{avg}^{i\,\Delta t}\right), \\ P_{avg}^{i\,k} = \left(w_{avg}^{i\,k}, \quad p_{avg}^{i\,k}\right), \quad 1 \leq \mathrm{k} \leq \quad \mathrm{t} \\ p_{avg}^{i\,k} = \left(x_{avg}^{i\,k}, \quad y_{avg}^{i\,k}, \quad z_{avg}^{i\,k}\right) \end{cases} \qquad (3)$$

Here $L^{\ i}_{\ avg}$ is Δt dimensional vector of objects, representing the average instance for each labeled activity. Every object, in its turn, is a combination of $w_{a\ v\ g}^{\ i\ \ k}$ non-negative weight value and $p^{i}_{\ a\ v}{}^{k}_{\ g}$ 3-dimensional point. Coordinates of the average point calculated as mean of proper coordinates of the same time frames along each of axes and the average weight of specified point calculates below using the formula (5):

$$p_{avg}^{i\,k} = \frac{1}{h}\sum_{m=1}^{s_i} p_m^{A_i^k} =$$

$$\left(\frac{1}{h}\sum\nolimits_{m=1}^{s_i} x_m^{A_i^k}, \frac{1}{h}\sum\nolimits_{m=1}^{s_i} y_m^{A_i^k}, \frac{1}{h}\sum\nolimits_{m=1}^{s_i} z_m^{A_i^k}\right) \qquad (4)$$

$$w_{avg}^{i\,k} = \frac{1}{h-1}\sum\nolimits_{m=1}^{s_i-1} sumw\left(p_m^{A_i^k}, \, p_m^{A_{i+1}^k}\right) \qquad (5)$$

Here sum of average weight for neighbors of activity in given *i* time frame calculates as follows:

$$sumw\left(p_m^{A_i^k}, \quad p_m^{A_{i+1}^k}\right) = \begin{cases} \xi_i, if\ dist\left(p_m^{A_i^k}, \quad p_m^{A_{i+1}^k}\right) > d_i \\ \qquad 1, \qquad otherwise \end{cases} \qquad (6)$$

where value of variable 0≤ ξ_i ≤1 is coefficient of weight decrease those two neighbor points which have distance more than feasible bound for given activity class A_i. Distance itself is calculated as Euclidean distance between vector points as follows:

$$dist\left(p_m^{A_i^k}, \quad p_m^{A_{i+1}^k}\right) = \qquad (7)$$

$$\sqrt{(x_m^{A_i^k} - x_m^{A_{i+1}^k})^2 + (y_m^{A_{i+1}^k} - y_m^{A_{i+1}^k})^2 + (z_m^{A_{i+1}^k} - z_m^{A_{i+1}^k})^2}$$

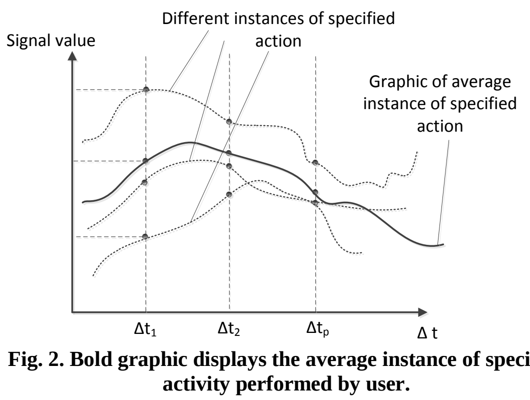

Fig. 2. Bold graphic displays the average instance of specified activity performed by user.

Figure 2 illustrates the average graphic for one specified activity, based on calculations of mean points for each time frame of different training instances. Result graphic was estimated according to values of signals of several labeled instances representing specified activity. After that unlabeled activity must be sequentially compared with average activities in order to find the most matching one:

$$weighted\_distance(L_{avg}^{i}, L_{unlabeled}) \rightarrow min \quad (8)$$

During calculations our method uses weights of signals and SVM algorithms in order to increase the answer accuracy and to minimize noise.

4. CONCLUSION

Smartphone is a new category of mobile phones that can perform computing just like a personal computer, but with smaller resources capability. Many sensors are already embedded smartphones which, thus, can be considered as a perfect tool for short-term physical activity recognition. Thus, these wearable devices can be used for unobtrusive activity recognition. Smartphones are also able to provide a wide range of connectivity option in one integrated device. Ultimately, these devices have been very personalized in human's daily life so that implementation using smartphone will relieve users to carry wearable sensors creating discomfort. Our approach introduces a method of classifying collected, stored and transferred from mobile device to server database signals. This approach constructs a set of average signal sequences for given activities (one average instance for each activity) and generates weights for each average point each instance, then algorithm implements SVM algorithm and uses point weights during execution for final classification. In our future work we intend to process the classification stage straight on mobile device and process real-time classification.

REFERENCES

[1] Oner M., Pulcifer-Stump J. A., Seeling P. and Kaya T., "Towards the Run and Walk Activity Classification through Step Detection – An Android Application", 34th Annual International Conference of the Engineering in Medicine and Biology Society, San Diego, CA, (2012).

[2] Vapnik, V.N., "Statistical Learning Theory", New York: John Wiley & Sons, 1998.

[3] Das, S., Green, L., Perez, B., Murphy, M., Perring, A., "Detecting user activities using the accelerometer on Android smartphones", (2010).

[4] Kaghyan, S., Sarukhanyan, H., "Activity Recognition Using K-Nearest Neighbor Algorithm on Smartphone with Tri-axial Accelerometer", International Journal of Informatics Models and Analysis (IJIMA), ITHEA International Scientific Society, Bulgaria, pp. 146-156, (2012).

[5] Parkka, J., Ermes, M., Korpipaa, P., Mantyjarvi, J., Peltola, J., Korhonen, I., "Activity Classification Using Realistic Data From Wearable Sensors", IEEE Transactions on Information Technology in Biomedicine, 10 (1), (2006).

[6] Mannini, A., Sabatini, A. M., "Machine Learning Methods for Classifying Human Physical Activities from on-body sensors", Sensors 10, pp. 1154–1175, (2010).

[7] He, Z., Jin, L., "Activity Recognition from Acceleration Data Based on Discrete Cosine Transform and SVM", Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics San Antonio, USA, (2009).

[8] Chernbumroong, S., Atkins A. S., Yu H., "Activity classification using a single wrist-worn accelerometer", 5th IEEE International Conference on Software, Knowledge, Information Management and Applications, Benevento, Italy, (2011).

[9] Yang, J., Schilit, B. N, McDonald, D. W., "Activity recognition for the digital home," IEEE Computing and Processing, (41), pp. 102–104, (2008).

[10] Efros, A. A., Berg, A., Mori, G., Malik, J., "Recognising Action at a Distance", Proceedings of the International Conference on Computer Vision, Nice, France, (2003)

[11] Tapia, E. M., Intille S., "Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor", International Symposium on Wearable Computers (ISWC), (2007).

[12] Lee, S.-W., Mase, K., "Activity and location recognition using wearable sensors", IEEE Pervasive Computing, 1 (3), pp. 24–32, (2002).

[13] Huynh, T., Schiele, B., "Unsupervised Discovery of Structure in Activity Data using Multiple Eigenspaces", 2nd International Workshop on Location- and Context-Awareness (LoCA), (3987), (2006).

[14] Ikizler, N., Duygulu, P., "Human Action Recognition Using Distribution of Oriented Rectangular Patches. In Human Motion", ICCV'07, (2007).

[15] Fikri Azli bin Abdullah, M., Fahmi Perwira Negara, A., Sayeed, S. Md., Choi, D., Sonai Muthu, K., "Classification Algorithms in Human Activity Recognition using Smartphones", International Journal of Computer and Information Engineering 6, pp. 77–84 (2012).

[16] Thompson, P. D., "Exercise and Physical Activity in the Prevention and Treatment of Atherosclerotic Cardiovascular Disease: A Statement from the Council on Clinical Cardiology (Subcommittee on Exercise, Rehabilitation, and Prevention) and the Council on Nutrition, Physical," Arteriosclerosis, Thrombosis, and Vascular Biology (23), 42–49 (2003).

[17] Kaghyan, S., Sarukhanyan, H., Akopian, D., "Human Movement Activity Classification Approaches that use Wearable Sensors and Mobile Devices", IS&T/SPIE Electronic imaging symposium, Conference on multimedia and mobile devices, Vol. 8667, Burlingame, CA, (2013).