Inertial Sensor Based Modelling of Human Activity Classes: Feature Extraction and Multi-sensor Data Fusion using Machine Learning Algorithms

Tahmina Zebin[1], Patricia J Scully[2], and Krikor B. Ozanyan[1,2]
[1]School of Electrical and Electronic Engineering, [2]Photon Science Institute
University of Manchester, Manchester, M13 9PL, UK
{tahmina.zebin; patricia.scully; k.ozanyan} @ manchester.ac.uk

Abstract. Wearable inertial sensors are currently receiving pronounced interest due to applications in unconstrained daily life settings, ambulatory monitoring and pervasive computing systems. This research focuses on human activity recognition problem, in which inputs are multichannel time series signals acquired from a set of body-worn inertial sensors and outputs are automatically classified human activities. A general-purpose framework has been presented for designing and evaluating activity recognition system with six different activities using machine learning algorithms such as support vector machine (SVM) and artificial neural networks (ANN). Several feature selection methods were explored to make the recognition process faster by experimenting on the features extracted from the accelerometer and gyroscope time series data collected from a number of volunteers. In addition, a detailed discussion is presented to explore how different design parameters, for example, the number of features and data fusion from multiple sensor locations- impact on overall recognition performance.

Keywords: Inertial Measurement Unit, Accelerometer Data, Feature Extraction, Data-Fusion, Machine Learning Algorithms, Human Activity Recognition

1 Introduction

Despite significant research efforts over the past few decades, activity recognition still remains a challenging problem. Wearable sensor based Human Activity Recognition (HAR) is currently playing a key role in the development of innovative human-machine interfaces and assistive technologies[2]. The information obtained from human physical activity is valuable in the long-term assessment of biomechanical parameters and physiological variables, which can then be used to support care of the elderly, the chronically ill and people with special needs [3]. Moreover, for accurate monitoring of physical activity, information on the type, intensity, and duration of the activities is of substantial interest to the research community [4].

Over several years, studies of gestures and activity recognition have been confined to clinical settings and conventional lab-based equipment, such as stationary and expensive 3D motion capturing systems and force plates[5]. For studying activities in unconstrained daily life settings, body-worn inertial sensors are emerging as a preferable research option in many cases [6, 7]. In addition, these systems are portable, more affordable than their laboratory counterparts. Hence, in this paper, we have developed an off-the-shelf lower body inertial sensor system. The system is designed and built as a set of 5 sensor units initially, each with an integrated MPU-9150 IMU to capture motion data. The system is specifically designed to study lower body motion. The sensors are connected via ribbon cables to a single control hub based on an Arduino board and an XBee transmitter. The data from the sensor (accelerometer and gyroscope) is post-processed to facilitate an automatic classification of the activities performed.

For modelling and evaluating physical activity, a general-purpose machine learning framework is presented in this paper. The framework comprises components for data acquisition and pre-processing, data segmentation, feature extraction and selection[1, 8], training and classification, decision fusion[9], and performance evaluation. It should be noted that, machine learning based algorithms for recognition of gestures and activities is a relatively new application area, and we provide a systematic insight on the use of classification algorithms (e.g. SVM, ANN) in MATLAB for some common physical activities.

2 Design of the Human Activity Recognition Chain

A typical Human Activity Recognition (HAR) system contains a stream of sensor data at the input stage acquired using multiple sensors worn on the body. The sensor data is then pre-processed to filter out signal variability or artefacts. The processed data is then segmented to isolate the region of interest of the activity or gesture. Afterwards, features that capture the activity characteristics are extracted from the signals within each segment[8]. In the training stage, the extracted features and corresponding ground truth class labels are used as input to train a classifier model in the training stage. In classification stage, the features and a previously trained model are used to calculate a score for each activity class and to map these scores into a single class label in the classification stage. If multiple sensors or classifiers are considered, the output of several classifiers may subsequently be fused. In addition, a performance evaluation stage allows the assessment of the performance of the recognition system[1]. In the rest of the paper, significant stages (shown in Figure 1) are used and the design decisions we made for the activity recognition task in hand is presented in detail.

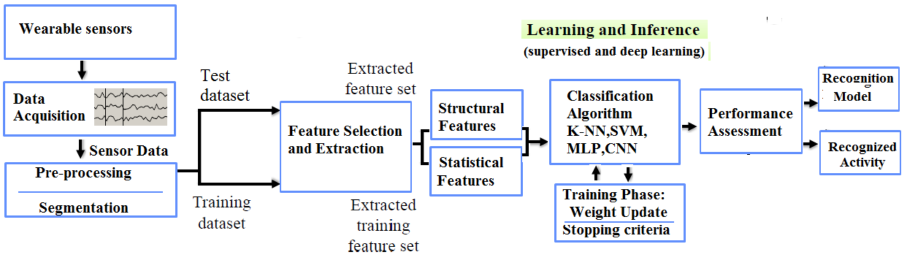


Fig. 1. Components of the human activity recognition chain[1])

2.1 Sensor Data collection and Preprocessing

In the first stage of a typical activity recognition system, raw data is acquired using several sensors attached to different locations on the body. In our research the activities were tracked using five sensing units (model: MPU-9150) placed at (a) Sensor 1: Pelvis/waist region, (b) Sensor 2 and 3: Left and right thigh, (c) Sensor 4 and 5: Left and right shank of the volunteer. A schematic diagram of the Inertial Measurement System used for this research is shown in Figure 2.

A single sensing unit is comprised of a 3-axis accelerometer and a 3-axis gyroscope recording timestamped motion data at a sampling rate of 50 Hz. All the recorded data was sent via XBee to a laptop placed in close proximity to the participant. Five volunteers performed a continuous sequence of six generic ambulatory activities[7] listed in Table 1. The activity was repeated 10 times for each participant, resulting in a dataset of about 120 minutes.

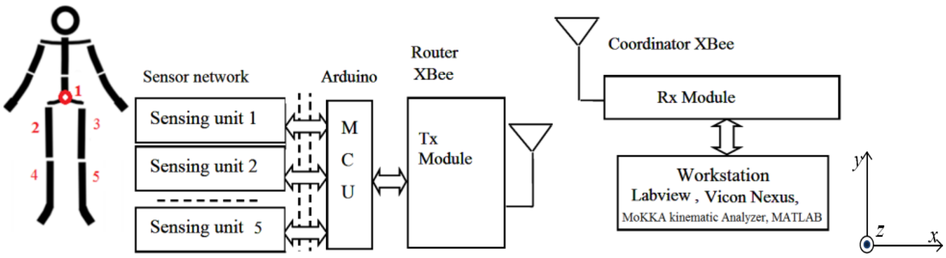


Fig. 2. Schematic of the wearable inertial sensor system and orientation of the sensing unit.

Table 1. Categorical Physical Activity and Activity ID for the activity recognition task

Activity	Activity ID
Walking	1
Walking_upstairs	2
Walking_downstairs	3
Sitting	4
Standing	5
Lying down	6

2.2 Extraction and Selection of Features

Manual selection of features is a difficult task. The higher the dimensionality of the feature space, the more training data is needed for model parameter estimation and the system becomes more computationally intensive. For real-time processing on embedded systems the objective is to minimize memory, computational power and bandwidth requirement. It is therefore important to use an optimum number of features that will still allow the system to achieve the desired target performance. Current literature uses a wide range of features such as signal based features [10] (e.g. mean, variance, FFT – coefficients, spectral entropy, and overall energy.) Other than that, body model based parameters (calculated from a 3D skeleton using multiple on body sensors) incorporating prior knowledge can lead to higher performance and increase robustness from person to person [3, 6].

By identifying the most salient features for learning, the most useful aspects of the data is used for analysis and future prediction. The hypothesis explored in this research is that feature selection for classification tasks can be accomplished on the basis of convolution[11] and pooling of features, and that such a feature selection process can be beneficial to a variety of common machine learning algorithms. Here, we have utilized the statistical and spectral features from segmented time series data as the features to be processed by the classification algorithm. The 66 features computed from the inertial sensor's accelerometer data are listed in table 2.

Table 2. List of extracted features from accelerometer data for each activity for the activity recognition scenario

Feature Name	Feature Number	Additional Information
Average value(1 each)	feature(1) feature(2) feature(3)	For all three acceleration components (x,y,z direction)
RMS value(1 each)	feature(4) feature(5) feature(6)	All three acceleration components
Autocorrelation features(3 each)	feature(7:9) feature(10:12) feature(13:15)	Height of main peak; height and position of second peak
Spectral peak features(12 each)	feature(16:27) feature(28:39) feature(40:51)	Height and position of first 6 peaks
Spectral power features (5 each)	feature(52:56) feature(57:61) feature(62:66)	Total power in 5 adjacent and pre-defined frequency bands

3 Training and Classification using Machine Learning Algorithms

The classifier itself influences the recognition performance of an activity recognition system. The decision for or against different classifier can be made either by having lower computational complexity or simply by superior performance. In our research, we have investigated the performance of several classifiers used in activity recognition to suggest an automated and alternative approach to hand-crafted feature extraction and classification techniques.

Classification techniques such as Nearest Neighbors, Naïve Bayes (NB), Support Vector machine (SVM) and Multi-layer Perceptron (MLP) based neural networks has been tested in this research. In the following sections, we will explore the capabilities and efficiency of two machine learning algorithms: Support vector Machine and multi-layer perception on inertial sensor based human activity recognition data. It should be noted that, machine Learning approaches such as SVM and MLP includes kernel based and random forest feature selection mechanism ensuring the generalization of the relevant features.

3.1 Neural Network based Classification

Neural networks are capable of performing pattern-recognition techniques useful in the analysis of gait dynamics [12]. In this section activity classification was performed with a MATLAB based multilayer perceptron (MLP) model as a neural network. The multilayer perceptron consists of three or more layers (an input and an output layer with one or more hidden layers) of nonlinearly-activating nodes. Since an MLP is a Fully Connected Network, each node in one layer connects with a certain weight w_{ij} to every node in the following layer. The weight of each node is adjusted in a manner so that minimize the error in the entire output.

a) b)

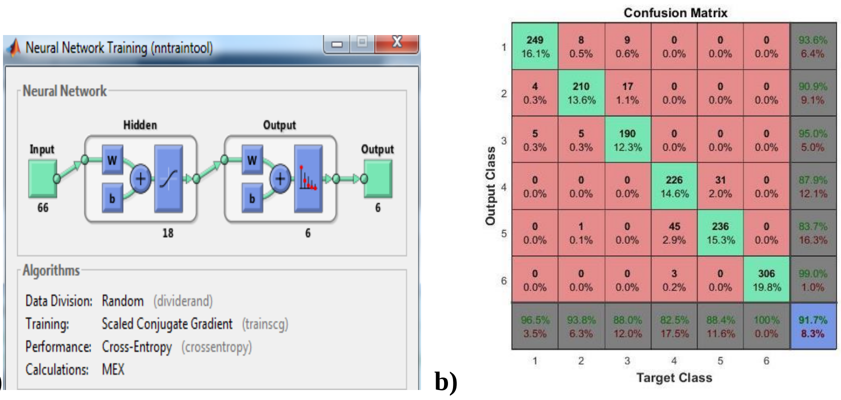
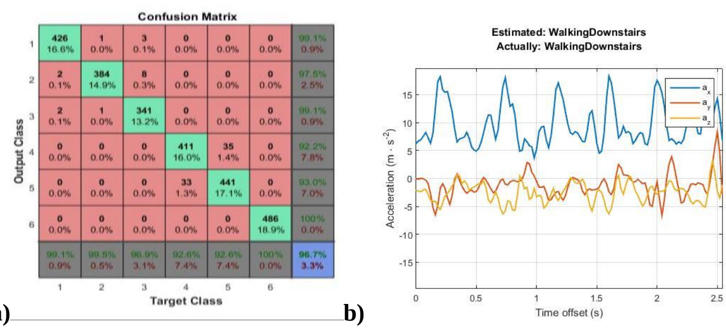


Fig. 3. (a) MATLAB neural network Train tool has been utilized to train the neural network. (b) The confusion matrix shows an accuracy rate of 91.7% for activity recognition for neural network based activity classification.

Learning occurs in the perceptron by changing connection weights after each piece of data is processed, based on the amount of error in the output compared to the expected result. The learning is carried out through backpropagation, a generalization of the least mean squares algorithm in the linear perceptron. To quantitatively assess the performance of a classification algorithm we have predicted the activities for a small test dataset, and compared them against the known class values. To visually represent the accuracy, a confusion matrix is used in this paper. The confusion matrix is a square matrix that summarizes the cumulative prediction results for all couplings between actual and predicted classes, respectively. As indicated in figure 3, it was observed that there has been above 12% misclassification of walking downstairs and sitting activity based on the accelerometer signal based features. Whether features from the gyroscope improve the accuracy, is yet to be explored. In addition, training the network with a bigger database from more volunteers is planned as a part of future research.

3.2 Support Vector Machine based Classification

The Support Vector Machine (SVM) technique is a powerful machine-learning algorithm based on its ability to find non-linear patterns. The classifier is trained at the first stage with a specific activity and their known classes. A MATLAB based ‘Classification Learner’ App[13] has been used here to auto generate functions to train a classifier based on the dataset. The returned arguments include information of how the dataset is partitioned during the training phase. The remaining samples of the dataset can be used for testing the accuracy of the classifier. The prediction result is visualized in a confusion matrix. Figure 4 shows the Confusion Matrix when the data is classified and tested using support vector machine. During this initial stage of testing 96.7% of the activities were classified correctly. The accelerometer based feature for walking downstairs and sitting down caused 7.4% of false hits which need further specification in classifying that activity. As can be observed from the results the best performance was obtained for SVM classifier because of the suitability of the kernels to the activities we chose to classify.



a) b)

Fig. 4. (a) Confusion Matrix when the data is classified using support vector machine; (b) Screenshot of an activity classified correctly by the recognition system.

4 Effect of Multi-Sensor Data Fusion

During our studies, we also experimented on the impact of different sensor modalities on the activity recognition performance. This part of the study is conducted using a the K-NN (nearest neighbor) classifier using MATLAB classification learner app[11].In order to quantitatively understand the recognition performance, some standard metrics such as accuracy, recall, precision and confusion matrices were used.

From the results presented in figure 5 and 6, a strong influence on the recognition accuracy can be observed with the combination of sensors. Figure 5 shows that, the precision of person dependent activity recognition changes from 90% to 94.1% when sensor data from the shank and thigh are also used along with the pelvic sensor data. It was observed that while some parts of the sensor data (e.g. single axis from the accelerometer or gyroscope) do contribute to a precise classification whilst some other axial data might introduce noise. Other than that, the classification performance is found to be 65.7% precision for a person independent scenario (where the classifier is trained with activities from multiple volunteers).

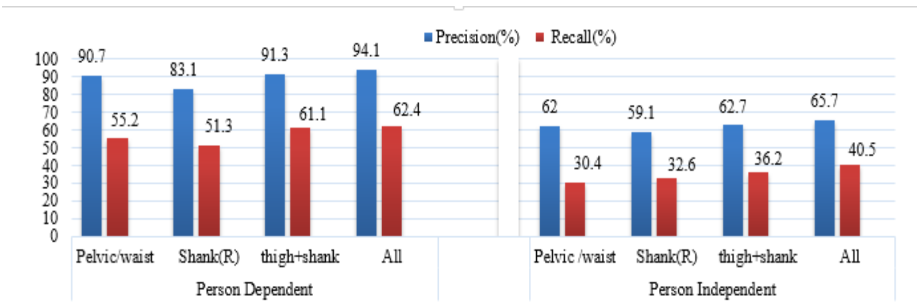


Fig 5. Activity recognition performance for different sensor position combinations.

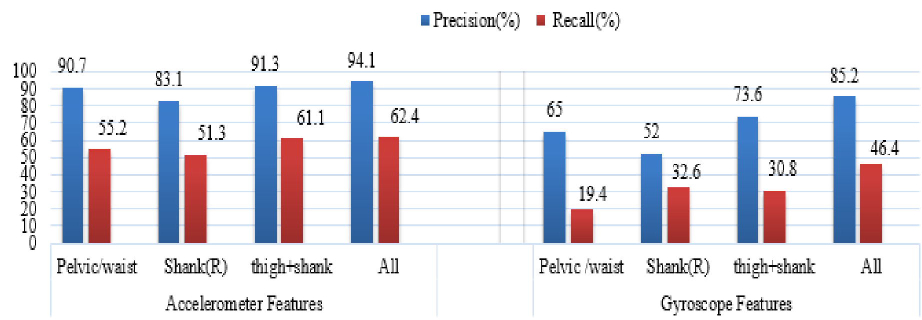


Fig 6. Recognition performance for features extracted from different sensors for person dependent evaluation

Figure 6 shows the impact of features processed from accelerometer and gyroscope separately and it is observed that even for the person dependent scenario, the gyroscope data contained far less useful features than the accelerometer data. However, features from a gyroscope improve the accuracy in the case where the activities are constrained and distinguished by translation and rotation of the joint angles. A combined accelerometer and gyroscope feature processing is a planned part of our future research.

5 Conclusion

The present work described the development of an IMU-based measurement system and investigated the feasibility of its use in human activity recognition and classification scenario. The activities of the system were selected to be of low complexity, which allowed us to compare algorithms in terms of overall recognition performance. For designing more complex activity recognition system, the procedural stages involved and studied in this research will infer some intuitive decisions. In addition, accurate information on the sensor model, positioning and orientation of sensors during different activities will provide generalization and will contribute to a open dataset for human activity recognition based research. Future research will include Composite activities, Concurrent and overlapping activities and also some multi-attribute classification approaches and deep learning approaches for activity recognition in a multi-sensor scenario.

Acknowledgments.

Tahmina Zebin would like to thank the Presidents Doctoral Scholar award scheme, University of Manchester for funding her PhD.

References

[1] A. Bulling, U. Blanke, and B. Schiele, "A tutorial on human activity recognition using body-worn inertial sensors," *ACM Comput. Surv.*, vol. 46, pp. 1-33, 2014.

[2] W. Ugulino, D. Cardador, K. Vega, E. Velloso, R. Milidiú, and H. Fuks, "Wearable Computing: Accelerometers' Data Classification of Body Postures and Movements," in *Advances in Artificial Intelligence - SBIA 2012: 21th Brazilian Symposium on Artificial Intelligence, Curitiba, Brazil, October 20-25, 2012. Proceedings*, L. N. Barros, M. Finger, A. T. Pozo, G. A. Gimenénez-Lugo, and M. Castilho, Eds., ed Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 52-61.

[3] A. Mannini and A. M. Sabatini, "Machine Learning Methods for Classifying Human Physical Activity from On-Body Accelerometers," *Sensors,* vol. 10, pp. 1154-1175, Feb 2010.

[4] L. Shaopeng, R. X. Gao, D. John, J. W. Staudenmayer, and P. S. Freedson, "Multisensor Data Fusion for Physical Activity Assessment," *Biomedical Engineering, IEEE Transactions on,* vol. 59, pp. 687-696, 2012.

[5] C. Bettini, O. Brdiczka, K. Henricksen, J. Indulska, D. Nicklas, A. Ranganathan, *et al.*, "A survey of context modelling and reasoning techniques," *Pervasive and Mobile Computing,* vol. 6, pp. 161-180, 2010.

[6] O. D. Lara and M. A. Labrador, "A Survey on Human Activity Recognition using Wearable Sensors," *IEEE Communications Surveys & Tutorials,* vol. 15, pp. 1192-1209, 2013.

[7] C. A. Ronao and C. Sung-Bae, "Human activity recognition using smartphone sensors with two-stage continuous hidden Markov models," in *Natural Computation (ICNC), 2014 10th International Conference on*, 2014, pp. 681-686.

[8] P. Turaga, R. Chellappa, V. S. Subrahmanian, and O. Udrea, "Machine Recognition of Human Activities: A Survey," *IEEE Trans. Cir. and Sys. for Video Technol.,* vol. 18, pp. 1473-1488, 2008.

[9] H. Fourati, Ed., *Multisensor Data Fusion: From Algorithms and Architectural Design to Applications (Book)*. United States: CRC Press, Taylor & Francis Group LLC, 2015, pp. 509-517

[10] L. Bao and S. S. Intille, "Activity Recognition from User-Annotated Acceleration Data," in *Pervasive Computing: Second International Conference, PERVASIVE 2004, Linz/Vienna, Austria, April 21-23, 2004. Proceedings*, A. Ferscha and F. Mattern, Eds., ed Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 1-17.

[11] M. Zeng, L. T. Nguyen, B. Yu, O. J. Mengshoel, J. Zhu, P. Wu, *et al.*, "Convolutional Neural Networks for human activity recognition using mobile sensors," in *Mobile Computing, Applications and Services (MobiCASE), 2014 6th International Conference on*, 2014, pp. 197-205.

[12] K. R. Kaufman, "Future Directions in Gait Analysis," *RRDS Gait Analysis in the Science of Rehabilitation,* pp. 85-112, 2011.

[13] http://uk.mathworks.com/help/stats/classificationlearner-app.html.