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

SVM-Based Multi-Sensor Fusion for Free-Living Physical Activity Assessment

Shaopeng Liu, Robert X. Gao*, Dinesh John, John Staudenmayer, and Patty S. Freedson

Abstract—**This paper presents a sensor fusion method for assessing physical activity (PA) of human subjects, based on the support vector machines (SVMs). Specifically, acceleration and ventilation measured by a wearable multi-sensor device on 50 test subjects performing 13 types of activities of varying intensities are analyzed, from which the activity types and related energy expenditures are derived. The result shows that the method correctly recognized the 13 activity types 84.7% of the time, which is 26% higher than using a hip accelerometer alone. Also, the method predicted the associated energy expenditure with a root mean square error of 0.43 METs, 43% lower than using a hip accelerometer alone. Furthermore, the fusion method was effective in reducing the subject-to-subject variability (standard deviation of recognition accuracies across subjects) in activity recognition, especially when data from the ventilation sensor was added to the fusion model. These results demonstrate that the multi-sensor fusion technique presented is more effective in assessing activities of varying intensities than the traditional accelerometer-alone based methods.**

I. Introduction

PHYSICAL activity (PA) is defined as bodily movement generated by skeletal muscles [1]. Engaging in physical activities on a regular basis by means of walking, jogging, or sport activities, is critical to maintaining health and preventing cardiovascular diseases, diabetes, and obesity. Accurate monitoring of PA under free-living conditions provides information on the realistic type and intensity of the activities that the person has been engaged in, thus is of significant interest to the research community [1].

The goal of PA assessment is to recognize the type, duration, and intensity of a broad range of activities and quantify the energy expenditure (PAEE) of the person during physical activities. Accelerometer-based PA assessment has recently become the device of choice, due to its low subject burden and non-invasive nature. Studies have shown good results using this method in monitoring PA types and intensities [2]. However, accelerometers by nature cannot distinguish different types of activities that produce similar acceleration profiles but have different energy expenditure [3]. For example, walking at a certain speed may result in acceleration outputs similar to that of walking at the same speed while carrying a load, although the energy expenditure is different.

To address the drawbacks of this method, researchers have investigated alternative techniques, e.g. by placing multiple accelerometers at different locations on the body [4], or combining accelerometers with other types of sensors, such as respiratory sensor or GPS [5], or investigating advanced computational techniques such as machine learning and sensor fusion [4], [5] to better differentiate the activities. These techniques, while having demonstrated promising results, are subject to limitations in that a priori knowledge and "intuitive modeling" of different activities are needed to build a classification model, e.g., a customized decision trees for activity classification. Besides, the output is either PA type or estimated PAEE, but not both simultaneously, which is important because either alone is not sufficient to correctly access the PA. Such limitations motivate research into PA assessment under free living conditions.

In recent years, the technique of Support Vector Machines (SVMs) has been increasingly investigated for medical and biomedical applications [6], [7], due to its classification and estimation capabilities [8]. Such prior applications make SVM an attractive candidate for PA assessment, where features extracted from the recorded multi-channel signals could be redundant due to the internal redundancy of the data. Based on such motivation, an SVM-based multi-sensor fusion technique has been developed to analyze data acquired from a subject-worn integrated measurement system (IMS) [9] to predict both the PA type and PAEE. The performance of the method is experimentally evaluated by human subjects performing free-living activities.

II. SVM-Based Multi-Sensor Fusion

A. SVM Framework

The objective of the study is to fuse data (e.g. bodily movement and ventilation) measured by different types of sensors to more accurately assess the types, intensities, and associated energy expenditure of the physical activities a test subject has engaged in, than using a single sensor. Such a data fusion problem can be addressed by the SVM algorithm, which formulates a decision boundary to separate one activity (or class of data) from another.

Manuscript received April 15, 2011. This work was supported by National Institutes of Health under Grant UO1 CA130783.

Shaopeng Liu and Robert X. Gao* are with the Electromechanical Systems Laboratory, Department of Mechanical Engineering, University of Connecticut, Storrs, CT 06269, USA. (phone: 860-486-7110; fax: 860-486-5088; e-mail: rgao@engr.uconn.edu).

Dinesh John and Patty S. Freedson are with the Physical Activity and Health Laboratory, Department of Kinesiology, University of Massachusetts Amherst, MA 01003, USA.

John Staudenmayer is with the Department of Mathematics and Statistics, University of Massachusetts Amherst, MA 01003, USA.

To formulate a SVM pairwise-model, assume a data set $\{x_i\}$ consisting of data measured by multiple sensors, where $x_i \in R^n$ (n is the dimension of the input vectors) and $i = 1,\ldots, N$ (N is the total number of data points). Within the duration of this data set, the subject is assumed to have engaged in two types of activities $\{y_i\}$, labeled as -1 and 1 ($y_i \in \{-1, 1\}$), respectively. Each data set $\{x_i\}$ can be associated with one of the two activity types $\{y_i\}$. To distinguish the type of activity which the data set $\{x_i\}$ is associated with, a function $f(x)$ is assumed to exist that draws a separation decision boundary of the two activities. Data points above the boundary (when $f(x)$ yields a value ≥ 0) belong to the activity labeled as 1, whereas data points below the boundary (when $f(x)$ produces a value < 0) are labeled as -1.

A drawback of such a separation boundary built in the original lower dimensional space is that it is often a complex and nonlinear function, which is computationally demanding for determining activity types for each new data point added. To overcome this limitation, the activities are assumed to be separable in a high-dimensional space where a linear and explicit decision boundary – a hyperplane can be formulated [8]. To achieve this, the SVM algorithm first transforms the data $\{x_i\}$ from the original lower dimensional space to a higher dimensional space via a transformation function – a hyperplane $f'(x') = w^T x' + b = 0$ is then built to separate the two activities. In this formulation, w and b are the weighing factors, and x' is the transformed high-dimensional data. The hyperplane function is expressed as:

$$\begin{cases} f'(x_i') = w^T x_i' + b \geq 0 \;\Rightarrow\; y_i = 1 \\ f'(x_i') = w^T x_i' + b < 0 \;\Rightarrow\; y_i = -1 \end{cases} \tag{1}$$

The hyper-plane is then built by solving the following optimization problem [8]:

$$\min_{w\in F, b\in R, \xi\in R^n} \left\{ \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N} \xi_i \right\}$$
$$\text{subject to } \xi_i \geq 0,\; y_i\left(w^T \phi(x_i) + b\right) \geq 1 - \xi_i \tag{2}$$

where ξ_i is the slack variable that accommodates overlapping (i.e. misclassified) data points in the feature space for multiple activities, as may happen in practical applications, C is the penalty parameter for those sample points misclassified by the optimal separating plane, and ϕ is the transformation function that coverts the sensor data in the lower dimensional space into a higher-dimensional space $\phi(x_i) = x_i'$. The decision function $f(x)$ is then determined as the following sign function (sgn(t) = 1 for $t \geq 0$, and sgn(t) = -1 for $t < 0$):

$$f(x) = \mathrm{sgn}\left(\sum_{i=1}^{N} y_i \alpha_i \phi(x_i)^T \phi(x) + b\right) \tag{3}$$

A kernel function $K(x_i, x) = \phi(x_i)^{\mathrm{T}}\phi(x_i)$ can replace the inner product in Eq. (3), and the decision function of the two activities can be rewritten as:

$$f(x) = \mathrm{sgn}\left(\sum_{i=1}^{N} y_i \alpha_i K(x_i, x) + b\right) \tag{4}$$

For the present study, the Gaussian kernel was selected due to its reported effectiveness in activity recognition [10].

B. PA Assessment

As illustrated in Fig. 1, the SVM-based multi-sensor data fusion system extracts features (e.g. the mean value, percentiles, and dominant frequency) from PA-relevant data measured on human subjects. These include breathing (or ventilation) from an abdominal (AB) ventilation sensor (1325 Piezo Crystal Sensor, Ambu Sleepmate), bodily motion from two triaxial accelerometers (MMA7260QT, Freescale) placed at hip and wrist locations, and environmental context from an ultraviolet sensor [9]. These features are then fused by the SVM algorithm to quantify the types of physical activity that the subject has engaged in, and the associated energy expenditure.

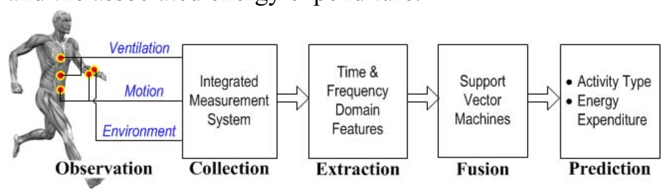

Fig. 1. Illustration of the SVM-based multi-sensor fusion algorithm.

In this study, a total of 51 time- and frequency-domain features were extracted from the hip and wrist accelerometers carried by the subjects plus the respiratory signals from the AB ventilation sensor. All these features were used as inputs to the SVM algorithm. The time-domain features included the mean value, standard deviation, 10th, 25th, 50th (median), 75th, 90th percentiles, and correlation between the vector magnitudes of the hip and wrist accelerometer readings. Mean and standard deviation of the PA signals are calculated, providing a general description of the activity intensity levels. The middle three percentiles (25th, 50th, and 75th) characterize signal distributions, and the 10th and 90th percentiles represent an estimate of the low and high values in each signal. The correlation between the vector magnitudes provides a measure for the coordination or variation between the upper limb and body during an activity. Frequency-domain feature was obtained from a spectral analysis, and is defined as the dominant frequency of the respiratory signal that is taken as the breathing frequency. The features were computed for every 30-second data segment, and linear scaling was then applied on the extracted features in the range of [0, 1], to avoid that features of greater numeric values would overwhelm those in the smaller numeric ranges.

A two-step procedure was taken for predicting the types of physical activity. First, a training data set that consists of features obtained from all 50 subjects but one was constructed for building the SVM model and selecting the penalty parameter C and Gaussian kernel parameter γ. The LIBSVM package [11] was implemented to build the model, and the parameters were selected through a "grid-search" with 5-fold cross validation. The parameters that yielded the highest recognition rate were chosen during the process.

Upon completion of training, the SVM model was applied to the feature set of the subject that was left out in the training process, to predict the activity type reflected in the 30-second data segments. Such a two-step procedure constitutes a "leave-one-subject-out" cross validation, and was executed on each subject data. In a similar fashion, energy expenditure associated with each activity was predicted by the regression version of the SVM, Support Vector Regression (SVR).

III. Experimental Study

Fifty test subjects (19 male and 31 female) were recruited. The group has the following characteristics (i.e., mean ± standard deviation): age = 32.6±29.9 years, weight = 67.7±12.3 kg, height = 171.2±8.6 cm, and body mass index = 23.2±4.6 kg/m^2.

The subjects were arranged to perform 13 types of activities of varying intensities (Table I) commonly seen in daily lives, and involve motions from different parts of the body, e.g. upper-limb-dominant activities such as computer work or filing papers, lower-limb-dominant activities such as cycling and treadmill running, and whole body activities such as basketball and tennis playing. These 13 activities were the target classes of the SVM and other classifiers.

TABLE I
PHYSICAL ACTIVITIES

Activities	Category	Abbr.	Routine *1*	Routine *2*
Computer work	Sedentary activity	*CW*	•	
Filing papers		*FP*		•
Vacuuming	Household and other	*VA*		•
Moving boxes		*MB*	•	
Self-paced walk		*SW*		•
Cycling with 1-kp resistance		*C1*	•	
Cycling with 2-kp resistance		*C2*		•
Treadmill at 3.0 mph	Moderate locomotion	*T3*	•	
Treadmill at 4.0 mph 0% grade		*T4-0*		•
Treadmill at 4.0 mph 5% grade	Vigorous activity	*T4-5*	•	
Treadmill at 6.0 mph 0% grade		*T6-0*		•
Basketball		*BA*		•
Tennis		*TE*	•	

For purpose of experimental organization, the activities were also classified into two routines, and each subject was asked to perform one group of activities. The specific activities are completed in a random order. Each activity, when performed, lasted for 7 minutes long, followed by a 2-minute rest period. Prior to the start of each test, subjects were asked to lie down on a bed (for consistency with previous calibration studies [3]) to rest for 10 minutes, in order to achieve a resting metabolic rate. All the tests were performed during the day, and the subjects were asked to eat four hours before the test, after which no food or drink was allowed to be taken, except for water. The total duration of each subject test session was 2 hours.

During the tests, the actual PA types performed by the test subjects were noted down, and the PAEEs were determined by a respiratory gas exchange system (Oxycon Mobile, Cardinal Health), which serves as a criterion measure. The respiratory gas exchange system, secured to the subject using an adjustable vest, provides physiological measurements such as the breathing rate, ventilation volume, and metabolic equivalent of task (MET). The MET quantifies the intensity and energy expenditure. The measured data were then wirelessly transmitted to a laptop. The clocks for the integrated measurement system and the respiratory gas exchange system were synchronized and the time was noted at the beginning of each activity.

IV. Results and Discussions

A. Recognition of Activity Type

The results of activity type recognition using the SVM method were compared with those obtained by commonly used methods: the k-nearest neighbor (kNN) and Naïve Bayes (NB) classifiers. Furthermore, the results were also compared among four different fusion models, listed below, to investigate the effect of number of sensors and their specific locations on PA assessment:

1) M_1: hip accelerometer only;
2) M_2: hip accelerometer, AB ventilation sensor;
3) M_3: hip, wrist accelerometers;
4) M_4: hip, wrist accelerometers, AB ventilation sensor.

Figure 2 shows a comparison of the activity type recognition accuracies for the above four models. Within each model, the recognition is also compared among the three classifiers: SVM, kNN, and NB. The recognition accuracies are expressed as the mean and standard deviation, computed over the 50 subjects.

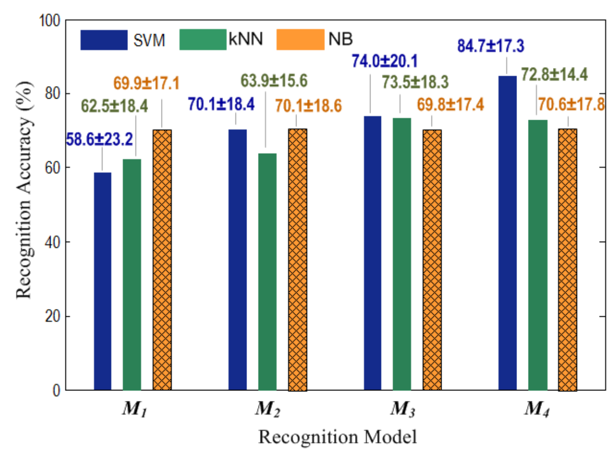

Fig. 2. Activity type recognition rates for different models.

It can be seen that, for both the SVM and kNN methods, the recognition performance is consistently enhanced when more sensor data are included in the models. For example, the average accuracy obtained by SVM has increased from 58.6% (achieved by the single-sensor model M_1) to 70.1% and 74.0% for the dual-sensor models (M_2 and M_3), and then to 84.7% for the multiple-sensor model (M_4), respectively.

The standard deviation of the recognition accuracies reflects upon the subject-to-subject variability. The variability in SVM has shown to have decreased from 23.2% when using a single hip accelerometer to 18.4% and 20.1%

when using two sensors (either the hip accelerometer with ventilation sensor, or hip accelerometer with wrist accelerometer), and further down to 17.3% when using all three sensors. This indicates that with ventilation sensor and wrist accelerometer data included in the fusion process, the subject-to-subject variability can be effectively reduced, making the fusion algorithm more generalized than using data from only two accelerometers. The results obtained by kNN have shown similar trends.

The results from the NB classifier have shown little variation among the four recognition models. This indicates that the choice of an appropriate classifier has an effect on the performance of the assessment. It is also interesting to note that, when only signals from the hip accelerometer (M_1) were involved in activity type estimation, the recognition accuracy obtained by SVM was less than those obtained by kNN and NB. While in the cases where signals from more sensors were fused, the results from SVM were consistently better than those from using the kNN and NB. This indicates that the SVM algorithm has better performance on fusing data from multiple sensors than the other two techniques.

B. Estimation of Metabolic Equivalent (MET)

The MET characterizes the energy expenditure associated with activities, and was predicted in the presented study by using the SVR fusion model. Specifically, each 30-second data segment was first placed into one of the four activity categories, based on the results of activity type classification. Subsequently, the energy expenditure during the 30-second period was estimated by the specific SVR model for that activity group. The estimated MET values were then compared with values measured by the respiratory gas exchange system.

Figure 3 shows a comparison between the measured and predicted MET values. The predicted MET values were computed by the SVR fusion model with both the hip and wrist accelerometers and the ventilation sensor (M_4). The measured and predicted average METs of each activity were plotted as the abscissa and ordinate, and the percentage difference between them, for each activity, was also shown.

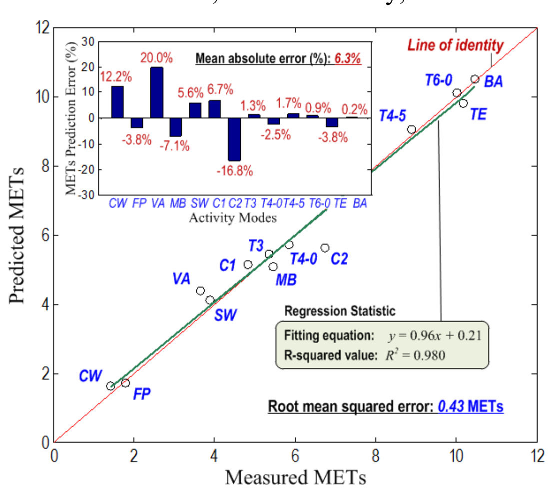

Fig. 3. Comparison of measured and predicted METs for different activities.

It is noted that the fitting equation of the data points is close to the line of identity (with a slope of 0.96), and follows a high coefficient of determination, reflected in the R^2 value of 0.980. This indicates a good agreement between the predicted METs by the SVR fusion model with the measured MET by the criterion device.

V. Conclusions

An SVM-based multi-sensor fusion method for physical activity assessment under free-living condition is presented. Experimental results have demonstrated that the system was able to recognize 13 types of activities of varying intensities (sedentary to rigorous) and estimate the corresponding energy expenditure with good accuracy. The results have demonstrated two advantages of the developed SVM-based multi-sensor fusion algorithm: (1) enhancing the assessment performance by improving the accuracy in estimation of the activity types and the corresponding energy expenditure, and (2) reducing the subject-to-subject variability (i.e. standard deviation) in activity type recognition by about 20%, especially when signals from the AB ventilation sensor were included in the fusion models. Future research will involve a large-population testing for the overall system performance.

References

[1] C. Caspersen, K. Powell and G. Christensen, "Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research," *Public Health Reports*, vol. 100, pp. 126-131, 1985.

[2] D. Hendelman, K. Miller, C. Baggett, E. Debold, and P. Freedson, "Validity of accelerometry for the assessment of moderate intensity physical activity in the field," *Medicine & Science in Sports & Exercise*, vol. 32, pp. 442-449, 2000.

[3] J. Staudenmayer, D. Pober, S. Crouter, D. Bassett, and P. Freedson, "An artificial neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer," *Journal of Applied Physiology*, vol. 107, pp. 1300-1307, 2009.

[4] L. Bao and S. Intille, "Activity recognition from user-annotated acceleration data," *Lecture Notes in Computer Science*, vol. 3001, pp. 1–17, 2004.

[5] M. Ermes, J. Parkka, J. Mantyjarvi, and I. Korhonen, "Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions," *IEEE Transactions on Information Technology in Biomedicine*, vol. 12, pp. 20-26, Jan. 2008.

[6] G. Naik, D. Kumar, and Jayadeva, "Twin SVM for gesture classification using the surface electromyogram," *IEEE Transactions on Information Technology in Biomedicine*, vol.14, no.2, pp.301-308, March 2010.

[7] S. Shao, K. Shen, C. Ong, E. Wilder-Smith, and X. Li, "Automatic EEG artifact removal: a weighted Support Vector Machine approach with error correction," *IEEE Transactions on Biomedical Engineering*, vol. 56, pp. 336-344, 2009.

[8] N. Cristianini and J. Shawe-Taylor, *Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, 2000.

[9] S. Liu, R. Gao and P. Freedson, "Design of a wearable multi-sensor system for physical activity assessment," in *Proc. IEEE/ASME Int. Conf. on Advanced Intelligent Mechatronics*, pp. 254-259, 2010.

[10] T. Zhang, J. Wang, L. Xu, and P. Liu, "Fall detection by wearable sensor and one-class SVM algorithm," *Intelligent Computing in Signal Processing and Pattern Recognition*, vol. 345, pp. 858-863, 2006.

[11] C. Chang and C. Lin, "LIBSVM: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1-27:27, 2011.