Feature Extraction and Analysis for Body Sensor Networks

SUPERB 2007

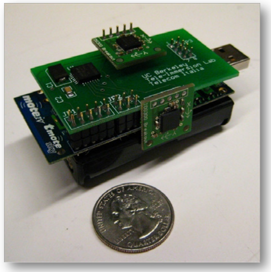
Amalia Viti, Columbia University, anv2106@columbia.edu
Operations Research: Engineering Management Systems

Faculty Mentors: Professor Roozbeh Jafari, Professor Ruzena Bajcsy
Graduate Mentor: Sameer Iyengar

Abstract

Wireless body sensor networks are networks composed of nodes placed on different parts of the human body. These nodes are able to communicate in a wireless manner and transmit data from sensors mounted on them to a base unit or a computer that analyzes the data received. Wireless body sensor networks make on-body and mobile health-care monitoring possible; such systems can integrate information from different sources, and can initiate actions or trigger alarms when needed.

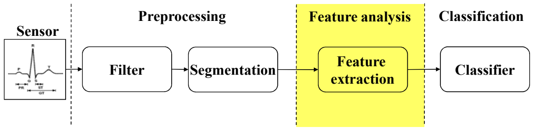
In this project we investigate a collaborative signal processing scheme for physical movement monitoring with motion sensors. The signal processing consists of preprocessing, feature extraction and classification. We define a measure on feature significance as well as features' correlations. Because these nodes have a limited amount of battery power, we extract features to characterize the data. We develop a metric to rank each feature by its significance with respect to a particular movement class. We validate our metric by analyzing the accuracy of classification when using a subset of features selected by our metric. Using the characteristics of the best features we find from the rankings, we develop new features to improve the system accuracy and then assess their effectiveness.



A picture of the Telos-B mote used in the system. Each mote is equipped with a processing unit, an accelerometer that records on the x, y, and z-axes, and a gyroscope that records on the x and y-axes.

System Overview

We propose a system composed of on-body motion sensors and a gateway system allowing it to connect to a base station (i.e. a laptop, cell phone, or PDA) for monitoring purposes. The system utilizes three Telos-B motes worn on the subject's ankle, waist, and wrist. We recorded the movements of seven different subjects with ages ranging between 20 and 60 years old. There are three main steps we take in classifying our subjects' movements:



- **Preprocessing –** Each mote gives us 5 readings of data (3 from accelerometer, 2 from gyroscope) that must each be segmented into periods of activity.
- **Feature Extraction –** More significant and robust features yield better classification accuracy.
- **Classification –** A k-nearest-neighbor (k-NN) algorithm is used to classify each of the eight movements established for testing purposes.

Feature Significance Metrics

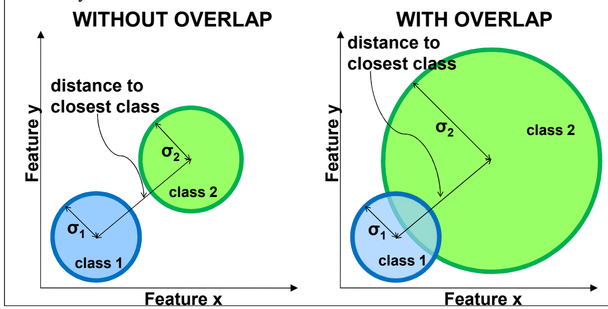
We use a metric in order to rank each of the features for classification. A metric assigns the significance value for a certain movement class. Those features ranked with the best rankings are the most significant features. Metrics we analyzed are as follows:

σ_1	the standard deviation of the data for a specific feature
$\dfrac{\sigma_1}{d_{1,2}}$	the standard deviation of the data for a specific feature divided by the distance to the closest class
$\dfrac{3\sigma_1 + 3\sigma_2}{d_{1,2}}$	the sum of three times the standard deviations of both the specific feature and the closest class, divided by the distance to the closest class

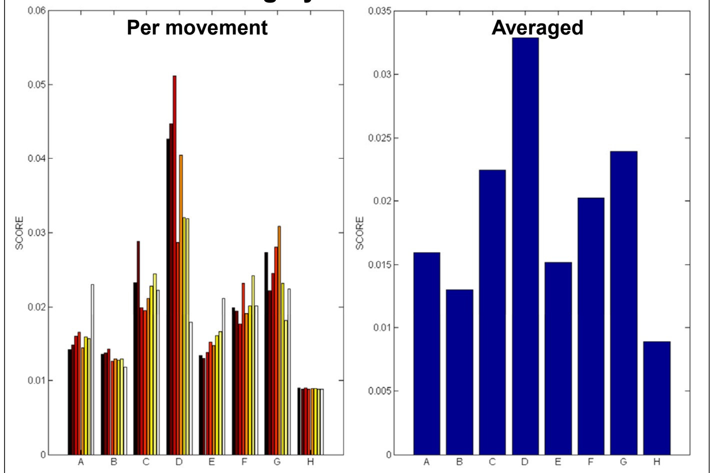
Metric Accuracy

σ_1	$\sigma_1/d_{1,2}$	$(3\sigma_1 + 3\sigma_1)/d_{1,2}$
86.7%	89.5%	88.4%

After comparing all three metrics we came to the conclusion that the third metric was the best for our data. This metric takes into account the nearest class and considers scenarios in which there might be overlapping classes while losing only a negligible amount of accuracy.



Ranking System for Features



LEGEND

A. Mean Value	E. RMS
B. Start-to-end Amplitude	F. Morphology
C. Standard Deviation	G. IQR
D. Peak-to-peak Amplitude	H. Peak-to-peak width/length(x)

Feature Analysis

We calculate our own features for extraction from the raw data. The features are as follows:

- **Mean Value**
- **Start-to-End Amplitude**
- **Standard Deviation**
- **Peak-to-Peak Amplitude**
- **Root Mean Square (RMS)**
- **Morphology**

From the graph above, one can see that standard deviation, peak-to-peak amplitude, and morphology rank higher than the others. Standard deviation and peak-to-peak amplitude are both related to the general distribution of data along the y-axis. We found that features with values based on the overall spread of the data along the y-axis were more significant, as well as, features relating to the shape or form of the data.

Feature Creation

Keeping in mind certain characteristics the existing significant features shared, we created two new features. As can be seen on the graph, IQR performed the best of the two:

- **Interquartile Range (IQR)**
- **Peak-to-Peak Width/Length(x)**

August 3, 2007

Special thanks to Professor Ruzena Bajcsy, Professor Roozbeh Jafari, Sameer Iyengar, Kristen Gates, Annarita Giani, Alvaro Cardenas, TRUST, and SUPERB 2007 staff and students.