Data Collection Protocol for Wearable Action Recognition Database (WARD)

Version 1.0

Allen Y. Yang
Email: yang@eecs.berkeley.edu

1 Overview

This article proposes a protocol to construct a benchmark database for human action recognition using a wearable motion sensor network, called *Wearable Action Recognition Database* (WARD). The purpose of WARD is two-fold: 1. A public and relatively stable data set provides a platform for quantitative comparison of the existing algorithms for human action recognition using wearable motion sensors. 2. The database should steer the development of future innovative algorithms in the area of distributed pattern recognition by bringing together the investigators from the pattern recognition and sensor networks communities. In order to achieve the two objectives, the proposed database should be carefully constructed with the following requirements:

1. The database needs to contain sufficient numbers of human subjects with a large range of age differences.
2. The designed action classes should be general enough to cover most typical activities that a common human subject is expected to perform in his/her regular daily life.
3. The locations of the wearable sensors should be selected to be consistent with practical scenarios for a potential commercial system to deploy the same type of sensor on its customers.
4. The sampled action data should contain sufficient variation, measurement noise, and outliers in order for existing and future algorithms to meaningfully examine and compare their performance to the full extent.

In this WARD database version 1.0, we will assume that all wearable sensors are firmly fastened at a list of body locations. The data collected in this version will serve as a baseline benchmark, as almost all existing algorithms in the literature adopt this assumption in their systems.

2 Architecture of the Wearable Motion Sensor Network

The wearable sensor network consists of sensor nodes placed at multiple body locations, which communicate with a base station attached to a base-station computer via a USB port. The sensor nodes and base station are built using the commercially available Tmote Sky boards. Tmote Sky runs TinyOS on an 8MHz microcontroller with 10K RAM and communicates using the IEEE 802.15.4 wireless protocol. Each custom-built sensor board has a triaxial accelerometer and a biaxial gyroscope, which is attached to Tmote Sky (shown in Fig 1). Each axis is reported as a 12bit value to the node, indicating values in the range of $\pm 2g$ and $\pm 500^\circ/s$ for the accelerometer and gyroscope, respectively. Each node is currently powered by a Li-ion battery housed between the two sensor boards.

The current hardware design of the sensor contributes certain amounts of measurement error. The accelerometers typically require some calibration in the form of a linear correction, as sensor output under $1g$ may be shifted up to 15% in some sensors. It is also worth noting that the gyroscopes produce an indication of rotation under straight line motions. Fortunately these systematic errors appear to be consistent across experiments for a given sensor board.

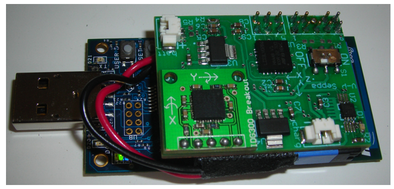

Figure 1: Illustration of a motion sensor node. The sensor board on the top is a custom-built motion sensor with triaxial accelerometer and biaxial gyroscope. The middle layer is a Li-ion battery. The sensor board on the bottom is a standard Tmote sensor board.

However, without calibration to correct them, the errors may affect the action recognition if different sets of sensors are used interchangeably in the experiment.

To avoid packet collision in the network, we use a TDMA protocol that allocates each node a specific time slot during which to transmit data. This allows us to receive sensor data at 20Hz with minimal packet loss. To avoid drift in the network, the base station periodically broadcasts a packet to resynchronize the nodes' individual timers. The code to interface with the sensors and transmit data is implemented directly on the mote using *nesC*, a variant of C.

3 WARD Data Format

We use the MATLAB environment to store the action data sampled from the wearable sensors. Each action sequence is given a trial number N, $N = 1, 2, \cdots$. The file name for storing the action data for the Nth trial is **WearableAction_$<$ N $>$.mat**, where $<$ N $>$ is to be replaced with the trial number.

Each trial file stores the complete sensor readings of the action sequence in one variable: **WearableData**. **WearableData** contains the following fields:

1. **Version**: A string indicating the version of the database. Currently the only value is: "**WARD 1.0**."
2. **Date**: A string indicating the date of the recording session. The string has the following formet: "**yyyy-mm-dd**".
3. **Class**: An integer value representing the action class number.
4. **Subject**: An integer value representing the subject number.
5. **Age**: An integer value representing the age of the subject.
6. **Reading**: A cell structure. The number of cells is equal to the number of sensors in the network. Each cell contains an array structure of dimension $5 \times t$, where 5 corresponds to the sensor readings from 3-axis accelerometer and 2-axis gyroscope on one sensor node, and t represents the length of the trial sequence.

NOTE: The database contains motion sequences with partial missing sensor data mainly due to battery failure and/or network packet loss. Such missing data are indicated as "Inf" in MATLAB. Verification needs to be in place to detect such exceptions when classification is performed.

4 WARD 1.0 Protocol: Action Sampling with Fixed Sensor Locations

In this section, we describe the protocol to construct WARD database version 1. The fundamental assumption is that all wearable sensors are firmly fastened at the selected body locations during the whole recording session.

For the WARD database, the enrollment of human subjects should take into account the range of age differences that will be reflected on the change of body motions for same types of actions. The experiment should enroll five to ten subjects with their ages ranging between 18 and 70. One example could be 18, 30, 45, 55, and 65 for the ages of five subjects, respectively.

We choose to deploy five motion sensors to sample and monitor the full body motion. At all sensor locations, the Tmote board should face inside the body, and the motion sensor board should face outside.

- Sensor 1: Outside center of the lower left forearm joint. The y-axis of the gyroscope points to the hand.
- Sensor 2: Outside center of the lower right forearm joint. The y-axis of the gyroscope points to the hand.
- Sensor 3: Front center of the waist. The x-axis of the gyroscope points down.
- Sensor 4: Outside center of the left ankle. The y-axis of the gyroscope points to the foot.
- Sensor 5: Outside center of the right ankle. The y-axis of the gyroscope points to the foot.

Finally, we define a list of human action/activity categories. Each subject should perform the actions based on his/her own interpretation and style. In WARD, we are interested in nontransient human actions. In order to sufficiently sample the continuous movement of a nontransient action, one subject should perform one trial of an action for more than 10 seconds, and the same action should be repetitively samples more than 10 times. The following is the list of action classes with the names in the parentheses denoting the synonyms of the actions.

1. Rest at Standing (ReSt): The subject stands still for more than 10 seconds.
2. Rest at Sitting (ReSi): The subject sits still for more than 10 seconds.
3. Rest at Lying (ReLi): The subject lies still for more than 10 seconds.
4. Walk forward (WaFo): The subject walks straight forward for more than 10 seconds.
5. Walk forward left-circle (WaLe): The subject walks in counter-clock circle for more than 10 seconds.
6. Walk forward right-circle (WaRi): The subject walks in clock circle for more than 10 seconds.
7. Turn left (TuLe): The subject stays at the same position and turn left for more than 10 seconds.
8. Turn right (TuRi): The subject stays at the same position and turn right for more than 10 seconds.
9. Go upstairs (Up): The subject goes up one flight.
10. Go downstairs (Down): The subject goes down one flight.
11. Jog (Jog): The subject jog straight forward for more than 10 seconds.
12. Jump (Jump): The subject stays at the same position and jump for more than 5 times.
13. Push wheelchair (Push): The subject pushes a wheelchair/walker for more than 10 seconds.