Preliminary Evaluation of Feature Level Compensation for Missing Data in Multi-sensor Activity Recognition

Ryoma Uchida
Toyohashi University of Technology
ryoma@usl.cs.tut.ac.jp

Ren Ohmura
Toyohashi University of Technology
ren@tut.jp

ABSTRACT
Activity recognition using multiple body-worn sensors can directly monitor the movement of each body part and can recognize various activities accurately. However, using multiple sensors increases the chance of sensor failure or communication failure, and most current activity recognition algorithms do not work when failure occurs due to the difference (reduction) of the dimension of the feature vector from that of complete sensor data expected in system design time. Therefore, we compared three possible techniques to solves this problem on the feature value level: a classifier trained with reduced feature values, feature value compensation with multiple regression, and feature value compensation with kernel regression, in a no failure situation. All of these techniques do not depend on classification algorithms. While creating a regression model, which is in the training phase, requires relatively high computational power, compensation itself can work with low computational power. As overall results, kernel regression had the best performance that was the closest to the no failure situation. Also, the results imply that each sensor position has its own effective method and more accurate coping can be viable with the appropriate choice of the method.

Author Keywords
activity recognition, feature value compensation, missing data compensation

ACM Classification Keywords
J4 : Social and Behavioral Sciences.

General Terms
Experimentation, Reliability

INTRODUCTION
Most current activity recognition techniques are based on machine learning algorithms. They basically compare feature values extracted from sensor data between training time and recognition time. Data loss from even only one sensor causes loss of some feature values and can degrade the recognition accuracy. Moreover, these algorithms cannot give any output with incomplete sensor data (feature values) in the worst case scenario. In our project, in which nursing activity is monitored, we can only obtain 60% of the whole experiment time due to Bluetooth communication failure between body-worn sensors and a wearable storage device. That is, 40% of activity data is necessary for coping with data loss.

While coping with missing data in classification has been studied in other fields, such as voice recognition, there are a few studies in activity recognition. Compensating for missing data is one method, and it can be done in three levels in activity recognition: the raw data level, the feature value level, and the classifier level. Hesam *et al.* proposed a classifier level compensation method[1]. However, their approach fixes the classification algorithm due to the classifier level compensation. Also, their approach requires matrix calculation, including obtaining the inverse matrix to find the compensating value, and it needs relatively high computational power.

Therefore, we propose and evaluate three coping techniques for the problem. First, we evaluate the no compensation situation in which multiple classifiers are trained with reduced feature values when supposing sensor error and select the appropriate classifier for recognition. Also, we propose two feature value level compensation models: multiple regression and kernel regression. The regression method depends on only calculated feature values and does not restrict feature calculation and the classification algorithm. Also, while creating a regression model, which is done in the training phase, sometimes requires relatively high computational power, compensation itself can be done by relatively easy computation.

EXPERIMENTS
We conducted experiments using six tri-axial accelerometers, which could measure $\pm$ 3G at 100 Hz, with five subjects who are university students. Each subject was asked to put the accelerometers on both his/her wrists, ankles, chest (in pocket), and back waist and to do "hand clapping," "arm folding," "sitting," "running," "skipping," "standing," and "walking" during thirty minutes without any instructions from the experimenter.

The feature values were mean and standard deviation. These values were computed from each axis of both accelerometers and gyro and were extracted from a sliding window

UbiComp '12, Sep 5-Sep 8, 2012, Pittsburgh, USA.

with a width of 2.56 seconds that shifted in steps of 1.28 seconds. As a result, we obtained 72 dimension feature vectors. Training data was studied by using 30 minutes sequence data (including the seven activities). A support vector machine (SMO in Weka[2]) was used as the classification algorithm. Here, we compared the following four situations.

Normal situation means no data loss, and six accelerometers and gyro data are always available. This can be thought of as giving the maximum recognition rate (”normal” in short in the remaining part).

Reduced sensor classifier is the condition in which multiple classifiers are trained with data when one sensor is removed, and recognition is done by selecting an appropriate classifier that uses the same sensors with the available sensors. That is, five sensors are used in training and recognition, and six cases (right wrist, left wrist, ...) are averaged to get the final result (”reduced classifier”).

Compensating by multiple regression is the condition in which data from one sensor data is lost and the missing feature value is compensated by using multiple regressions. The regression model is a linear model. One sensor has two three-axis accelerometer and gyro, and three feature values are extracted from each axis data. Therefore, twelve missing features are calculated from the available sixty values. The result is averaged from six cases. (”multi-reg”).

Compensating by kernel regression is the condition in which data from one sensor is lost and the missing feature value is compensated by kernel regression. Kernel regression is able to generate a regression model in higher-dimensional space. Therefore, this regression model is a higher accuracy linear model than is multiple regression. Processing the data is the same as in the multiple regression case, expect the difference of the regression model (”kern-reg”).

RESULTS

We compared the F-measure values, which were calculated by $\frac{2*Precision*Recall}{Precision/Recall}$ in this experiment.

First, the overall result showed that the normal F-measure value was 75%, the reduced classifier was 70.4% (-4.6%), the multi-reg was 71.1%(-3.9%), and the kern-reg was 72.0% (-3.0%). Thus, kern-reg showed the closest performance with the normal situation, although that was a little difference.

Figure 1 shows the difference of the result of each activity as a result of normal being zero. This figure shows that the reduced classifier and both regression methods had better performance in standing, skipping, and sitting. It can be thought that the number of sensors was too high for these activities and that redundant sensor data was eliminated by the missing sensor. Also, this figure shows a different tendency of the reduced classifier and both regression methods.The reduced classifier showed better performance for walking and running, both moving activities. The regression method performed better for hand clapping and arm folding, both hand-

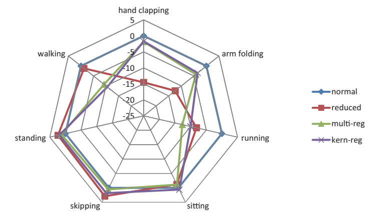

Figure 1. Result Differences from Normal Situation (Activity)

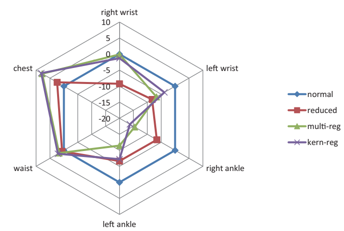

Figure 2. Result Differences from Normal Situation (Sensor Position)

related activities. This implies that the effective method depends on the position of the missing sensor. In more detailed data, we confirmed that sensors on the right and left wrists for hand clapping and arm folding could not get good results in the reduced classifier. The same occurred in the regression method for those sensors on the right and left ankles for running and walking. This fact can be seen in Figure 2, which shows the difference of the result of each sensor position as a result of normal being zero. This suggests that choosing the method depending on the position of the missing sensor can provide better performance to cope with the problem.

SUMMARY

For coping with sensor data loss in accelerometer and gyro based activity recognition, we evaluated the reduced classifier, multiple regression, and kernel regression methods. Both regression methods compensated for missing feature values. In overall performance, compensation by both regression models produced better results than did the reduced classifier method. However, some reduced classifier methods resulted in better performance than did both regression methods. This result implies the list position where compensation is effective or not. Therefore, we need to choose the compensation position. Then, we can achieve a high accuracy activity recognition technique.

REFERENCES

1. Hesam Sagha, José del R. Millán, and Ricardo Chavarrlaga.: A Probabilistic Approach to Handle Missing Data for Multi-Sensory Activity Recognition, *UbiComp ’10 joint workshop on Opportunistic Ubiquitous Systems*, W14 (2010).

2. Ian H. Witten and Elbe Frank.: Data Mining: Practical Machine Learning Tools and Techniques (Second Edition). Morgan Kaufmann (2005) ISBN 0-12-088407-0.