Proceedings of
IEEE International Conference on
Multisensor Fusion and Integration for Intelligent Systems
Seoul, Korea, August 20 - 22, 2008

Recognizing Human Activities from Accelerometer and Physiological Sensors

Sung-Ihk Yang and Sung-Bae Cho

Abstract—**Recently the interest about the services in the ubiquitous environment has increased. These kinds of services are focusing on the context of the user's activities, location or environment. There were many studies about recognizing these contexts using various sensory resources. To recognize human activity, many of them used an accelerometer, which shows good accuracy to recognize the user's activities of movements, but they did not recognize stable activities which can be classified by the user's emotion and inferred by physiological sensors. In this paper, we exploit multiple sensor signals to recognize to user's activity. As Armband includes an accelerometer and physiological sensors, we used them with a fuzzy Bayesian network for the continuous sensor data. The fuzzy membership function uses three stages differed by the distribution of each sensor data. Experiments in the activity recognition accuracy, have conducted by the combination of the usages of accelerometers and physiological signals. For the result, the total accuracy appears to be 74.4% for the activities including dynamic activities and stable activities, using the physiological signals and one 2-axis accelerometer. When we use only the physiological signals the accuracy is 60.9%, and when we use the 2 axis accelerometer the accuracy is 44.2%. We show that using physiological signals with accelerometer is more efficient in recognizing activities.**

I. Introduction

THE recognition of human activity is a concerning problem to provide interactive service with the user in various environments. Research about activity recognition is emerging recently, using various sensory resources. There are studies using cameras or GPS based on the user's movement by using pattern recognition techniques. Many of the other research use accelerometers as the main sensor placing the accelerometer on the user's arm, leg or waist. They recognize activities like walking or running, which has movements in the activity and can be optimized for the recognition based on the accelerometer.

Accelerometers show a high accuracy in recognizing activities with lots of movements. But they are weak to recognize the activities with little movements. Accelerometers also provide continuous sensory data which is hard to separate to the exact boundary of several states. Other research use physiological sensors to recognize the user's context, as the user's body status can represent the user's activity and also the user's emotion which depends on the activity. Therefore using the physiological sensor with accelerometers will help to recognize the user's activity.

In this paper, we use not only an accelerometer but also physiological sensors to help the recognition of activities. Accelerometers have advantages to measure the user's movement and the physiological signals have advantages to measure the user's status of the body. Most of these sensors calibrate continuous sensor data, and if the data is near the boundary of a specific state, the evidence variable will show a radical difference in small changes. To lessen these changes in this situation we propose to preprocess the data with fuzzy logic. Fuzzy logic can represent ambiguous states in linguistic symbols, which is good in continuous sensory data. As sensory data include uncertainty of calibrating data and also human activity itself has it too, we use Bayesian network to do the inference, and modify the learning and inference methods that fit to the fuzzy preprocessing.

S. -I. Yang is with Yonsei University, Republic of Korea, Seoul (e-mail: unikys@sclab.yonsei.ac.kr).

S. -B. Cho.,is with Yonsei University, Republic of Korea, Seoul (e-mail: sbcho@yonsei.ac.kr).

II. Background

A. Related works

There are many research groups studying about human activity recognition using various sensors like cameras, GPS or accelerometers. Tapia used a simple state-change sensor to detect the objects which the user is using at home [1], and Han used an infrared camera to contrast the silhouette of the user and recognized the activity by using the sequence of the images [2]. When using these kinds of sensors such as state-change sensor, cameras or microphones, the research uses pattern recognition and tracks the user's position to recognize the user's activity.

Using motion detection sensors or cameras can only collect log data about the user's activities in an abstract way, and using a camera needs a large consumption of calculation. There are other research using sensors which can represent the user's activity like accelerometers or physiological sensors. These are sensors that mostly use continuous data values. As these sensors' measurements are continuous, the research using these sensors are using various methods to quantize the continuous measurements. Meijer used a motion sensor and accelerometers for measurements, calculated the difference with each activity and compared it with a rule [3]. Ravi uses three axes accelerometers with several classifiers, naïve Bayes, decision tree and SVM [4]. Parkka also used three axes accelerometers, and in addition, they added physiological sensors and used a decision tree for the classifier [5].

Subramanya used a GPS and a light sensor to detect the location and an accelerometer to recognize the related activity with the location [6]. They used binning for preprocessing and dynamic Bayesian network for classification. These research use sensors which collect continuous measurements and classify activities by using binning or decision tree. Binning and decision tree are useful to recognize activities but there can be some problems of segmenting the continuous

978-1-4244-2144-2/08/$25.00 ©2008 IEEE

measurements, especially if the measurements are ambiguous between activities. In this paper, we use fuzzy logic which has advantages to represent continuous data in symbolic states for preprocessing [7], and a fuzzy Bayesian network, which is compromised with the preprocessed fuzzy data, to solve the problem between ambiguous measurements. We use a sensor, Armband, which has an accelerometer and physiological sensors as well, and collect the user's activity log information by using a PDA.

B. Physiological sensor

Various sensors are used to recognize the user's activity such as GPS, cameras, microphones, accelerometers and physiological sensors. Among those various sensors, Bodymedia's Armband is a sensor which can measure the user's physiological signals. It has five kinds of sensory resources inside, a two axes accelerometer, a heat flux sensor, a galvanic skin response sensor, a skin temperature sensor, and a thermometer [8]. With these five sensors it calibrates the sensory data and combines to 24 kinds of data.

Thus, the Armband can recognize not only the dynamic activities by the accelerometer but also the stable and static activities by the physiological signals, too. The Armband has a maximum of 32 Hz sampling rate, and uses the Innerview professional 5.0 [8] to collect the data. Innerview professional 5.0 is able to show the collected data in a graph or convert the data to several formats of files. Figure 1 shows the appearance of the Armband when it is worn on the upper left arm and the screen of the Innerview professional 5.0 showing the data in a graph.

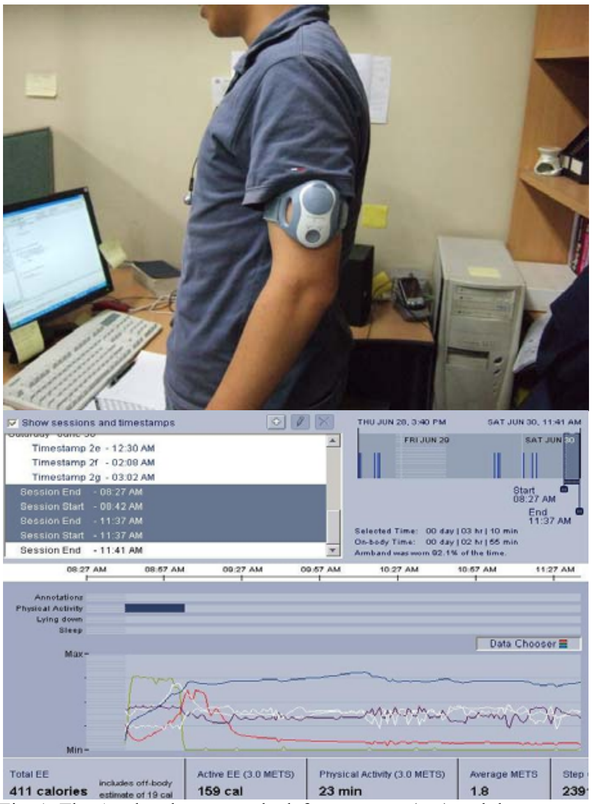

Fig. 1. The Armband worn on the left upper arm (*top*) and the screen of Innerview professional 5.0 (*bottom*)

In this work, from the 24 kinds of measurements by the Armband, we use the two axes accelerometer values, galvanic skin response, heat flux, skin temperature, energy expenditure, metabolic equivalants, and step counts.

C. Fuzzy logic

Fuzzy logic has been used in a wide range of problem domains in process control, management and decision making, operations research, and classification. The conventional crisp set model makes decision in black and white, yes or no, and it has a typical boundary in the classification of several stages. The representation of the low stage will have a value of 1 until the upper bound, and the moment the value goes over the upper bound, the value of the low stage will suddenly change into 0. This crisp set model of classification is simple to implement but when it is used in continuous values in real number, like sensory data, it is a hard problem to decide the boundary. It is also a hard problem for the inference models when the data is nearby the boundary, keeping the robustness of a little change of the data. By using fuzzy logic, the decision becomes flexible and can keep each stage's representation near by the boundary of the data. It helps to represent the vagueness of human intuition in a linguistic modeling which is hard in the crisp model [7].

A fuzzy membership function calculates the fuzzy membership for each stage with a specific value. The most popular function type is a trapezoidal membership function and a triangular membership function which the graph is shown in figure 1. These functions are easy to implement, have low consuming calculations like formula (1) and (2) [9]. The triangular membership function(MF_{tri}) requires 3 parameters and the trapezoidal membership function(MF_{trap}) requires 4. The value is simple divided with a rule of the range of data x. The parameter can be chosen and modified by a direct view of a graph. There are also a membership using the Gaussian distribution and a sigmoidal membership function.

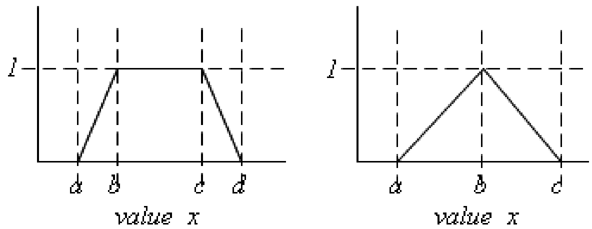

Fig. 2. The graph of the trapezoidal (*left*) and triangular (*right*) fuzzy

$$MF_{tri}(x) = \begin{cases} 0 & , x < a \\ \dfrac{x-a}{b-a} & , a \le x < b \\ \dfrac{x-c}{b-c} & , b \le x < c \\ 0 & , c \le c \end{cases} \tag{1}$$

$$MF_{trap}(x) = \begin{cases} 0 & , x < a \\ \dfrac{x-a}{b-a} & , a \le x < d \\ 1 & , b \le x < c \\ \dfrac{x-d}{c-d} & , c \le x < d \\ 0 & , d \le x \end{cases} \quad (2)$$

D. Bayesian networks

A Bayesian network is a graph with probabilities for representing random variables and their dependencies. It efficiently encodes the joint probability distribution of a large set of variables. This representation consists of two components. The first component is a directed acyclic graph(DAG) and the second component describes a conditional distribution for each variable. The nodes of the graph represent random variables and its arcs represent dependencies between random variables with conditional probabilities at nodes. As the graph is a directed acyclic graph all edges are directed and there is no cycle when edge direction are followed.

Each node has a initial probability table which can be measured by a statistical method or by a expert's knowledge, so that the network can make a robust result even if there is a missing value in a uncertain environment. This characteristic is especially important in the usage of sensory resources. Calculating the probability of a node is based on the Bayes rule. If a node does not have a parent the probability is as the initial probability table, and if it has, it is calculated by adding the multiplication of each probability of the state in the variable node and the conditional probability of the variable node's state in the parent's probability table. The calculation of the probability is like formula (3) [10].

$$P(A) = \sum_i P(A \mid B_i) P(B_i) \quad (3)$$

The joint probability of random variables $\{x_1, \dots, x_n\}$ in a Bayesian network is calculated by the multiplication of local conditional probabilities of all the nodes. Let a node x_i denote the random variable x_i in the Bayesian network, and $parent_i$ denote the parent nodes of x_i, from which dependency arcs come to node xi, Then, the joint probability of $\{ x_1, \dots, x_n\}$ is given as the following formula (4).

$$P(x_1, \dots, x_n) = \prod_i P(x_i \mid parent_i) \quad (4)$$

It is not a simple task to get the exact conditional probability distribution when the variables have continuous values and high order dependencies. As conditional probability table is not suited for continuous values because the values should be quantized and the table size will grow larger with the dependency order.

So we used fuzzy Bayesian network which can make more flexible inferences by preprocessing the continuous variable data in fuzzy logic, and train the conditional probability table by a fuzzy training method which can be differed with the conventional discrete training model.

III. FUZZY BAYESIAN NETWORK FOR ACTIVITY RECOGNITION

As directly using the sensor data is an issue to consider, we need a framework to quantize the data. The flow of the system is like figure 3. The Armband collects the log data based on the user's activity, and a PDA is used for the user to annotate the current activity he or she is doing. These two log data are collected simultaneously and save the same format of the current time. Then, the log integrator will merge these data into entire integrated log data. The preprocessor will use this data and generate a fuzzy integrated data for each sensor log. With this data the fuzzy Bayesian network will train the conditional probability tables, and the inference also uses this fuzzy integrated data.

A. Preprocessing with fuzzy logic

As measurements from Armband are continuous a step of preprocess is necessary. Figure 4 shows the distribution of the measurements and the continuous measurements with the discrete function results.

Segmenting continuous data is a considerable issue when the data lay on one side, like figure 4(*top*), because a little difference of segmentation makes a huge difference of the result. We made a fuzzy membership function depending on the distribution using the mean and standard deviation of the each sensory measurement.

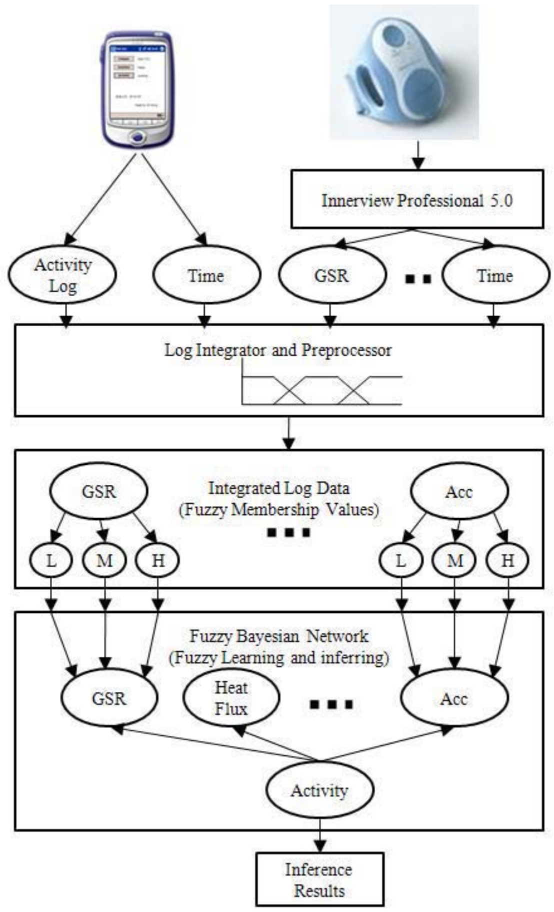

Fig. 3. The overview of the system flow

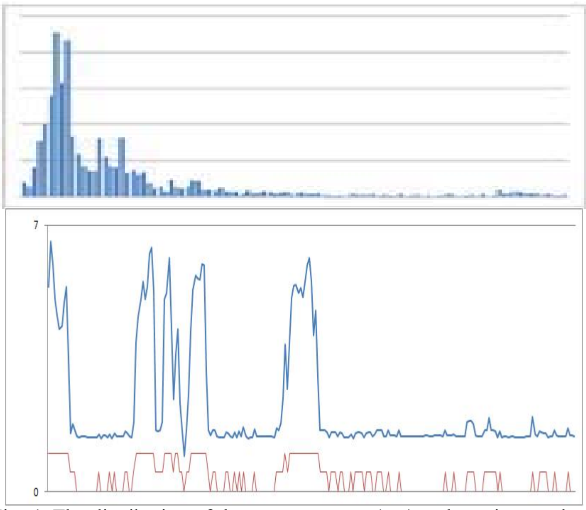

Fig. 4. The distribution of the measurements (*top*) and continuous data with the discrete function results (*bottom*)

We separated the data in half by the mean value, and calculated the distribution of each side with the standard deviation in three stages. If the distribution of a side is wide it uses a Gaussian membership function, or if it's distribution is moderate it uses a trapezoidal, and if the distribution is very narrow it uses a triangular fuzzy membership function. Each side can have these three kinds of different fuzzy membership function, so by combining the fuzzy membership of each side, there can be nine kinds of combinations of a full complete membership function. For the discrete function we used a function based on the fuzzy membership function which segments the data into three states. The state of the discrete function chooses the state which has the maximum fuzzy membership value between other states. But this function can have a problem when the data is like figure 4(*bottom*). The curve above represents the continuous data from the sensor and the curve below shows the results of a discrete function. The higher values have no problem, but the lower values' discrete function result shows a shaking result with a small transition. This can be a problem in the inferring phase, because the evidence will change with just a small amount of transition.

B. Training of fuzzy Bayesian network

Using the same training method with the discrete model will not work perfectly as the input of each variable are not the same as the discrete model which only chooses a single state of the variable, but it contains the fuzzy membership value of each state. As we used fuzzy logic to preprocess the measurements, we used a learning method which fits to the fuzzy data, differed with the discrete model of training based on the Bayes theorem. The training method of the discrete model only counts a state of the data, and the fuzzy method will count all of the state which has a fuzzy membership value.

As sensory resources are independent to each other without any dependencies, we used the structure of the naïve Bayes classifier, and make the variables to give no influence or have any interactions between each other [11]. Naïve Bayes classifier is based on the Bayes theorem calculating P(B|A). This can be calculated by formula (5) in a discrete model. In formula (5), μ stands for the discrete function(D) result of an input of the evidence variable(A) about the state of the evidence($state$) with the input x. The discrete function only allows one state to be counted for the conditional probability.

$$\mu_{state} = \begin{cases} 1 & , D(A) \in state \\ 0 & , D(A) \notin state \end{cases} \qquad P(state \mid A) = \sum_x \frac{\mu_{state}}{\sum_A \mu_{state}^A} \tag{5}$$

Differed with the discrete model the fuzzy learning method allows more than one states to be counted for the conditional probability. The fuzzy membership function(MF) will produce the fuzzy membership value of each state of the input evidence, and all of the value is considered to be used.

$$P(state \mid A) = \sum_x \frac{(MF(x) = \mu_{state})}{\sum_A \sum_x (MF(x)^A = \mu_{state}^A)} \tag{6}$$

This makes the values, on the edge of the discrete function which is ambiguous of determining the state, to be more flexible and give probability to both of the ambiguous states.

C. Inference of fuzzy Bayesian network

When we use the fuzzy Bayesian network to infer the status of a variable, we use a discriminant which chooses the state which has the largest probability. The discriminant function f is like formula (7) in a naïve Bayes classifier. It will multiply each evidences of each variable's conditional probability. In this formula, E is the evidence values from the environment, and k is the inferring status of the Bayesian network. v_{ik} means the value of the evidence variable A_j [11].

$$f_i(E) = P(C_i) \prod_j P(A_j = v_{jk} + C_i) \tag{7}$$

In the discrete method of the naïve Bayes classifier, formula (7), only one of the evidence conditional probability of each variable only gives influence to the result of the discriminant. As a very little difference of data will change the results of the inference of the Bayesian network, in the discrete method there are problems when the data is like figure 4(*bottom*). So we modified the discriminant function like formula (8).

$$f_i(E) = P(C_i) \prod_j \sum_k (MF(E_j)_k \times P(A_j = v_{jk} + C_i) \tag{8}$$

In formula (8), when the result of the membership function(MF) is not zero, which means the data is near the boundary of the status, the conditional probability each of the status' above and below the boundary gives influence to the result. This dissolves the radical difference when the data moves over the boundary and helps the classifier to keep the changes calm about the data which is ambiguous to divide into bins

IV. Experiment

A. Experiment method

We collected the data using the Armband sensor for the physiological signals which was worn on the right upper arm and a PDA for the labeling of the activity by the user. The user selects an activity saved in the PDA. The program for the labeling was programmed by Embedded Visual C++ based on the PocketPC standard development kit. As physiological signals change moderately the data collection term of the Armband and the PDA was set for a data once a minute. After the log collected the data from the Armband was converted by the Innerview professional 5.0 program to an excel format. The converted data is integrated with PDA labeling data by a integration program, and it is used for the training and testing of the fuzzy Bayesian network. There are nine kinds of activity which can occur in a real life and office environment. The activities which can be differed by the movements were walking, running, and exercising. The other activities were eating, reading, studying, playing, resting and sleeping. Totally nine activities were collected, in the real life, freely with no restricts to the user.

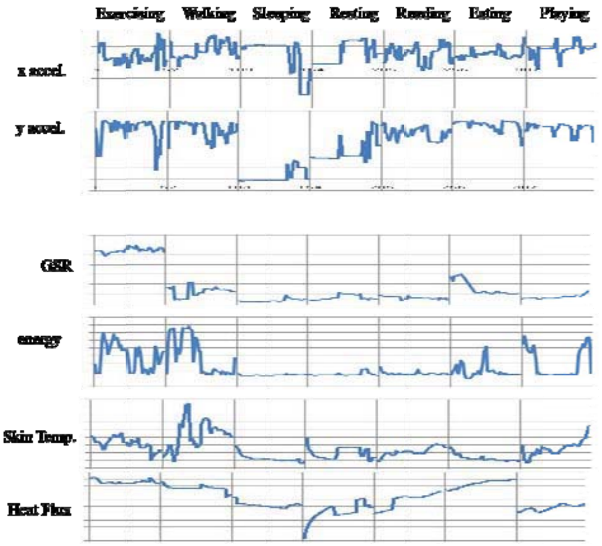

Fig. 5. The raw sensor data from the accelerometer and significant physiological signals of each activity

The raw sensor data was collected like figure 5. The graph shows the sequence of the sensor signals. By only using the accelerometer signals, the activities which are dependent of the user's movements show they are significantly different with the others, but other activities, which are stable, are hard to recognize the difference between each other. As the physiological signals in figure 5 shows, using these additional signals can lessen the difficulty only using the accelerometer signals.

The first experiment was in four kinds of combinations of the discrete model and fuzzy model for the training and testing method. We compared the training method between the Bayes theorem based method and the fuzzy logic based method. The testing method was compared between the naïve Bayes classifier and the naïve Bayes classifier using the fuzzy inference. The total data set has a size of 5,500 samples and placed randomly for the cross validation which is divided to 10 folds.

The second experiment was to show the accuracy of the activities when using different sensor data. The first set uses all of the sensor data, both the physiological signals and accelerometer signals. The second set only used the physiological signals, and the third set used the 2 axis accelerometer signals. Each set was trained and tested in the fuzzy Bayesian network and was compared the accuracy between activities.

B. Experiment results

Table 1 is the result of the first experiment's average of the ten folds cross validation. The discrete methods had 70.0% of accuracy and the combination of the discrete training and the fuzzy inferring had 71.3% of accuracy. Even though the training method is a discrete model the fuzzy inference had a higher and more stable accuracy than the discrete inference. The fuzzy methods show 74.4% of accuracy and when the discrete inference is used with the fuzzy training the accuracy is low by 62.3%. This shows that the discrete inference is not capable for the fuzzy training, but the combination of fuzzy methods has higher accuracy than the discrete methods.

TABLE I
THE AVERAGE RESULTS OF THE EXPERIMENT

Learning	Discrete		Fuzzy	
Inferring	Discrete	Fuzzy	Discrete	Fuzzy
Accuracy	70.0(±2.3)	71.3(±2.0)	62.3(±2.7)	74.4(±1.4)

The following figure 6 is a chart of the result of the cross validation with ten folds. The bar represents the combinations of training and inferring methods, discrete training and discrete inferring, discrete and fuzzy, fuzzy and discrete, and fuzzy and fuzzy from the left. As the chart shows, the fuzzy method has higher accuracy in all ten folds.

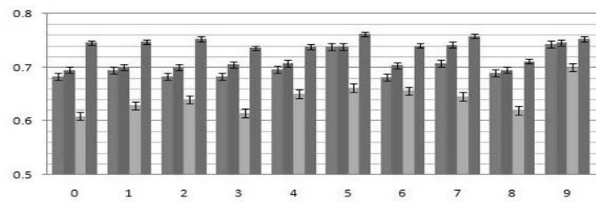

Fig. 6. The results of the cross validation with ten folds. The bars from the left are discrete learning-discrete inferring, discrete-fuzzy, fuzzy-discrete, fuzzy-fuzzy

Table 2, 3, 4 shows the result of the second experiment, in three different data sets. Table 2 shows the confusion matrix when using all of the sensor data, and table 3 and 4 each shows when using only the physiological signals and accelerometer signals. The rows show the activity and the columns show the recognition results.

TABLE II
CONFUSION MATRIX USING PHYSIOLOGICAL SIGNALS AND ACCELEROMETER

	E1	P	R1	R2	R3	S1	S2	W	E2
E1	43.7	12.1	10.4	0	0	0	14.2	18.9	0
P	0	92.7	0	0	0	0	0	0.07	6.6
R1	16.3	1.0	73.5	5.1	0	0	3.1	0	1.0
R2	4.0	0	0.5	86.3	0	0	7.6	0	1.6
R3	0	0	1.6	0	84.1	0	0	12.7	1.6
S1	2.4	0.8	9.9	0	0	85.8	0	0.3	0.8
S2	11.7	0	1.3	5.2	0	0	74.7	2.6	34.5
W	6.0	2.7	0.7	0	2.4	0	3.1	75.8	9.3
E2	4.7	9.9	8.3	0	1.1	0	1.6	31.2	43.2

E1=Eating, P = Playing, R1 = Reading, R2 = Resting, R3 = Running, S1 = Sleeping, S2 = Studying, W = Walking, E2 = Exercising (Total accuracy = 74.2%).

TABLE III
CONFUSION MATRIX USING PHYSIOLOGICAL SIGNALS

	E1	P	R1	R2	R3	S1	S2	W	E2
E1	30.9	13.3	0	17.1	0	0	18.3	20.4	0
P	0	90.6	0	0	0	0	0	0	9.4
R1	78.6	1.0	1.0	13.3	0	0	6.1	0	0
R2	3.0	0	0	55.9	0	0	41.1	0	0
R3	1.6	0	0	0	58.7	0	0	38.1	1.6
S1	13.1	1.1	0	0	0	85.5	0	0.3	0
S2	13.6	0	0	7.1	0	0	73.5	5.8	0
W	1.3	2.2	0	3.5	15.3	0	3.8	57.7	16.2
E2	14.6	10.9	0	1.0	3.1	0	0	29.2	41.2

E1=Eating, P = Playing, R1 = Reading, R2 = Resting, R3 = Running, S1 = Sleeping, S2 = Studying, W = Walking, E2 = Exercising (Total accuracy = 60.9%).

TABLE IV
CONFUSION MATRIX USING ACCELEROMETER

	E1	P	R1	R2	R3	S1	S2	W	E2
E1	24.2	32.5	0.8	0	0	0	19.2	20.0	3.3
P	17.1	67.9	0	2.1	0	0	3.5	3.5	5.9
R1	5.1	70.4	12.3	2.0	0	0	6.1	3.1	1.0
R2	1.0	44.7	0.5	48.8	0	0	2.0	0.5	2.5
R3	0	1.6	0	0	82.5	0	0	14.3	1.6
S1	0.8	47.4	1.9	42.2	0	0	6	0.3	1.4
S2	5.8	7.8	0.6	12.3	0	0	39.7	2.6	31.2
W	5.5	3.8	0	0	4.2	0	2	76.3	8.2
E2	7.3	13.5	3.1	0	0.5	0	2.1	28.6	44.9

E1=Eating, P = Playing, R1 = Reading, R2 = Resting, R3 = Running, S1 = Sleeping, S2 = Studying, W = Walking, E2 = Exercising (Total accuracy = 44.7%).

The accuracy of table 2 is all better than the other tables, table 3, which only used the physiological signals, show that it is more efficient when recognizing the stable activities than table 4, and in the opposite table 4 shows that the accelerometer is more efficient when recognizing the dynamic activities. Because there is only one 2-axis accelerometer worn on the upper right arm the results are not so good, but the dynamic activity recognition accuracy is as good as or better than table 2's results. As table 3's accuracy of each activity is a little lower than table 2's accuracy, because that even an activity is stable, the accelerometer helps the recognition. The dynamic activities in table 3 shows that they are confused with each other because the there is no accelerometer information. Table 4's accuracy of walking is much higher than table 2 and running and exercising is similar with table 2. This means that dynamic activities can be recognized only with the accelerometer. The reading activity was confused with the eating activity and resting activity, but when the physiological signals and accelerometer signals are combined it showed 73.5 percent of accuracy.

V. CONCLUSION

In this paper, we showed that using a fuzzy Bayesian network is more efficient than the discrete model when using continuous data for recognizing the user's activity. We used the Armband sensor which calibrates physiological signals and includes a 2-axis accelerometer. The first experiment results are, when we used the discrete model of the naïve Bayes classifier has shown 70.0 percent of accuracy and the fuzzy Bayesian network has shown 74.4 percent of accuracy. The second experiment has shown that each activity has an efficient sensor to recognize using the physiological signals or the 2 axis accelerometer depend on the vitality, and combining these two kinds of sensor helps to recognize all of those activities. In the future work, we will need to integrate more sensors for the context-aware service about the user activity, change the frequency of the data collection time, and improve the classifier for a temporal inference model to analyze the sensory data's alterations.

REFERENCES

[1] E. Tapia, S. Intille and K. Larson, "Activity recognition in the home using simple and ubiquitous sensors," Lecture Notes in Computer Science, Int. Conf. on Pervasive Computing, vol. 3001, pp. 158–175, 2005.

[2] J. Han and B. Bhanu, "Human activity recognition in thermal infrared imagery," IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, vol. 3, pp. 17–25, 2005.

[3] G. Meijer, K. Westererp, F. Verhoeven, H. Koper and F. Hoor, "Methods to asses physical activity with special reference to motion sensors and accelerometers," IEEE Trans. Biomedical Enginerring, vol. 38, no. 3, pp. 221–229, 1991.

[4] N. Ravi, N. Dandekar, P. Mysore and M. Littman, "Activity recognition from accelerometer," Proc. Seventeenth Innovative Applications of Artificial Intelligence Conf., pp. 11-18, 2005.

[5] J. Parkka, M. Ermes, P. Korpipaa, J. Mantyjarvi, J. Peltola and I. Korhonen, "Activity classification using realistic data from wearable sensors," IEEE Trans. Information Technology in Biomedicine, vol. 10, no. 1, pp. 119-128, 2006.

[6] A. Subramanya, A. Raj, J. Bilmes, and D. Fox, "Recognizing activities and spatial context using wearable sensors," In Proc. of Conf. on Uncertainty in Artificial Intelligence, 2006.

[7] E. Cox, "Fuzzy fundamentals," IEEE Spectrum, vol. 29, No. 10, pp. 58-61, 1992.

[8] Bodymedia©, "Sensewear behavior prescribed," http://www.sensewear.com/images/bms_brochure.pdf

[9] J. Jang, C. Sun and E. Mizutani, Neuro-fuzzy and soft computing, Prentice Hall, 1997.

[10] K. Korb and A. Nicholson, Bayesian artificial intelligence, Chapman and Hall/Crc Computer Science and Data Analysis, 2004.

[11] P. Domingos and M. Pazzani, "On the optimality of the simple Bayesian classifier under zero-one loss," Machine Learning, vol. 29, No. 2-3, pp. 103-130, 1997.