Proceedings of the 2011 IEEE
International Conference on Robotics and Biomimetics
December 7-11, 2011, Phuket, Thailand

Classification of Long-term Motions Using a Two-layered Hidden Markov Model in a Wearable Sensor System

Si-Jung Ryu and Jong-Hwan Kim

Abstract— **This paper proposes a classification system for long-term motions in a wearable sensor system with 3-axis accelerometers. The long-term motion is defined as a sequence of short-term motions so that the overall classification algorithm processes short-term motions in the first layer and then classifies long-term motions in the second layer. The hidden Markov model is employed in each layer as a classification algorithm. The wearable sensor system consists of two 3-axis accelerometers which are attached to both forearms. Raw data from the accelerometers are pre-processed and forwarded to the classification algorithm designed with the hidden Markov model. For comparison, other algorithms such as artificial neural networks, support vector machine, k-nearest neighbor algorithm and k-means clustering, are tested. In experiments, eight kinds of short-term motions are randomly selected from daily life to test the performance of the proposed system and to compare its performance with that of existing algorithms. Also, three long-term motions which consist of short-term motions are selected and tested to demonstrate the effectiveness of the proposed algorithm.**

I. INTRODUCTION

Computer-related technological advances have led a new era of computing environments for years. Accordingly, an interest of the intelligent human-computer interface has grown recently. A convenient and efficient interface for computer will improve the quality of life. Generally, wearable computing covers a wide range including not only intelligent computer that is worn to the body, but also just sensors attached to the body. These days, wearable computing technology has been applied in a variety of fields including sports, medical care, the game, etc. These wide demands for wearable computing cause many researches about wearable computing, and there are two big issues recently. One is a new computer environment that is combined with human-friendly interface. The other is a wearable health care system that can help patients or elderly people at long distance.

These researches focus on the wearable computing based on human-friendly interface. To develop the human-friendly interface, many devices have been employed. Firstly, the devices such as accelerometers, gyro sensors or combinations of two sensors are mostly used to measure physical activities [1]. In addition, the number of sensors used in many reasearches is also various. There are pros and cons in recognizing human motions whether many sensors are used or not. The benefit of using many sensors is that sensor system could improve the accuracy of classification of the complex motions and the disadvantage is that overfitting problems might occur. Therefore, the appropriate number of sensors should be determined considering complexity of motions and robustness. Secondly, there are the studies using bio-signal measuring devices which can measure brain signal or cardiovascular signal like EEG or ECG [2]. These kinds of studies have grown rapidly in recent years and are regarded that feature extraction of bio-signal is a key point. Thirdly, there are researches using the combination devices of the various devices to improve the recognition accuracy of human activities [3]-[4]. Generally, a variety of combination methods are employed such as Bayesian network, neural network, or simple weighted summation [5].

As mentioned above, human motion classification using sensor systems is widely used. Hardware platforms based on accelerometers are most popular. The sensors are attached to various locations such as forearms, wrists, head, waist, legs, etc [6]. As classification algorithms, a variety of methods have been developed such as classifier such as neural network, support vector machine, k-means clustering, k-nearest neighbor, etc. [7].

However, compared to many researches for the motion that occur over a short time, researches for the activities that is occur over a long time have not been actively studied. This is because accuracy rate for classifying long-term motions is not enough to be used. Generally, it is assumed that long-term motion is a sequence of several short-term motions. Therefore, high classification accuracy for short-term motions must be guaranteed to classify the long-term motion efficiently [8].

This paper proposes a long-term motion classification for a wearable sensor system with 3-axis accelerometers considering a human-friendly interface. To demonstrate the effectiveness of the proposed algorithm, experiments are carried out for eight kinds of motions that are randomly selected from daily life. Also, its short-term performance is compared with that of existing algorithms such as neural network, k-nearest neighbor, support vector machine, and k-means clustering. After then, long-term motions are classified using a proposed algorithm.

The rest of this paper is organized as follows. In Section 2, motion classification with two-layered hidden Markov model is proposed. Section 3 describes the wearable sensor system. The experimental results are discussed in Section 4 and concluding remarks follow in Section 5.

II. Motion Classification with Two-layered Hidden Markov Model

A. Hidden Markov Model

A hidden Markov model (HMM) is a statistical Markov model in which the system being modeled is assumed to be a Markov process with unobserved hidden states. A HMM

978-1-4577-2138-0/11/$26.00

can be considered as a simplest dynamic bayesian network. The hidden Markov models are popularly used for pattern recognition such as speech, handwriting, gesture recognition because they have the advantages for spatio-temporal data.

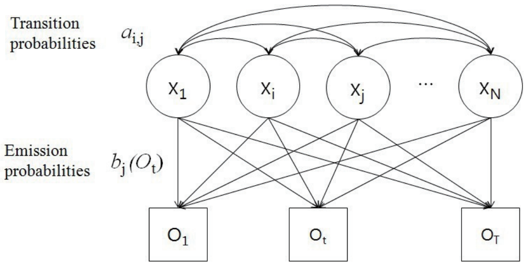


Fig. 1: Hidden Markov model

There are two important topics for HMM: forward-backward algorithm and Viterbi path. The forward-backward algorithm iteratively re-estimates the parameter of the HMM $\lambda = (\pi, a, b)$ to train HMM for a given sequential data.

π, a, and b denotes initial state transitions distribution, state transition probability matrix, and emission probabilitiy matrix, respectively. Initial state transition distribution π_i is the probability of the first state for the state i. Transition probability $a_{i,j}$ is the probability of the state transition from the state i to the state j. Emission probability $b_j(O_t)$ is the probability of emitting the observation O_t from the state j. Parameter of the HMM λ is arbitrarily defined in the initialization procedure.

1) Forward-backward Algorithm: The forward probability $\alpha_t(i)$ for the state i at time t, means the probability that HMM outcomes a sequence of observaions $(O_1, O_2, ..., O_t)$ and ends in the state i. Each forward probability is calculated through the three steps, i.e. initialization, induction and termination. T he forward probability is calculated of as follows:

a) Initialization:

$$\alpha_t(i) = \pi_i b_i(O_t) \qquad 1 \le i \le N \tag{1}$$

where N is the number of the states.

b) Induction:

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^{N} \alpha_t(i) a_{i,j}\right] b_j(O_{t+1}) \qquad \begin{array}{l} 1 \le j \le N \\ 1 \le t \le T \end{array} \tag{2}$$

where T is the length of the obervations.

c) Termination: The probability of generating a sequence of the states for a given input data M_S is evaluated by

$$P(O|M_S) = \sum_{i=1}^{N} \alpha_T(i). \tag{3}$$

Evaluation of (3) needs a summation of N^T possible sequences. However, it is not needed to compute all the forward probabilities. Instead, once the HMMs have been trained, the forward-backward algorithm produces a sequence of the states with a largest likelihood $P(O|M_S)$ for a given observation.

If the observation O_t has L dimensions, the probability of the partial observation sequence is evaluated as follows:

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^{N} \alpha_t(i) a_{i,j}\right] \prod_{k=1}^{L} b_j(O_{t+1}(k)). \tag{4}$$

2) Viterbi Path: The most probable sequential states that forward-backward algorithm finds is called Viterbi path. If the observation O_t has L dimension, the Viterbi path is generated as follows:

$$\underset{j}{\operatorname{argmax}} \left[\sum_{i=1}^{N} \alpha_t(i) a_{i,j}\right] \prod_{k=1}^{L} b_j(O_{t+1}(k)). \tag{5}$$

The acceleration values are discretized into the same interval and become a sequence of the observations in HMM. In this paper, independent HMM models are designed as many as the number of motions, so each HMM model indicates one motion. One HMM model is trained with acceleration data of one motion. After the forward probability of each input data is calculated via all HMM models, a motion that has the most probable model indicates a specific motion.

B. Classification of Long-term Motion

It is assumed that long-term motion is a sequence of several short-term motions. Therefore, classification of the long-term motion is also a sequence of classifications of the short-term motions. The whole data is divided into equal parts, and each part is applied into the HMM over a sliding window. Then, the forward probability is calculated for a given input data in the sliding window. If the forward probability is smaller than some threshold ξ, it means meaningless motion has occurred. Otherwise, one motion with the highest probability is selected by the HMM. The value of ξ is determined with the value which induces the highest classification accuracy of the short-term motions. The value of ξ is set to 10^{38} for all experiment unless otherwise stated. The number of the sliding windows, $N_{sliding}$ is computed as follows:

$$N_{sliding} = L_{total} - L_{sliding} + 1 \tag{6}$$

where the length of sliding window is $L_{sliding}$, and the length of whole data is L_{total}.

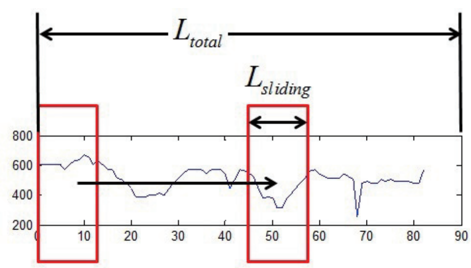


Fig. 2: Sliding window

A sliding window moves to the right following the time axis as shown in Fig 2. Classification is performed by computing the forward probability of each HMM model over a sliding window, and assigning to the data the label of the motion with the highest probability. If the meaningful short-term motions $(ml_1, ml_2, ..., ml_{N_S})$ have appeared, they are stored in the set M_L. After all classifications are finished through the sliding windows, the set M_L is occupied with meaningful short-term motions, where M_L denotes one long-term motion, as follows:

$$M_L = \{ml_1, ml_2, ..., ml_{N_S}\} \tag{7}$$

where N_S is the number of the short-term motions detected by HMM in the first layer. Then, the set M_L is forwarded to another HMMs in the second layer to classify the long-term motion. The overall procedure is depicted in Fig 3.

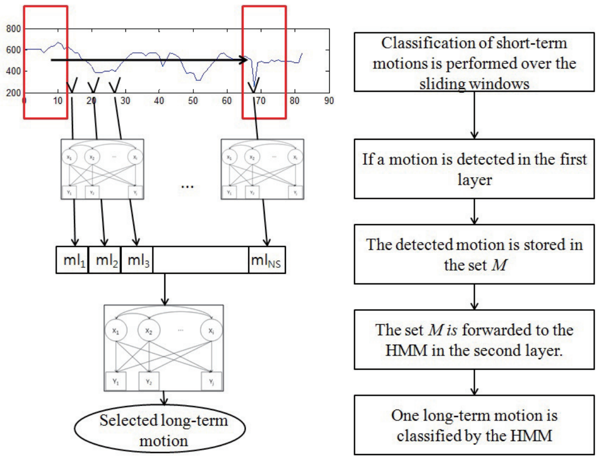


Fig. 3: Two-layered hidden Markov model

III. WEARABLE SENSOR SYSTEM

A. Hardware Structure

Hardware platform of the wearable sensor system mainly consists of accelerometers and a communication module. Freescale MMA7260Q triaxial accelerometer is selected as a measurement device, which has four different sensitivities, 1.5g, 2.0g, 4.0g, and 6.0g. In this paper, 1.5g is adopted because most of the activities from daily life are occurred in the range of $\pm$1.5g. Two accelerometers are attached to forearms, and a communication module to a waist, as shown in Fig. 4. Considering both of the accelerometer and micro-controller, a 5V dry cell type battery is used and connected to an ATmega128 and two accelerometers. For communication, Zigbee module is used, which is connected to UART in ATmega128 board. The internal structure of the wearable sensor system is shown in Fig. 5.

B. Data Collection

The output of the accelerometer is analog data, so AT-mega128 converts it into 10-bit digital data. Also, acceleration data from the two accelerometers are collected for each sampling time of 20ms for 3 seconds. In this case, the amount

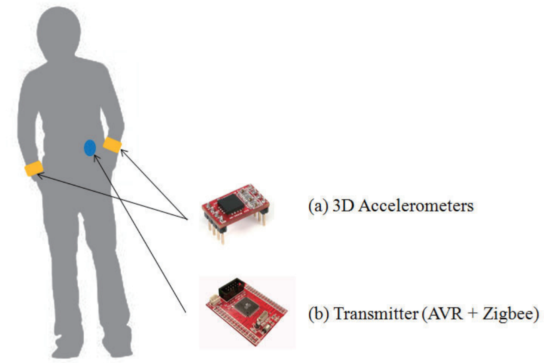


Fig. 4: Sensors and a communication module

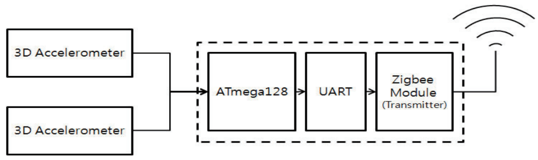


Fig. 5: Internal structure of the wearable sensor system

of acceleration data is too big to be used as an input data of classification. Also, there are two drawbacks: over-fitting problem and heavy computational complexity. Thus, a series of pre-processing steps is needed to use the accelerometer data as an input of the classification algorithms. The overall procedures are executed by data sampling, removal of the gravity acceleration, quantization and arrangement in order.

C. Pre-processing

1) Data Sampling: Raw acceleration data are obtained at 50Hz from the both accelerometers, but the mean value of every five acceleration data is used to reduce the computation complexity of the algorithm.

2) Removal of Gravity Acceleration: Gravity acceleration is included in the acceleration data at the beginning. However, motions are classified using the trajectory of the acceleration data, so the initial gravity acceleration value should be subtracted from the acceleration data. After the removol of the gravitation value, all initial values are set to 0.0 g.

3) Quantization: The initial ADC values vary 0 to 1024. If raw ADC values are applied to the HMM without quantization, enormous number of the states are needed to carry out the classification. In this paper, ADC values are qunatized to be an integer from 0 to 4.

4) Arrangement: There are six acceleration signals from two 3-D accelerometers that can measure x,y, and z-axis accelerations. This arrangement procedure is needed to use these six signals as an input of the proposed algorithm, because it uses one dimensional input data. In this step, six signals are transformed into one dimenisonal string by arranging signals of the x,y, and z-axis signal from left arm to the right arm sequentially.

TABLE I: Short-term motions

Motion No	Motion
1	Halt
2	Swing two arms
3	Shake two arms back and forth
4	Stretch two arms forward
5	Put the hands behind the head
6	Wave one's left hand
7	Raise a right arm
8	Intend to hit something with left arm

IV. EXPERIMENTS

A. Experimental Setup

Eight kinds of motions were randomly selected from daily life. Each motion is defined in Table I. A training set was obtained by repeating ten times for each motion, and a test set was obtained by repeating five times for each motion. One motion data was composed of six accelerometer values for 3 seconds.

B. Experimental Results for Short-term Motions

In the HMM, five states were used and transition values were randomly initialized. The multi-layered perceptron (MLP) trained with backpropagation was employed for neural networks. Three hidden layers were used, and each hidden layer contained five neurons. In the case of k-NN, motions were classifed based on the closest training examples in the 8-dimensional Euclidean space. Finally, a 3rd order polynomial kernel function was used for SVM because input data was non-linear.

Table IIa and IIb indicate the results of classification for the short-term motions using HMM. Table IIIa and IIIb show that the performances of HMM, k-NN and SVM are better than those of the others. However, there are drawbacks in k-NN and SVM. First, the performance of SVM is heavily influenced by kernel functions which transform support vector. Actually, SVM had the worst performance when inappropriate kernel functions were used. Second, the performance of k-NN is dependent on k and types of motions. In other words, k-NN might have poor performance for the classification problem of complex motions. In case of neural network, its structure is similar to that of the HMM because both algorithms have randomness initially and users cannot know how to classify data internally. In addition, both of them can classify the data without specific information about data, so such a kernel function is not needed. It means that performances of classification is less affected by the input data. However, HMM performed far better compared to ANN.

C. Experimental Results for Long-term Motions

The long-term motion data were obtained by moving two arms, to each of which an accelerometer is attached, during 30 seconds. Five motions among the short-term motions in Table I were performed according to a specific order. When the specific short-term motions were not enacted, experimenter moved both arms randomly. Therefore, one long-term motion data consisted of short-term motion of 15 seconds and meaningless data of 15 seconds. The HMM evaluated probabilities of each sliding window and indicated a label of motions among the eight short-term motions. Each long-term motion was performed by repeating 10 times, and 10 data sets of each motion were tested in this paper. Each long-term motion is defined in Table IV.

It was verified that each long-term motion was composed of a series of the short-term motions as shown in Fig 6. The y-axis denotes the label of short-term motions in Table I. For the long-term motion, several short-term motions were

TABLE II: Classification accuracy of the short-term motions

(a) Training data

Motion \ Result	1	2	3	4	5	6	7	8
1	10	0	0	0	0	0	0	0
2	0	10	0	0	0	0	0	0
3	0	0	10	0	0	0	0	0
4	0	0	0	10	0	0	0	0
5	0	0	0	0	10	0	0	0
6	0	0	0	0	0	10	0	0
7	0	0	0	0	0	0	10	0
8	0	0	0	0	0	0	0	10

(b) Test data

Motion \ Result	1	2	3	4	5	6	7	8
1	5	1	0	0	0	0	0	0
2	0	4	0	0	0	0	0	0
3	0	0	5	0	0	0	0	0
4	0	0	0	5	0	0	0	0
5	0	0	0	0	5	0	0	0
6	0	0	0	0	0	5	0	0
7	0	0	0	0	0	0	5	1
8	0	0	0	0	0	0	0	4

TABLE III: Accuracy for training and test data

(a) Training data

Classification Algorithm	Training Data Accuracy
ANN	68.75%
k-NN	-
SVM	96.25%
k-Means Clustering	83.75%
HMM	100.00%

(b) Test data

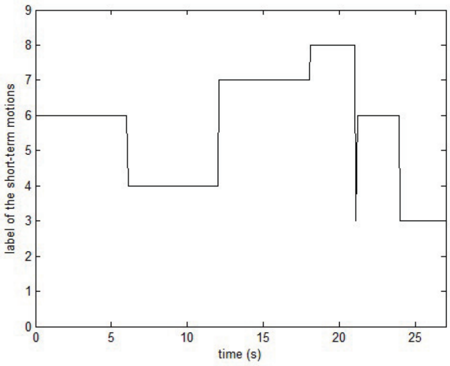
Classification Algorithm	Test Data Accuracy
ANN	46.00%
k-NN	95.00%
SVM	85.00%
k-Means Clustering	47.50%
HMM	95.00%

TABLE V: Accuracy for long-term motions

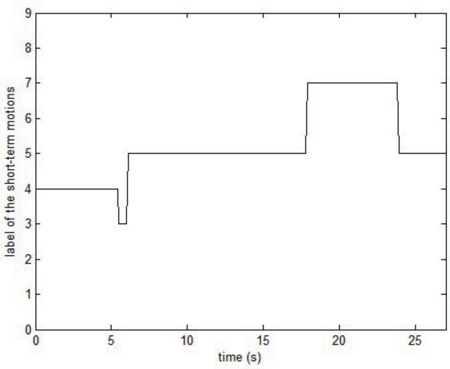
Long-term motion	1	2	3
Accuracy	80.00%	70.00%	70.00%

TABLE IV: Long-term motions

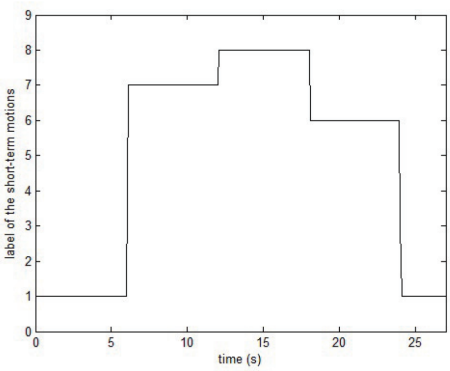
Long-term motion No.	1	2	3
1st short-term motion	Wave one's left hand	Stretch two arms forward	Halt
2nd short-term motion	Stretch two arms forward	Put the hands behind the head	Raise a right arm
3rd short-term motion	Raise a right arm	Halt	Intend to hit something with left arm
4th short-term motion	Intend to hit something with left arm	Raise a right arm	Wave one's left hand
5th short-term motion	Shake two arms back and forth	Put the hands behind the head	Halt



(a) Long-term motion No. 1



(b) Long-term motion No. 2



(c) Long-term motion No. 3

Fig. 6: Classification of long-term motions

detected. Each of the short-term motions was captured at the time when experimenter did one of the short-term motions. Then, these motions were combined in a set M_L. Then, the set M_L was forwared to the HMM in the second layer and the HMM models indicated the classification results. classification results of the long-term motions are shown in Table V. Classification accuracy of the long-term motions decreased compared to that of the short-term motions. The accuracy rate for the long-term motions cannot help dropping to about the products of the accuracy rate for the short-term motions, because the long-term motion is the sequential of the short-term motions. However, there were also the case that the HMM in the second layer classified correctly with one or two incorrect short-term motions.

V. CONCLUSIONS

This paper proposed a novel motion classification system using a two-layered hidden Markov model in the long-term motion data and applied in a wearable sensor system. The system classified the long-term motions of the human attaching two acceleromters to both forearms. The long-term motion was defined as a sequence of the short-term motions so that classification was performed for the short-term motions firstly and the long-term motions secondly. The hidden Markov model was employed to classify eight short-term motions and three long-term motions. To test the performance of the HMM, eight short-term motions were randomly selected from daily life. The performance of the classification for the short-term motion using a hidden Markov model was better than that of other algorithms such as neural network, support vector machine, k-nearest neighbor algorithm and k-means clustering. The long-term motions that consist of seqeunce of the short-term motions were also classified by proposed two-layered hidden Markov model. For the future work, it is expected that the classification accuracy of the long-term motions will improve, if the system is independent on the length of the motions.

VI. ACKNOWLEDGMENTS.

This research was supported by the MKE (The Ministry of Knowledge Economy), Korea, under the National Robotics Research Center for Robot Intelligence Technology support program supervised by the NIPA (National IT Industry Promotion Agency) (NIPA-2010-N02100128)

REFERENCES

[1] N. Kern, B. Schiele, and A, Schmidt, "Multi-sensor activity context detection for wearable computing", *Ambient Intelligence Lecture Notes in Computer Science*, vol. 2875, 2003, pp 220-232.

[2] N. Bu, Member, M. Okamoto, and T. Tsuji, "A Hybrid Motion Classification Approach for EMG-Based HumanRobot Interfaces Using Bayesian and Neural Networks", *IEEE TRANSACTIONS ON ROBOTICS*, vol. 25(3), 2009, pp 502-511.

[3] C. Zhu, and W. Sheng, "Multi-sensor Fusion for Human Daily Activity Recognition in Robot-assisted Living", *3rd European Medical and Biological Conference*, California, USA, 2009, pp. 303-304.

[4] J. Parkka, M. Ermes, P. Korpipaa, J. Mantyjarvi, J. Peltola, and I. Korhonen, "Activity Classification Using Realistic Data from Wearable Sensors", *IEEE Transactions on Information Technology in Biomedicine*, vol. 10(1), 2003, pp. 1-17.

[5] L. Bao, and S.S. Intille, "Activity Recognition from User-Annotated Acceleration Data", *Pervasive Computing Lecture Notes in Computer Science*, vol. 3001, 2003, pp 1-17.

[6] Z.L. Husz, and A.M. Wallace, and P.R. Green, "Human activity recognition with action primitives", *IEEE 4th Conference on Advanced Video and Signal Based Surveillance*, London, UK, 2007, pp. 330-335.

[7] H. Junker, O. Amft, P. Lukowicz, and G. Troster, "Activity Gesture spotting with body-worn inertial sensors to detect user activities", *Pattern Recognition*, vol. 41(6), 2008, pp. 2010-2024.

[8] R.W. DeVaul, and S. Dunn, "Real-Time Motion Classification for Wearable Computing Applications", *Technical report*, MIT Midea Lab.

[9] J.S.I. Caros, O. Chetelat, P. Celka, and S. Dasen, "Very Low Complexity Algorithm for Ambulatory Activity Classification", *3rd European Medical and Biological Conference*, Prague, Czech Republic, 2005.

[10] N. Ravi, N. Dandelkar, P. Mysore, and M.L. Littman, "Activity Recognition from Accelerometer Data", *17th conference on Innovative applications of artificial intelligence*, Pittsburgh, Pennsylvania, USA, 2005, pp. 1541-1546.

[11] X. Long, N. Yin, and R.M. Aarts, "Single-Accelerometer-Based Daily Physical Activity Classification", *31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Minneapolis, Minnesota, USA, 2006, pp. 119-128.

[12] T. Huynh, U. Blanke, and B. Schiele, "Scalable Recognition of Daily Activities with Wearable Sensors", *3rd international conference on Location-and contextawareness*, Oberpfaffenhofen near Munich (DLR), Germany, 2007.

[13] T. Huynh, and B. Schiele, "Analyzing Features for Activity Recognition", *14th joint conference on Smart objects and ambient intelligence*, Grenoble, France, 2005, pp. 159-163.

[14] P. Keir, J. Elgoyhen, M. Naef, J. Payne, M. Hornet, and P. Anderson, "Gesturerecognition with Non-referenced Tracking", *IEEE symposium on 3D User Interfaces*, Virginia, USA, 2006, pp. 151-158.

[15] C. Zhu, Q. Cheng, and We. Sheng, "Human Intention Recognition in Smart Assisted Living Systems Using A Hierarchical Hidden Markov Model", *4th IEEE Conference on Automation Science and Engineering*, Key Bridge Marriott, Washington DC, USA, 2008, pp. 254-258.

[16] D. Aarno and D. Kragic, "Layered HMM for Motion Intention Recognition", *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Beijing, China, 2006, pp. 5130-4135.

[17] D. Aarno, and D. Kragic, "Motion intention recognition in robot assisted applications", *Robotics and Autonomous Systems*, vol. 56, 2008, pp. 692-705.

[18] W. Yu, R. Alqasemi, R. Dubey, and N. Pernalete, "Telemanipulation Assistance Based on Motion Intention Recognition", *IEEE International Conference on Robotics and Automation*, Barcelona, Spain, 2005, pp. 1121-1126.