November 2005

Human Motion Recognition through Fuzzy Hidden Markov Model

X. Zhang
University of Wollongong

F. Naghdy
University of Wollongong, fazel@uow.edu.au

Abstract
A new type of Hidden Markov Model (HMM) developed based on the fuzzy clustering result is proposed for identification of human motion. By associating the human continuous movements with a series of human motion primitives, the complex human motion could be analysed as the same process as recognizing a word by alphabet. However, because the human movements can be multi-paths and inherently stochastic, it is indisputable that a more sophisticated framework must be applied to reveal the statistic relationships among the different human motion primitives. Hence, based on the human motion recognition results derived from the fuzzy clustering function, HMM is modified by changing the formulation of the emission and transition matrices to analyse the human wrist motion. According to the experimental results, the complex human wrist motion sequence can be identified by the novel HMM holistically and efficiently.

Human Motion Recognition through Fuzzy Hidden Markov Model

Xucheng Zhang
School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, NSW 2522, Australia, xz24@uow.edu.au, tel: +61242213455, fax: +61 242213236

Fazel Naghdy
School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, NSW 2522, Australia, fazel@uow.edu.au, tel: +61242213398, fax: +61 242213236

Abstract

A new type of Hidden Markov Model (HMM) developed based on the fuzzy clustering result is proposed for identification of human motion. By associating the human continuous movements with a series of human motion primitives, the complex human motion could be analysed as the same process as recognizing a word by alphabet. However, because the human movements can be multi-paths and inherently stochastic, it is indisputable that a more sophisticated framework must be applied to reveal the statistic relationships among the different human motion primitives. Hence, based on the human motion recognition results derived from the fuzzy clustering function, HMM is modified by changing the formulation of the emission and transition matrices to analyse the human wrist motion. According to the experimental results, the complex human wrist motion sequence can be identified by the novel HMM holistically and efficiently.

1. Introduction

Acquisition of the behavioural skills of a human operator and recreating them in an intelligent autonomous system has been a critical but rather challenging step in the development of complex intelligent autonomous systems. In the recent years, the most popular approach is to learn the human skill through observing the demonstrator. However, due to inherent stochastic characteristics of the human psychomotor movements[1], the actions applied in each instance of performing the task can be different. They all, however, represent the same skill set for the task. A framework is required to define the inherent characteristics of that skill, independent from the actions applied each time the task is performed. This is achieved by considering the most likely human action in performing the task as the skill employed to perform the task.

Hidden Markov model (HMM) is deployed as a stochastic method to define the skills and model their uncertainties. HMM can easily represent the two stochastic processes associated with skill learning including the mental state – the hidden process and resultant state – the observed process. In addition, HMM is a parametric model and its parameters can be optimized by Baum- Welch algorithm for accurate estimation[2] In addition to speech recognition[2], [3], [4], force analysis and mobile robot path planning simulations[5], [6], HMM has been successfully applied to skill acquisition from the human. Yang et al[7] have applied HMM-based learning scheme to the teleoperation control of the Self-Mobile Space Manipulator. At the evaluation stage, the observation symbols were the exchange of an Orbit Replaceable Unit (ORU), while the state symbols were the trajectories of robot manipulator in Cartesian space. By using Viterbi and Baulm-welch algorithms, the optimal pattern matrix – transition probability matrix A, and the emission matrix B have been extracted and applied as the representations of different skills based on different specific tasks. Similar to Yang's work, Tetsunari et al[8], Amit[9] and Hovland et al[10] have also applied HMM to robotic imitation. Contrary to Yang's work, the two stochastic sequences of HMM in the Tetsunari's project are related to the trajectories of robot manipulators and the observation symbols of the motion path.

In the Fuzzy-C-Mean algorithm[11] employed in this work, the original data recorded from sensors is partitioned and transformed into observation symbols. HMM can be used to integrate the observation symbols

Proceedings of the 2005 International Conference on Computational Intelligence for Modelling, Control and Automation, and International Confe Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'05)
0-7695-2504-0/05 $20.00 © 2005 IEEE

over time, and to identify and perceive the motion trajectories. HMM should, however, be modified to adapt the observation symbols to include fuzzy partition measures. This is achieved in this work by modifying the emission matrix of HMM.

The conceptual study of the project is conducted based on the movement of the human wrist. The training data is acquired from the sensors mounted on the wrist.

In the course of the paper, a brief introduction to HMM is provided in Section 2. Based on the theory of Fuzzy-C-Mean algorithm, the proposed HMM model is explained in Section 3. The results on human motion recognition are presented in Section 4 and some conclusions are drawn in section 5.

2. Hidden Markov Model

Hidden Markov Model is a finite state machine generating random observation sequences. At each step t, the mental state of HMM is changed to j. It generates an observation vector – the resultant state o_t, with the probability density

$$b_j(o_t) = P(o_t | x(t) = j) \tag{1}$$

Where *x(t)* denotes the model state at time *t*. Transition between states i and j is governed by the discrete probability

$$a_{ij} = P(x(t+1) = j | x(t) = i) \tag{2}$$

Thus the transition probability is dependent on the mental state denoted as S, and is independent of time t. Figure 1 shows an example of this process for a five-state model $S = \{s_i\}_{i=1\cdots5}$, where each state has three probabilities associated with it including initial, transition and output probabilities, and three observation symbols $b = \{b_i(j)\}_{i=1\cdots5; j=1\cdots3}$. The initial probability $\pi_i = P(x(1) = i)$, and the complete model can be defined in terms of its transition probability A={a_{ij}} and state output probability B={$b_j(o_t)$}. Moreover, the transition probabilities should satisfy $\sum_{j=1}^{N} a_{ij} = 1$.

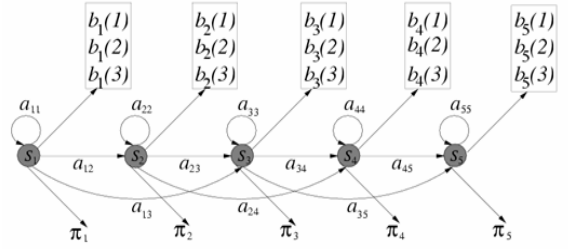

Figure 1 **–The architecture of a typical Hidden Markov Model**[12]

The aim of HMM training is to obtain a set of models $\lambda = (A, B, \pi)$ that, according to an appropriate criterion, will fully match the available training data. The primary goal in the development of HMM is to optimise A and B. Currently, the most popular approach deployed in the optimisation process is the Maximum Likelihood (ML). In practice, three basic problems are addressed in real time pattern recognition based on HMM[13]:

1. Given observation sequence O=$(o_1 ... o_T)$ and $\lambda = (A, B, \pi)$, how efficiently $P(O/\lambda)$-- the observation sequence probability can be computed based on the current model λ.
2. Given O and λ, how an optimal state sequence $X = (x_1 ... x_T)$ can be obtained.
3. How the parameters $\lambda = (A, B, \pi)$ can be tuned to maximise.

In order to resolve these problems efficiently, the following forward-backward algorithm is introduced:

$$\alpha_j(t) = \begin{cases} \pi_i b_i(o_1) & t = 1 \\ [\sum_{i=2}^{N-1} \alpha_i(t-1) a_{ij}] b_j(o_t) & t > 1 \end{cases} \tag{3}$$

$$\beta_j(t) = \begin{cases} 1 & t = T \\ \beta_j(t) = \sum a_{ij} b_j(o_{t+1}) \beta_j(t+1) & t < T \end{cases} \tag{4}$$

Therefore, the total likelihood of observation sequence $P(O/\lambda)$ is given by (5)

$$P(O/\lambda) = \alpha_N(T) = \beta_1(T) = \sum_{j=1}^{N} \alpha_j(t) \beta_j(t) \tag{5}$$

In order to solve the second problem, the most likely state x(t) at time t is chosen as the optimal state. The state x(t) can be estimated by (6), where $\gamma_i(t)$ is the probability of being in state i at time t with model λ and observation sequence o, and is estimated by (7)

$$x(t)^* = \arg(\min_{1\le i\le N}(\gamma_t(i))) \qquad 1 \le t \le T \tag{6}$$

$$\gamma_i(t) = P(x(t)=i|o,\lambda) = \frac{\alpha_i(t)\beta_i(t)}{\sum_{i=1}^{N}\alpha_i(t)\beta_i(t)} \tag{7}$$

In order to simplify the calculation for a long observation sequence, it is necessary to apply another efficient algorithm called Viterbi algorithm which finds the most likely state sequence in the model for the given observation vectors, $P(X|O,\lambda)$.

The third problem is solved by applying Baum-Welch Re-estimation algorithm. Since the full likelihood of each observation sequence is based on the summation of all possible state sequences, each observation vector *o(t)* contributes to the computation of the maximum likelihood parameter values for each state j. Therefore, it is preferable to use soft state boundaries in which, instead of assigning each observation vector to a specific state, it is assigned to every state in proportion to the likelihood of the model being in that state when the vector is observed. All the parameters of HMM are re-estimated by (8) and (9).

$$\tilde{a}_{i,j} = \frac{\sum_{t=1}^{T-1}\xi_{i,j}(t)}{\sum_{t=1}^{T-1}\gamma_{i,j}(t)} \tag{8}$$

$$\tilde{b}_j(k) = \frac{\sum_{t=1}^{T-1}\gamma_{i,j}(t)\Big|_{o(t)=v_k}}{\sum_{t=1}^{T-1}\gamma_{i,j}(t)} \tag{9}$$

Where, v_k denotes a specific observation vector k and $\xi_{i,j}(t)$ is the likelihood of being in state j after state i at time t based on the model λ observation sequence o, and is calculated by (10).

$$\xi_{i,j}(t) = P(x(t)=i, x(t+1)=j|O,\lambda). \tag{10}$$

3. Fuzzy Hidden Markov Model

The data derived from the sensors is partitioned into a series of clusters by applying the Fuzzy-C-Mean algorithm. The partition takes place based on the Euclidean distances calculated for each cluster in the multidimensional space of the sensor data. Based on these clusters, any data can be endued with fuzzy partition measures which denote the likely rate the data belong to each cluster. For example, if there are four clusters, then the data will have four values, and each of them denotes a likely rate that the data belongs to the cluster.

Hence, if the number of clusters is C, the number of data is N, i.e., the length of the observation sequence, and the number of groups that the data are partitioned into is M, then the sequence of clustering results can be expressed as:

$$U = \{u_{ij}\}_{i=1\cdots C, j=1\cdots N} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1N} \\ u_{21} & u_{22} & \cdots & u_{2N} \\ \vdots & \vdots & \vdots & \vdots \\ u_{C1} & u_{C2} & \cdots & u_{CN} \end{bmatrix} \tag{11}$$

Also, the formulation of the emission matrix and transition matrix can be defined as:

- Emission Matrix: $\hat{B} = [\hat{b}_i(O_t)]$, where $\hat{b}_i(O_t)\big|_{t=1\cdots N, i=1\cdots M}$ is the conditional fuzzy measure on the space of observation vectors – Ω –with respect to state I at time t.
- Transition Matrix: $\hat{A} = [\hat{a}_{ij}]$, where $\hat{a}_{ij}\big|_{i,j=1\cdots M}$ is the fuzzy measure on the state j given the previous state i [14].

Therefore, according to the fuzzy clustering results which are tagged with relevant groups, given the statistical independence, the symbol fuzzy density $\hat{b}_i(O_t)$ and transition fuzzy density $\hat{a}_{ij}$ can be calculated directly based on the fuzzy clustering results by using following equations:

- $$\hat{b}_m(\Omega_c) = \frac{\sum_{j\in m} u_{cj}}{\sum_{i=1}^{C}\sum_{j\in m} u_{ij}}; \tag{12}$$

  here, $c \in C$ is the observation symbol c in the space of observation vectors $-\Omega$, while $m \in M$ is the mth state in the state space $S = \{S_1 S_2 \cdots S_M\}$.

- $$\hat{a}_{mn} = \frac{\sum_{S_i=m,S_{i+1}=n} \max_{i=1\cdots C}(u_{ij})}{\sum_{S_i=m,S_{i+1}=n} \sum_{i=1}^{C} (u_{ij})}; \qquad (13)$$

Therefore, the fuzzy forward and backward algorithm can be derived from studying the equations below:

$$\hat{\alpha}_{t+1}(j) = \int \hat{a}_{ij} \circ \hat{\alpha}_{\Omega(t)}(\{o_1 o_2 \cdots o_t\},) \quad \wedge \hat{b}_j(o_{t+1})$$

(14)

$$\hat{\beta}_{\Omega(t+1),T}(\{o_{t+1}o_{t+2}\cdots o_T\}|S_t) = \int [\hat{\beta}_{\Omega(t+2),T}(\{o_{t+2}\cdots o_T\}|S_{t+1}) \wedge \hat{b}_j(o_{t+1})] \circ \hat{a}_{ij}$$

(15)

On the basis of $\hat{\alpha}$ and $\hat{\beta}$, the fuzzy Viterbi algorithm[14] can be applied and the optimal sequence of the states can be derived.

4. Human Motion Recognition

The way a particular skill is applied varies from one individual to another and is performed differently by the same person from one time to another. It is desirable that HMM can identify the inherent characteristics of a particular motion independent from such variations.

The state transition probability matrix A of different discrete states denoted as motion primitives could be a good representation of the inherent characteristics of a particular movement. In order to identify patterns of different human movements, variant state transition probability matrices should be classified as a series of groups which can describe different skill patterns successfully.

The study is currently focused on study and perception of different types of wrist movements. The dual axis accelerometer and the gyroscope are mounted to the hand as shown in Figure 2.

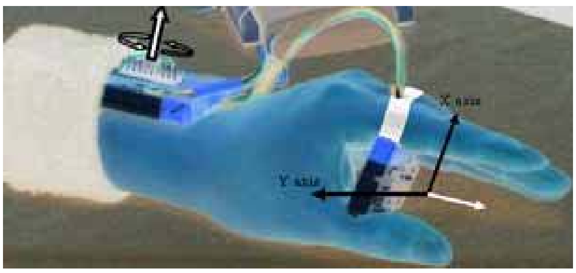

Figure 2 **– Mounting of the sensors on hand**

The gyroscope, Analog Devices ADXRS300 is a complete angular rate sensor (gyroscope). The output signal, RATEOUT (1B, 2A), is a voltage proportional to angular rate about the axis normal to the top surface of the package. A single external resistor can be used to lower the scale factor. An external capacitor is used to set the bandwidth.

The data produced by the sensors are fed into a PMD-1208LS data acquisition system. This is a USB low-speed device usually used for data acquisition and control applications. The device has eight analog inputs, two 10-bit analog outputs, 16 digital I/O connections and one 32-bit external event counter. The device is powered by the +5 volt USB supply and does not require any external power. An example of the four wrist motion primitives recorded by the experimental rig is illustrated in Figure 3.

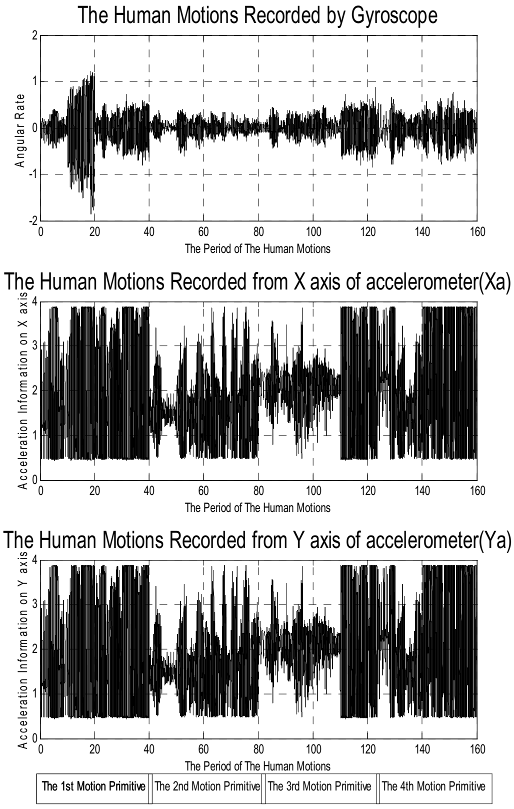

Figure 3 **– Data produced by the sensors**

With this experimental rig, it is possible to identify 4 different hand motion primitives which represent variant fundamental human wrist movements. In building the pattern recognition system, a number of features have been derived from the data obtained from the sensors. This includes the covariance of data obtained from the accelerometers and gyroscope, the mean of those data, and the covariance of the positions of the hand obtained from the gyroscope signal. A typical covariance of the position of the hand is illustrated in Figure 4.

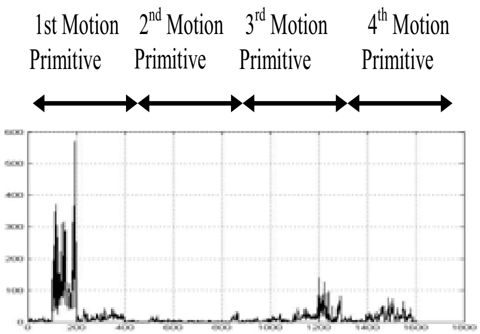

Figure 4 – **Covariance of the positions of the hand**

The fuzzy clustering algorithm was applied to the data obtained from the sensors. Following this, the Eigen Space algorithm[15] was used to derive four different covariance matrices from the fuzzy clustering results. By studying the eigenvalue and eigenvector of the four different matrices, the four clusters could be described as shown in Figure 5.

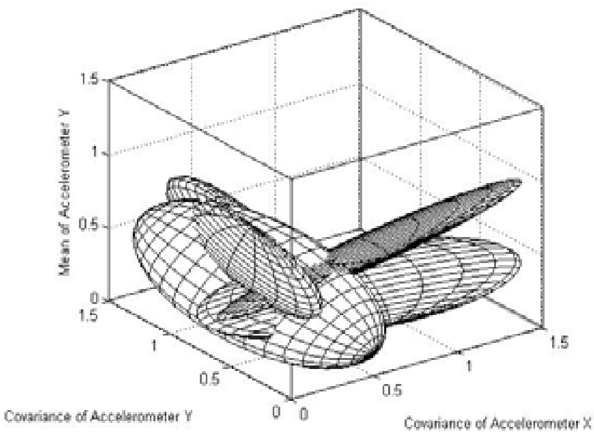

Figure 5 – **The Templates representing the wrist shaking primitives.**

In this work, by defining each cluster as one motion primitive, the spaces of states and observation are superimposed. Hence, the sequence of the human wrist movements can be derived from selecting the symbols with the maximum fuzzy measures. The sequence of observation symbols or the motion primitives are illustrated in Figure 6.

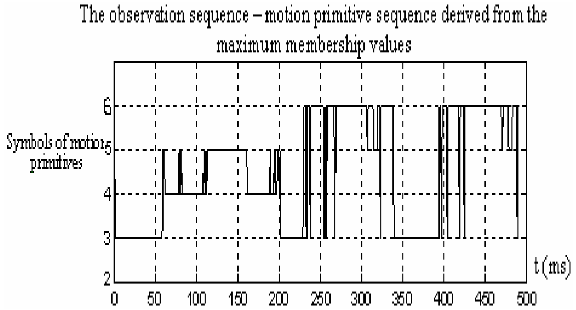

Figure 6 – **Sequence of observation symbols**

In this diagram, the numbers 3, 4, 5, and 6 represent the four different motion primitives respectively. Due to transients, it is rather difficult to confidently assign a motion primitive to a particular segment of the motion,

This result can be improved by applying the fuzzy Viterbi algorithm. Based on the emission and transition matrices calculated by (12) and (13), the recognition result derived from the Fuzzy Hidden Markov Model is significantly enhanced as shown in Figure 7.

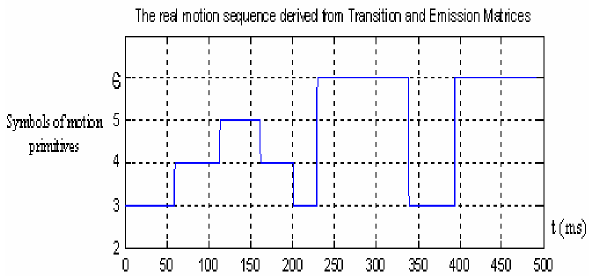

Figure 7 – **Sequence of motion primitives**

Comparison between the Figures 6 and 7 reveals that the Fuzzy HMM has greatly improved the recognition rate of the motion identification system by removing the transients. Furthermore, the transition matrix $\hat{A} = \left[\hat{a}_{ij}\right]$ has effectively described another property of the sequence of the human wrist movements – the grade of certainty of transitions among different motion primitives. In this case, the fuzzy transition matrix is:

$$\hat{A} = \begin{bmatrix} 0.9716 & 0.0071 & 0 & 0.0213 \\ 0.0108 & 0.9785 & 0.0108 & 0 \\ 0 & 0.0192 & 0.9423 & 0.0385 \\ 0.0098 & 0 & 0.0098 & 0.9804 \end{bmatrix}.$$

5. Conclusion

The design and development of a novel Hidden Markov Model for deriving the human wrist motion sequence from the fuzzy clustering results was reported. A new approach for calculating transition and emission matrices was introduced. The observation symbols with fuzzy measures can be applied in the stochastic model to calculate the real motion primitive sequence. The approach was applied to the human wrist motion to identify the sequence of motion, The results show that the proposed algorithm has a performance superior to normal HMM when applied to the detection of them human wrist motion.

Reference

[1] A. J. Harrow, *A Taxonomy of the Psychomotor Domain*, David McKay Company, p.16. Inc, New York 1972.

[2] Moon, S. Hwang, J.N., 1995, 'Noisy speech recognition using robust inversion of hidden Markov models', International Conference Acoustics, Speech, and Signal Processing, vol: 1, pp: 145 – 148.

[3] Bilmes, J.A., 1999, 'Buried Markov models for speech recognition', *Processing of 1999 IEEE International Conference on Acoustics, Speech, and Signal (ICASSP '99).* vol: 2 , pp:713 – 716.

[4] Jacob, B. & Senac, C., 1996, 'Hidden Markov models merging acoustic and articulatory information to automatic speech recognition', *Proceedings of Fourth International Conference on Spoken Language (ICSLP 96),* vol: 4, pp: 2313 – 2315.

[5] Hannaford, B. & Lee,P., 1991, 'Hidden Markov Model analysis of force/torque information in telemanipulation', *Int. J. Robotics Res,* vol: 10, no: 5, pp: 528-539.

[6] Zhu, Q., 1991, 'Hidden Markov model for dynamic obstacle avoidance of mobile robot navigation', *IEEE Trans. Robotics Automat,* vol: 7, no: 3, pp: 390-397.

[7] Yang, J. Xu, Y. Chen, C.S., 1993, 'Hidden Markov model approach to skill learning and its application in telerobotics', *Proceedings of 1993 IEEE International Conference on Robotics and Automation,* vol.1, pp:396 – 402.

[8] Inamura, T. Nakamura, Y. Toshima, I. & Ezaki, H., 2001, 'Mimesis embodiment and proto-symbol acquisition for humanoids'*, Proceedings of International Conference on Advanced Intelligent Mechatronics (2001 IEEE/ASME),* vol: 1, pp: 159 - 164.

[9] Amit, R. & Matari, M., 2002, 'Learning movement sequences from demonstration', *Proceedings of the 2nd International Conference on Development and Learning,* pp. 203 – 208.

[10] Hovland, G.E. Sikka, P. & McCarragher, B.J., 1996, 'Skill acquisition from human demonstration using a hidden Markov model', *Proceedings of IEEE International Conference on Robotics and Automation,* vol: 3, pp. 2706 – 2711.

[11] Bezdek, J.C., 1981, *Pattern Recognition with Fuzzy Objective Function Algorithm,* New York: Plenum.

[12] Definitions of Hidden Markov Model (HMM) [Online]. (No date). Available: http://www.tu-chemnitz.de/physik/KSND/abb/ node55.html [Accessed 2004, Aug. 11].

[13] Rabiner, LR ., 1989, 'A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition', *Proc. of the IEEE*,vo1:77, no:2, pp:257—286.

[14] Mohamed, M.A., & Gader, P., 2000, 'Generalized hidden Markov models. I. Theoretical frameworks', *IEEE Transactions on Fuzzy Systems*, vol: 8, no: 1, pp:67-81.

[15] Rahman, M.M. Nakamura, K. & Ishikawa, S., 2002, 'Recognizing human behavior using universal eigenspace', *Proceedings of 16th International Conference on Pattern Recognition,* vol: 1, pp: 295 – 298.