

Dynamic Hand Gesture Recognition : A Literature Review

Deepali N. Kakade

*Department of Electronics, Bharati
Vidhyapeeth University, Pune*

Prof. Dr. J.S. Chitode

*Department of Electronics, Bharati
Vidhyapeeth University, Pune*

Abstract

Vision based real time gesture recognition system received a great attention in recent years because of its manifoldness applications and the ability to interact with system efficiently through human computer interaction. In this paper, a review of recent hand gesture recognition systems is presented. This paper includes a brief review on camera interface. Image processing, hand gestures, colour detection and hand gesture recognition. Advantages and drawbacks of the system are mentioned finally.

1. Introduction

The essential aim of building hand gesture recognition system is to create a natural interaction between human and computer where the recognition gestures can be used for controlling a robot or conveying meaningful information [1]. The aim of this technique is the proposal of a real time vision system for its application within visual interaction environments through hand gesture recognition, using general purpose hardware and low cost sensors, like a simple personal computer and an USB Webcam, so any user could make use of it in his/her office or home [2].

The development in computer vision makes it possible to approach the interface problems from a human perspective, making the communication between computer and humans more natural [6]. When we as humans communicate, we use our voice and parts of our body such as our face and arms in making gestures. Recently, many attempts have been made to create system that through computer vision could be able to understand gestures [6].

There are two main characteristics that should be deemed when designing a HCI system as mentioned in [3] : Functionality and Usability. System functionality referred to the set of functions or services that the system equips to the users [3],

while system usability referred to the level and scope that the system can operate and perform specific user purposes efficiently [3]. The system that attains a suitable balance between these concepts considered as influential performance and powerful system [3].

Compared to many existing interfaces, hand gestures have the advantages of being easy to use, natural and intuitive. Successful applications of hand gesture recognition include computer game control, human robot interaction and sign language recognition to name a few [4].

The research on visual interpretation of hand gestures for Human-computer Interaction has been a topic of research for many years. The main attempt of the HCI is to communicate naturally with the machines. On similar lines, there is a tremendous progress in speech recognition.

To exploit the use of gestures in HCI, it is necessary to provide the means by which they can be interpreted by the computers [7]. The HCI requires a static/dynamic movements of human hand as an input. Apart from human hand, HCI can pick up the gesture by the movement of human arm, face or even other parts of the human body. Initially, attempts for this HCI using hand gestures required to use hand gloves. These were known as Glove-Based Devices [7][8][9][10]. Glove based gestural interfaces require the user to wear a cumbersome device and generally carry a load of cables that connect the device to a computer. This hinders the ease and naturalness with which the user can interact with the computer controlled environment[7].

Eventually the awkwardness and unnaturalness of using gloves was overcome with an advent in HCI where the video based non contact interaction techniques required a video cameras or the computer visions as an input device which would take human hand gestures as an input. These vision based non contact interaction made the gesture inputs more natural.

2. System components

A low cost computer vision system that can be executed in a common personal computer equipped with an USB webcam is one of the main component of the system. The system should be able to work under different degrees of scene background complexity and illumination conditions.

3. Hand gestures

Outside the HCI framework, hand gestures cannot be easily defined. The definition if they exist, are particularly related to communicational aspect of human hand and body movement. Gestures are expressive, meaningful body motions involving physical movements of the fingers, hands, arms, head, face or body with the intent of : 1. Conveying meaningful information of 2. Interacting with the environment[11]. However in the domain of HCI, the notion of gestures is somewhat different. In a computer controlled environment, one wants to use the human hand to perform tasks that mimic both the natural use of the hand as a manipulation and its use in human machine communication (Control of Computer/Machine functions through gestures.)[12]

All hand movements are classified into two major classes:

- Gestures
- Unintentional Movements

The unintentional movements must be differentiated from the gestures to avoid the meaningless output from the system as the unintentional movements do not convey meaningful information. [7]

Gestures can be static (the users assume a certain pose or configuration) or dynamic. Some gestures also have both static and dynamic elements as in sign language.[11]

3.1. Modelling of gesture

A more appealing approach suitable to real time computer vision lies in the use of simple 3D geometric structures to model in human body [13]. The 3D hand and arm models have often been a choice for hand gesture modelling. They can be classified in two large groups:

- Volumetric Models
- Skeletal Models

Volumetric models are meant to describe the 3D visual appearance of the human hand and arms. They are commonly found in the field of computer animation but recently also have been used in computer vision application.[14]

While static hand gestures are modelled in terms of hand configuration, as defined by the flex angles of the fingers and palm orientation, dynamic hand gestures include hand trajectories and orientation in addition to these. So, appropriate interpretation of dynamic gestures on the basis of hand movement in addition to shape and position is necessary for recognition [5]

4. Image preprocessing

Hand gesture inputs is taken from the web cam and stored in visual memory which is created in a start up step.

A frame from the web cam is captured and each frame is processed separately before its analysis [2]. Two generally sequential tasks are involved in the analysis. The first task involves “detecting” or extracting relevant image features from raw image or image sequence. The second task uses these image features for computing the model parameters [7]. In detection process, it is first necessary to localize the gesturer. Once the gesturer is localized; the desired set of features can be detected.

In dynamic hand gesture recognition, after tracking the hand motion from the gesture video sequence, subsequently, trajectory is estimated through which the hand moves during gesticulation [5]. While trajectory estimation is quite simple and straight forward in Glove based hand gesture recognition system [16] [17] that provide spatial information directly, trajectory estimation in vision based system may require to apply complex algorithm to track hand and fingers using silhouettes and edges.

4.1. Localization

Gesture localization is a process in which the person who is performing the gestures is extracted from the rest of the visual image.

4.2. Segmentation

Segmentation is the next process for recognizing hand gestures. It is the process of dividing the input image (in this case hand gesture) into regions separated by boundaries. The segmentation process depends on the type of gesture if it is dynamic gesture then the hand gesture need to be located and tracked, if it is static gesture (posture), the input image have to be segmented only.

5. Colour detection

[18] Skin colour detection plays an important role in a wide range of image processing applications ranging from face detection, face tracking, gesture analysis, content-based image retrieval (CBIR) systems and to various human computer interaction

domains. The common helpful cue used segmenting the hand is the skin colour. Skin colour detection in visible spectrum can be a very challenging task as the skin colour in an image is sensitive to various factors such as :

- Illumination
- Camera Characteristics
- Ethnicity
- Individual Characteristics of the gesturer
- Other factors like background colours, shadows & motions also influence skin colour appearance.

The major drawback of colour based localization techniques is the variability of the skin colour footprint in different lighting conditions. This frequently results in undetected skin regions or falsely detected non skin textures. The common solution to this problem is the use of restrictive backgrounds and clothing [7].

It has been demonstrated [19] that regardless of ethnicity, the chrominance of skin is quiet consistent. The major difference between skin tones is intensity, as dark skin people have greater skin saturation then light skin people [20].

[18] The HSV space defines colour as hue, saturation and brightness (Lightness or value). The transformation of RGB to HSV is invariant to high intensity at white lights, ambient light and surface orientations relative to the light source and hence can form a very good choice for skin detection methods. [20] However, the YUV colour system is employed for separating chrominance and intensity. The symbol 'Y' denotes intensity while the 'UV' components specify chrominance components. Human skin colours are distributed over a very small region in the chrominance plane.

6. Blobs analysis

Blobs, binary linked objects are groups of pixels that share the same label due to their connectivity in a binary image [2]. After a blob analysis, all those pixels that belong to a same object share a unique label so that every blob can be identified with a label.

7. Gesture recognition

Good segmentation process needs to leads to perfect feature extraction process and the later plays an important role in a successful recognition process [3]. Gesture recognition is an ideal example of multi-disciplinary research. There are different tools for gesture recognition, based on approaches ranging from statistical modelling, computer vision and pattern recognition, image processing etc. Most of the problems have been addressed based on statistical model such as HMM's ,ANN's etc[18]. We are going to see some learning algorithms in brief.

7.1 Learning Algorithms

Here we describe the use of three of the most common learning algorithms used to recognize hand gestures. These algorithms all stem from the artificial intelligence community, and their common trait is that recognition accuracy can be increased through training.[24]

7.1.1 Neural Networks

This section presents a brief introduction into the concepts involved in neural networks.[25] A neural network is an information processing system loosely based on the operation of neurons in the brain. While the neuron acts as the fundamental functional unit of the brain, the neural network uses the node as its fundamental unit; the nodes are connected by links, and the links have an associated weight that can act as a storage mechanism. Each node is considered a single computational unit containing two components. The first component is the input function which computes the weighted sum of its input values; the second is the activation function, which transforms the weighted sum into a final output value. Many different activation functions can be used; the step, sign, and sigmoid functions being quite common. Since they are all relatively simple to use. For example, using the step function, if the weighted sum is above a certain threshold, the function outputs a one indicating the node has "fired" otherwise it outputs a zero indicating the node has not fired. The other two activation functions act in a similar manner. Neural networks generally have two basic structures or topologies, a feed-forward structure and a recurrent structure. A feed-forward network can be considered a directed acyclic graph, while a recurrent network has an arbitrary topology. The recurrent network has the advantage over a feed-forward network in that it can model systems with state transitions Training is an important issue in neural networks and can be classified in two different ways. First, supervised learning trains the network by providing matching input and output patterns; this trains the network in advance and as a result the network does not learn while it is running. The second learning mechanism is unsupervised learning or self-organization which trains the network to respond to clusters of patterns within the input. There is no training in advance and the system must develop its own representation of the input, since no matching output is provided. Neural networks are a useful method for recognizing gestures, yield increased accuracy conditioned upon network training, and work for both glove based and vision-based solutions. However, they have distinct disadvantages. First, different configurations of a given network can give very different results, and it is difficult to determine which configuration is best

without implementing them. Another disadvantage is the considerable time involved in training the network. Finally, the whole network must be retrained in order to incorporate a new gesture. If the gesture set is known beforehand this is not an issue, but if gestures are likely to change dynamically as the system develops, a neural network is probably not appropriate.

Strengths:

1. Can be used in either a vision- or glove-based solution
2. Can recognize large posture or gesture sets
3. With adequate training, high accuracy can be achieved

Weaknesses:

1. Network training can be very time consuming and does not guarantee good results
2. Requires retraining of the entire network if hand postures or gestures are added or removed.

7.1.2 Hidden Markov models

In the Markov model, the state sequence is observable. The output observable event in any given state is deterministic, not random. This will be too constraining when we use it to model the stochastic nature of the human performance, which is related to doubly stochastic processes, namely human mental states (hidden) and human actions (observable). It is necessary that the Observable event is a probabilistic function of the HMM is a representation of a Markov process and is a doubly embedded stochastic process with an underlying stochastic process that cannot be directly observed, but can only be observed through another set of stochastic processes that produce the sequence of observable symbols .

We define the elements of an HMM as follows. N is the number of states in the model. The state of the model at time t is $q_t=1, 1 \leq q_t \leq N$ and $1 \leq t \leq T$. where T is the length of the output observable symbol sequence M is the size of the codebook or the number of distinct observable symbols for each state at time t , then $0 \leq o_t \leq M - 1$. π_N is an N -element vector indicates the initial state probability. $\pi_N = \{\pi_i\}$. Where $\pi_i = p(q_t = i)$, $1 \leq i \leq N$. $A_{N \times N} = \{a_{ij}\}$, where $a_{ij} = P(q_t = j / q_{t-1} = i)$, $1 \leq i, j \leq N$ and $a \geq 0$, $\sum_{j=1}^N a_{ij} = 1$.

$B_{M \times N}$ is an $M \times N$ matrix specifying that the system will generate the observable symbol o_t at state j and at time t . $B_{M \times N} = \{b_j\}$, where $b_j = P(o_t = o_t / q_t = j)$, where $1 \leq i \leq N$, $0 \leq o_t \leq M - 1$ and $b_j(o_t) \geq 0$, $\sum_{o_t=0}^{M-1} b_j(o_t) = 1$.

The complete parameter set λ of the discrete HMM is represented by one vector π and two matrices A and B . To accurately describe a real-world process such as gesture with an HMM, we need to appropriately select the HMM parameters. The parameter selection process is called the HMM 'training.' This parameter

set can be used to evaluate the probability $P(O/\lambda)$; that is to measure the maximum likelihood performance of an output observable symbol sequence O ; where T is the number of frames for each image sequence. For evaluating each $P(O/\lambda)$; we need to select the number of states N ; the number of observable symbols M (the size of codebook), and then compute the results of probability density vector π and matrices A and B by training each HMM from a set of corresponding training data after VQ.

There are three basic problems in HMM design:

1. Probability evaluation: How do we efficiently evaluate $P(O/\lambda)$; the probability (or likelihood) of an output observable symbol sequence O given an HMM parameter set λ .

Probability evaluation using the forward-backward procedure. We compute the output probability $P(O/\lambda)$ with which the HMM will generate an output observable symbol sequence $O = \{o_1, o_2, \dots, o_t\}$; given the parameter set $\lambda = \{\pi, A, B\}$. The most straightforward way to compute this is by enumerating every possible state sequence of length T ; so there will be N^T possible combinations of state sequence where N is the total number of states.

2. Optimal state sequence: How do we determine an optimal state sequence $q = \{q_1, q_2, \dots, q_t\}$; which is associated with the given output observable symbol sequence O ; by given an HMM parameter set λ .

Optimal state sequence using the viterbi algorithm: We use a dynamic programming method called the

Viterbi algorithm to find the single best state sequence $q = (q_1, q_2, \dots, q_T)$ (or the most likely path) given the observable symbol sequence $O = \{o_1, o_2, \dots, o_t\}$; and the HMM parameter set λ in order to maximize $P(q/\lambda, O)$ Since

$$P(q/\lambda, O) = \frac{P(q, O)}{P(O/\lambda)}$$

Maximizing $P(q/\lambda, O)$ is equivalent to maximizing $P(q/\lambda)$ using the Viterbi algorithm.

3. Parameter Estimation. How do we regulate an HMM parameter set λ to maximize the output probability $P(O/\lambda)$ of generating the output observable symbol sequence.

Parameter estimation using the baum-welch method: We can use a set of training observable symbol sequences to adjust the model parameters in order to build a signal model that can be used to identify or recognize other sequences of observable symbols. There is, however, no efficient way to optimize the model parameter set that globally maximizes the probability of the symbol sequence. Therefore, the Baum-Welch method is used to choose the maximum likelihood model parameter set

$\lambda = \{\pi, A, B\}$ such that its likelihood function $P(O/\lambda)$ is locally maximized using an iterative procedure.

Strengths:

1. Can be used in either a vision- or glove-based solution
2. Can recognize large posture or gesture sets
3. With adequate training, high accuracy can be achieved
4. Well discussed in the literature

Weaknesses:

1. Training can be time consuming and does not guarantee good results
2. As with multi-level neural networks, the hidden nature of HMMs makes it difficult to observe their internal behavior

7.1.3 Instance-Based Learning

Instance-based learning is another recognition technique that stems from work done in machine learning. The main difference between instance-based learning and other learning algorithms such as neural networks and hidden Markov models is the way in which the training data is used. With supervised neural networks, for example the training data is passed through the network and the weights at various nodes are updated to fit the training set. With instance-based learning, the training data is simply used as a database in which to classify other "instances". An instance, in general, is a vector of features of the entity to be classified. For instance, in posture and gesture recognition, a feature vector might be the position and orientation of the hand and the bend values for each of the fingers.

Instance-based learning methods include techniques that represent instances as points in Euclidean space, such as the K-Nearest Neighbor algorithm, and techniques in which instances have a more symbolic representation, such as case-based reasoning[24].

In the K-Nearest Neighbor algorithm, an instance is a feature vector of size n with points in n -dimensional space. The training phase of the algorithm involves storing a set of representative instances in a list of training examples. For each new record, the Euclidean distance is computed from each instance in the training example list, and the K closest instances to the new instance are returned. The new instance is then classified and added to the training example list so that training can be continuous. In the case of hand posture recognition, the training set would be divided into a number of categories based on the types of recognizable postures. As a new posture instance is entered, its K nearest neighbors are found and used to determine the category in which the instance should be placed (thus recognizing the instance as a particular posture).

Another type of instance-based learning technique is case-based reasoning, in which instances have more elaborate descriptions

Instance-based learning techniques have the advantage of simplicity, but they have a number of

disadvantages as well. One major disadvantage is the cost of classifying new instances. Another disadvantage of these methods is that not all of the training examples may fit in main memory, and thus will also increase response time. Unfortunately, very little work has been done on instance-based learning in recognizing hand postures and gestures. More research is needed to determine whether the technique can be applied to hand gestures and if the accuracy can be improved.

Strengths

1. Except for case-based reasoning, instance-based learning techniques are relatively simple to implement
2. Can recognize a large set of hand postures with moderately high accuracy
3. Provides continuous training

Weaknesses

1. Requires a large amount of primary memory as training set increases
2. Response time issues may arise due to a large amount of computation at instance classification time.
3. Only a little reported in the literature on using instance-based learning with hand postures and gestures.

7.2. Model evaluation process

In this process, a gesture trajectory is tested over the set of trained HMM's in order to decide which gesture it belongs to from the database, however, the larger the number of trained HMM's (gestures), the more the computationally demanding the recognition procedure.

8. Applications of hand gesture

Gesture recognition has wide range in applications such as following: [18]

- Developing aids for the hearing impaired
- Enabling very young children to interact with computer
- Designing techniques for forensic identification
- Recognizing sign language
- Medically monitoring patients
- Navigating and/or manipulating virtual environments
- Communicating in video conferencing
- Distance learning / tele-teaching assistance
- Graphic editor control

9. Drawbacks

Although there are numerous advantages of dynamic hand gesture recognition system, there are

few drawbacks or limitations to it which are as follows:

- The number of cameras used
- Their speed and latency
- Structure of environment (Restrictions such as lighting or speed of movement)
- Any user requirements (Whether user must wear anything special)
- The low level features used (Edges, regions, Silhouettes moments, histograms)
- Whether 2D or 3D representation is used.

There is however an inherent loss of information whenever 3D image is projected to a 2D plane. A tracker also needs to handle changing shapes and sizes of the gesture generating object, other moving objects in the background.

10. Conclusion and discussion

This survey has provided a comprehensive overview of hand gesture recognition techniques. Hand gestures are an interesting interaction insight in a variety of computer applications. Two principal questions must be answered when using them: The first question is what technology to use for collecting raw data from the hand. Generally, two types of technologies are available for collecting this raw data. The first one is a glove input device and the second way of collecting raw data is to use computer vision. Accuracy of a glove input device depends on the type of sensor technology used; usually, the more accurate the glove is, the more expensive it is. In a vision-based solution, one or more cameras placed in the environment record hand movement. Both types of solutions have their own advantages and disadvantages, and the question that which solution to use is a difficult one. However, when using a hand posture or gesture-based interface, the user does not want to wear the device and be physically attached to the computer. If vision-based solutions can overcome some of their difficulties and disadvantages, they appear to be the best choice for raw data collection.

The second question to be answered when using hand gestures is what recognition technique will maximize accuracy and robustness. A number of recognition techniques are available and in some cases, where hand gesture recognition is to be used, the possibilities are narrowed down, as some algorithms are best fitted only for gesture recognition. This review has categorized these techniques into three broad categories: Feature extraction, statistics, and model learning algorithms.

There are a number of interesting areas for future research in hand gesture recognition. The field is not yet very matured – we have a long way to go before this type of concept is robust enough to be seen in commercial, main stream applications. Research into

better hardware for data collection is important and HMM technique for gesture recognition was found to be more accurate compared to neural networks and instance based model. This survey has provided a comprehensive overview of hand gesture recognition techniques.

11. Bibliography

- [1] Pragati Garg, Naveen Aggarwal and Sanjeev Sofat, "Vision Based Hand Gesture Recognition" World Academy of Science, Engineering and Technology 49 2009
- [2] William T. Freeman, P. A. Beardsley, H. Kage, K. Tanaka, K. Kyuma, C. D. Weissman, "Computer vision for computer interaction" mitsubishi electric research laboratories <http://www.merl.com> TR99-36 October 1999
- [3] Fakhreddine Karray, Milad Alemzadeh, Jamil Abou Saleh and Mo Nours Arab, "Human-Computer Interaction: Overview on State of the Art" Pattern Analysis and Machine Intelligence Lab., Department of Electrical and Computer Engineering University of Waterloo, Waterloo, Canada international journal on smart sensing and intelligent systems, vol. 1, no. 1, march 2008
- [4] Mr. Chetan A. Burande, Prof. Raju M. Tugnayat, Prof. Dr. Nitin K. Choudhary, "Advanced Recognition Techniques for Human Computer Interaction" Information Technology Dept. Jawaharlal Darda Institute of Engg & Tech Yavatmal, India & Shri Bhagavati College of Engg. Nagpur, India
- [5] M.K. Bhuyan, D. Ghosh and P.K. Bora, "Feature Extraction from 2D Gesture Trajectory in Dynamic Hand Gesture Recognition" Department of Electronics and Communication Engineering Indian Institute of Technology Guwahati, India 781039
- [6] Henrik Birk, Thomas B. Meslund, Claus B. Madsen, "Real Time Recognition of Hand Alphabet Gestures Using Principal Component Analysis" Aalborg University, Laboratory of Image Analysis, Fredrik Bajersvej 7, Building D-1 DK-9220 Aalborg East email: [tbm,cbm]vision.auc.dk
- [7] Vladimir I. Pavlovic, Rajeev Sharma, Thomas S. Huang, "Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review" IEEE transactions on pattern analysis and machine intelligence, vol. 19, no. 7, July 1997 Page 677 to 695
- [8] Mohamed Alsheakhali, Ahmed Skaik, Mohammed Aldahdouh, Mahmoud Alhelou, "Hand Gesture Recognition System" Computer Engineering Department, The Islamic University of Gaza Gaza Strip, Palestine, 2011
- [9] S. Sidney Fels and Geoffrey E. Hinton, "Glove-Talk: A Neural Network Interface Between a Data-Glove and a Speech Synthesizer" IEEE Transactions on Neural Networks, Vol. 3, No. 6, November 1992 Page 1 to 7
- [10] M. Ali Qureshi, Abdul Aziz, Muhammad Ammar Saeed, Muhammad Hayat, "Implementation of an Efficient Algorithm for Human Hand Gesture Identification" Dept. Electronic Engineering, University College of Engineering & Technology, The Islamia University of Bahawalpur, Bahawalpur Pakistan
- [11] Sushmita Mitra, Tinku Acharya, "Gesture Recognition: A Survey" IEEE transactions on systems, man, and cybernetics—part c: applications and reviews, vol. 37, no. 3, May 2007 page 311 to 324
- [12] Sanshzar Kettebekov, Rajeev Sharma, "Toward Natural Gesture/Speech Control of a Large Display"

Department of Computer Science and Engineering
 Pennsylvania State University 220 Pond Laboratory,
 University Park PA 16802 Engineering for Human-
 Computer Interaction (EHCI'01). Lecture Notes in
 Computer Science, Springer Verlag, 2001

[13] Satoshi be, Kozi akamura, Mamoru Maekawa,
 Takanobu Endo, Nobuyuki sugiura, "Full-Passive Human
 Recognition from Image Sequences" Department of
 Information Science, Faculty of Science, The University of
 Tokyo 7-3-1 Hongo, Bunkyo-ku, Tokyo, 1 13 Japan and
 Graduate School of Information Systems, University of
 Electro-Communications 1-5-1 Chofugaoka, Chofu-shi,
 Tokyo, 182 Japan MVA '92 IAPR Workshop on Machine
 Vision Applications Dec. 7-9.1992, Tokyo

[14] Nadia Magnenat Thalmann, Daniel Thalmann,
 "Computer Animation" MIRALab, University of Geneva
 Geneva, Switzerland and Swiss Federal Institute of
 Technology (EPFL) Lausanne, Switzerland

[15] Rafiqul Zaman Khan, Noor Adnan Ibraheem, "hand
 gesture recognition: a literature review" department of
 computer science, a.m.u. aligarh, india international journal
 of artificial intelligence & applications (ijaia), vol.3, no.4,
 july 2012 page 161 to 174

[16] M.K. Bhuyan, P.K. Bora, and D. Ghosh, "Trajectory
 Guided Recognition of Hand Gestures having only Global
 Motions" World Academy of Science, Engineering and
 Technology 21 2008 Page 832 to 843

[17] David J. Sturman, David Zeltzer, "A Survey of Glove-
 based Input" Medialab and MIT Research Lab of
 Electronics Page 30 to 39

[18] P. Kakumanu, S. Makrogiannis, N. Bourbakis, "A
 survey of skin-colour modeling and detection methods"
 ITRI/Department of Computer Science and Engineering,
 Wright State University, Dayton OH 45435, USA Pattern
 Recognition 40 (2007) 1106 – 1122 Published by Elsevier
 Ltd on behalf of Pattern Recognition Society.
 doi:10.1016/j.patcog.2006.06.010

[19] Zhenyao Mo, J. P. Lewis, Ulrich Neumann,
 "SmartCanvas: A Gesture-Driven Intelligent Drawing Desk

System" CGIT Lab, USC 3740 McClintock Ave, EEB131
 Los Angeles, CA 90089 USA

[20] Nianjun Liu, Brian C. Lovell, Peter J. Kootsookos,
 "Evaluation of HMM Training Algorithms for Letter Hand
 Gesture Recognition Intelligent real time imaging and
 sensing (IRIS) Group School of Information Technology
 and Electrical Engineering The University of Queensland,
 Brisbane, Australia 4072

[21] James Davis, Mubarak Shah, " Recognizing Hand
 Gestures" Computer Vision Laboratory University of
 Central Florida, Orlando FL 32816 USA To appear in
 ECCV-94, Stockholm, Sweden, May 2-6, 1994

[22] Mario Ganzeboom, "How hand gestures are
 recognized using a dataglove" Human Media Interaction
 MSc University of Twente The Netherlands
m.s.ganzeboom@student.utwente.nl

[23] Christopher Lee, Yangsheng Xu, "Online, Interactive
 Learning of Gestures for Human / Robot Interfaces" The
 Robotics Institute Carnegie Mellon University Pittsburgh
 Pennsylvania 15213 USA

[24] Joseph J. La viola Jr. "A survey of hand posture and
 gesture recognition techniques and technology" Department
 of computer science Brown University Providence, Rhode
 Island CS-99-11 June 1999.

[25] carlos gershenson "Artificial neural network for
 beginners-" (C.gershenson@sussex.ac.uk)

[26] Feng-Sheng Chen, Chih-Ming Fu, Chung-Lin Huang
 "Hand gesture recognition using a real-time tracking
 method and hidden Markov models" Institute of Electrical
 Engineering, National Tsing Hua University, Hsin Chu 300,
 Taiwan, ROC Received 15 January 2001; received in
 revised form 2 January 2003; accepted 20 March 2003