ISSN: 2348 9510
International Journal Of Core Engineering & Management (IJCEM)
Volume 2, Issue 4, July 2015

Hand Posture Detection: Principle, Need and Algorithms

Manjinder Kaur
M.Tech Scholar, Department of CSE
GIMET, Amritsar, Punjab, India
mkwahla@gmail.com

Prabhjit Singh
AP, Department of CSE
GIMET, Amritsar, Punjab, India
prabhjitsingh85@gmail.com

Abstract

Hand posture identification research has gained much attention because of its applications for interactive human–machine interface and virtual environments. Hand postures are frequently used as intuitive and convenient communications in our daily life, and the recognition of hand postures can be widely applied in human computer interfaces, robot control, and augmented reality, etc. Therefore an efficient, fast and reliable hand posture system is the need for today. In this paper we present the study, need type and various algorithms of a hand posture identification system.

Keywords— ***Hand Posture, Point matching, SIFT, ASL***

I. Introduction

Hand posture identification system, is very useful for deaf & dumb people. It is also used to replaces input devices like keyboard and mouse with static and dynamic hand postures, for interactive computer applications Research on hand postures can be classified into three categories. The first category, glove based analysis, employs sensors (mechanical or optical) attached to a glove that transducer finger flexions into electrical signals for determining the hand posture. The relative position of the hand is determined by an additional sensor. This sensor is normally a magnetic or an acoustic sensor attached to the glove. For some data-glove applications, look-up table software toolkits are provided with the glove to be used for hand posture Identification.

The second category, vision based analysis, is based on the way human beings perceive information about their surroundings, yet it is probably the most difficult to implement in a satisfactory way. Several different approaches have been tested so far. One is to build a three-dimensional model of the human hand. The model is matched to images of the hand by one or more cameras, and parameters corresponding to palm orientation and joint angles are estimated. These parameters are then used to perform posture classification. A hand posture analysis system based on a three-dimensional hand skeleton model with 27 degrees of freedom was developed by Lee and Kunii. They incorporated five major constraints based on the human hand kinematics to reduce the model parameter space search. To simplify the model matching, specially marked gloves were used. The third category, analysis of drawing postures, usually involves the use of a stylus as an input device. Analysis of drawing postures can also lead to Identification of written text. The vast majority of hand posture Identification work has used mechanical sensing, most often for direct manipulation of a virtual environment and occasionally for symbolic communication [1, 2]. Sensing the hand posture mechanically has a range of problems, however, including reliability, accuracy and electromagnetic noise. Visual sensing has the potential to make postural interaction more practical, but potentially embodies some of the most difficult problems in machine vision. The hand is a non-rigid object and even worse self-occlusion is very usual.

II. Colours & Shapes

Skin colour is an important image feature to localize and track human hands. However, colour-based algorithms face the difficult task of distinguishing objects which have the similar colour with the hand such as human arm and face. In order to solve this problem, users are often required to wear long sleeve shirts and restrictions are imposed on the colours of other objects in the observed scene. Colour based algorithms are also very sensitive to lighting variations. When the lighting does not meet the special requirements, colour based algorithms usually fail. To solve the problem of colour distributions change under different lighting conditions. Manresa implemented a real-time system that aims for the control of a video game based on hand posture recognition. Their system is based on three main steps: hand segmentation, hand tracking and posture recognition.

ISSN: 2348 9510
International Journal Of Core Engineering & Management (IJCEM)
Volume 2, Issue 4, July 2015

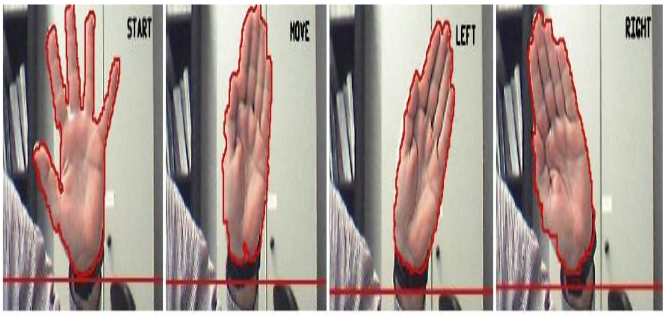

Figure 1: Angles of hand

To prevent errors from hand segmentation, they add a second step: hand tracking. Tracking is performed assuming a constant velocity model and using a pixel labelling approach.

Several hand features are extracted and fed to a finite state classifier to identify the hand configuration. The hand can be classified into one of the four posture classes or one of the four movement directions. For shape-based algorithms, global shape descriptors such as Zernike moments and Fourier descriptors are used to represent different hand shapes [3, 4].

Most shape descriptors are pixel-based and the computation cost is usually too high to implement real-time systems. Another disadvantage for shape-based approaches is the requirement for noise-free image segmentation, which is a difficult task for the usually cluttered background images.

III. Hand Features

Hand feature based algorithms extract certain local image features such as fingertips or hand edges, and use some heuristics to find configurations or combinations of these features specific to an individual hand posture. Oka developed an augmented desk interface depending on accurate, real-time hand and fingertip tracking for seamless integration between real objects and associated digital information. They introduce a method for locating finger tip positions in image frames and measuring fingertips trajectories across image frames. By using an infrared camera, their method can track multiple fingertips reliably even on a complex background under changing lighting conditions without invasive devices or color markers. A mechanism for combining direct manipulation and symbolic postures based on multiple fingertip motions was proposed. Zhang presented a vision-based interface system named “Visual Panel”, which employs an arbitrary quadrangle-shaped panel (e.g. an ordinary piece of paper) and a fingertip pointer as an intuitive input device.

ISSN: 2348 9510
International Journal Of Core Engineering & Management (IJCEM)
Volume 2, Issue 4, July 2015

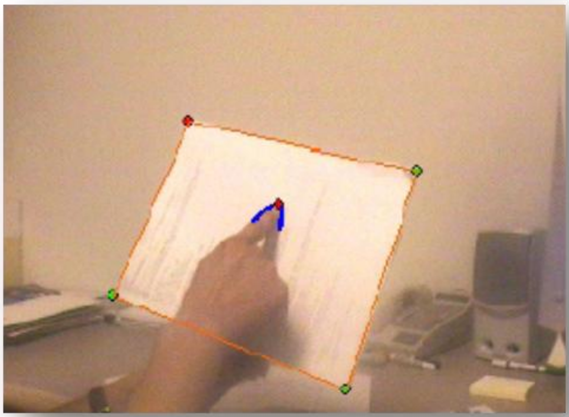

Figure 2: Visual Panel system

The system can accurately and reliably track the panel and the fingertip pointer. By detecting the clicking and dragging hand actions, the system can fulfill many tasks such as controlling a remote large display, and simulating a physical keyboard. Users can naturally use their fingers to issue commands and type text. Furthermore, by tracking the 3D position and orientation of the visual panel, the system can also provide 3D information, serving as a virtual joystick to control 3D virtual objects. Malik designed a plane-based augmented reality system that tracks planar patterns in real-time, onto which virtual 2D and 3D objects can be augmented. Interaction with the virtual objects is possible via a fingertip-based posture recognition system. As illustrated in Figure 3, the basis of the mechanism to capture the posture is the number of detected fingertips.

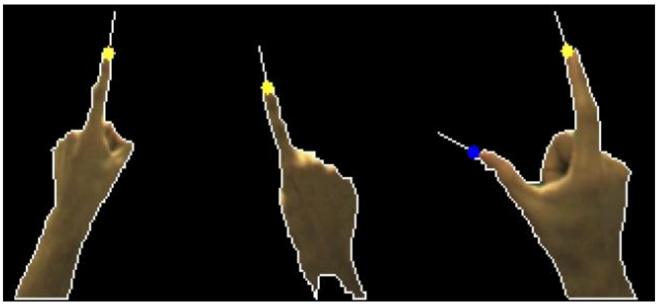

Figure 3: posture recognition based on fingertips

A single detected fingertip represents the posture of pointing, whereas multiple detected fingertips represent the posture of selecting. The fingers of the hand are detected by background subtraction and scanning the binary image for pixels of full intensity. Each time

ISSN: 2348 9510
International Journal Of Core Engineering & Management (IJCEM)
Volume 2, Issue 4, July 2015

such a pixel is found, a subroutine is called to perform a neighborhood flood-fill operation to collect all neighboring pixels. The orientation of the finger can be calculated using the central moments the detected finger bob. The axis line is then defined by forcing it through the blob's centroid. The fingertip location is recovered by finding the farthest point in the blob from the root point, and is used as the pointer location. Huang introduced a model-based hand posture recognition system, which consists of three phases: feature extraction, training, and recognition In the feature extraction phase, a hybrid technique combines hand edges and hand motions information of each frame to extract the feature images. Then, in the training phase, they use the principal component analysis (PCA) to characterize spatial shape variations and the hidden Markov models (HMM) to describe the temporal shape variations. Finally, in recognition phase, with the pre-trained PCA models and HMM, the observation patterns can be generated from the input sequences, and then apply the Viterbi algorithm to identify the posture. For feature-based based approaches, a clean image segmentation is generally a must step to recover the hand features. This is not a trivial task when the background is cluttered [5, 6]. On the other hand, for the highly articulated human hand, it is sometimes difficult to find local hand features and heuristics that can handle the large variety of hand postures. It is not always clear about how to correlate local hand features with different hand postures in an efficient manner. Mean shift Comaniciu proposed a method for real-time tracking of non-rigid objects based on mean shift iterations and found the most probable target position in the current frame. As illustrated in Figure 3.4, the mean shift algorithm is a simple iterative procedure that shifts the center of the region of interest to the center of mass. The data could be visual features such as colors or textures.

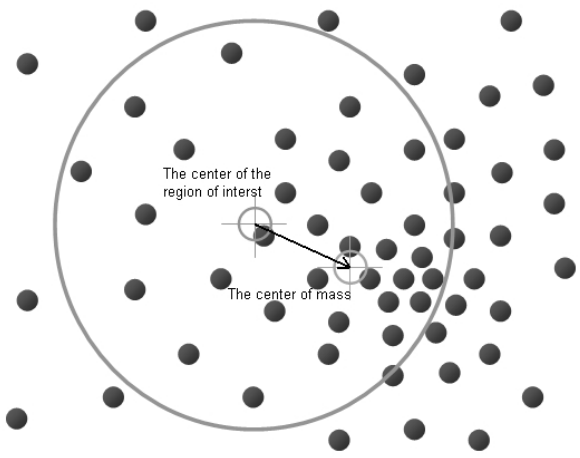

Figure 4: The mean shift algorithm.

ISSN: 2348 9510
International Journal Of Core Engineering & Management (IJCEM)
Volume 2, Issue 4, July 2015

The statistical distributions characterize the object of interest. Bradski proposed a modified mean shift algorithm named Continuously Adaptive Mean Shift (CamShift), which primarily intended to perform efficient head and face tracking in a perceptual user interface. The CamShift algorithm finds the center and the size of the object on a color probability image frame. The probability is created via a histogram model of a specific color. Figure 5 shows some results for hand tracking using this algorithm.

Figure 5: Hand tracking use the Cam Shift algorithm

The tracker moves and resizes the search window until its center converges with the center of mass. Compared with regular mean shift that uses static distributions (i.e. the distributions are not updated unless the target has significant changes in shape, size or color), CamShift can effectively track dynamically changing probability distributions\ in a visual scene. The CamShift algorithm depends on image features such as colored blobs for fast tracking. However, it is a difficult task for the algorithm to effectively classify different hand postures that might have similar center of mass.

IV. Methodology based classification

From the point of view of the methodology used to describe hand postures, vision based hand posture recognition algorithms can be grouped into statistical approaches and syntactic approaches. For statistical approaches, hand postures are represented by a number of features and each hand posture is viewed as a point in the feature space. The goal is to select appropriate features that allow hand postures belonging to different regions in the feature space so that the classifier can identify them correctly. For syntactic approaches, hand postures are decomposed into a set of simpler primitives, and each hand posture is represented by the interrelationships of these primitives using a string, a tree or a graph. The classification is achieved by parsing the corresponding representation according to a given grammar.

ISSN: 2348 9510
International Journal Of Core Engineering & Management (IJCEM)
Volume 2, Issue 4, July 2015

A. Statistical Algorithms

Hand posture recognition systems based on statistical approaches can be decomposed into two components: the training component and the recognition component. In the training component, the feature selection module finds appropriate features to represent the training samples for hand postures and a classifier is obtained by the learning module based on the training samples. The feedback path between the feature selection module and the learning module allows the user to optimize the training process by adjusting the training strategies and parameters. In the recognition component, the feature measurement module extracts the features from the test sample and the trained classifier recognizes them based on the measured features and the corresponding decision rules. There are four different methods to design a classifier:

- Similarity measurement: This is the most intuitive and simplest method. With this approach, patterns that are similar should belong to the same class. Template matching and minimum distance classifiers are typical methods of this category.
- Probabilistic classifiers: The most popular probabilistic classifier is the Bayes classifier that uses the Bayes rule to estimate the conditional probability of a class.
- Geometric approach: The classifier of this category constructs decision boundaries by optimizing certain error criterion. The classic example of this type of classifier is Fisher's linear discriminant, which can project high dimensional data onto a line and performs classification in this one-dimensional space. The projection maximizes the distance between the means of the two classes while minimizing the variance within each class. One example of the application of Fisher's linear discriminant is human face recognition. Human faces include a large number of image features. Fisher's linear discriminant can reduce the large number of features to a more manageable number for easier and more accurate classification. Each of the new dimensions is a linear combination of pixel values, which form a new template.
- Decision tree classifiers: This type of classifiers can deduce the conclusion using iterative multistage decisions based on individual features at each node of the decision tree. The basic idea is to break up a complex decision problem into a series of several simpler decisions, so that the final conclusion would resemble the desired solution. The most popular decision tree classifiers are binary decision tree classifiers, which make true or false decisions at each stage based on the single corresponding feature at the node.

To improve the overall classification accuracy, different classifiers can be combined so that the overall performance can be optimized. A classifier combination can achieve better performance especially when the individual classifiers are largely independent .A typical example is the boosting algorithm, which combines a series of weak classifiers (whose

ISSN: 2348 9510
International Journal Of Core Engineering & Management (IJCEM)
Volume 2, Issue 4, July 2015

accuracies are only slightly better than 50%) into a strong classifier which has a very small error rate on the training data .In the boosting algorithm, individual classifiers are invoked in a linear sequence. The inaccurate but cheap classifiers (low computational cost) are applied first, followed by more accurate and expensive classifiers. The number of mistakenly classified samples is reduced gradually as more individual classifiers have been invoked and added to the sequence. The final strong classifier (whose accuracy meets the requirement) is a linear combination of the invoked individual classifiers. Besides the discussed statistical models, there is a special model need to be introduced: the neural networks, which have become an important tool in computer vision. Neural networks are parallel computing systems consisting of a large number of interconnected neurons (elementary processors), which mimics the complex structure of neurons in human brains. Neural networks are able to learn complex nonlinear input-output relationships using sequential training procedures and adapt themselves to the data. In spite of the seemingly different mechanisms, there is considerable overlap between neural networks and statistical models in the field of pattern recognition. Most of the well known neural network models are implicitly equivalent or similar to classical statistical approaches .Neural networks implement pattern recognition as black boxes by concealing the complex statistics from the user [7, 8].

B. Syntactic Approaches

In many computer vision problems involving complex patterns and activities, statistic approaches and numeric measurements might not be enough to represent the complex structures of these patterns and activities. Under this situation, it is more appropriate and effective to use a syntactic approach to describe these patterns and activities with their simpler sub-patterns and elementary parts .The elementary parts used to syntactically describe a complex pattern or an activity are called primitives [9]. For computer vision applications, the principles to identify the primitives are:

- The number of primitive types should be small.
- The primitives selected must be able to form an appropriate object representation.
- Primitives should be easily segmental from the image.
- Primitives should be easily recognizable using some statistical pattern recognition method.
- Primitives should correspond with significant natural elements of the object structure being described.

ISSN: 2348 9510
International Journal Of Core Engineering & Management (IJCEM)
Volume 2, Issue 4, July 2015

C. Statistical vs. Syntactic

When the patterns and activities under research are complex and include explicit structural information and relationships among sub-patterns and primitives, it is more effective to use a syntactic approach to decompose the complex pattern into their simpler sub-patterns and primitives, which are easier to process. Statistical approaches and syntactic approaches are not absolute disjoint categories .In many situations, statistical approaches and syntactic approaches can complement with each other, and a combination of statistical and syntactic approaches can be used to construct a more efficient hybrid system: the statistical approach is responsible for primitives extraction and identification; the syntactic approach is responsible for analyzing the structural relationship among the identified primitives so that the whole pattern can be recognized. Most current posture recognition systems treat the hand posture as a whole element without considering its hierarchical composite property and breaking it into its simpler constituent components that would be easier to process. This results in a rather slow and inefficient system unsuited for real-time applications. To solve the problem, the advantages brought by syntactic approaches need to be considered [10, 11]. An appropriate combination of statistical and syntactic approaches can result in an efficient and effective hand posture recognition system.

V. Conclusions

This paper gives the overview of hand posture detection system. In this paper, we present the basic idea, need and various algorithms of the hand posture system. Each system has benefits and drawbacks. Every hand posture system is useful for certain application.

References

[1] Liu Yun, Zhang Lifeng, Zhang Shujun ,"A Hand Posture Recognition Method Based on Multi-Feature Fusion and Template Matching", Elsevier journal's procedia engineering,2012.

[2] Siddharth S. Rautaraya, Anupam Agrawala ,"Real Time Posture Recognition System for Interaction in Dynamic Environment", Indian Institute of Information Technology, Allahabad, 21 October 2012.

[3] Helen Cooper, Eng-Jon Ong, Nicolas Pugeault, Richard Bowden,"Sign Language Recognition using Sub-Units",Journal of Machine Learning Research , July 2012.

[4] Ms. Rashmi D. Kyatanavar, Prof. P. R. Futane , "Comparative Study of Sign Language Recognition Systems", Department of Computer Engineering, Sinhgad College of Engineering, Pune, India, Volume 2, Issue 6, June 2012.

ISSN: 2348 9510
International Journal Of Core Engineering & Management (IJCEM)
Volume 2, Issue 4, July 2015

[5] Włodzimierz Kasprzak, Artur Wilkowski , Karol Czapnik∗,”Hand posture recognition based on free–form contours and probabilistic inference”, int. J. Appl. Math. Comput. Sci., vol. 22, no. 2, 437–448, 2012
[6] Siddharth S. Rautaraya, Anupam Agrawal ,“Real Time Posture Recognition System for Interaction in Dynamic Environment” , Elsevier journal, 2011.
[7] Nuwan Gamage a, Ye Chow Kuang a,⇑, Rini Akmeliawati b, Serge Demidenko , “Gaussian Process Dynamical Models for hand posture interpretatio in Sign Language”, 6 September 2011.
[8] Xiaohui Shen , Gang Hua , Lance Williams , Ying Wu , “Dynamic hand posture recognition: An exemplar-based approach from motion divergence fields”, 4 November 2011.
[9] Deng-Yuan Huang , “Gabor filter-based hand-pose angle estimation for hand posture recognition under varying illumination”, Elsevier journal of Expert Systems with Applications,2010.
[10] Zhiquan Feng , ’Features extraction from hand images based on new detection operators”, Elsevier journal, 2010.
[11] Deng-Yuan Huang , Wu-Chih Hub, Sung-Hsiang Chang , “Gabor filter-based hand-pose angle estimation for hand posture recognition under varying illumination”,Department of Electrical Engineering, Dayeh University, 1., Dacun, Changhua 515, Taiwan, 2010