© 2018 IJSRST | Volume 4 | Issue 8 | Print ISSN: 2395-6011 | Online ISSN: 2395-602X
Themed Section: Science and Technology

A Survey on Hand Gesture Using Imageprocessing

Binu Ruby Sunny
PG Scholar, Akshaya College of Engineering and Technology, Coimbatore, TamilNadu, India

ABSTRACT

As technology becomes the part of human life for decades, the relationship between human and computer called human-computer interaction (HCI) is important to study for improving the system to serve the human need. HCI can be applied in various areas including medical system which is valuable for the elder who is not able to walk or express the feelings by words. The intuitive approach is the development of algorithm by using hand gestures. The proposed system called dynamic hand gesture recognition algorithm can be applied for elder people. The algorithm implements in a vision-based hand gesture recognition using optical flow and blob analysis to track six dynamic hand gestures and classify their meanings. The experiment provided good results for all six hand gestures in detection, tracking and classification procedures.

Keywords : Human-Computer Interaction, Hand Gesture Recognition, Digital Color Image, Grayscale Images, Binary Images

I. INTRODUCTION

1.1OBJECTIVE

The main objective is to implement hand gestures using contour technique. The aim is classifying the image segmentation and threshold image for hand gestures and deal with the preprocessing.

1.2 IMAGE PROCESSING

The image processing is process in which a two dimensional image is treated as an input and the specified output image is obtained by setting some parameter over the two dimensional input image.This process would probably start with image processing techniques such as noise removal, followed by (low-level) feature extraction to locate lines, regions and possibly areas with certain textures. The clever bit is to interpret collections of the shapes as a single objects, e.g. cars on a road, boxes on a conveyor belt or cancerous cells on a microscope slide. One reason in the AI problem is that an object can appear very different when viewed from different angles or under different lighting. Another problem is deciding what features belong to what object and which are background or shadows etc.

The human visual system performs these tasks mostly unconsciously but a computer requires skillful programming and lots of processing power to approach human performance. Manipulating data in the form of an image through several possible techniques. An image is usually interpreted as a two-dimensional array of brightness values, and is most familiarly represented by such patterns as those of a photographic print, slide, television screen, or movie screen. An image can be processed optically or digitally with a computer.

1.3 CLASSIFICATION OF IMAGES:

There are 3 types of images used in Digital Image Processing. They are

- Binary Image
- Gray Scale Image
- Color Image

1.3.1 BINARY IMAGE

A binary image is a digital image that has only two possible values for each pixel typically the two colors used for a binary image are black and white though any two colors can be used. The color used for the object(s) in the image is the foreground color while the rest of the image is the background color.

Binary images are also called bi-level or two-level. This means that each pixel is stored as a single bit (0 or 1).This name black and white, monochrome or monochromatic are often used for this concept, but may also designate any images that have only one sample per pixel, such as grayscale images.

Binary images often arise in digital image processing as masks or as the result of certain operations such as segmentation, thresholding, and dithering. Some input/output devices, such as laser printers, fax machines, and bi-level computer displays, can only handle bi-level images

1.3.2 GRAY SCALE IMAGE

A grayscale image is digital image is an image in which the value of each pixel is a single sample, that is, it carries only intensity information. Images of this sort, also known as black-and-white, are composed exclusively of shades of gray (0-255), varying from black (0) at the weakest intensity to white (255) at the strongest. Grayscale images are distinct from one-bit black-and-white images, which in the context of computer imaging are images with only the two colors, black, and white (also called bi-level or binary images). Grayscale images have many shades of gray in between. Grayscale images are also called monochromatic, denoting the absence of any chromatic variation.

Grayscale images are often the result of measuring the intensity of light at each pixel in a single band of the electromagnetic spectrum (e.g. infrared, visible light, ultraviolet, etc.), and in such cases they are monochromatic proper when only a given frequency is captured.

1.3.3 COLOUR IMAGE:

A digital color image is a digital image that includes color information for each pixel. Each pixel has a particular value which determines it's appearing color. This value is qualified by three numbers giving the decomposition of the color in the three primary colors Red, Green and Blue. Any color visible to human eye can be represented this way. The decomposition of a color in the three primary colors is quantified by a number between 0 and 255. For example, white will be coded as R = 255, G = 255, B = 255; black will be known as (R,G,B) = (0,0,0); and say, bright pink will be : (255,0,255).

In other words, an image is an enormous two-dimensional array of color values, pixels, each of them coded on 3 bytes, representing the three primary colors. This allows the image to contain a total of 256x256x256 = 16.8 million different colors. This technique is also known as RGB encoding, and is specifically adapted to human vision.

II. LITERATURE SURVEY

2.1 INTRODUCTION:

This chapter deals with the previous work done by various authors in the field of hand gesture recognition system and the techniques and algorithms have been used for hand recognition system.

2.2 BACKGROUND STUDIES:

HUMAN COMPUTER INTERACTION AND MEDICAL DEVICES (C. Acharya, H. Thimbleby, P. Oladimeji)

C.Acharya Et Al. shows that some healthcare devices fall far short and thus identifies a gap in applied Human Computer Interaction(HCI). The authors are collaborating with design team, medical school, users and manufacturers to redesign the hospital bed, by applying HCI and human factors and fixing programming bugs. To achieve dependable, usable, and well-engineered interactive devices in healthcare requires applied Human Computer Interaction (HCI) research and awareness of HCI issues throughout the lifecycle, from design through to procurement, training and use. Basic, interactive hospital bed as a case study, arguably so routine and simple enough that there should have been very few problems. However, the bed's interactive control panel design violates standard HCI principles. It is also badly programmed by the manufacturer. Evidently, something has gone wrong, somewhere from design to procurement and most of the problems would have been managed or avoided by conventional HCI processes. Driven by the case study, this paper explores the problems and makes recommendations.

There are many similarly poorly designed medical devices. Manufacturers and healthcare purchasing groups should adhere to HCI processes and guidelines, as well as those provided by regulatory agencies for the design, regulation, and procurement of devices, products or systems that contribute to patient safety. The challenge is to make HCI knowledge and priorities available to and effective in this important domain in any places that can make a difference. One of the possibilities is that HCI is complex and that medical devices and environments are complex. It is clearly important to consider how the system is trained and motivate HCI practitioners to engage or avoid important areas such as healthcare. One might assume with the paucity of healthcare HCI literature that there is no significant problem to be addressed.

COMBINATION OF HIDDEN MARKOV MODELS WITH DYNAMIC TIME WARPING FOR SPEECH RECOGNITION (S. Axelrod and B. Maison)

S.Axelrod Et Al. proposes a combine Hidden Markov Models of various topologies and Nearest Neighbor classification techniques in an exponential modeling framework with a model selection algorithm to obtain significant error rate reductions on an isolated word digit recognition task. This work is a preliminary investigation of large scale modeling techniques to be applied to large vocabulary continuous speech recognition. Increases in computational power, storage capacity and training data available for use by automatic speech recognition (ASR) systems, combined with the perception that the performance of such systems has reached a plateau motivate us to consider modeling strategies for speech recognition that, while more resource intensive, have the potential to obtain significant reductions in error rate.

Advantages

- It can be applied to large vocabulary continous speech recognition.
- The potential to obtain significant reduction in error rate.
- Increases in computational power , storage capacity and training data.

A DYNAMIC GESTURE RECOGNITION AND PREDICTION SYSTEM USING THE CONVEXITY APPROACH (P. Barros, N. T. Maciel-Junior, B. J. Fernandes, B. J. Bezerra, and S. M. Fernandes)

Real-time recognition of dynamic gestures is a problem for most of the applications nowadays. The

prediction approach can be used as a solution for the above problem. This approach uses an incomplete gesture input and it tries to predict which gesture the given input is represented. This study proposes a system for dynamic gesture recognition and prediction using a innovation feature extraction technique called the Convexity Approach. The proposed system generates a smaller feature vector to describe the hand shape with a minimal amount of data. Two data sets one for data recognition and another for prediction. Dynamic gesture recognition and prediction system is implemented using two independent modules based on Hidden markov modules and Dynamic time wraping.

The paper presents the application of the dynamic gesture feature extraction technique called Convexity Local Contour Sequence (CLCS) as the extractor for the prediction task. Two predictor systems are used to achieve this task and results are compared and discussed.

A HEALTH MONITORING SYSTEM FOR ELDER AND SICK PERSONS (A. Chaudhary and J. Raheja)

A.Chaudhary Et Al. discusses a vision based health monitoring system which would be very easy in use and deployment. The system works on the principles of Computer Vision and it uses image processing for acquiring gesture and preprocessing. The system have 2D camera only and its cost is very much affordable. Elder and sick people who are not able to talk or walk, they are dependent on other human beings for their daily needs and need continuous monitoring. The developed system provides facility to the sick or elder person to describe his or her need to their caretaker in lingual description by showing particular hand gesture with the developed system.

The system uses fingertip detection technique for gesture extraction and artificial neural network for gesture classification and recognition. The system is able to work in different light conditions and can be connected to different devices to announce user's need on a distant location. Hand Gesture Recognition is very easy and natural way to interact with the machines where no training for the user is required. So future of this technique is very bright in systems for handicap and disabled persons.

The technique can be applied in the field of disaster management so that lives of more people could be saved. In the damage situations i.e. in mining where, workers are always in danger zone. This system will be placed on the body of each miner or persons staying or working in such dangerous places and will put the transceiver on the nearby wall. So in any case of disaster, if the user is in danger and can't go to ground level, he can show predefined gesture syntax to system that will interpret it and send a signal to transceiver nearby and it will forward the signal further to the rescue team in the control room.

In this scenario an advanced version of this system is required, which would be GPS enabled to detect the exact location of the user in the mine. It will work in the case of suffocation, gas poisoning or fire hazards where the person is not able to shout or to tell even a single word. This system is based on human to machine interaction, in which machine would be able to do action according to the predefined syntax of the gesture made by user. In gesture recognition color based methods are applicable because of their characteristic color footprint of human skin. The developed system helps the elder/sick person to express their wish or need in predefined lingual description to a specified place. This place can be hospital staff room or living room in home, where generally people stay.

VISION BASED HAND GESTURE RECOGNITION FOR HUMAN COMPUTER INTERACTION(S. S. Rautaray and A. Agrawal)

S.S. Rautaray Et Al. explains about the human– computer interaction (HCI) will have a positive impact on their use. Hence, there has been growing interest in the development of new approaches and technologies for bridging the human– computer barrier. The ultimate aim is to bring HCI to a regime where interactions with computers will be as natural as an interaction between humans, and to this end, incorporating gestures in HCI is an important research area. Gestures have long been considered as an interaction technique that can potentially deliver more natural, creative and intuitive methods for communicating with our computers.

The paper provides an analysis of comparative surveys done in this area. The use of hand gestures as a natural interface serves as a motivating force for research in gesture taxonomies, its representations and recognition techniques, software platforms and frameworks which is discussed briefly in this paper. It focuses on the three main phases of hand gesture recognition i.e. detection, tracking and recognition. Different application which employs hand gestures for efficient interaction has been discussed under core and advanced application domains. It further discusses the advances that are needed to further improvise the present hand gesture recognition system for further perspective that can be widely used for efficient human computer interaction. Gestures have long been considered as an interaction technique that can potentially deliver more natural, creative and intuitive methods for communicating with our computers. The main goal of this survey is to provide researchers in the field of gesture based HCI with a summary of progress achieved to date and to help identify areas where further research is needed. This paper also provides an analysis of existing literature related to gesture recognition systems for human computer interaction by categorizing it under different key parameters.

Yee Yong Nor Azman Ismail Et Al. proposes about the human-computer interaction (HCI) has been an interest research in recent years which witnessed the development from text-based like using a keyboard to graphic user interface (GUI) based on a mouse, from cumbersome data gloves and tracking devices to visual-based computer application. One of the interest fields is by using hand gestures to interact with computer. Gesture recognition is one of the popular methods for Human Computer Interaction. This paper is mainly focused on the application of this technology in computer environment. The idea is to construct such a system which can take gesture inputs and on the basis of that controlling and commanding of the computer is performed. This module basically performs the finger counting and then on the basis of which actions are performed.

These actions are used to control various functions of operating system. One of the interest fields is by using hand gestures to interact with computer. However, the complexity of a hand set a lot of challenges to be tracked. In real-time, the application requires high accurate detection and recognition. In additional the real and clutter environments have a big impact on recognition process because it included with irrelevant information from the application point of view. In this paper, a real time vision based hand gesture interaction prototype was proposed. Currently a prototype has built for controlling the desktop cursor and concerned the tasks involving in navigation the desktop cursor by using hand gesture input modality. Results reveal that the proposed technique works well with the robust condition.

OPTICAL FLOW HAND TRACKING AND ACTIVE CONTOUR HAND SHAPE FEATURES FOR CONTINOUS SIGN LANGUAGE RECOGNITION WITH ARTIFICIAL NEURAL NETWORKS (P. Kishore and M. Prasad)

P.Kishore Et Al. proposes about the Horn Schunck optical flow (HSOF) extracts tracking features and

Active Contours (AC) extract shape features. To extract hand tracks and hand shape features from continuous sign language videos for gesture classification using back propogation neural network. A feature matrix characterizes the signs in continuous sign videos. A neural network object with back propagation training algorithm classifies the signs into various words sequences in digital format. Digital word sequences are translated into text with matching and the suiting text is voice translated using windows application programmable interface (Win-API).

Ten signers, each doing sentences having 30 words long tests the performance of the algorithm by computing word matching score (WMS). Learning skills of a hearing impaired person are seriously hampered because of the missing hearing sense. A mute person has to depend largely on visual sense and any learning and communication aids will help them learn faster and communicate better. Usually human interpreter trained in sign language understanding acts as a bridge between normal people and mute people.

RULE-BASED APPROACH TO RECOGNITION HUMAN BODY POSES AND GESTURES IN REAL TIME(T. Hachaj and M. Ogiela)

T Hachaj Et Al. proposed a classifier capable of recognizing human body static poses and body gestures in real time. The method is called the gesture description language (GDL). The proposed methodology is intuitive, easily thought and reusable for any kind of body gestures. The very heart of our approach is an automated reasoning module. It performs forward chaining reasoning (like a classic expert system) with its inference engine every time new portion of data arrives from the feature extraction library. All rules of the knowledge base areorganized in GDL scripts having the form of text files that are parsed with a LALR-1 grammar . Gestures might also be recognized using a neural network and fuzzy sets.

The main novelty of this paper is a complete description of our GDL script language, its validation on the large dataset and the presentation of its possible application. The recognition rate for examined gestures is within the range of 80.5 – 98.5 %. There is also implementation of the application that uses this method: it is a 3D desktop visualizing 3D dataset that is controlled by gestures recognized by the GDL module. In GDL, the letter case does not matter. The GDL script is a set of rules. Each rule might have an unlimited number of premises that are connected by conjunction or alternative operators.

In GDL, premises are called logical rules. A logical rule can take two values: true or false. Apart from logical rules, the GDL script also contains numeric rules (3D numeric rules) which are simply some mathematical operations that return floating-point values (or floating three-dimensional points). A numeric rule might become a logical rule after it is combined with another numeric rule by a relational operator. The brackets in logical and numeric (3D) rules are used to change the order in which instructions are executed. The GDL script description of this gesture is very straightforward. GDL has the ability to detect and classify many techniques described by the GDL script rule in one recording. GDL also eliminates the problem of body proportions between users. This approach has proven to be reliable tool for recognizing human body static poses and body gestures. In this approach, static poses the so-called key frames form components of dynamic gestures.

HAND SEGMENTATION USING SKIN COLOR AND BACKGROUND INFORMATION (Wei Wang and Jing Pan)

Wei wang Et Al. proposed about the Skin color based

hand segmentation using skin color models shows poor performance in complex background where similar colors of the skin and non-uniform illumination exist. A new method has been proposed for hand segmentation by using an adaptive skin color model and the background information around the hand. Precise hand segmentation is crucial for gesture based Human-machine interaction. Recognizing hand signs and tracking hand motion is crucial for human machine interface which has always been a hot topic in recent years. The purpose of hand signals recognizing and hand fingers tracking is to make the machine get our instructions in a non-contact way. Other applications such as helping the disabled people communicate with the normal people also show the importance of the research in this area. Hand segmentation is the first and also the critical step for recognizing hand signs and tracking hand motion. The excellence of the hand segmentation affects the accuracy of its following applications in a straight way. Firstly, our method captures pixel values of the hand and the background then converts them into YCbCr color space.

Secondly, skin and background Gaussian models based on the color space of CbCr are proposed. Lastly, these models are taken to segment the whole image respectively, and then required for the intersection. The main contribution of the paper is that the background information is taken into account to split image in reversed side to enhance the performance. Hand segmentation is the first and also the critical step for recognizing hand signs and tracking hand motion. Experimental results show that our method outperforms the method that uses the skin color model only.

DETERNIMING OPTICAL FLOW (B. Horn and B.Schunck)

B. Horn Et Al. described about Optical flow cannot be computed locally, since only one independent measurement is available from the image sequence at a point, while the flow velocity has two components. The optical flow cannot be computed at a point in the image independently of neighboring points without introducing additional constraints, because the velocity field at each image point has two components while the change in image brightness at a point in the image plane due to motion yields only one constraint.

Optical flow is the distribution of apparent velocities of movement of brightness patterns in an image. Optical flow can arise from relative motion of objects. Consequently optical flow can give important information about the spatial arrangement of the objects viewed and the rate of change of this arrangement. Discontinuities in the optical flow can help in segmenting images into regions that correspond to different object. A second constant is needed for the above calculation. A method for finding the optical flow pattern is presented which assumes that the apparent velocity of the brightness pattern varies smoothly almost everywhere in the image.

An iterative implementation is shown which successfully computes the optical flow for a number of synthetic image sequences. Algorithm is robust in that it can handle image sequences that are quantized rather coarsely in space and time. It is also insensitive to quantization of brightness levels and additive noise. A method has been developed for computing optical flow from a sequence of images. It is based on the observation that the flow velocity has two components and that the basic equation for the rate of change of image brightness provides only one constraint. Smoothness of the flow was introduced as a second constraint. An iterative method for solving the resulting equation was then developed. The computed optical flow is somewhat inaccurate since it is based on noisy, quantized measurements. Proposed methods for obtaining information about the shapes of objects using derivatives (divergence and curl) of the

optical flow field may turn out to be impractical since the inaccuracies will be amplified.

AN AGING NATION:THE OLDER POPULATION IN THE UNITED STATES (J. M. Ortman, V. A Velkoff and H. Hogan)

J M. Ortman Et Al. proposes one of the interest fields is by using hand gestures to interact with computer. The aging of the population will have wide-ranging implications for the country. The age structure of the U.S. population is expected to change over the coming decades and focuses on the older population in terms of age, sex, race, and Hispanic origin. The size and structure of the older population is important to public and private interests, both socially and economically. The aging demographers often mean that the proportion of the population in the older ages increases.

As the United States ages over the next several decades, its older population will become more racially and ethnically diverse. The projected growth of the older population in the United States will present challenges to policy makers and programs, such as Social Security and Medicare. It will also affect families, businesses, and health care providers. However, the complexity of a hand set a lot of challenges to be tracked. In real-time, the application requires high accurate detection and recognition. In additional the real and clutter environments have a big impact on recognition process because it included with irrelevant information from the application point of view. In this paper, a real time vision based hand gesture interaction prototype was proposed.

HAND GESTURE RECOGNITION WITH MULTI-SCALE WEIGHTED HISTOGRAM OF CONTOUR DIRECTION (MSWHCD) NORMALIZATION FOR WEARABLE APPLICATIONS (Yiyi Ren, Xiang Xie, Guolin Li and Zhihua Wang)

Y. Ren Et Al. proposes a static hand gesture recognition method with low computation and memory consumptions for wearable applications. The hand contour is chosen as the hand gesture feature and SVM is used to classify the feature. Multi-Scale Weighted Histogram of Contour Direction (MSWHCD) based direction normalization is proposed to ensure a good recognition performance. In order to improve efficiency, the proposed histogram only counts the direction of contour point to focus on the most significant hand feature in the first person view of wearable devices.

Scanline-based stereo matching is used to find the scene's depth in order to lower the computation and memory consumptions. The stereo matching algorithm works scanline by scanline. First, each scanline is segmented into small intervals by edge detector, such as Canny edge detector. Stereo correspondences are found among these intervals. Until now, only intra-scanline information is used for the stereo matching. To utilize the inter-scanline information, maximum spanning forest (MSF) is constructed between adjacent scanlines. The edges of MSF are weighted by the consistencies of the interval's color and position. Therefore, inter-scanline disparity information (in stereo matching, depth is represented by disparity) can propagate along the edges of MSF to refine the disparity between adjacent intervals and to get rid of horizontal disparity streaks. At last, the hand is segmented by using depth information.

Based on hand anatomy, the proposed histogram is weighted by considering each contour point's position and direction jointly using Direction-Angle Map (DAM), so as to ensure the robustness. Experimental results show that the proposed method can give a recognition accuracy of 97.1% with a frame rate of 30fps on PC.

RULE-BASED APPROACH TO RECOGNIZING HUMAN BODY POSES AND GESTURES IN REAL

TIME(Tomasz Hachaj , Marek R. Ogiela)

A method called the gesture description language (GDL) is introduced to recognize human body static poses and body gestures in real time. The proposed methodology is intuitive, easily thought and reusable for any kind of body gestures. The very heart of our approach is an automated reasoning module. It performs forward chaining reasoning with its inference engine every time new portion of data arrives from the feature extraction library. All rules of the knowledge base are organized in GDL scripts having the form of text files that are parsed with a LALR-1 grammar. The main novelty of this method is a complete description of our GDL script language, its validation on a large dataset (1,600 recorded movement sequences) and the presentation of its possible application. The recognition rate for examined gestures is within the range of 80.5–98.5 %.

GDL scripts contain many numerical rules. A numeric rule might become a logical rule after it is combined with another numeric rule by a relational operator. The brackets in logical and numeric (3D) rules are used to change the order in which instructions are executed. The first rule checks if the right elbow and the right hand are situated to the right of the torso, if the right hand is above the right elbow and if the vertical coordinates of the right hand and the right elbow are no more than 50 mm different. The last part of the rule is the premise that checks if the horizontal coordinates of the right shoulder and the right elbow are no more than 50 mm different. The second rule is similar to first one, but it describes the left arm, shoulder and elbow. The last rule checks if both previous rules are satisfied. This is done by checking the logical conjunction of both previous conclusions.

The last very important ability of GDL scripts to check the presence of particular sequences of body joints that appeared in a constrained time range. A gesture is defined in GDL as a series of static poses (so-called key frames) appearing one after another within given time constraints. All gestures are based on arm movements, which make the gestures easier to make for test participants, as a result of which they may be more relaxed and make the gestures naturally.

It has been proven to be reliable tool for recognizing human body static poses and body gestures. In the above approach, static poses, the so-called key frames form components of dynamic gestures. The recognition rate for all of the tested gestures ranges from 80.5 to 98.5 %.

COMBINATION OF HIDDEN MARKOV MODELS WITH DYNAMIC TIME WARPING FOR SPEECH RECOGNITION (Scott Axelrod and Benoit Maison)

The above method combines the hidden markov models of various technologies and neighbourhood algorithm to avoid error reductions in recognitions. Increases in computational power, storage capacity and training data available for use by automatic speech recognition (ASR) systems, combined with the perception that the performance of such systems has reached a plateau to consider modeling strategies for speech recognition that, while more resource intensive, have the potential to obtain significant reductions in error rate.

The acoustic modeling uses Dynamic Time Warping (DTW) techniques to match segments of test utterances to stored training data. DTW systems can capture long-range dependencies in the acoustic data and can potentially adapt to differences in gender, speaker, and accent by pinpointing, at decode time, similar data in the training set.

Hidden Markov models are created with 5000 words in the training set. Models are built with the condition in which the number of states in each branch was fixed independently of the data as well as models where the number of states varied with the branch.

The models in the fixed length case as
F(B, N,S)
where B = number of branches,
N= number of Gaussian components for each of the state models and
S = training set size

All the models were trained from a flat start without reference to any auxiliary data such as baseforms or a reference model. For the seed models we use uniform transition probabilities and a choice of initial hard state alignments. For efficiency, the seed Gaussian mixture models for each states was obtained by first applying a few rounds of k-means to the data aligned to the state and then training with the usual Expectation Maximization algorithm.

Dynamic Time Warping(DTW) is a technique that allows the computation of distances between sequences of speech frames of different lengths. The symmetric DP algorithm permits the matching of sequences of any length. However, since each frame of each sequence must be matched to some frame of the other sequence, the total DTW cost is dominated by the length of the longest of the two sequences. This effect makes it harder to compare the matching cost of a test sequence against templates of different durations.

Nearest Neighbor techniques can be significantly enhanced by distance modeling, smoothed voting, and normalization. The error rate reductions are done by combining DTW and multi-branch HMM models using maximum entropy techniques and a greedy model selection algorithm.

SUPPORT VECTOR MACHINE (SVM) CLASSIFICATION THROUGH GEOMETRY (Michael E. Mavroforakis, and Sergios Theodoridis)

M.E. Mavroforakis Et Al proposes the geometric framework for the support vector machine (SVM) classification problem. It provides an intuitive ground for the understanding and the application of geometric optimization algorithms, leading to practical solutions of real world classification problems.

In this work, the notion of reduced convex hull is employed and supported by a set of new theoretical results. These results allow existing geometric algorithms to be directly and practically applied to solve not only separable, but also non separable classification problems both accurately and efficiently. As a practical application of the new theoretical results, a known geometric algorithm has been employed and transformed accordingly to solve separable problems successfully.

Support Vector Machine (SVM) formulation of pattern recognition problems brings along a bunch of advantages over other approaches. Some of the problems are:
1) the assurance that once a solution has been reached, it is the unique solution,
2) good generalization properties of the solution,
3) sound theoretical foundation based on learning theory and Optimization theory,
4) common ground / formulation for the class separable and the class non-separable problems as well as for linear and non-linear problems and
5) clear geometric intuition on the classification task.

Due to the above properties, SVM have been successfully used to a number of applications. The SVM approach to machine learning is known to have both theoretical and practical advantages. Among these, are the sound mathematical foundation of SVM, overcoming of the “curse of dimensionality” and the intuition that is displayed.

Reduce convex hull is represented by R(C, μ) is calculated based on upper bound μ. The overlapping convex hulls are overcome by the proper selection of μ. Once the upper bound is properly fixed the

overlapping becomes separable and hence geometric algorithm gets applied.

The algorithm presented here does not use any heuristics and provides a clear understanding of the convergence process and the role of the parameters used. Furthermore, the penalty factor μ can be set different for each class, reflecting the importance of each class.

III. EXISTING SYSTEM

THE VISION-BASED HAND GESTURE RECOGNITION USING BLOB ANALYSIS

Image acquisition is done by importing file with dynamic hand gestures. The video frames are separated into single frame. RGB image frame is converted into grayscale image frame because grayscale level provides effective processing on pixels more than RGB channels. It is easier and faster to implement on grayscale image as it only requires the gray levels which have less colors than the RGB channels and do not affect to other pixels. Before applying the optical flow for motion analysis, the sequence of inputs or image frames set with specified value must be calculated by an object to find the mean. This process is based on two dimensional mean block which is able to calculate over single or multiple inputs and track the mean value of each channel in a sequence of image frames over a period of time. Hence each element is treated as a channel.

After calculating the mean of each channel in a time sequence of image frames, the median filter process calculates the two dimensional moving median filtering of the image frames along each channel independently over time by using the sliding window method. In this method, a window of specified length moves over each channel sample by sample and computes the median of the data. The Morphological Close technique is used for extracting interested region and remove out noises in the image frames. It is applied with an intensity image to improve the quality of the input. Intensity or gray level closing consists of the gray level dilation followed by the gray level erosion.

Dilation operation enlarges the boundaries of the hand shape from the background and contains the original boundary of the hand shape in the intensity image. It also removes undesirable objects such as small holes on the hand as well. Blob Analysis is used to calculate statistics for labeled regions in an intensity image. It proceeds quantities such as the centroid, bounding box, label matrix, and blob count. The goal of this method is to detect corresponding regions in scaled versions of the same image. According to the goal, the scale selection mechanism is needed for finding characteristic region size that is covariant with the image transformation. For spatial selection, the magnitude of the Laplacian response will achieve a maximum at the center of the blob in which the scale of the Laplacian and blob is matched.

The dynamic six hand gestures are defined including the close fist, one finger, two finger, three finger, four finger, and open palm gestures. The main methods include the use of optical flow and blob analysis. The optical flow detects the motion vector of the hand movement in detection process. Additionally, the blob analysis detects and tracks the hand area in tracking and classification processes. The algorithm performs good results at detection, tracking and classification for all hand gestures even it is blurred because of the motion.

IV. CONCLUSION AND FUTURE WORK

The main goal is to introduces multi finger touch input based authentication system for mobile devices the characteristics hand movement based five number of gestures based on the hand movements have been

taken into account from the users. The main method includes the uses of DWT and Contour detection. DWT detects the approximate image and the detailed images using the low pass and high pass filter. Additionally Contour detection detects the boundaries in an image. Users data patterns of the gestures are taken and features are sensed based on DWT algorithm and then the five finger position are sensed and based on the location features of the five position hand gestures are sensed. Based on the location features of the five finger, already defined functions of hand gestures and the given input ratio is compared and the Equal Error Rate (EER) is sensed over the images. Finally the images are compared with the contour counts and send the respective messages to the caretaker.

The future research can be applied to monitor many immobilized people at the same time using skin color detection. It will focus more on pre-processing process to improve the algorithm to be more robust and practical.

V. REFERENCES

1. C Acharya, H. Thimbleby, P. Oladimeji ?Human computer interaction and medical devices?, In: the 24th BCS Interaction Specialist Group Conference, British Computer Society, pp. 168-176, 2010.
2. S Axelrod and B. Maison, "Combination of hidden Markov models with dynamic time warping for speech recognition, 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, vol.1, pp. I- 173-6, 2004.
3. P Barros, N. T. Maciel-Junior, B. J. Fernandes, B. J. Bezerra, and S. M. Fernandes, "A dynamic gesture recognition and prediction system using the convexity approach", Computer Vision and Image Understanding.
4. A Chaudhary and J. Raheja, "A Health Monitoring System for Elder and Sick Persons", International Journal of Computer Theory and Engineering, pp. 428-431, 2013.
5. T Hachaj and M. Ogiela, "Rule-based approach to recognizing human body poses and gestures in real time", Multimedia Systems, vol. 20, no. 1, pp. 81-99, 2013.
6. B Horn and B. Schunck, "Determining optical flow", Artificial Intelligence, vol. 17, no. 1-3, pp. 185-203, 1981.
7. C Hsieh and D. Liou, "Novel Haar features for real-time hand gesture recognition using SVM", Journal of Real Time Image Processing, vol. 10, no.2, pp. 357-370, 2015.
8. P Kishore and M. Prasad, "Optical Flow Hand Tracking and Active Contour Hand Shape Features for Continuous Sign Language Recognition with Artificial Neural Networks", International Journal of Software Engineering and Its Applications, vol. 9, no. 12, pp. 231-250, 2015.
9. M E. Mavroforakis and S. Theodoridis, "Support Vector Machine (SVM) classification through geometry, 2005 13th European Signal Processing Conference, Antalya, 2005, pp. 1 4. 41
10. T. Pringsheim, N. Jette, A. Frolkis and T. Steeves, "The prevalence of Parkinson's disease: A systematic review and meta-analysis", Movement Disorders, vol. 29, no. 13, pp. 1583-1590, 2014.
11. S S. Rautaray and A. Agrawal, Vision based hand gesture recognition for human computer interaction: a survey. In: Artificial Intelligence Review, vol. 43, no.1, pp. 1-54, 2015.
12. Y. REN, X. Xie, G. Li and Z. Wang, "Hand Gesture Recognition with Multi-Scale Weighted Histogram of Contour Direction (MSWHCD) Normalization for Wearable Applications", IEEE Transactions on Circuits and Systems for Video Technology, pp. 1-1, 2016.
13. J P. Wachs, M. Kolsch, H. Stern and Y. Edan, ?Vision-based handgesture applications?, Communications of the ACM, vol. 54, no. 2, pp. 60-71, 2011.
14. Wei Wang and Jing Pan, "Hand segmentation using skin color and background information, 2012 International Conference on Machine Learning and Cybernetics, Xian, pp. 1487-1492 , 2012.