Hand gesture recognition using support vector machine

Nagashree R N[1] , Stafford Michahial[1], Aishwarya G N[2], Beebi Hajira Azeez[2],
Jayalakshmi M R[2], R Krupa Rani[2]

[1]*Asst. Professor, Dept of Telecommunication, GSSSIETW, Mysuru.*
[2]*Student of 8th sem, Dept of Telecommunication, GSSSIETW, Mysuru.*

--**ABSTRACT**---
The images contain measurement information of key interest for a variety of research and application areas. On the other hand, computers have become an inseparable part of our society, influencing many aspects of our daily lives in terms of communication and interaction. The main motive is to develop a system that can simplify the way humans interact with Computers. The system is designed using Canny's edge detection for edge detection and Histogram of gradients for feature extraction and the Support Vector Machine (SVM) Classifier which is widely used for classification and regression testing. SVM training algorithm builds a model that predicts whether a new example falls into one category or other. And the classifier learns from the data points in examples when they are classified belonging to their respective categories.

KEYWORDS – *SVM, Canny's Edge detection, Feature extraction, HOG*.

I.   INTRODUCTION

With the continued growth of multimedia and communication systems, the instrumentation and measurement fields have seen a steady increase in the focus on the image data. Digital image processing deals with manipulation of digital images through a digital computer. It is a subfield of signals and systems but focus particularly on images. DIP focuses on developing a computer system that is able to perform processing on an image. The input of that system is a digital image and the system process that image using efficient algorithms, and gives an image as an output [1].

A gesture is a form of non-verbal communication in which visible bodily actions communicate particular messages, either in place of speech or together and in parallel with spoken words. Gestures can be static (posture or certain pose) which require less computational complexity or dynamic (sequence of postures) which are more complex but suitable for real time environments. Different methods have been proposed for acquiring information necessary for recognition gestures system. Some methods used additional hardware devices such as data glove devices and color markers to easily extract comprehensive description of gesture features. Other methods based on the appearance of the hand using the skin color to segment the hand and extract necessary features, these methods considered easy, natural and less cost comparing with methods mentioned before. It includes movement of the hands, face, or other parts of the body. Gesture recognition can be seen as a way for computers to begin to understand human body language, thus building a richer bridge between machines and humans than primitive text user interfaces or even Graphical User Interfaces (GUIs), which still limit the majority of input to keyboard and mouse [1]. It enables humans to interface with the machine (HMI) and interact naturally without any mechanical devices. The main purpose of developing such a system lies in the fact that gesture recognition has implemented in motion analysis to machine learning. It also serves many applications from virtual reality to sign language recognition. A lot of research has been carried out in last 2-3 decades on hand gesture recognition approach. These can be roughly divided into two categories, namely, Glove Based Gesture Recognition and Vision Based Gesture Recognition. Glove based Hand Gesture Recognition hinders the naturalness as cumbersome devices are required to wear. A Vision Based approach uses features extracted from visual appearance of the input image model of the hand, comparing these modelled features with features extracted from input camera(s) or video input [4]. Here we are emphasising on different aspects of Hand Gesture Recognition with Support Vector Machine as a classifier.

II.   OBJECTIVE

The SVM is a popular pattern recognition technique with supervised learning. Since it divides the feature space for each class, the SVM can handle unknown data well, although it is not suited to grouping sample data. Support Vector Machines (SVM) is a linear machine with some very good properties. The main idea of a SVM in the context of pattern classification is to construct a hyper plane as the decision surface. The hyper plane is constructed in such a way that the margin of separation between positive and negative examples is maximized. The SVM uses a principled approach based on statistical learning theory know as structural risk minimization, which minimizes an upper bound on the generalization error(i.e. a parameter estimation heuristic that seeks parameter values that minimize the 'risk' or 'loss' that the model incurs on the training data).

III.   DESIGN METHODOLOGY

The project involves the use of web camera which takes the input; it will also sense the depth of an image. The Hand gesture classification is done in two phases. The first is the training phase where we train the system with certain predefined symbols and we store them in the database of the system for further classification. The second phase is the testing phase where the system is given the symbols and it compares it with the symbols in the database and compares the most matched feature and gives the output.
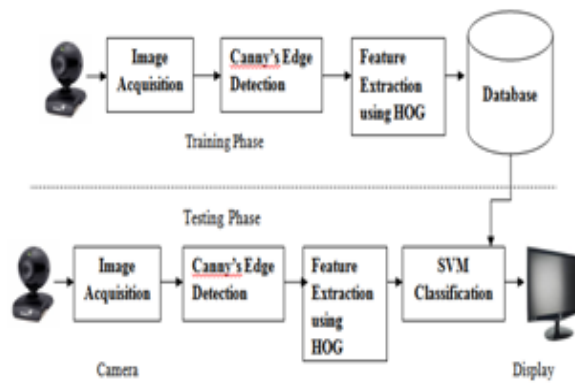

Fig 1 Block Diagram

The processing involves acquiring the hand gesture using a web camera. In the per-processing stage a bounding box is created to identify the hand for the separation of the foreground from the background. This is done by making use of difference in consecutive frames. The acquired image is subjected to resizing in order to have faster computation. The edge of the resized image is detected using Canny's edge detection. The extraction of the feature of the image is done using Histogram of gradients (HOG). The extracted features are stored in database. SVM compares the image with the most matched feature in the database and gives the output on the screen.

a)   CANNY'S EDGE DETECTION

Edge detection is the name for a set of mathematical methods which aim at identifying points in a digital image at which the image brightness changes sharply or more formally, has discontinuities. The point at which image brightness changes sharply are typically organized into a set of curved line segments termed as edges. The purpose of Edge detection in general is to significantly reduce the amount of data in an image, while preserving the structural properties to be used for further image processing. Canny's Edge Detection one of the edge detection method which was developed by John F Canny (JFC) in 1986. This method makes use smoothing concept which is applied in the Gaussian operation, this makes error detection more effective. This also improves the signal with respect to the noise ratio and this is established by non-maxima suppression method as it results in one pixel wide ridges as the output. Another advantage is better detection of edges especially in noise state. Even though it is quite old, it has become one of the standard edge detection methods and it is still used in research. The algorithm runs in five different steps
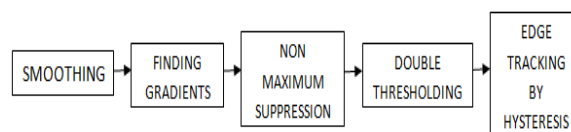

Fig 2 Canny's Edge Detection Block Diagram

i) Smoothing: The image is then converted from RGB scale to Grey scale and is blurred in order to remove the noise that is accompanied during the image acquisition. The noise is removed using the Gaussian filter by specifying double threshold value.

ii) Finding gradients: The edges should be marked where the gradients of the image has large magnitudes.

iii) Non maximum suppression: Only local maxima should be marked as edges.

iv) Double thresholding: Potential edges are determined by thresholding.

v) Edge tracking by hysteresis: Final edges are determined by suppressing all edges that are not connected to a very certain edge.

b) HISTOGRAM OF GRADIENTS
Feature extraction is a type of dimensionality reduction that efficiently represents interesting part of an image as a compact feature vector. Detecting humans in images is a challenging task owing to their variable appearance and the wide range of poses that they can adopt. The first need is a robust feature set that allows the human form to be discriminated cleanly, even in cluttered backgrounds under difficult illumination. We study the issue of feature sets for human detection, showing that locally normalized Histogram of Oriented Gradient (HOG) descriptors provide excellent performance relative to other existing feature sets including wavelets. This approach is useful when the image sizes are large and a reduced feature representation is required to quickly complete tasks. This involves reducing the amount of resources required to describe a large set of data. HOG features were introduced by Navneed Dalal and Bill Triggs who have developed and tested several variants of HOG. HOG deals with moving body parts and are better compared to other feature extractors.
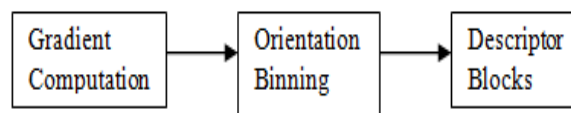The HOG algorithm includes the following steps:

Fig 4 HOG Block Diagram

i)   Gradient Computation: The most common method is to apply the 1D centred point discrete derivative mask in both the horizontal and vertical directions. Specifically this method requires filtering the gray scale image.
ii) Orientation Binning: This involves creating cell histograms. Each pixel within the cell casts a weighted vote for an orientation based histogram channel based on the values found in the gradient computation.
iii) Descriptor Blocks: The gradient strengths must be locally normalized, which requires grouping the cells together into larger, spatially connected blocks. Two main block geometrics exist: Rectangular HOG(R-HOG) blocks, Circular HOG(C-HOG).

c)   CLASSIFICATION USING SUPPORT VECTOR MACHINE
Classification is an ordered set of related categories used to group data according to its similarities. It consists of codes and descriptors and allows survey responses to be put into meaningful categories in order to produce useful data. It is a useful tool for developing statistical surveys. Classifier is an abstract metaclass which describes (classifies) set of instances having common features. Support Vector Machine (SVM), is one of the best machine learning algorithms, which was proposed in 1990's by Vapnik. SVM's are a set of related supervised learning methods used for classification and regression. A supervised learning is the machine learning task of inferring from supervised training data. A supervised learning algorithm analyzes the training data and produces an inferred function which is called a classifier. SVM has high accuracy, nice theoretical guarantees regarding over fitting, and with an appropriate kernel which can work well even when the data is not linearly separable in the base feature space. The support vector network is a new learning machine for two group classification problems. The machine conceptually implements the following idea: input vectors are non-linearly mapped to a very high dimensional feature space. In this feature space a linear decision surface is constructed. Special properties of decision surface ensures high generalisation ability of the learning machine. The idea behind support vector was previously implemented for the restricted case where the training data can be separated without errors.

Support vector machine constructs a hyperplane or set of hyperplanes in a high- or infinite-dimensional space, which can be used for classification, regression, or other tasks. A hyperplane is a subspace of one dimension less than its ambient space. A good separation is achieved by the hyperplane that has the largest distance to the nearest training-data point of any class. It is explicitly told to find the best separating line. It searches for the closest points which is called the "support vectors" (the name "support vector machine" is due to the fact that points are like vectors and that the best line "depends on" or is "supported by" the closest points). Once it has found the closest points, the SVM draws a line connecting them. It draws this connecting line by doing vector subtraction (point A - point B). The support vector machine then declares the best separating line to be the line that bisects -- and is perpendicular to -- the connecting line.

The support vector machine is better because when you get a new sample (new points), you will have already made a line that keeps B and A as far away from each other as possible, and so it is less likely that one will spill over across the line into the other's territory.
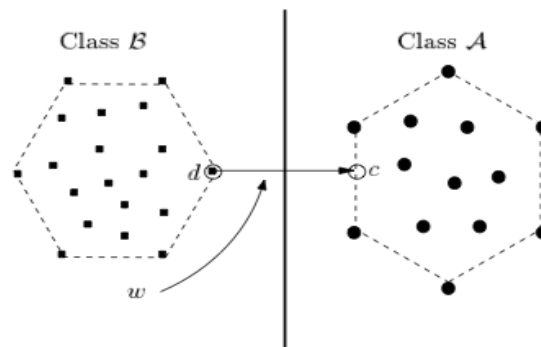

Fig 3 SVM classification.

DATASET FOR CLASSIFICATION

Fig 4 Dataset for classification

IV. EXPIRIMENTAL RESULT

The software analysis is carried out using MATLAB, that facilitates matrix manipulations, plotting of functions and data, implementation of algorithms, creation of user interfaces, and interfacing with programs. In this section, the performance of real time static hand gesture recognition system in complex background is evaluated for each of the 20 hand gestures. And running input image is used for the performance evaluation. The images were acquired in true colour using a webcam. The images in the testing phase are compared with the database and the result is evaluated.
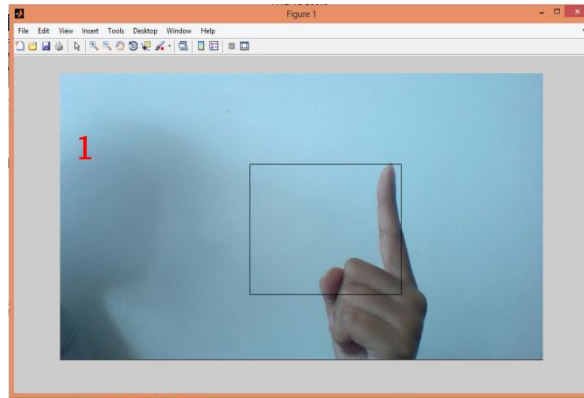

Fig 4 Gesture Recognised

V. CONCLUSION

Area of Hand gesture based computer human interaction is very vast. This project recognizes hand gesture off-line so work can be done to do it for real time purpose. Hand recognition system can be useful in many fields like robotics, computer human interaction and so make this offline system for real time will be future work to do. Local contour sequence (LCS) is our prime feature in this project. LCS can be used with other features so that it can be optimized to achieve higher recognition accuracy. Support Vector Machine can be modified for reduction of complexity. Reduction of complexity leads us to a less computation time. Reduced complexity provides us less computation time so we can make system to work real time. The applications of hand gesture recognition is for non-verbal communication between human and computer, general fit person and physically challenged people, 3D gaming, virtual reality etc.

The increase in applications, the gesture recognition system requires lots of research in different directions. A large number of research works carried out during last twenty years have been reviewed. The different sub-components, methodologies used for recognition of mainly hand gestures in those works have been described. A brief comparison of backgrounds, segmentation techniques, features used and the recognition methods have been done and presented. In today's digitized world, processing speeds have increased dramatically, with computers being advanced to the levels where they can assist humans in complex tasks. Yet, input technologies seem to cause a major bottleneck in performing some of the tasks, under-utilizing the available resources and restricting the expressiveness of application use. Hand Gesture recognition comes to rescue here. Computer Vision methods for hand gesture interfaces must surpass current performance in terms of robustness and speed to achieve interactivity and usability. A review of vision-based hand gesture recognition methods has been presented. Considering the relative infancy of research related to vision based gesture recognition remarkable progress has been made. To continue this momentum it is clear that further research in the areas of feature extraction, classification methods and gesture representation are required to realize the ultimate goal of humans interfacing with machines on their own natural terms.

REFERENCES

[1] Aseema Sultana, T Rajapuspha , "**A vision based Hand gesture recognition**" in the proceeding of IJCSET(2012), Vol 3,No 7.
[2] S Nagarajan, T S Subashini, "**Static Hand Gesture Recognition for Sign Language Alphabets using Edge Oriented Histogram**" in the proceeding of IJCA(2013), Vol 82, No 4.
[3] S. Naidoo, C.W. Omlin M. Glaser "**Vision based Static Hand gesture recognition**", South Africa.
[4] S. Mohamed Mansoor Roomi, R. Jyothi Priya and H. Jayalakshmi "**Hand Gesture Recognition for Human-Computer Interaction**," in the proceedings of Journal Of Computer Science (2009).