Human Computer Interaction Through Hand Gesture Recognition Technology

Nikhat Parveen, Arpita Roy, D. Sai Sandesh, J.Y.P.R.Sai Srinivasulu, N.Srikanth

Abstract: Communication is the main channel to interact between individuals with each other. Due to birth defects, accidents and oral disorders, there has been a drastic increase in the number of deaf and dumb victims in recent years. Since deaf and dumb people are unable to communicate with ordinary people, they must rely on some kind of visual communication. Throughout the world, most languages are spoken and interpreted. The people, that is, those who find it difficult to speak and hear "The Dumb" and "The Deaf," It's hard to understand exactly what the other person is trying to communicate with the deaf, and so on. Key recognition issues for hand gesture are consistent with the complexities of the gesture process. Also offered methods for evaluating the process of identifying recent postures and gestures. Hand gesture recognition is becoming an increasingly popular field of research in human computer interaction. Considering the similarity of human hand form with four fingers and a thumb, this paper aims to present a real-time hand gesture recognition system focusing on recognizing some important shape-based features such as orientation, center of mass (centroid), finger position, thumb in terms of raised or folded hand fingers and their respective location in the image. The solution in this paper depends entirely on the shape parameters of the hand gesture. This does not include any other means of identification of hand gesture such as skin color, texture, as these image-based characteristics vary greatly from different light conditions to other variables. This basic form-based approach to hand gesture recognition proposed in this paper can identify around 1-5 different gestures based on the performance of this algorithm. This proposed implemented algorithm was evaluated over 50-60 images and provides approximately 94 percent recognition level. However, the recognition rate remains to be improved at the identification level.

Index Terms: (deaf, dumb, gestures, HGR, SVM, HCI, SIFT, HOG)

1 INTRODUCTION

We still hear of new technology nowadays that enhances our lifestyle and makes life easier for us. Society has been revolutionized by innovation. Human race has engineered a gear and is not in a mood to move the pedals out of the frame. Different areas of technology are undergoing major research, including artificial intelligence, smart phones, and much more. Such work leads to new developments and one's life is promoted. But much less research has been done for the deaf and dumb people. Compared to other fields, this subject has received less publicity. The main challenges this individual faces are the communication gap between the particular person and the normal person. Those who are deaf and dumb are always unable to communicate with normal people. With this enormous challenge, we are dissatisfied and feel discriminated against in society. Thanks to miscommunication, deaf and dumb people believe they don't talk, so they can never express their feelings. HGRVC (Hand Gesture Recognition and Voice Conversion) program locates and records the hand gestures of the dumb and deaf to create a communication channel with others. It is possible to use web camera to display hand gestures. The frames are then converted to standard size using pre-processing. The goal of this paper is to develop a system in which hand gestures can be translated into text. This paper aims to place the objects in the database and to use the corresponding database to translate the image into text. Detection means tracking the motion of the hand. The solution includes text format creation that helps reduce the communication gap between deaf-mute and individuals. Concepts of understanding of gesture and gesture are widely present in human machine interaction. Management is the client's body movement or form of physical activity to communicate some meaningful information.

The Histogram of Gradient Orientation (HOG) features extraction algorithm have been implemented in this paper and these features are used to transfer a training model in a Support Vector Machine (SVM) The test images received by the client should be further categorized according to the respective function vector. Using two feature vectors namely HOG and Scale Invariant Feature Transform (SIFT) features, we developed a system to recognize alphabet characters and checked both to deliver the optimum result. The real-time images will be captured first by webcam, then stored in the file, and feature extraction will be performed on the recently captured image to determine what sign the user has expressed via MATLAB's algorithm. In contrast, the output will be produced to the image already stored in the database for a specific document in accordance with the input image. There are typically two types of gestures. Depending on the level of communication required, different types are used and it is not always appropriate to use one over the other, but it makes things easier to understand. There are two different types of gestures:

Online gestures– Simply put, these are movements that control a computer system or a real-time machine.

Offline gestures– It is known to be offline movements that are viewed after completion.

2. PROCESSION DIGITAL IMAGE:

An image can be defined as a two-dimensional function f(x, y), wherex & y are spatial coordinates, and the amplitude of f at any pair of coordinates (x, y) is called the intensity or gray

- *Nikhat Parveen, Associate Professor in Department of Computer Science and Engineering in Koneru Lakshmaiah Education Foundation, Guntur, A.P., India. PH-9935158356 .E-mail: nikhat0891@gmail.com*
- *Arpita Roy, Assistant Professor in Department of Computer Science and Engineering in Koneru Lakshmaiah Education Foundation, Guntur, A.P., India. PH-9851654138 .E-mail: arpitaroy@kluniversity.in*
- *D. Sai Sandesh is currently pursuing Bachelor's Degree program in Computer Science and Engineering in Koneru Lakshmaiah Education Foundation, Guntur A.P, India. PH- 8096838052. Email: saisandesh3@gmail.com*
- *J.Y.P.R.Sai Srinivasulu is currently pursuing Bachelor's Degree program in Computer Science and Engineering in Koneru Lakshmaiah Education Foundation, Guntur, A.P, India. PH- 9951051536. Email: jyasasvi@gmail.com*
- *N Srikanth is currently pursuing Bachelor's Degree program in Computer Science and Engineering in Koneru Lakshmaiah Education Foundation, Guntur, A.P, India. PH-7660904061 . Email:srikanthnavuluri846@gmail.com*

level of the image at that point. If all finite discrete quantities arex, y & f are amplitude values, we find the object to be a digital image. The DIP field refers to the image processing of digital computers. Digital image consists of a finite number of elements for each element with a specific location and meaning. The pixels are the components. Sight is our system's most complex, so it's no wonder which individual plays the most important role in human sight. Yet, like humans, who are limited to the EM spectrum imaging devices ' visual band span nearly the entire EM spectrum, from gamma to radio waves. We can also work on images created from sources not typically associated with the face of humans. Authors generally disagree where image processing ends and other similar fields such as image analysis and computer vision start. In describing image processing as a procedure in which images are both the input and output of a device, a distinction is sometimes made. An image is defined as a two-dimensional function f(x, y) where x and y are spatial coordinates and the amplitude of ' f ' is called the frequency of the image at that point at any pair of coordinates (x, y).

2.1 Picture of a gray scale:

A gray image is a function I (xylem) of the two spatial co-ordinates of the target plane.

I(x, y) is the frequency of the image at the point (x, y) of the object plane.

I(xylem) assumes that a rectangle[0, a]*[0, b]I:[0, a]*[0, b]**[0, info] I(xylem) assumes non-negative values.

2.2 Coordinate convention:

The sampling and quantization product is a real number matrix. We use two main ways of classifying digital images. Suppose an image f(x, y) is sampled with M rows and N columns in the resulting image. We say the size of the image is M X N. Coordinates (xylem) are numbers that are distinct. We use integer values for these discrete co-ordinates for notational accuracy and versatility. The image source is defined in many image processing books at (xylem)=(0,0). The next coordinate values along the first row of the frame are (xylem)=(0,1). The notation (0,1) is used along the first line to represent the second sample, it is important to remember. It doesn't mean that when the object is measured, these are the real values of spatial coordinates. The following figure shows the structure of the coordinate. Note that x varies in integer proportions from 0 toM-1 and y from 0 toN-1.

2.3 Picture as Matrices:

The previous discussion leads to a digitized picture function representation as follows:

f (0,0) f(0,1)......F(0,N-1) f(1,0) f(1,1)............f(1,N-1) f(xylem)=.F(M-1,0) f(M-1,1)................ f(M-1,N-1)

A digital image can naturally be represented as a matrix of MATLAB: f(1,1) f(1,2)....... f(1,N) f(2,1) f(2,2)........ F(2,N) Member. F= f(M,1) f(M,2)....... f(M, N) Where f(1,1)= f(0,0) (note MATLAB amounts using a monoscope font). The two representations are exactly similar, with the exception of the initial shift. The notation f(p, q) refers to the element in the line p and the line q. For example, f(6,2) is the element in the sixth and second column of matrix f. Usually letters M and N are used to denote, respectively, the number of rows and columns in a matrix. A row vector is called a matrix of 1xN while a vector of the column is called a matrix of Mx1. A 1x1 matrix is a scalar. Matrices in MATLAB are stored in variables with names such as A, a, RGB, real set, etc. Variables must begin with a letter that only includes letters, numerals and underscores. As stated in the previous paragraph, all MATLAB amounts are written using monoscope characters. We use Roman, Italian norm for mathematical expressions such as f(x, y).

2.4 Reading Photos:

Images are read in the MATLAB environment using imread function whose syntax is "imread('filename ')" This is a spring containing the whole image file.

For example, the command line > > f= imread('8.jpg'); reads the JPEG image chest xray in the image array f. For example, > > f= image('D:\myimages\chestxray.jpg'); reads the image on D from a folder called my images, whereas > > f= imread('.\myimages\chestxray.jpg'); reads the photo of the current working folder in my photo subdirectory. The current directory window on the MATLAB desktop toolbar shows MATLAB's current working directory and offers a quick, manual way to change it. Some of the popular formats of image / graphics supported by imread and imwrite.

Picture Types: Four picture types are provided by the toolbox: 1.Image frequency, 2. Binary images, 3. Pictures indexed 4. Color images indexed and RGB.

3. MODELLING

3.1 Hand Detection:

Hand Detection Locating the hand in the video frame is applying the static gesture recognition method. A segmented skin mask of the original image is given at this point. While the segmented skin mask provides us with the regions where the hand might be, the hand is just part of everything that is segmented from the mask. Since the mask captures skin regions, particularly the skin of the arm and face, further processing is needed to find the hand. In particular, we want to find an interchangeable function for the side, which is a simple feature that we can always use. We squeezed the palm of the hand with that mentality for that continuity. The size variance is a lot of the difficulty in finding the palm's location. We didn't have any details about the user's hand location and size (i.e. the closer the user's hand to the display the larger hand appears). We were therefore more interested in the fingers which played a crucial role in the static hand gesture recognition algorithm. We used the hand geometry to solve this problem. The largest cross-section area of any human hand lies somewhere near the middle of the palm, regardless of size. We use the centroid formula with this property to find the arithmetic mean location of the hand corresponding to the middle of the palm. Upon reaching the middle of the hand, we can easily distinguish between top and bottom of the hand. We given a robust algorithm to detect the center of any human hand by using this tool.

3.2 Static Hand Gesture Recognition:

The task becomes to identify different hand gestures after the position of the hand is identified. Which type of features to use in distinguishing different static movements is the main difficulty in this. We concluded that most of the static movements depended on the fingers of the hand. Although a simple inference, our algorithm was designed around this very deduction, while more complicated algorithms for contour detection and gesture recognition.

3.3 Matlab :

MATLAB is a high-performance computing language. This combines calculation, visualization, and programming in a user-friendly environment where common mathematical notation communicates problems and solutions. Typical implementations include:

- Math and computation
- Algorithm development
- Data acquisition
- Modelling, simulation, and prototyping
- Data analysis, exploration, and visualization
- Scientific and engineering graphics

Application development, including graphical user interface building.

3.3.1 Why Matlab?

MATLAB is a popular computing and programming language environment. Matlab's definition applies to the entire package, including the IDE. The standard library does not include as many generic programming features, but contains matrix algebra and extensive data processing and plotting software.You will need the NumPy, SciPy and Matplotlib packages to get similar functionality in Python. Scipy is a kit that aims to provide all of Matlab's other functions, including those in the Matlab toolboxes. Nonetheless, Simulink is one example that is not protected by Python. You will probably stick to Matlab if you rely on it. Perhaps a Python in the future alternative will be created.

4. RELATED WORK

Using MATLAB, a method to derive the characteristics of hand gestures from the angle and peak measurement approach and then use MATLAB integrated control to convert the recognized gesture into expression. Unlike the American sign language used, an Indian hand sign language device is used to make a gesture using both hands. Their system is implemented using MATLAB without the use of any other external hardware for the user, where the live image runtime is captured after image frames are extracted and image processing is applied using HIS model. For most of the hand signs, the results of this template are found to be satisfactorily perfect. A SVM-based recognition system to extract the feature in which the feature is extracted and processed after the pre-processing part of the image in the Nx2 matrix. This matrix is also used to suit the server's photo. The software has some disadvantages. The very light brown history to somewhat dark brown gives error as it is considered in the skin segmentation value spectrum. But that's going to be the outcome. The system works on MATLAB for hand gesture recognition, they took a real-time image as input and used the correlation-based approach after the image processing level was introduced for mapping purposes. We reviewed existing papers related to this subject in the Literature Review and try to understand the behavior of the existing system. SEMG-based leadership identification with integrated digital hand positions and antagonistic learning by Mohan Kankanhalli and Weidong Geng [1] reported that We present a novel hybrid approach combining actual SEMG signals with corresponding digital hand poses to improve the accuracy of gesture recognition based on surface electromyography (SEMG). Virtual hand poses are created to capture the intrinsic relationship between sEMG signals and hand poses through a proposed cross-modal association model based on adversarial learning. The experimental results show that major improvements were made compared to existing playbacks with the new method Recognition of SEMG movements. The accuracy of recognition of the proposed model is improved by an average of 5.2 percent in sparse multichannel for frame-based sEMG gesture recognition. sEMG databases and by an average of 6.7% in high density, the average enhancement of the proposed window-based approach method is + 2.5%.Based on Zhengjie Wang's Active Ultrasonic Sensing of Smartphone, Yushan Hou said [2] Hand gesture recognition has gained broad interest in the area of omnipresent computing with the rapid development of the Internet of Things as it provides us with a simple and natural form of human-computer interaction. Next, we're investigating the new acoustic signals-based hand gesture recognition work. We then present and describe the basis of the ultrasonic signal characteristics. We then focus on the standard methods used in these studies and provide a detailed analysis of methods for signal generation, extraction of features, pre-processing and recognition. First, we use adaptive gesture recognition and hand tracking smartphones to explore and examine in-depth the state-of - the-art hand gesture recognition technologies. To order to gain some insight into the development of the ultrasonic gesture recognition system, both systems are then discussed from signal generation, signal processing, and performance evaluation. Finally, on the basis of the ultrasonic signal of the device, we conclude by addressing the potential barriers, perspective and open hand gesture recognition issues. Recognition of static and adaptive hand management in detail using dynamic time warping Guillaume Plouffe and Ana-Maria Cretu [7] said this paper explored the design of a natural gesture user interface that monitors and recognizes real-time hand gestures based on Kinect sensor depth information. A new algorithm is proposed to improve the scanning time to identify the first pixel on the contour of the hand within this area. A directional search algorithm allows the detection from this pixel of the entire contour of the hand. The k-curvature algorithm is then used to locate the fingertips across the contour, and dynamic time warping is used to select gesture candidates and also to distinguish gestures by contrasting an observed gesture with a set of pre-recorded reference gestures. In the area of interest, the method discusses both static and dynamic gestures and multiple hands. A total identification level of 92.4 percent is reached between 55 static and dynamic movements. Two possible applications of this work have been discussed and evaluated: one for the interpretation of sign digits and movements for a more relaxed interaction between humans and machines and the other for the natural regulation of sign digits. Tiny Hand Management Recognition without Localization by Peijun Bao, Ana I. Maqueda, Carlos R. [3] Strong Convolutionary Network. Del-Blanco and Narciso García claimed that for human-computer communication interfaces, Visual hand-gesture recognition is increasingly required. Hands occupy only about 10 percent of the picture in many applications, while most of it includes the context, human face, and human body. Spatial hand position in such situations could be a challenging task, and ground-based truth bounding boxes need to be given for learning that is not normally available. Nevertheless, the position of the hand is not a necessity when the criteria are merely a gesture to operate a consumer electronics device, such as mobile phones and televisions. A deeply convolutionary neural network is proposed in this paper to classify hand gestures directly into artifacts without any stage of segmentation or detection that could discard non-hand areas. The designed hand-gesture recognition network can classify seven types of hand gestures in a user-independent manner and in real time, achieving 97.1 percent accuracy with simple data set backgrounds and 85.3 percent with complex data set backgrounds. Cooperative Sensing and Wearable Computing for Sequential Hand Management Recognition Xiaoliang Zhang, Ziqi Yang, Taiyu Chen, Diliang Chen, Ming-Chun Huang [4] reported that Hand gestures recognition (HGR) was considered one of the main human-computer interaction (HCI) research areas. By training and testing the collected IMU (Inertial Measurement Unit), Electromyographic (EMG), finger and palm pressure data, a LSTM algorithm for deep learning was calculated to create an effective model for recognizing hand gestures. Our experimental results showed the proposed LSTM algorithm's outstanding classification efficiency. Findings have positive implications for the identification of sequential hand gesture and the status of HCI science. The author included the algorithm in which the video was first captured and then divided into different frames and the frame with the image was extracted from which various features such as Gaussian Difference (DOG) was measured and the author used a combination of shape-based (2D) and size-based features to identify the hand configuration in the scene. With a 3D camera, the researcher obtained the tracking data over time, And the author establishes a SVM-based system for the identification of movements between one finger and two fingers. The researchers introduced Hidden Markov models, widely used in handwriting recognition, to achieve a 97 percent + accuracy in classifying 40 words in American Sign Language. A program that developed a model for gesture recognition using SVM and obtained better results using Cross Validation. Using libSVM, movements using Multiclass SVM were categorized. Current solutions include a variety of machine learning techniques for hand-classification. As a function for a cross-validated 7030 SVM, the author used a Histogram Pyramid of Oriented Gradients, which was able to distinguish between a hand and non-hand.

4.1 Applications Involved
Besides being practical, it is not possible to argue against the cool appearance of gesture control. Like the Minority Report hit movie in 2002, a time will probably come when all is managed by gestures. While it may seem like a technology that will only increase our lethargy, the truth is that besides making life easier, it also has a wide range of applications in nearly every field. Some of the requests are:

- Health applications
- Alternative computer interfaces
- Entertainment applications
- Automation systems

4.2 Challenges and Issues faced
During the study of Gesture recognition , we have came across different challenes faced by Human Computer Interaction. The challenges and issues are listed below:

4.2.1 Usability challenge
We describe the challenges encountered by the use of gesture recognition software by the user.

4.2.2 Training cycle:
User training is a time-consuming task, the user is taught to manage the program as well as monitor his / her movements in the preliminary stage, while the learned system is used to learn the classifier used during the configuration phase.

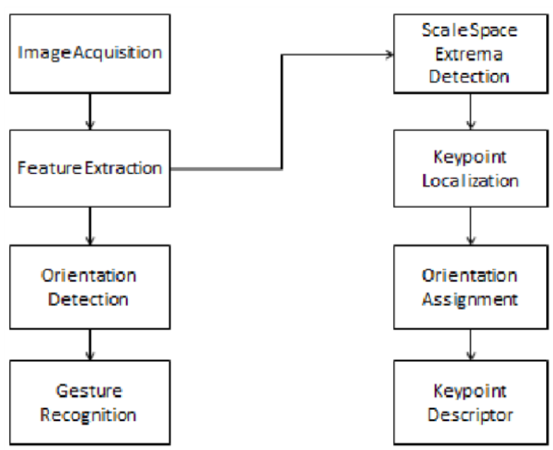
4.2.3 Technical challenges:
These are issues related to the recorded movements of the device, including non-linearity, noise and non-stationarity, small training sets, and the company's dimensionality curse.

4.2.4 Non-linearity:
The problem of classification of form in computer vision is considered to be more complicated than any other problem of classification, as the features derived from shapes can sometimes lead to very large dimensions. Human beings are able to visualize the classification question up to three dimensions at a time, very high dimensions usually leave the human consumer without any insight into the high dimensionality.

4.2.5 Non-stationary and Noise:
The gestures ' non-stationary and noise characteristic of non-stationarity is a major problem in the design of an HCI process. It creates a constant switch between or within the recording sessions of the used movements over time.

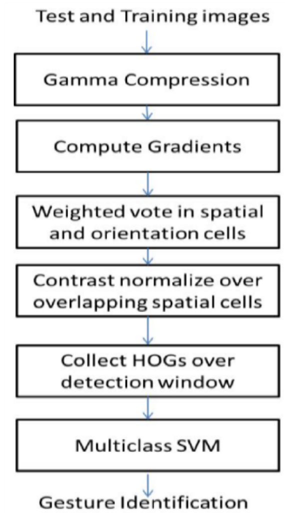
5. ALGORITHMS

5.1 SIFT
Scale-invariant feature transform (SIFT) is an algorithm for computer vision to classify and describe local object features.

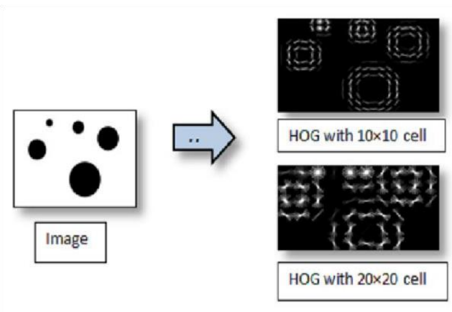


Flowchart of SIFT Technique

5.2 HISTOGRAM OF GRADIENTS:
HOG is a function descriptor used to identify artifacts in image processing. In localized portions of an image, the technique counts occurrences of gradient orientation.



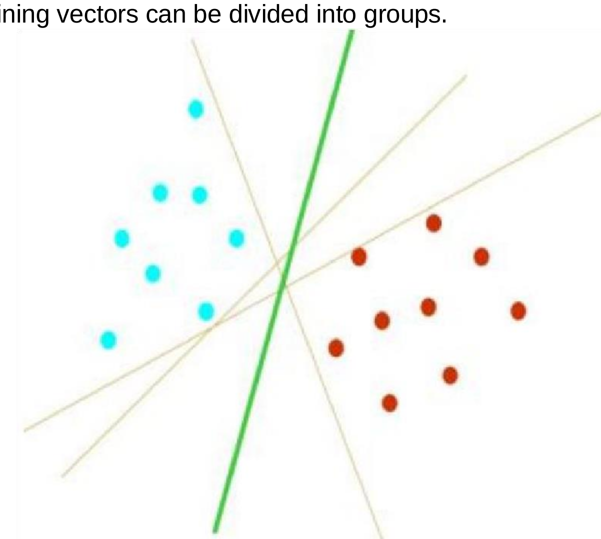
Flowchart of Histogram of Oriented Gradients



Visualization of HOG

5.3 SUPPORT VECTOR MACHINE:

For classification and regression testing, the Support Vector Machine (SVM) Classifier is commonly used. SVM learning algorithm generates a template predicting whether a new example falls into one or the other class. When marked as belonging to their respective categories, the SVM classifier learns from the data points in instances. The SVM is a well-known classification problem learning algorithm. SVM's main goal is to develop an ideal hyper plane to isolate so that training vectors can be divided into groups.



SVM hyper-plane

SVM then separates the specified binary marked training data from the hyperplane, which is also known as the total margin hyperplane as far away as possible from them. We can operate in conjunction with the kernel technique in cases where no linear separation is feasible, which immediately performs a non-linear mapping to a function space.

5.4 IMPLEMENTATION OF ALGORITHM:

Implementing machine learning and image processing techniques, we hope to achieve a high degree of accuracy in the distinction between the letters in the Indian sign language alphabet.

5.4.1 Gesture Segmentation: By implementing color-based object detection, we use blue glove to isolate the gesture from the background to differentiate hand gesture from the background image.

Steps:-

a) Input test images are further limited to a resolution of 320x240 pixels.

B) The 3D image's blue plane is subtracted from the original image's gray scale. Later median filtering is achieved while retaining the edges to reduce the salt and pepper noise.

Original image

Threshold image

Output with Otsu

Canny Output

5.4.2 Feature vectors: To measure feature vectors, as described above, HOG and SIFT algorithms have been implemented. Feature vectors provide the vector-shaped behavior of all the pixels.

Cell size=8, block size=2, orientation bins=9, oriented gradients=1 fixed dimension feature length was obtained for further calculation for the HOG parameters included in the code. In a pad, these vectors have been saved. Data. The data. Another function vector SIFT has been introduced to compare the output. Such two function vectors have been used separately to define and check the output. Similarly feature vectors are calculated for different gestures that we used to train the model for classification using the dataset available on Cambridge Hand Gesture Dataset. Training images were segmented into different files and the listed feature vectors were stored in a matrix folder with the correct gesture marking.

5.4.3 How does the SVM work?

In n-dimensions, the general formula of a plane is w. Fx= b wherex is a nx1 vector. One has minimal distance d min perpendicular from the hyperplane to the origin of all the points on the plane.

Patterns of training(x1,y1),. If xi is a d-dimensional vector, and y I= 1 if xi is a group 1, then y I= -1 if xi is a group 2 and so on.
There is a d-dimensional vector w and a scalar b if the data are linearly separable.

Multiclass SVMs : The basic SVMs are for a query of two groups. However, to classify it into more than two categories, it should be expanded to multiclass. There are two basic methods for addressing SVMs q-class issues.
Multi-class SVMs: One to Another Take the samples of learning with the same tag as one class and the other class, and then it becomes a question with two classes.
Multi-class SVMs: Pair wise SVMs q2machines are equipped for q-class issues in the pair wise method. The pair of wise classifiers are arranged in trees where a SVM is represented by each tree node. The one-to-other approach is better with regard to the learning effort, since only q SVMs have to be trained in the pair wise approach compared to q2 SVMs. All approaches, however, include evaluation of q-1 SVMs at runtime. Recent experiments on identification of people show similar performance in the classification of the two strategies. The figure below gives a brief idea of implementing multi-class SVM using binary classifiers in the top-down or bottom-up approach

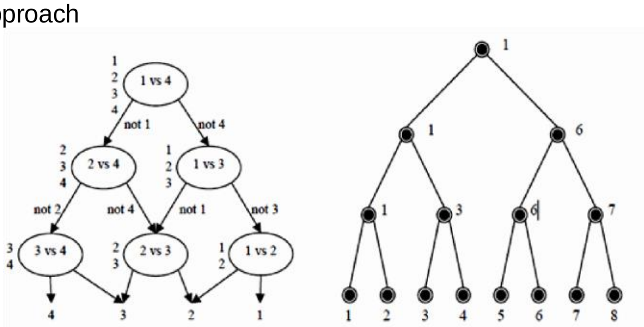
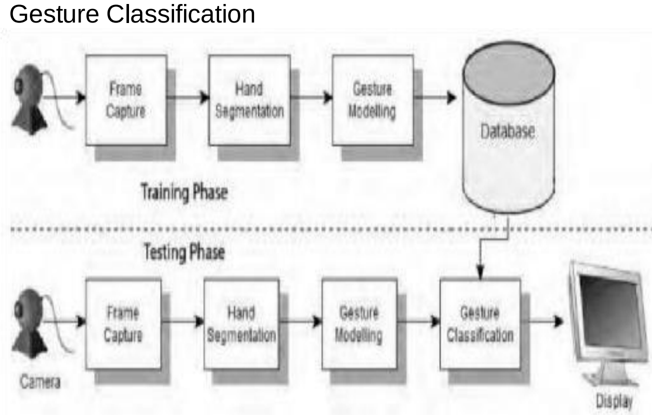


Fig Example of top-down approach Fig. Example of bottom-up approach

6. METHODOLOGY

In general, hand gesture recognition based on vision is made up of using these basic processing stages:
A. Data Collection
B. Image Segmentation
C. Gesture Modelling
D. Gesture Classification

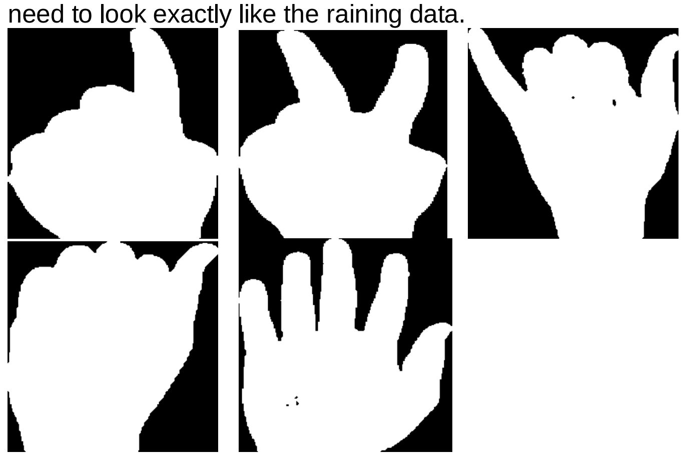


Project Model

In two stages, namely the learning stage and the evaluation process, the diagram above indicates both phases. The training stage only comprises the first two processing stages, while the test phase covers all three processing phases.

A. DATA COLLECTION

Our dataset consists of over 5 separate directories for robust classification with selection of hand gestures of different orientations. The test images were taken at resolutions of 320x 240pix from a standard laptop webcam. The hand gestures are shifted or subtly rotated when collecting photographs in order to avoid taking sets of images that were too close. In adding some variation into the dataset, we ensure that in order to be correctly categorized, any new data we check does not need to look exactly like the raining data.



Our Database collection of images

B. IMAGE SEGMENTATION

The key technique for hand segmentation, is to identify and distinguish hand regions from backgrounds in the image captured by hand gesture. The accuracy of hand gesture recognition is closely linked to the accuracy of hand segmentation. Most conventional methods of manual segmentation use color signals. Nonetheless, due to the presence of similar skin colored objects, the accuracy of hand gesture segmentation appears to be easily influenced by several factors such as skin color variations between humans, color sensitivity to lighting, and particularly the circumstance. We used the color detection scheme here to detect only a specific color.

C. GESTURE MODELLING

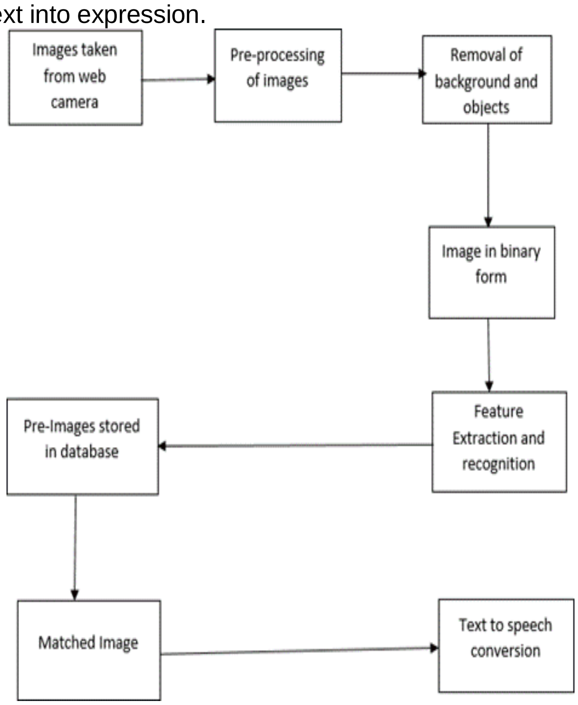
Hand postures as well as movement patterns are determined from the hand gesture frame series during the hand gesture analysis process, and the hand gesture template is generated accordingly, in the training and testing phases.

D. GESTURE CLASSIFICATION

The final stage is hand gesture recognition in which the output from the second stage of the current gesture model is compared to each model in the hand gesture database in which the most matched hand gesture is selected as the final result of recognition. This was shown in testing phase. Different methods of modeling hand gesture have different approaches to recognition. Through counting the number of active fingers, the hand gesture is remembered. The movement of the hand is modelled as the skeleton of the star, and distance signature is used to recognise it. Certain characteristics such as position and direction of the hand, position and direction of the finger and distance between the fingers are always used establishing spatial model of hand gesture.

7. SYSTEM ARCHITECTURE:

The figure below shows the system architecture of our system, which basically shows which component of the system, how the system works, and the process flow, etc. Web camera images go through pre-processing stages to boost the quality of an image. Then there is a deletion of target and context representations which later turns into binary form. Extracting and reorganizing functionality helps to suit the images stored in the database and we get the desired text output and convert the text into expression.



7.1 IMPLEMENTATION OF THE SYSTEM

7.1.1 Image Pre-Processing

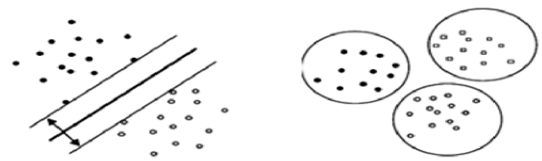
To boost the resources of an image, the captured objects are being pre-processed. In essence, pre-processing is done to remove the object and meaning of an image and focus only on the movements of the face. The pre-processed image is then depicted as black and white pixels, which essentially means binarized image.

7.1.2 Feature Extraction and Recognition

An algorithm is used to determine an image's characteristic. To retrieve the best featured image from the database, the algorithm applied to the captured images. It then transforms the objects into some separate linear set of variables corresponding to the information in the original data referred to as the main components.

7.1.3 Hand Gesture Recognition using Support Vector Machine:

Because different hand movements are needed to classify (and not just two), as shown in the figure below, multiple binary SVMs need to be combined in a Classifier Tournament structure. Training to a binary classifier is necessary for each pair of classes to be categorized between them. In the pre-processing stage, computed image features will be entered in all binary classifiers. Each binary classifier returns a number of classes representing a gesture. The motion that has the most classifier votes wins and is the product of the SVM multiclass. On the large-scale classification of data, MEB-SVM (Minimum Enclosure Ball-SVM) was implemented. Compared to others, the algorithm proved to be successful with less computational time. Several techniques have been applied in that article, such as mean shift and Fourier descriptor. Such approaches improved the efficacy of hand gesture recognition. The MEB-SVM key algorithm demonstrates strong when dealing with the identification of multiple classes.



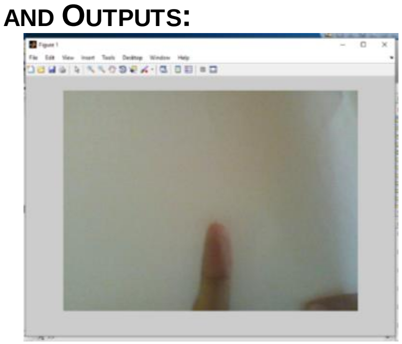
Typical SVM model ***MEB model***

The following steps are used to remove an image's function using the algorithm:

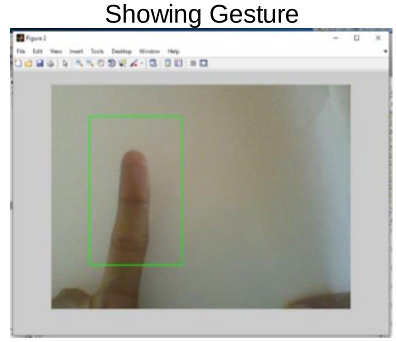
Phase 1: Transform all images into the matrix of the column.
Phase 2: Assess the matrix of the line matrix in the mean column.
Phase 3: For each set of vectors, measure the difference.
Phase 4: Calculate a matrix of covariance.
Phase 5: Calculate the value of the origin and mean value of the origin for the matrix of covariance.
Phase 6: Order the value of Eigen.
Phase 7: Calculate the mapping of your own vectors and project the data to suit.

We get the reduced features of extracted images after applying all of the above steps and then calculate the result.

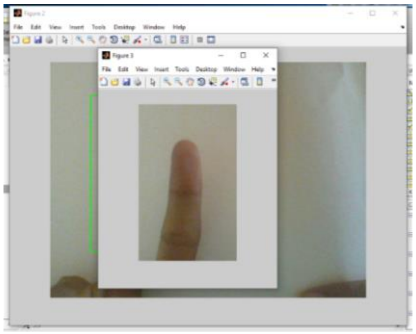
8. RESULT AND OUTPUTS:



Showing Gesture



Captured Gesture Image



Masked Image

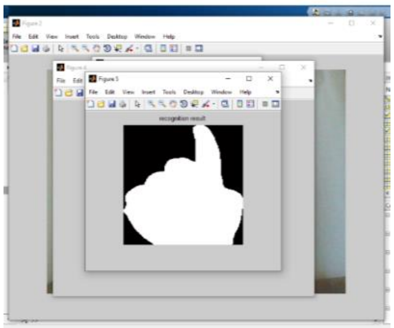
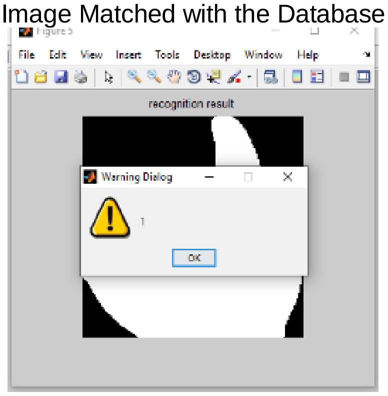


Image Matched with the Database



Recognized Gesture Value

There is a continuous video frame running once the process started. On Showing the Real Time Gesture before the web cam, the gesture will be captured by the system. The system captures the gesture in the prescribed frame. The image in the frame is masked and the background noise is removed. The masked image is matched with one of the grayscale images that is similar to the gesture which we have shown before the web cam in the database. Finally the recognized gesture value is displayed in the dialogue box.

9. CONCLUSION

This paper used images of depth to discuss the creation of a new method of recognition of gestures. The hand contour's initial pixel is found using an improved block search scheme starting from the scene's closest pixel and a spatial search is performed to find the entire hand contour. The machine achieves an average output of 92.4 percent to deal with one or more static and dynamic hand movements at the same time. The skin's context, lighting conditions, clothes, and colour have little impact on the results of the proposed solution. There's no absolute conclusion because for all the alphabets we didn't implement the algorithm. The implemented framework is effective in adjusting the lighting as well as the gesture orientation. Many implementations using SVMs have shown that a SVM-based problem-solving approach is better than a one-to - one comparison method. SVM is a new technique, but it has been applied to a wide range of machine learning tasks and has been used to produce a lot of learning code with the correct choice of kernel function.

REFERENCES

[1] Yu-Hu, Yongkang Wong(Member of IEEE), Embedded Digital Hand Poses And Adversarial Training, Aug 2019.
[2] Zhngjie wang, Yushan Hou, Kangkang Jiang, Wenwen Dou, Chenming Zhang, Zehu Huang, and Yinjing Guo, “Recognition based on Active sonic sensing” Aug 26,2019.
[3] Peijun Bao, Ana L Maqueda, Carlos R. del-Blanco and Narciso Garci, “Tiny Gestures without localization, a deep convolutional network”, 2017 IEEE.
[4] Xiaoliang Zhang, Ziqi Yang, Taiyu chen, Diliang chen, Ming-chun Huang, “Cooperative Sense and Wearable Computing for Sequential Recognition”, 2018 IEEE.
[5] Shinde, Shweta S., M. Autee Rajesh, and K. Bhosale Vitthal. "Real time two way communication approach for hearing impaired and dumb person based on image processing." International IEEE Conference 2016 on Computational Intelligence and Computing Research (ICCIC). IEEE, 2016.
[6] Sood, Anchal, Mishra and Anju. ' AAWAAZ: a deaf and dumb communication system. ‘Reliability, Infocom Technology and Optimization (Trends and Future Directions)(ICRITO), 2016 5th International Conference. IEEE, this year.
[7] Recognition of static and adaptive hand management in detail by Guillaume Plouffe and Ana-Maria Cretu, Member of the IEEE 2015.
[8] Computing Communication Control and Automation (ICCUBEA), 2015 International Conference on "Two Ways Communicator between Deaf and Dumb People and Normal People."
[9] R. K., V. Valliammai, and S. Shangeetha. Padmavathi. "Indian Sign Language Character Recognition computer vision-based approach." Machine Vision and Image Processing (MVIP), 2012 International Conference on. IEEE, 2012.
[10] Kumari, Sonal, K. Mitra, Suman. Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), 3rd National Conference on 2011. IEEE, 2011.
[11] Shariz, S. and Kulkarni, A. "Identifying Hand Con_gurations with Low Resolution Depth Sensor Data", CS229 Final Project Paper, 2009.
[12] Ruslan Kurdyumov, Phillip Ho, Justin Ng, “Sign Language Classification Using Webcam Images”, December 16, 2011.
[13] Chih-Wei Hsu, Chih-Chung Chang, and Chih-Jen LinDepartment of Computer Science National Taiwan University, Taipei106, Taiwan, Last updated: April 15, 2010.
[14] David G. Lowe, “Distinctive Image Features from ScaleInvariant Keypoints” January 5, 2004.
[15] Starner, T. and Pentland, A. "Real-Time American Sign Language Recognition Using Desk and Wearable Computer Based Video", IEEE Transactions on Pattern Analysis and Machine Intelligence Vol. 20, No. 12, December 1998.

[16] Nikhat Parveen, Arpita Roy, Yagna Srilatha, Chaitanya, Hemanth kumar, “A Fuzzy Logic Based System for Secure Scrum” Ijrte, November 2019.
[17] Arpita Roy, Nikhat Parveen, P.Ramya Bhargavi, A.Navya, A.Pavan Kumar,”Disciplinary Control System Based on Naïve Bayes Classification Technique using JSP Servlets”, December 2019.