International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

Real Time Control System Based on Hand Gesture Detection and Recognition

Shalini[1], Dr. Rekha Patil[2]

[1]Department of Computer Science and Engineering, PDA College of Engineering, Kalaburagi, Karnataka, India

[2]M.Tech in Computer Science and Engineering, PDA College of Engineering, Kalaburagi, Karnataka, India

Abstract: Hand gesture recognition techniques have been studied for more than two decades. Several solutions have been developed; however, little attention has been paid on the human factors, e.g. the intuitiveness of the applied hand gestures. This study was inspired by the movie Minority Report, in which a gesture-based interface was presented to a large audience. In the movie, a video-browsing application was controlled by hand gestures. Nowadays the tracking of hand movements and the computer recognition of gestures is realizable; however, for a usable system it is essential to have an intuitive set of gestures. The system functions used in Minority Report were reverse engineered and a user study was conducted, in which participants were asked to express these functions by means of hand gestures. We were interested how people formulate gestures and whether we could find any pattern in these gestures. In particular, we focused on the types of gestures in order to study intuitiveness, and on the kinetic features to discover how they influence computer recognition. We found that there are typical gestures for each function, and these are not necessarily related to the technology people are used to. This result suggests that an intuitive set of gestures can be designed, which is not only usable in this specific application, but can be generalized for other purposes as well. Furthermore, directions are given for computer recognition of gestures regarding the number of hands used and the dimensions of the space where the gestures are formulated.

Keywords: Hand gesture recognition, Local binary pattern(LBP), K nearest neighbour algorithm(KNN), Eigen classifier**.**

1. Introduction

Several successful approaches to spatio-temporal signal processing such as speech recognition and hand gesture recognition have been proposed. Vision based gesture recognition system is the attractive solution for human computer interaction and machine vision application like robotic application. Most of them involve time alignment which requires substantial computation and considerable memory storage.

Due to congenital malfunctions, diseases, head injuries, or virus infections, deaf or non- vocal individuals are unable to communicate with hearing persons through speech. They use sign language or hand gestures to express themselves, however, most hearing persons do not have the special sign language expertise. Hand gestures can be classified into two classes: (1) static hand gestures which relies only the information about the angles of the lingers and (2) dynamic hand gestures which relies not only the fingers' flex angles but also the hand trajectories and orientations. The dynamic hand gestures can be further divided into two subclasses. The first subclass consists of hand gestures involving hand movements and the second subclass consists; of hand gestures involving fingers' movements but without changing the position of the hands. That is, it requires at least two different hand shapes connected sequentially to form a particular hand gesture. Therefore samples of these hand gestures are spatio-temporal patterns. The accumulated similarity associated with all samples of the input is computed for each hand gesture in the vocabulary, and the unknown gesture is classified as the gesture yielding the highest accumulative similarity.

Developing sign language applications for deaf people can be very important, as many of them, being not able to speak a language, are also not able to read or write a spoken language. Ideally, a translation system would make it possible to communicate with deaf people. Compared to speech commands, hand gestures are advantageous in noisy environments, in situations where speech commands would be disturbing, as well as for communicating quantitative information and spatial relationships. A gesture is a form of non-verbal communication made with a part of the body and used instead of verbal communication (or in combination with it).

Most people use gestures and body Language in addition to words when they speak. A sign language is a language which uses gestures instead of sound to convey meaning combining hand-shapes, orientation and movement of the hands, arms or body, facial expressions and lip-patterns. Similar to automatic speech recognition (ASR), we focus in gesture recognition which can be later translated to a certain machine movement. The goal of this project is to develop a program implementing real time gesture recognition. At any time, a user can exhibit his hand doing a specific gesture in front of a video camera linked to a computer. However, the user is not supposed to be exactly at the same place when showing his hand. The program has to collect pictures of this gesture thanks to the video camera, to analyze it and to identify the sign. It has to do it as fast as possible, given that real time processing is required. In order to lighten the project, it has been decided that the identification would consist in counting the number of fingers that are shown by the user in the input picture. We propose a fast algorithm for automatically recognizing a limited set of gestures from hand images for a robot control application. Hand gesture recognition is a challenging problem in its general form. We consider a fixed set of manual commands and a reasonably structured environment, and develop a simple, yet effective, procedure for gesture recognition. Our approach contains steps for segmenting the hand region, locating the fingers and finally classifying the gesture. The algorithm is in variant to translation, rotation, and scale of the hand .We can even demonstrate the effectiveness of the technique on real imagery.

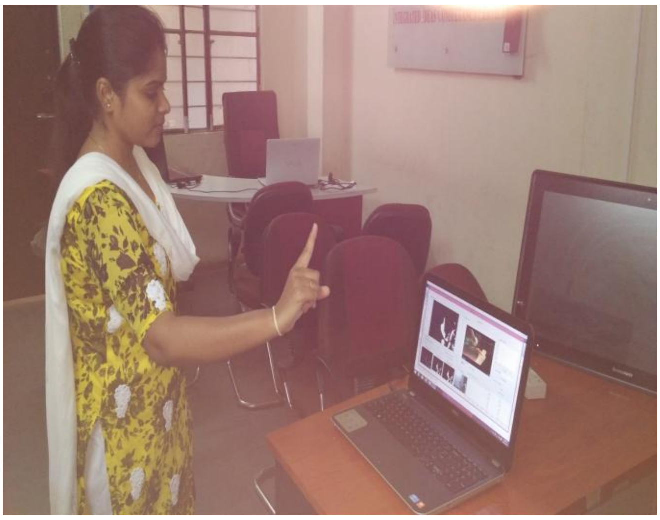

Figure 1: Real time gesture recognition

This paper deals with identification of gesture in a real time with an application like Slide Show Control/ Windows media player control. Figure1 shows gesture recognition by background separation.

The paper is organized as: section 1 discusses Introduction, section 2 discusses the Related Work, section 3 discusses Proposed Work, section 4 discusses Results and section 5 discusses Conclusion.

2. Related work

Jaroslaw Szewinski, Wojciech Jalmuzna, [1] deals with the description of the various algorithms used in Neural Networks viz •feed-forward (FF) •feedback (FB) •adaptive feed-forward (AFF). In this paper, the adaptive GPC algorithm is extended when the disturbance measurement signal is available for feed forward control. First, the adaptive feedback and feed forward GPC algorithm is presented when the disturbance is stochastic or random. Second, the adaptive algorithm is further extended when the disturbance is deterministic or periodic. Asanterabi Malima, Erolozgur, and Mujdatcetin [2] The above approach contains steps for segmenting the hand region, locating the fingers ,and finally classifying the gesture. The algorithm is invariant to translation, rotation, and scale of the hand.This algorithm can be extended in a number of ways to recognize a broader set of gestures. The segmentation portion of algorithm is too simple, and would need to be improved if this technique would need to be used in challenging operating conditions. Reliable performance of hand gesture recognition techniques in a general setting require dealing with occlusions, temporal tracking for recognizing dynamic gestures, as well as 3D modeling of the hand, which are still mostly beyond the current state of the art. Mark Batcher[3] Gripsee is the name of the Robot of whose design is discussed in the paper ,it is used for identifying an object, grasp it, and moving it to a new position. It serves as a multipurpose Robot which can perform a number of tasks, it is used as a Service Robot. Kevin Gabayan, Steven Lansel [4] This paper deals with the dynamic time warping gesture recognition approach involving single signal channels. Exemplar, a sensor interaction prototyping software and hardware environment, currently uses a dynamic time warping gesture recognition approach involving single signal channels. Author use a five channel accelerometer and gyroscope combination board to sample translational and rotational accelerations, and a microcontroller to perform analog to digital conversion and relay incoming signals. Template matching via linear time warping (LTW) and dynamic time warping (DTW) are performed offline, as well as reinforcement learning via Hidden Markov Models (HMM) in real-time. M. Ebrahim Al-Ahdal & Nooritawati Md Tahir [5]This paper presents an overview of the main research works based on the Sign Language recognition system, and the developed system classified into the sign capturing method and recognition techniques is discussed. The strengths and disadvantages that contribute to the system functioning perfectly or otherwise will be highlighted by invoking major problems associated with the developed systems. Next, a novel method for designing SLR system based on combining EMG sensors with a data glove is proposed. This method is based on electromyography signals recorded from hands muscles for allocating word boundaries for streams of words in continuous SLR. Iwan Njoto Sandjaja and Nelson Marcos [6] Sign language number recognition system lays down foundation for handshape recognition which addresses real and current problems in signing in the deaf community and leads to practical applications. The input for the sign language number recognition system is 5000 Filipino Sign Language number video file with 640 x 480 pixels frame size and 15 frame/second. The color-coded gloves uses less color compared with other color-coded gloves in the existing research. The system extracts important features from the video using multi-color tracking algorithm which is faster than existing color tracking algorithm because it did not use recursive technique. Next, the system learns and recognizes the Filipino Sign Language number in training and testing phase using Hidden Markov Model. The system uses Hidden Markov Model (HMM) for training and testing phase. The feature extraction could track 92.3% of all objects. The recognizer also could recognize Filipino sign language number with 85.52% average accuracy. Noor Adnan Ibraheem and Rafiqul Zaman Khan [7] In this paper a survey on various recent gesture recognition approaches is provided with particular emphasis on hand gestures. A review of static hand posture methods are explained with different tools and algorithms applied on gesture recognition system, including connectionist models, hidden Markov model, and fuzzy clustering. Challenges and future directions are also highlighted. Archana S. Ghotkar, Rucha Khatal , Sanjana Khupase, Surbhi Asati & Mithila Hadap [8] In this paper, some historical background, need, scope and concern of ISL are given. Vision based hand gesture recognition system have been discussed as hand plays vital communication mode. Considering earlier reported work, various techniques available for hand tracking, segmentation, feature extraction and classification are listed. Vision based system have challenges over traditional hardware based approach; by efficient use of computer vision and pattern recognition, it is possible to work on such system which will be natural and accepted, in general. Paulraj M P, Sazali Yaacob, Mohd Shuhanaz bin Zanar Azalan, Rajkumar Palaniappan [9] presents a simple sign language recognition system that has been developed using skin color segmentation and Artificial Neural Network. The moment invariants features extracted from the right and left hand gesture images are used to develop a network model. The system has been implemented and tested for its validity. Experimental results show that the average recognition rate is 92.85%. Nasser H. Dardas and Emil M. Petriu[10] presents a real time system, which includes detecting and tracking bare hand in cluttered background using skin detection and hand postures contours comparison algorithm after face subtraction, and recognizing hand gestures using Principle Components Analysis (PCA). Divya Deora1, Nikesh Bajaj,k [11] Every Sign Language Recognition (SLR) System is trained to recognize specific sets of signs and they correspondingly output the sign in the required format. These SLR systems are built with powerful image processing techniques. The sign language recognition systems are capable of recognizing a specific set of signing gestures and output the corresponding text/audio. Most of these systems involve the techniques of detection, segmentation, tracking, gesture recognition and classification. This paper proposes a design for a SLR System. Solomon Raju Kota, J.L.Raheja [12] Author present an approach for the detection and identification of human gestures and describe a working, near gesture recognition system and then recognize the person by comparing characteristics of the gesture to those of known individuals. The author approach treats gesture recognition as a two dimensional recognition problem, taking advantage of the fact that gestures are normally upright and thus may be described by a small set of 2-D characteristics values. With minimal additional effort PCA provides a roadmap for how to reduce a complex data set to a lower dimension to reveal the sometimes hidden, simplified structure that often underlie it. Yikai Fang, Kongqiao Wang, Jian Cheng and Hanqing Lu [13] Author proposed a robust real-time hand gesture recognition method. In this method, firstly, a specific gesture is required to trigger the hand detection followed by tracking; then hand is segmented using motion and color cues; finally, in order to break the limitation of aspect ratio encountered in most of learning based hand gesture methods, the scale-space feature detection is integrated into gesture recognition. Applying the proposed method to navigation of image browsing, experimental results show that our method achieves satisfactory performance.J. H. Kim, N. D. Thang, and T. S. Kim [14] author have developed a 3-D hand motion tracking and gesture recognition system via a data glove (namely the KHU-1 data glove consisting of three tri-axis accelerometer sensors, one controller, and one Bluetooth). The KHU-1 data glove is capable of transmitting hand motion signals to a PC through wireless communication via Bluetooth. Also we have implemented a 3-D digital hand model for hand motion tracking and recognition. The implemented 3-D digital hand model is based on the kinematic chain theory utilizing ellipsoids and joints. Finally, author have utilized a rule-based algorithm to recognize simple hand gestures namely scissor, rock, and paper using the 3-D digital hand model and the KHU-1 data glove. Some preliminary experimental results are presented in this paper.

J. Weissmann and R. Salomon [15] This paper explores the use of hand gestures as a means of human-computer interactions for virtual reality applications. For the application, specific hand gestures, such as "fist", "index finger", and "victory sign", have been defined. Most exisiting approaches use various camera-based recognition systems, which are rather costly and very sensitive to environmental changes.

3. Proposed work

For developing the code, and the whole algorithm, it was preferable to use Matlab. Indeed, in this environment, image displaying, graphical analysis and image processing turn into a simple enough issue concerning the coding, because Matlab has a huge and the fact that Matlab is optimized for matrix-based calculus make any image treatment more easier given that any image can be considered as a matrix. The starting point of the project was the creation of a database with all the images that would be used for training and testing. The image database can have different formats. Images can be either hand drawn, Digitized photographs or a 3D dimensional hand. Photographs were used, as they are the most realistic approach. Here the images are taken with a web camera. This meant that they have different sizes, different resolutions and sometimes almost completely different angles of shooting. Images belonging to the last case were very few but they were discarded, as there was no chance of classifying them correctly. Two operations were carried out in all of the images. They were converted to gray scale and the background was made uniform.

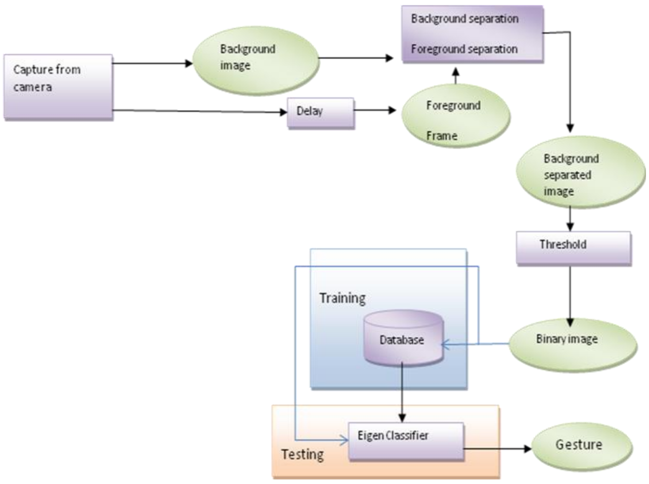

Figure 2: Proposed system of sign recognition

Block diagram of proposed system is shown in the figure 2. The database itself was constantly changing throughout the completion of the project as it was it that would decide the robustness of the algorithm. Therefore, it had to be done in such way that different situations could be tested and thresholds above which the algorithm didn't classify correct would be decided. The construction of such a database is clearly dependent on the application. If the application is a crane controller for example operated by the same person for long periods the algorithm doesn't have to be robust on different person's images. In this case noise and motion blur should be tolerable.

Next after completing the preprocessing stage training and testing steps are carried out, the training is performed by taking few samples from the database and the training is done and then testing is done which is started only after training is completed if we start testing before training is completes then it results as error, even for testing few samples or frames from the database is taken and tested and then compared with rest of the frames from the database and the percentage of match is observed, only those frames will be used for further progress in the project that matches with highest percentage and rest frames are being rejected from use.The testing of frames is done using KNN algorithm, then subtract Test image Itest from each of the mean KNN Images and calculate KNN vectors and we need to find minimum KNN vector. At the end of this loop Min will have the KNN distance of test image from each of training classes and T will be the detected class and then we use gesture to do some operations.

In our proposed system, we have considered few gesture, each with 10 samples these are the images captured by camera. Some of the database images have been shown for each gesture in Figure 3.

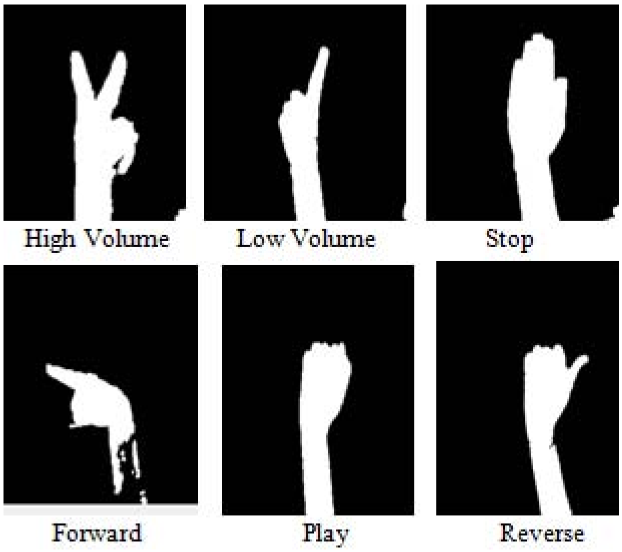

Figure 3: Some of the database images considered for proposed system

K-Nearest Neighbor (KNN) Algorithm: is an non parametric lazy learning algorithm. That is a pretty concise statement. When you say a technique is non parametric , it means that it does not make any assumptions on the underlying data distribution. This is pretty useful , as in the real world , most of the practical data does not obey the typical theoretical assumptions made (eg gaussian mixtures, linearly separable etc) . Non parametric algorithms like KNN come to the rescue here. It is also a lazy algorithm. What this means is that it does not use the training data points to do any generalization. In other words, there is no explicit training phaseor it is very minimal. This means the training phase is pretty fast . Lack of generalization means that KNN keeps all the training data. More exactly, all the training data is needed during the testing phase. (Well this is an exaggeration, but not far from truth). This is in contrast to other techniques like SVM where you can discard all non support vectors without any problem. Most of the lazy algorithms – especially KNN – makes decision based on the entire training data set (in the best case a subset of them).
The dichotomy is pretty obvious here – There is a non existent or minimal training phase but a costly testing phase. The cost is in terms of both time and memory. More time might be needed as in the worst case, all data points might take point in decision. More memory is needed as we need to store all training data.

Local Binary Pattern (LBP): is a simple yet very efficient texture operator which labels the pixels of an image by thresholding the neighborhood of each pixel and considers the result as a binary number. Due to its discriminative power and computational simplicity, LBP texture operator has become a popular approach in various applications. It can be seen as a unifying approach to the traditionally divergent statistical and structural models of texture analysis. Perhaps the most important property of the LBP operator in real-world applications is its robustness to monotonic gray-scale changes caused, for example, by illumination variations. Another important property is its computational simplicity, which makes it possible to analyze images in challenging real-time settings.

The value of the LBP code of a pixel (x_c, y_c) is given by:

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p \qquad s(x) = \begin{cases} 1, & if\ x \geq 0; \\ 0, & otherwise. \end{cases}$$

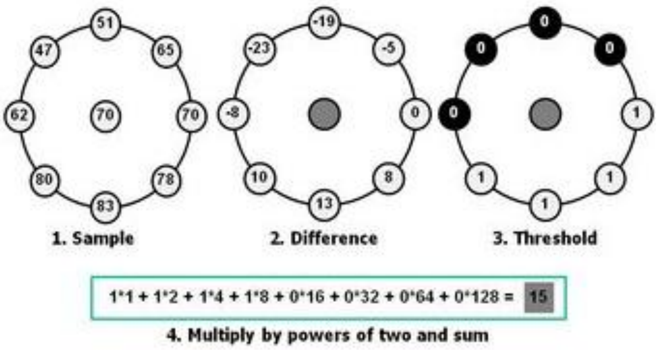

Figure 4: An example of LBP computation.

The following notation is used for the LBP operator: $LBP_{P,R}^{u2}$. The subscript represents using the operator in a (P,R) neighborhood. Superscript u2 stands for using only uniform patterns and labeling all remaining patterns with a single label. After the LBP labeled image $f_l(x,y)$ has been obtained, the LBP histogram can be defined as

$$Hi = \sum x,yI\{ fl(x,y)=i\},i=0,\ldots,n-1, \qquad \text{----------(1)}$$

$$Ni=Hi\sum n-1j=0Hj. \qquad \text{------------- (2)}$$

in which n is the number of different labels produced by the LBP operator, and I{A} is 1 if A is true and 0 if A is false.
When the image patches whose histograms are to be compared have different sizes, the histograms must be normalized to get a coherent description.

4. Results

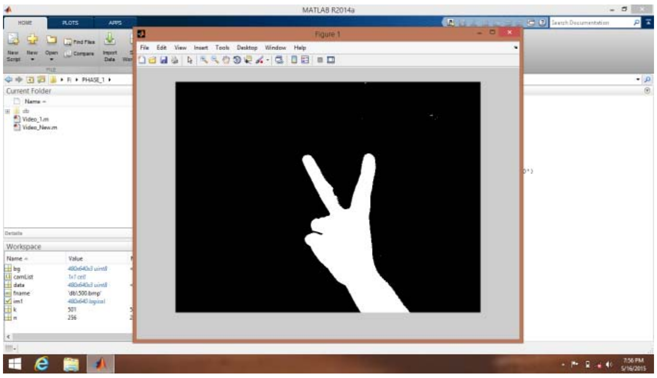

Figure 5

In the Figure 5 the image is captured by webcam in the matlab software, and the captured image is separated by the background and then converted into binary image.

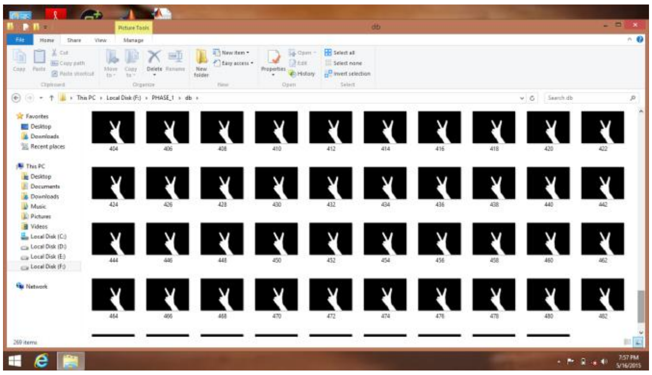

Figure 6

The above Figure 6 shows how the captured image is stored in the database, these images are continuously captured till where the loop limit is set once it reaches the limit all the captured images are stored in the database as shown in the Figure above.

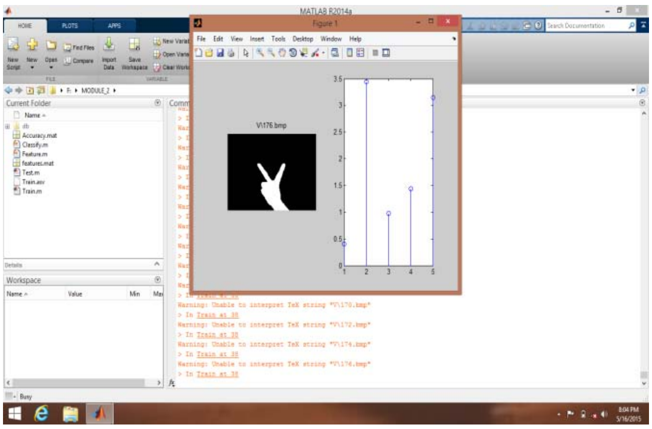

Figure 7

In the above Figure 7 the training step is carried out, once all the images are stored in the database few images from the database are choosen for training, these images are trained with rest of the images in the database.

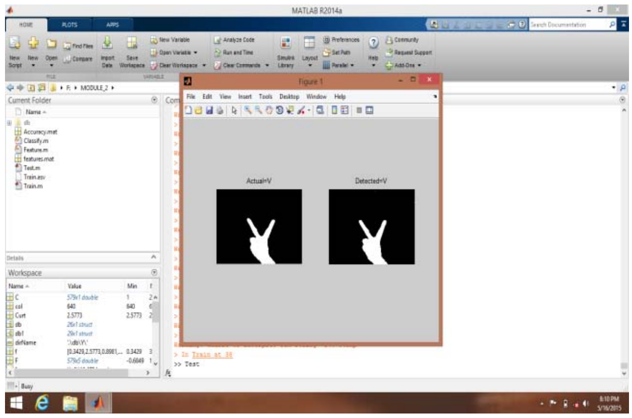

Figure 8

The above Figure 8 deals with testing stage where few images from the database are selected for testing and are tested with those images present in the database, where the actual image and the detected image is shown whether they both matches or no.

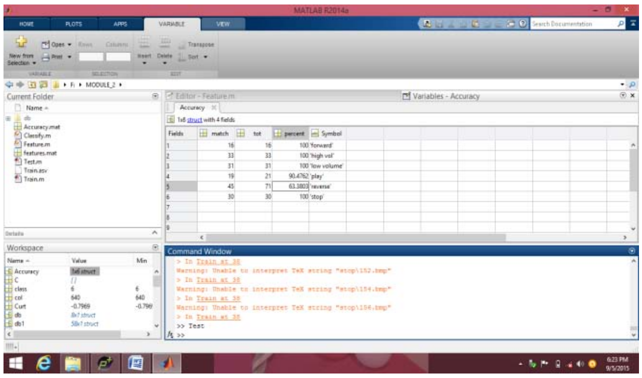

Figure 9

The above Figure 9 shown is the accuracy table which is the result of testing and training step, it has three columns containing match, total, percent and symbol where match is the number of gesture that are matched among total number of gesture and percent is to what percent it matches then the symbol column.

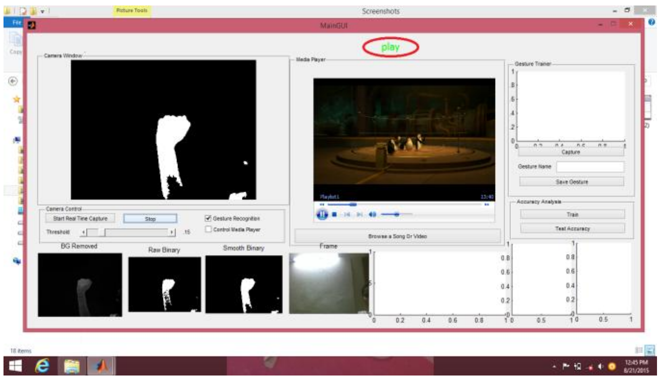

Figure 10

In the above Figure 10 the video is browsed in the media player and the gesture is captured in a camera window in real time where the above gesture indicates play symbol that is done in a plane background. When the video is browsed the gesture that indicates play is captured that controls the video in a play mode.

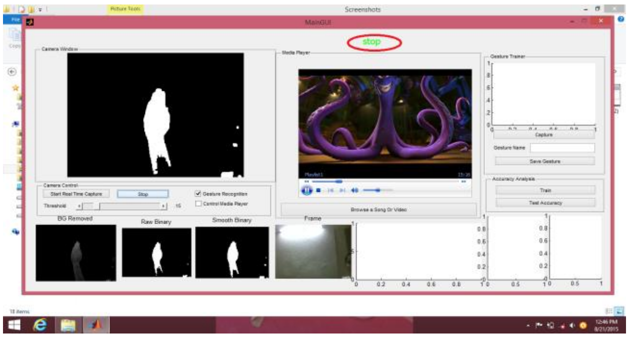

Figure 11

In the above Figure 11 the video is browsed in the media player and the gesture is captured in a camera window in real time where the above gesture indicates stop symbol that is done in a plane background. When the video is browsed the gesture that indicates stop is captured that controls the video in a stop mode.

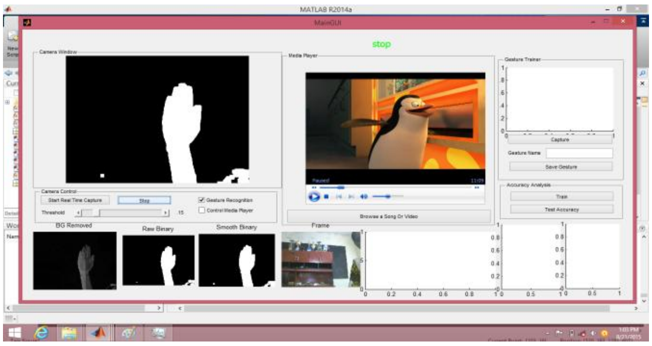

Figure 12

In the above Figure 12 the video is browsed in the media player and the gesture is captured in a camera window in real time where the above gesture indicates stop symbol that is done in a presence of background. When the video is browsed the gesture that indicates stop is captured that controls the video in a stop mode.

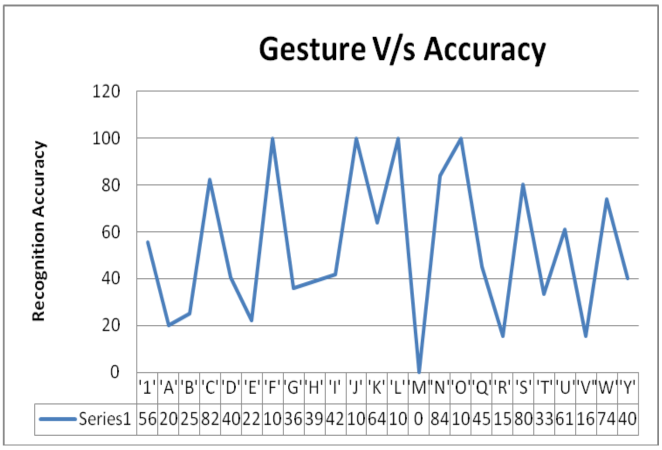

Figure 13: Graphical Analysis of Gesture

The graph analysis shown in Figure 13 is gesture v/s accuracy which tells the accuracy of each gesture, by which we can use those gesture in future that has maximum accuracy.

5. Conclusion

Hand gesture detection and recognition is a recent area of research in computer vision, several past work has been done in hand gesture recognition. Many new service providers are providing real time hand gesture detection and recognition devices and technologies. Most of these commercially available technologies are expensive and non-affordable.

The objective of this work was to develop a real time hand gesture recognition alternative that depends upon the normal webcam.

Through the analysis process we show that the propose system can detect and recognize gesture with about 80% accuracy. The proposed system can be further improved by improving the tracking methods by using model based tracking.

References

[1] Jaroslaw Szewinski, Wojciech Jalmuzna, University of Technology," Implementation of adaptive feed-forward algorithm" Institute of Electronic Systems, Warsaw, Poland.

[2] Asanterabi Malima, Erolozgür, and Mujdatcetin, "A Fast Algorithm For Vision-Based Hand Gesture Recognition For Robot Control" Faculty of Engineering and Natural Sciences, Sabancı University, Tuzla, Istanbul, Turkey.

[3] Mark Batcher," A Gesture controlled robot for object perception and Manipulation" Institute of Neuron informatics, Germany.

[4] Kevin Gabayan, Steven Lansel," Programming-By-Example Gesture Recognition".

[5] M. Ebrahim Al-Ahdal & Nooritawati Md Tahir,'' Review in Sign Language Recognition Systems'' Symposium on Computer & Informatics(ISCI),pp:52-57, IEEE ,2012

[6] Iwan Njoto Sandjaja, Nelson Marcos,'' Sign Language Number Recognition'' Fifth International Joint Conference on INC, IMS and IDC, IEEE 2009

[7] Noor Adnan Ibraheem and Rafiqul Zaman Khan," Survey on Various Gesture Recognition Technologies and Techniques" International Journal of Computer Applications (0975 – 8887) Volume 50 – No.7, July 2012

[8] Archana S. Ghotkar, Rucha Khatal , Sanjana Khupase, Surbhi Asati & Mithila Hadap,'' Hand Gesture Recognition for Indian Sign Language'' International Conference on Computer Communication and Informatics (ICCCI),pp:1-4.IEEE,Jan 2012.

[9] Paulraj M P, Sazali Yaacob, Mohd Shuhanaz bin Zanar Azalan, Rajkumar Palaniappan,'' A Phoneme Based Sign Language Recognition System Using Skin Color Segmentation" 6th International Colloquium on Signal Processing & Its Applications (CSPA), pp:1-5,IEEE,2010.

[10] Nasser H. Dardas and Emil M. Petriu'' Hand Gesture Detection and Recognition Using Principal Component Analysis" international conference on computational

intelligence for measurement system and application (CIMSA), pp:1-6, IEEE,2011 .

[11] Divya Deora1, Nikesh Bajaj ,k" Indian Sign Language Recognition" 1st International Conference on Emerging Technology Trends in Electronics, Communication and Networking, pp:1-5,IEEE ,2012.

[12] Solomon Raju Kota1, J.L.Raheja1,'' Principal Component Analysis for Gesture Recognition using System C", 2009 International Conference on Advances in Recent Technologies in Communication and Computing, pp:732-737, IEEE ,2009

[13] Y. Fang, K. Wang, J. Cheng, and H. Lu, "A Real-Time Hand Gesture Recognition Method", IEEE ICME, 2007, pp. 995-998.

[14] J. H. Kim, N. D. Thang, and T. S. Kim, "3-D Hand Motion Tracking and Gesture Recognition Using a Data Glove", IEEE International Symposium on Industrial Electronics (ISIE), July 5-8, 2009, Seoul Olympic Parktel, Seoul , Korea, pp. 1013-1018.

[15] J. Weissmann and R. Salomon, "Gesture Recognition for Virtual Reality Applications Using Data Gloves and Neural Networks", IEEE, 1999, pp. 2043-2046.