A Support System for Speech Impaired People using the Indian Sign Language

Pushpendra Kumar Tiwari[#1], Sithara Kamalakkannan[#2],S.V. Karthigaipriya[#3],Logeshwari R[#4]

Students, [4]Assistant Professor, Computer Science and Engineering, SRM University
C block, No.1, Jawaharlal Nehru Salai, 100 Feet Road, Vadapalani, TamilNadu 600026, India

Abstract* —Sign Language Recognition is a rapidly growing field of research. Several techniques have been developed recently. In this paper, we propose a system that uses Support Vector Machine (SVM) with image feature extractionas a classification technique for the recognition of the Indian Sign Language. The system comprises of four parts: Image capture,Background Subtraction, Feature Extraction and Classification. 26 signs were considered in this paper, each having over 200 samples to train the data. An accuracy of 98% was achieved during testing*.

Keywords — *Indian sign language, Support Vector Machine, feature extraction, image classification.*

I. INTRODUCTION

There are several strategies that may be employed for effective communication between speech impaired people and others. Some of these include writing, lip reading and sign language. Sign language which is one of the most common techniques has the pitfall of working only when both persons involved in the dialogue are familiar with the sign language.

The study of sign language recognition systems has been popular among linguists and engineers since the 1990s. It has been gaining momentum in the recent years with the development of image processing techniques and machine learning algorithms.

Two of the most popular sign language systems for which research has been carried out pertain to the American and British Sign Language Systems. Comparatively, less work has been contributed towards Indian Sign Language recognition systems.

Finding experienced and qualified interpreters at the business front for speech impaired people is quite a difficult task and also unaffordable. Even for day to day communication amongst people, unless everybody makes an effort to learn the sign language, communication comes to a standstill. This becomes a cause of isolation for speech impaired people. However, if a computer can be programmed to translate the sign language into textual or audio format, the differences between normal people and the speech impaired community can be minimized to a great extent.

Sign languages can be represented in the form of several techniques including finger spelling, gesture recognition, etc. In this paper, we concentrate on the technique of finger spelling. We have analysed and tested several techniques and discuss the limitations faced in the different algorithms. Finally, we have proposed a system which can recognize the various alphabets of Indian Sign Language using Support Vector Machine as the classification technique. This method gives accurate results at least possible time. The technique will not only benefit communication between speech impaired people and others in India, but can also be applied to other image processing problems.

II. LITERATURE REVIEW

Different approaches have been used by researchers for the recognition of various hand gestures. Some of the approaches are vision based approaches, glove based approaches, soft computing approaches like Artificial Neural Network, Fuzzy logic, Genetic Algorithm and others like Principal Component Analysis, Canonical Analysis, etc.

In vision based approaches, the initial idea revolves around feature extraction from an image of the hand. Most researchers start with a skin filtering technique for the segmentation of hand. This technique separates the skin coloured pixels from the non-skin coloured pixels, thus extracting the hand from the background. Fang [4] used Adaptive Boost algorithm which could not only detect a single hand but also overlapped hands. This approach is typical of any sign language recognition system. The problem arises when we get to sign recognition with a pre-defined database. The technique is helpful only in the detection of the hand.

Therefore, many researchers employ external aids like data gloves with sensors that can recognise which finger is being stretched or not with the help of sensors. Saengsri [15] in his paper for Thai Sign Language Recognition, used a '5DT Data Glove 14 Ultra' data glove which was attached with 14 sensors- 10 sensors on fingers and rest the rest between the fingers to measure flexures and abductions respectively. Still, the accuracy rate was limited to 94%. Kim [7] used 'KHU-1' data glove which comprises of 3 accelerometers sensor, a Bluetooth and a controller which extracted features like joints of hand. He performed the experiment for only 3 gestures and the process was very slow. Weissmann [6] used Cyberglove which measured features like thumb rotation, angle made between the neighbouring fingers and wrist pitch. The demerits of the system were that it recognized onlysingle hand gestures. Others use unique colour gloves that help in the segmentation process. Since skin colour detection for detecting the hand has a lot of interferences from the background such as the user's face itself or other light coloured backgrounds.

Principal Component Analysis (PCA) for extracting features like position of the finger, shape of the finger and direction of the image described by the means of Eigenvalues and Eigenvectors respectively have been tested by researchers like Lamar [10]. The limitations were that the accuracy rate obtained was 93% which was low and that the system could recognize gestures of only a single hand. Kapuscinski [17] proposed Hit-Miss transform for extracting features like orientation, hand size by computing the central moments. Accuracy rate obtained was 98%. But it lacked proper Skin filtering with changes in illumination.

Several gesture recognition approaches like Support Vector Machines, Artificial Neural Networks (ANN), Fuzzy Logic, Euclidean distance, Hidden Markov Model (HMM), etc. have been implemented by different researchers.

Saengsri [15] used Elman Backpropagation Neural Network (ENN) algorithm which consisted of an input layer with 14 nodes similar to the sensors in the data glove, an output layer with 16 nodes equal to the number of symbols, and a hidden layer with 30 nodes which is just the total of input and output nodes. Gestures were recognized by identifying the maximum value class from ENN. Recognition rate obtained was 94.44%. Difficulty faced in this paper was it considered only single gestured signs. Lamar [10] used ANN that comprised of an input layer with 20 neurons, a hidden and output layer each with 42 neurons. Back propagation algorithm was used, and after the training of the neural network, one output neuron was achieved, thus giving the proper recognized gesture. Gopalan [14] used Support Vector Machine for classification purpose. The linearly non separable data becomes separable when SVM was used, as the data was projected to a higher dimensional space, thus reducing error. Kim [7] in his paper of Recognition of Korean Sign Language used Fuzzy logic. Fuzzy sets were considered where each set represented the various speeds of the moving hand. They were mathematically given by ranges like small, medium, negative medium, large, positive large, etc. Accuracy rate obtained was 94%, but the difficulty faced by them was heavy computation.

Apart from these, several image comparison techniques were analysed and tested. The simplest being 'mean square evaluation' between an image in the data set and the test image.The structural similarity index (SSIM) technique was also used for comparison. It was found that the limitation in these techniques was that the index was quite incoherent for images that were similar.

Similarly, when techniques such as Scale-Invariant Feature Transform (SIFT) or Speeded up Robust Features as used in [13] or Orb, the unpatented version were tested, it too did not work well for images that were not distinct. The algorithm worked better for double handed sign language but not quite as well for single handed sign language.

III. PROPOSED SYSTEM AND IMPLEMENTATION

"Support Vector Machine" (SVM) is a supervised machine learning algorithm which can be used for both classification and regression challenges. However, it is mostly used in classification problems. Computing the SVM classifier amounts to minimizing an expression of the form

$$\left[\frac{1}{n}\sum_{i=1}^{n}\max\left(0,1-y_i\left(\omega\cdot x_i+b\right)\right)\right]+\lambda\|\,\omega\,\|^2$$

The formula represents the idea that the width of the hyperplane between two distinct groups must be maximised, whereas the error rate must minimised. Choosing a sufficiently small value for λ yields a hard-margin classifier for linearly classifiable input data.

In this technique, each data item that is an imagematrix, is plotted as a point in n-dimensional space (where n represents the number of features). The value of each feature being the value of a particular coordinate. We then carry out the classification of the different signs by determining the hyperplane that can differentiate the twenty six signs distinctly from one another. The main advantage of this system is that in addition to using SVM, background subtraction helps in recognising the Indian signs at dynamic backgrounds. The system is developed in python using the following steps:

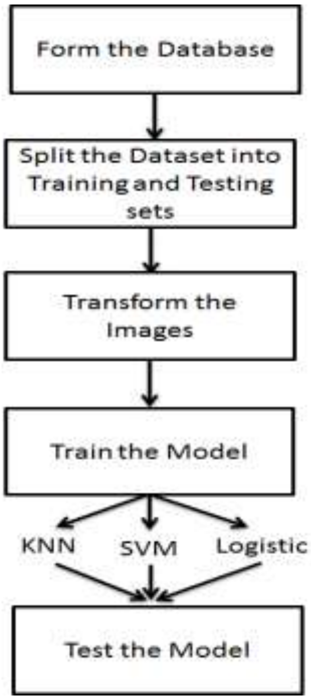

Fig. 1 A flow diagram of the steps followed in the proposed sign language recognition system

A. Form the database

Images of the letters of the Indian Sign Language are captured using a webcam on the laptop, with the help of the OpenCV library. Over two hundred images for each of the twenty six alphabets are stored as such in a folder.

B. Split Dataset into Training and Testing sets

The pre-defined database is traversed through and is classified into images that constitute the data to be read and their corresponding labels. The labels represent the output that is to be displayed during testing. In this case, it refers to the letter corresponding to the sign. Next, the data is split into the training set and the testing set. 2/3rd of the data is allocated for training and 1/3rd is allocated for testing, along with their labels.

C. Transform the Images

All the images are resized into 100x100 pixels. An image is captured for every key press. A picture of the current background is taken initially. For the next key press, the hand with the sign is introduced. The pixels that are not present in the first image are blacked out. This results in background subtraction, with only the hand depicting the sign in focus as represented in Fig 2.

The images are smoothened using the Gaussian Blur filter, after which the contour of the hand is determined. A convex hull is drawn over the fingers of the image. The image of the hand is centred and the portion of the arm till little below the palm is eliminated by establishing a bounding region of interest. This results in the final form the images in the data set are converted into.

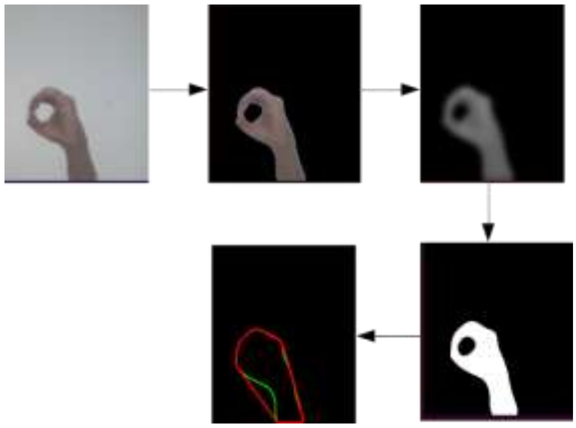

Fig. 2 Original image -> Image after background subtraction -> Smoothened image -> Binary Image -> Image with contours

D. Train the Model

The data set is trained with a classifier according to the user's preference. The different classifier options being k-nearest neighbours (KNN), a logistic and a support vector machine (SVM) classifier. The training set of the data is trained with the three classifiers using sklearn libraries. Once the model is trained, a comparison is made between the classifiers on the testing set of the data. The comparison indicated that the SVM based classifier yielded the highest accuracy during testing, resulting in 98.6% precision. The KNN and logistic classifiers resulted in 97.4% and 97.9% precision rate. Even though the differences in the precision rates seem minute, SVM resulted in higher accuracy for every test. It is believed that KNN will perform better with larger data sets. Therefore, SVM is preferable for languages that are not yet well established. The trained model is stored to disk.

Fig. 2 A graph representing the precision rates for every letter according to the type of classifier

E. Test the Model

A live input of the user doing the sign language is captured from the webcam. At every triggered interval, a frame is captured. The captured image or frame is tested using the trained SVM classifier model. The captured frame and the prediction of the SVM model, which is the label generated for the corresponding sign, is displayed as output. The process is repeated at regular intervals for every captured frame until the user wishes to quit using the appropriate key press. The background may be changed anytime as mentioned in section III. B.

IV. CONCLUSIONS

A sign language recognition system for the letters of the Indian Sign Language was created after careful analysis of several techniques and algorithms. The effectiveness of the proposed method was tested using different classifiers. It was established that the SVM model yields the highest accuracy.

The past few years have seen a huge improvement in sign language recognition systems. The major motivation behind our proposed system is to bring about a cost efficient, easy to operate, accurate sign language recognition system for the Indian Sign Language.

With the lack of popularity of the Indian Sign Language, especially regional sign languages in India, schools for special kids have resorted to adopting lip reading instead of sign language. This is a much more difficult approach for speech impaired people.

The creation of such systems will not only promote the Indian Sign Language, but can also be modified for the development of new sign languages such as Tamil.

REFERENCES

[1] A. S. Ghotkar, R. Khatal, S. Khupase, S. Asati, and M. Hadap, “Hand Gesture Recognition for Indian Sign Language”, IEEE International Conference on Computer Communication and Informatics (ICCCI), Jan. 10-12, 2012, Coimbatore, India.

[2] D.Karthikeyan, Mrs.G.Muthulakshmi., “English Letters Finger Spelling Sign Language Recognition System”, International Journal of Engineering Trends and Technology (IJETT) – Volume 10 Number 7 - Apr 2014

[3] D. Y. Huang, W. C. Hu, and S. H. Chang, “Vision-based Hand Gesture Recognition Using PCA+Gabor Filters and svm”, ieee Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2009, pp. 1-4.

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

[5] J. Rekha, J. Bhattacharya, and S. Majumder, “Shape, Texture and Local Movement Hand Gesture Features for Indian Sign Language Recognition”, IEEE, 2011, pp. 30-35.

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

[7] 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.

[8] Kumara Maruthi M. Mullur, “Indian Sign Language Recognition System”, International Journal of Engineering Trends and Technology (IJETT) – Volume 21 Number 9 – March 2015

[9] L. K. Lee, S. Y. An, and S. Y. Oh, “Robust Fingertip Extraction with Improved Skin Color Segmentation for Finger Gesture Recognition in Human-Robot Interaction”, WCCI 2012 IEEE World Congress on Computational Intelligence, June, 10-15, 2012, Brisbane, Australia.

[10] M. V. Lamar, S. Bhuiyan, and A. Iwata, “Hand Alphabet Recognition Using Morphological PCA and Neural Networks”, IEEE, 1999, pp. 2839-2844.

[11] Manigandan M. and I. M Jackin, “Wireless Vision based Mobile Robot control using Hand Gesture Recognition through Perceptual Color Space”, IEEE International Conference on Advances in Computer Engineering, 2010, pp. 95-99.

[12] O. B. Henia and S. Bouakaz, “3D Hand Model Animation with a New Data-Driven Method”,Workshop on Digital Media and Digital Content Management (IEEE Computer Society), 2011,pp. 72-76.

[13] Cheok Ming Jin, Zaid Omar, Mohamed HishamJaward, “A Mobile Application of American Sign Language Translation via Image Processing Algorithms”, 2016 IEEE Region 10 Symposium (TENSYMP), Bali, Indonesia

[14] R. Gopalan and B. Dariush, “Towards a Vision Based Hand Gesture Interface for RoboticGrasping”, The ieee/rsj International Conference on Intelligent Robots and Systems, October 11-15, 2009, St. Louis, usa, pp. 1452-1459.

[15] S. Saengsri, V. Niennattrakul, and C.A. Ratanamahatana, “TFRS: Thai Finger-Spelling Sign Language Recognition System”, IEEE, 2012, pp. 457-462.

[16] S. K. Yewale and P. K. Bharne, “Hand Gesture Recognition Using Different Algorithms Basedon Artificial Neural Network”, IEEE, 2011, pp. 287-292.

[17] T. Kapuscinski and m. Wysocki, “Hand Gesture Recognition for Man-Machine interaction”,Second Workshop on Robot Motion and Control, October 18-20, 2001, pp. 91-96.

[18] Dr. P.M. Mahajan, Jagruti S. Chaudhari, “Review of Finger Spelling Sign Language Recognition”, International Journal of Engineering Trends and Technology (IJETT) – Volume 22 Number 9-April 2015