Research Article Impact Factor: 0.621 ISSN: 2319-507X
Neha Tavari, IJPRET, 2014; Volume 2 (9): 106-115 IJPRET

INTERNATIONAL JOURNAL OF PURE AND APPLIED RESEARCH IN ENGINEERING AND TECHNOLOGY

A PATH FOR HORIZING YOUR INNOVATIVE WORK

A FRAMEWORK FOR INDIAN SIGN LANGUAGE RECOGNITION

NEHA V. TAVARI[1], PROF. ANIL V. DEORANKAR[2]

1. M. Tech. Scholar, Department of Computer Science and Engineering, Government College of Engineering, Amravati, Maharashtra, India.
2. Associate Professor, Department of Computer Science and Engineering, Government College of Engineering, Amravati, Maharashtra, India.

Accepted Date: 27/02/2014 ; Published Date: 01/05/2014

Abstract: Sign language is the language of communication for deaf and dumb people. Most of these physically impaired communities are dependent on sign language translators to express their thoughts to rest of the world. This causes isolation of these people in society. Hence, Sign Language Recognition is one of the most growing fields of research today. A sign language is composed of various gestures formed by physical movement of body parts i.e. hand, arms or facial expressions. In this paper, a method is proposed that makes the use of hand gestures for recognition of Indian Sign Language. Hand Gesture recognition system provides us an innovative, natural, user friendly way of interaction with the computer which is more familiar to the human beings. The proposed method is able to identify the images of the signer which are captured dynamically during testing phase. To implement this approach we have utilized a simple web camera to capture hand gesture images. Artificial neural network is used for recognizing different signs and translate them into text and voice format.

Keywords: Deaf and dumb, Hand gesture, Human computer interaction, Sign language.

Corresponding Author: MS. NEHA V. TAVARI

Access Online On:

www.ijpret.com

How to Cite This Article:

Neha Tavari, IJPRET, 2014; Volume 2 (9): 106-115

PAPER-QR CODE

INTRODUCTION

The expectation of widely extensive range of computer systems with the rapid development of information technology in our life , would be inter in our environments. These environments shold me made simple, natural and easy to use interfaces for human computer-interaction (HCI) . The user interface of any personal computer has evolved from primitive text user interfaces to a graphical user interfaces (GUIs) which still limited to keyboard and mouse input , however, they are inconvenient, unnatural, and not suitable for working in virtual environments. With the use of the hand gestures an efficient alternative would be provided to these onerous interface devices for human-computer interaction.

Sign language is widely used by physically impaired people who cannot speak and hear or who can hear but cannot speak and is the only medium of communication for those people. It is nothing but the combination of various gestures formed by different hand shapes, movements and orientations of hands or body, facial expressions and lip-patterns for conveying messages. These gestures are widely used by the deaf and dumb people to express their thoughts[1]. Usually physically impaired people needs the help of sign language interpreters for translating their thoughts to normal people and vice versa . But it becomes very difficult to find a well experienced and educated translator for the sign language every time and everywhere in daily life, but human-computer interaction system for this can be installed anywhere possible. So a system recognizing the sign language gestures automatically is necessary which will help to minimize the gap between deaf people and normal people in the society. As sign language is well structured code gesture, each gesture has a meaning assigned to it [2]. There are number of sign languages spreaded across the world. The sign language used by those deaf and mute at a particular place is dependent on the culture and spoken language at that place. American Sign. Language (ASL), British Sign Language (BSL), Japanese Sign Language family (Japanese, Taiwanese and Korean Sign Languages), French Sign Language family (French, Italian, Irish, Russian and Dutch Sign Languages), Australian Sign Language, etc.[3] are the examples of regionally different sign languages. Indian sign language (ISL) is used by the deaf and dumb community in India.

Since ISL got standardized only recently and also since tutorials on ISL gestures were not available until recently, there are very few research work that has happened in ISL recognition[4].Here we propose a method for hand gesture recognition of Indian sign language alphabet and numerals. The signs considered for recognition include 26 letters of the English alphabet and the numerals from 0-9. Indian sign language alphabet and numerals are shown in Fig.1 and Fig.2 respectively.

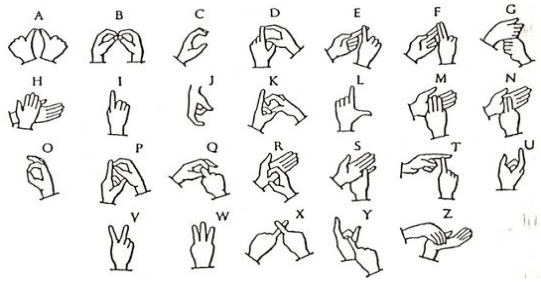

Fig 1. Representation of ISL alphabet

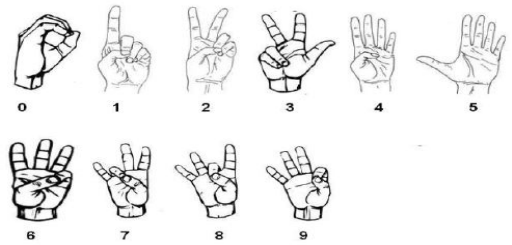

Fig 2. Representation of ISL numerals

RELATED WORK

Recognition based on 'UP' and 'DOWN' positions of fingers[2] was proposed by a Rajam, P. Subha and Dr G Balakrishnan. In this method, set of 32 signs in which each represents the binary 'UP' & 'DOWN' positions of the five fingers, the most significant bit represents the 'little' finger and the least significant bit represents the 'thumb' finger. A right hand palm images are loaded at run time having 32 combinations of binary number signs. The tip of "UP" fingers is identified by measuring their heights with respect to a reference point at the bottom of palm. The feature points are determined using one of the two scan modes i.e. left-right scan and right-left scan. The feature points located by left-right scan are marked as green in color, those located by Right-Left are marked as blue and reference point is marked as red color.

Deepika Tewari, Sanjay Kumar Srivastava[3] proposed an algorithm for hand gesture recognition system in ISL which is based on vision-based approach. An intensity (grayscale) representation of the segmented image is used for further processing. This grayscale version, also called a "skin map," contains intensity values for skin pixels and the background is represented as black. Then, the Two-Dimensional Discrete Cosine Transform (2D-DCT) for each region is computed, and feature vectors are formed from the DCT coefficients.The DCT can be extended to the transformation of 2D signals or images. This can be achieved in two steps:

i) computing the 1D-DCT of each of the individual rows of the two dimensional image,

ii) computing the 1D-DCT of each column of the image.

DCT-based feature vectors are classified to check whether sign mentioned in the input image is "present" or "not present" in the ISL database using self-organizing map (SOM)[3] with unsupervised learning technique in Artificial Neural Network (ANN). As SOM is based on unsupervised learning, no mediation is needed during the learning and little need to be known about the characteristics of the input data which makes it to be used for clustering data without knowing the class memberships of the input data. The SOM is also known as SOFM, the Self-Organizing Feature Map (SOFM) as it can be used to detect features belonging to the problem. The particular kind of SOM known as a Kohonen Network is used which have feed-forward structure with a single computational layer arranged in rows and columns.

Transition movement models (TMMs) [6] is proposed by Gaolin Fang, Wen Gao, and Debin Zhao to handle transition parts between two adjacent signs in large-vocabulary continuous SLR. For large-vocabulary continuous SLR, TMMs were proposed for continuous Chinese SLR. Sign samples taken from input devices are fed into the feature extraction unit and then input into two related parts i.e. TMM training and recognition based on TMMs. In the TMM training part, sign/sentence samples are trained into sign models and TMMs by the model training module (no TMMs in the first run). Then, these models are used to segment continuous sentence samples into sign parts and transition parts. Transition parts are clustered using the temporal clustering algorithm. We iterate this process until the convergence criterion is met. In the recognition part based on TMMs, the estimated TMMs and sign models obtained from the training part are viewed as candidates of the Viterbi search algorithm, together with language models (Bigram) for recognizing large-vocabulary continuous sign language. An approach is made to recognize alphabet characters dynamically from color image sequences using "Continuous Adaptive Mean Shift Algorithm (CAMSHIFT)" tracking algorithm stated in[7] by Sulochana M. Nadgeri, Dr. S. D. Sawarkar, Mr. A. D. Gawande. The algorithm used here is based on a robust nonparametric technique for climbing density gradients to find the mode(peak) of probability distributions called the mean shift algorithm. Here, it is to find the mode of a color distribution within a video scene. The color image data has to be represented as probability distribution by using color histogram for tracking colored objects in video frame sequences. Hence to deal with dynamically changing color probability distributions derived from video frame sequence, the mean shift algorithm is modified to Continuously Adaptive Mean Shift(CAMSHIFT) algorithm. The tracking accuracy of this algorithm is compared against tolerance to noise, distracters and performance is studied.

Data from five-channel surface electromyogram and 3-D accelerometer from the signer's dominant hand were analyzed by Vasiliki E. Kosmidou, and Leontios J. Hadjileontiadis using intrinsic mode entropy (IMEn) for the automated recognition of Greek sign language (GSL) isolated signs. As the gesture is directly connected to hand movement, measurement of the latter could contribute to the gesture representation in the recognition space. This was the motivation to explore the capabilities of sEMG and 3-D-Acc data in SL recognition. To this end, the movements of the arm can be captured with the help of a 3-D Accelerometer, whereas the motions of the wrist and the fingers can be obtained by the corresponding muscles on the arm. After experimentation on the exact placement and type of the sEMG electrodes, a position that provides with high signal quality and discrimination among the performed motions per channel was identified [8].

A novel technique is proposed by Dipak Kumar Ghosh , Samit Ari to obtain a rotation invariant gesture image which coincides the 1st principal component of the segmented hand gestures with vertical axes The shape of the contour is an important property that can be used to distinguish of the static hand gestures from one class to another. The localized contour sequence (LCS), which has been confirmed to be a very efficient representation of contours , is selected as a feature set of the hand gesture. A contour tracking algorithm is proposed to track the contour of a gesture in the clockwise direction and the contour pixels are numbered sequentially starting from the topmost left contour pixel. After successfully extracting a normalized LCS feature vector of the static hand gesture, The classification job is done via k-mean based radial basis function neural network (RBFNN) [9]. Ravikiran J, Kavi Mahesh, Suhas Mahishi, Dheeraj R, Sudheender S, Nitin V. Pujari stated an efficient algorithm[10] to identify the number of fingers opened in a gesture representing an alphabet of the American Sign Language and introduces a very effective and efficient technique for finger detection.The method has three main phases of processing viz., Edge Detection, Clipping and Boundary Tracing. The first phase having Canny edge operator produces an edge detected image which reduces the number of pixels to be processed at runtime. The second phase clips the undesirable portion of the edge detected image for further processing and the final phase traces the boundary of the image and detects finger tips which aid in finger detection.

PROPOSED SYSTEM

The proposed work is aimed to develop a sign language education and recognition platform for hearing impaired peoples and communication system for dumb people to convey their message.

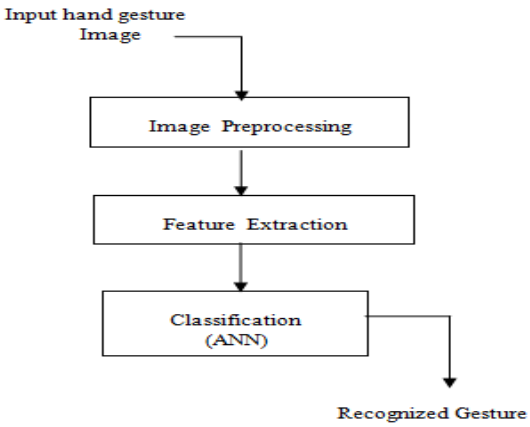

Fig 3. System block diagram

The proposed hand gesture recognition method translates the finger spelling in Indian sign language to textual and audio form.

- Image Acquisition
- Preprocessing and Segmentation
- Feature Extraction
- Classification

Image Acquisition

The video sequence of signer, i.e. the person conveying in the sign language, can be obtained by using a web camera. Image acquisition is the process to capture the hand gesture images which represents different signs. The image database is created for training and testing phase. As the image dataset of Indian sign language alphabet and numerals are not available from any resources, it is made available with suitable lighting and environmental setup.

Image capturing is a random process. The resolution of various image capturing devices may not be the same. This results in different resolution of the captured images. For accurate comparison of the features and to reduce the computational effort needed for processing, all the images should be scaled to a uniform size[1]. The interface of the application is provided with the button START and STOP. When the user clicks on the first

button it works up to open up the integrated webcam and the button changes its status to STOP and when the user is ready with the gesture it can click on the second button so that frame is captured and stored in the directory[5].

Preprocessing and Segmentation

The image scene and information should not be altered by local changes due to noise and digitization error. Hence to satisfy the environmental scene conditions, preprocessing of the raw data is highly important. The objective of gesture segmentation is to extract the gesture region from the background of the image. Hand segmentation is the process of extracting the hand sign from the captured image and also gesture region is extracted from the background of the image. Efficient hand segmentation has a key role in sign language recognition task. The segmentation process depends on the type of gesture, if it is dynamic gesture then the hand gesture need to be located and tracked , if it is static gesture the input image have to be segmented only[11].

The result of segmentation produces a binary image with the skin pixels in white color and background in black color. The resulting binary image may contain noise and segmentation errors. Filtering and morphological operations are performed on the input image to decrease noise and segmentation errors if any. Image preprocessing includes the set of operations on images whose goal is the improvement of the image data that suppresses undesired distortions or enhances some image features important for further processing.

Fig 4. Removal of background

Feature Extraction

Good segmentation process leads to perfect features extraction process and the later play an important role in a successful recognition process[11]. There are many interesting points on every object which can be extracted to provide a "feature" description of the object. Features vector of the segmented image can be extracted in different ways according to particular application. Under different scene conditions, the performance of different feature detectors will be significantly different. The nature of the background, existence of other objects

(occlusion), and illumination must be considered to determine what kind of features can be efficiently and reliably detected.

Classification

feature vector obtained from the feature extraction step is used as the input of the classifier that recognizes the sign[1]. Training and generalizing are the most basic and important properties of the neural networks. Hence ,Artificial neural network is used as the classification tool. Different network models exist for training the neural net and depending on the feature vectors, the best neural net training method is chosen. An artificial neural network processes information by creating connections between artificial neurons and they are widely used to model complex relationship between inputs and outputs. Training or learning is used to configure a neural network such that the application of a set of inputs produces a set of desired outputs. Many different algorithms exist to train an artificial neural network. Feed forward back propagation neural network is used widely for classification purpose. In feed forward neural network, each neuron receives a signal from the neurons in the previous layer, and each of those signals is multiplied by a separate weight value. The weighted inputs are summed, and passed through a limiting function which scales the output to a fixed range of values.

CONCLUSION

A neural network based method for automatically recognizing the hand gestures of Indian sign language is proposed in this work. In this ,we will develop a vision based gesture recognition using a simple system connected with a web camera. As the method implements completely by using digital image processing technique so the user does not have to wear any special hardware device to get features of hand shape. The features extracted from the sign image are used to train a neural network that recognizes the sign. The system will be able to recognize 36 hand gestures which represents the alphabets from A to Z and numbers from 0 to 9.Developing such system translating sign language to text/voice format will prove very useful for physically imapaired people of India. In future, there can be many possible improvements that will broaden the scope of this work.

REFERENCES

1. Adithya V., Vinod P. R., Usha Gopalakrishnan , "Artificial Neural Network Based Method for Indian Sign Language Recognition" , IEEE Conference on Information and Communication Technologies (ICT) ,2013, pp. 1080-1085.

2. Rajam, P. Subha and Dr G Balakrishnan , "Real Time Indian Sign Language Recognition System to aid Deaf and Dumb people", 13th International Conference on Communication Technology(ICCT) ,2011, pp 737-742.

3. Deepika Tewari, Sanjay Kumar Srivastava , "A Visual Recognition of Static Hand Gestures in Indian Sign Language based on Kohonen Self- Organizing Map Algorithm", International Journal of Engineering and Advanced Technology (IJEAT), Vol.2, Issue 2, Dec 2012, pp. 165-170,.

4. Geetha M, Manjusha U C , "A Vision Based Recognition of Indian Sign Language Alphabets and Numerals Using B-Spline Approximation", International Journal on Computer Science and Engineering (IJCSE), Vol. 4 ,No. 03,March 2012, pp. 406-415.

5. Sakshi Goyal, Ishita Sharma, Shanu Sharma , "Sign Language Recognition System For Deaf And Dumb People", International Journal of Engineering Research & Technology (IJERT), Vol. 2, Issue 4,April 2013, pp. 382-387.

6. Gaolin Fang, Wen Gao, and Debin Zhao, "Large-Vocabulary Continuous Sign Language Recognition Based on Transition-Movement Models", IEEE transactions on systems, man, and cybernetics—part a: systems and humans, Vol. 37, No. 1, Jan 2007.

7. Sulochana M. Nadgeri, Dr. S.D.Sawarkar, Mr. A. D. Gawande, "Hand gesture recognition using CAMSHIFT Algorithm", Third International Conference on Emerging Trends in Engineering and Technology, IEEE ,2010, pp.37-41.

8. Vasiliki E. Kosmidou, and Leontios J. Hadjileontiadis, "Sign Language Recognition Using Intrinsic-Mode Sample Entropy on sEMG and Accelerometer Data", IEEE transactions on biomedical engineering, Vol. 56, No. 12, Dec 2009.

9. Dipak Kumar Ghosh , Samit Ari , "A Static Hand Gesture Recognition Algorithm Using K-Mean Based Radial Basis Function Neural Network", IEEE,2011.

10. Ravikiran J, Kavi Mahesh, Suhas Mahishi, Dheeraj R, Sudheender S, Nitin V. Pujari, "Finger Detection for Sign Language Recognition", Proceedings of the international Multi Conference of Engineers and Computer Scientists, Vol. 1, March 2009.

11. Rafiqul Zaman Khan and Noor Adnan Ibraheem, "Hand Gesture Recognition: a Literature Review", International Journal of Artificial Intelligence & Applications (IJAIA), Vol.3, No.4, July 2012, pp.161-174.