Imperial Journal of Interdisciplinary Research (IJIR)
Vol.2, Issue-1 , 2016
ISSN : 2454-1362 , www.onlinejournal.in

Sign Language Hand Configuration Recognition

Aditya Baravkar
Department of Computer Engineering, Smt. Kashibai Navale College of Engineering, Vadgaon(Bk.), Pune-41

Rounak Burad
Department of Computer Engineering, Smt. Kashibai Navale College of Engineering, Vadgaon(Bk.), Pune-41

Menka Chandani
Department of Computer Engineering, Smt. Kashibai Navale College of Engineering, Vadgaon(Bk.), Pune-41

Rucha Parekh
Department of Computer Engineering, Smt. Kashibai Navale College of Engineering, Vadgaon(Bk.), Pune-41

Shyam Kosbatwar
Department of Computer Engineering, Smt. Kashibai Navale College of Engineering, Vadgaon(Bk.), Pune-41

Abstract: *In today's world, we can see that it is becoming very difficult for the disabled to communicate and to be able to convey whatever they want to. In this project we propose a method using projections of hand to create a sign language and to help convey what the disabled wants to say. The main reason behind this is that every country has a different sign language and it is not possible for a disabled person to be able to understand every language. So we can create a database of certain hand gestures of our own and personally store its meaning so it becomes easy for any disabled person to communicate without having to know the regional language. The users can then decide what the particular hand gesture means and thus create a database of their own version of language which they can use through their devices to make people understand what they want to say.*

Keywords: Communication**,** Sign Language, Hand Configuration, User defined Language

[1] Introduction

Sign language is widely used by the vocally disabled. Sign language is a combination of hand shapes, orientation and movement of the hands, arms or body, and facial expressions to express a speaker's thoughts. Sign language differs from region to region and for a normal person it's not feasible to learn all of them so there must be a medium of communicating his expressions to other people. There are various institutes which help them learn the standard language but it depends from person to person and obviously his grasping and willingness to learn the language. In this work, we will make a medium which will help them to communicate without learning any language; that medium will be a mediator between normal people and the disabled ones. So, there will be no need of learning the standard sign language for the disabled and even the normal person for understanding purpose.

1.1 Motivation

The motivation for developing such helpful application came from the fact that it would prove to be of utmost importance for socially aiding people and it would help increasingly for social awareness as well. Further, if we keep aside this world of computers and just take into consideration human-human interaction, we can realize that we are utilizing a broad range of gesture in personal communication. In fact gesturing is so deeply rooted in our communication that people often continue gesturing when speaking on the telephone. The significant use of gestures in daily life motivates the use of gestural interface in modern era.

[2] Literature Survey

2.1 Hand Gesture Technology

In the recent years, there has been tremendous research on the hand sign recognition. The technology of gesture recognition is divided into two categories-

[1] **Vision-based**

In vision-based methods, computer camera is the input device for observing the information of hands or fingers. In Vision Based method, the requirement for system is only the camera(s) to capture the image required for the natural interaction between human and computers without

the use of any extra devices. In this technique, a video sequence is captured with the help of camera and the input video is partitioned into frames.A set of features is extracted for each extracted frame sequence. Under image preprocessing the hand object is localized and segmented and the necessary features are extracted and stored in the computer as a trained set. Then each input image pass through the previous steps to extract its features, and classification algorithms are applied by comparing the extracted features from input image with the training set, to interpret the gesture meaning according to a specific application.

[2] **Glove-based methods**

In glove based systems data gloves are used which can archive the accurate positions of hand gestures as its positions are directly measured. The Data-Glove based methods use sensor devices for digitizing hand and finger motions into multi-parametric data. The extra sensors make it easy to collect hand configuration and movement. However, the devices are quite expensive and bring much cumbersome experience to the users Some of the earlier gesture recognition systems attempted to identify gestures using glove-based devices that would measure the position and joint angles of the hand. However, these devices are very cumbersome and usually have many cables connected to a computer. This has brought forth the motivation of using non-intrusive, vision-based approaches for recognizing gestures Also the sensors used for the detection of the sign language and the gesture recognition in the system that are available in the market are quite costly.

2.2 Related Work

Recognition based on "UP" and "DOWN" positions of fingers 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 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. 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) 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 special modified white color woolen hand gloves have been used to simplify the process of gesture identification. Taking into consideration the features of sign language, is shows that each finger in a gesture conveys a particular message and hence each and every finger has to be identified uniquely. Dhruva N. and Sudhir Rao Rupanagudi, Sachin S.K., Sthuthi B., Pavithra R. and Raghavendra developed a novel segmentation algorithm to meet this requirement in which the woolen hand gloves were modified by replacing and sewing each finger of the glove with a colored cloth for each digit of the hand. Thus, it utilises a unique color coding for each finger of our hand in order to assist in identifying the fingers. Therefore, segmentation based on various color spaces would be a viable option.

Transition movement models (TMMs) 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.

An approach is made to recognize alphabet characters dynamically from color image sequences using "Continuous Adaptive Mean Shift Algorithm (CAMSHIFT)" tracking algorithm stated in by Sulochana M. Nadgeri, S. D. Sawarkar, 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.

[3] Proposed System Work

The proposed work is aimed to develop a sign language education and recognition platform for hearing impaired people and communication system for dumb people to convey their message. The main approaches for analyzing and classifying hand gestures for Human Computer Interaction (HCI) include Glove based techniques and Vision based techniques. The objective of the this work is to build a system that uses natural hand gestures as a modality for recognition in the vision-based setup.

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

[1] Image Acquisition

[2] Preprocessing and Background Subtraction

[3] BLOB Extraction

[4] HSV Model - Skin color detection

[5] Template matching

3.1 Image Acquisition

The video sequence of 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. Image acquisition in image processing can be broadly defined as the action of retrieving an image from some source, usually a hardware-based source, so it can be passed through whatever processes need to occur afterward. Performing image acquisition in image processing is always the first step in the workflow sequence because, without an image, no processing is possible. The image that is acquired is completely unprocessed and is the result of whatever hardware was used to generate it.

3.2 Preprocessing and Background Subtraction

Image preprocessing, also called image restoration, involves the correction of distortion, degradation, and noise introduced during the imaging process. Image pre-processing can significantly increase the reliability of an optical inspection. Several filter operations which intensify or reduce certain image details enable an easier or faster evaluation.

Background subtraction (BS) is a common and widely used technique for generating a foreground mask by using static cameras. As the name suggests, BS calculates the foreground mask performing a subtraction between the current frame and a background model, containing the static part of the scene or, more in general, everything that can be considered as background given the characteristics of the observed scene. The objective of background subtraction is to extract the gesture region from the background of the image. Background subtraction is the process of extracting the hand sign from the captured image and also gesture region is extracted from the background of the image.

3.3 BLOB Extraction

Blob Analysis is a fundamental technique of machine vision based on analysis of consistent image regions. As such it is a tool of choice for applications in which the objects being inspected are clearly discernible from the background. **The purpose of BLOB extraction is to isolate** the BLOBs (objects) in a binary image. As mentioned above, a BLOB consists of a group of connected pixels. Whether or not two pixels are connected is defined by the connectivity, that is, which pixels are neighbors and which are not.

3.4 HSV Model - Skin color detection

HSL stands for hue, saturation, and lightness, and is often also called HLS. HSV stands for hue, saturation, and value, and is also often called HSB (B for brightness). A third model, common in computer vision applications, is HSI, for hue, saturation, and intensity.

3.5 Template matching

Template matching is a technique for finding areas of an image that match (are similar) to a template image (patch). Our goal is to detect the highest matching area. We need two primary components:

[1] Source image (I): The image in which we expect to find a match to the template image.

[2] Template image (T): The patch image which will be compared to the template image.

[4] Technology

Software Requirements

[1] OpenCV

[2] JDK

[3] Eclipse

Hardware Requirements

[1] 500GB HD

[2] 8GB RAM

[5] Feasibility Study

NP type problem -
A problem is assigned to the NP (nondeterministic polynomial time) class if it is solvable in polynomial time by a nondeterministic Turing machine. A is always also NP.

NP-Complete problem -
A given problem is said to be NP-Complete, if it can be solved within the determined polynomial time. The attribute values should completely satisfy the given polynomial expression. In computational complexity theory, a decision problem is NP-complete when it is both in NP and NP-hard.

This is considered with specifying equipment and software
that will successful satisfy the user requirement the technical needs of the system may vary considerably but might include
[1] The facility to produce outputs in a given time.
[2] Response time under certain conditions.
The image processing algorithms are NP type. Because we can get and verify the solution set. Hence the problem statement involving it are NP Complete.

[6] Mathematical Module

Set Theory Analysis:

Let 'S' be the | Signs of the hand configurations as the final set

S = {……
Identify the inputs as V, I
S = {V, I, …

V = {V1, V2, V3, V4, …| 'V' gives video input}
I = {I1, I2, … |'I' gives image input}

Identify the outputs as O, A
S = {V, I, O, A, …
O= {O1, O2, O3, … | 'O' get translated text}
A= {A1, A2, A3, … | 'A' get translated audio}

Identify the functions as 'F'

S = {V, I, O, A, F, …
F = {F1(), F2(), F3(), F4(), F5()}

F1(V): Video streaming
F2(V): Feature Extraction (3D)
F3(I): Image Input
F4 (V): Feature Extraction (2D Image)
F5 (O): Translated text or sound

Hence the functionality can be shown as,

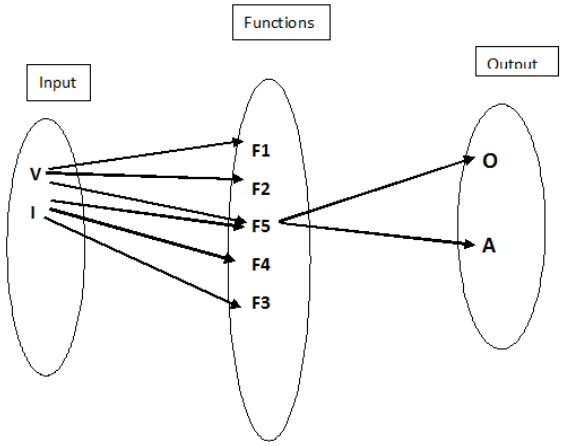

Figure : Functionality diagram

[7] System Features

[8] Advantages

[9] Challenges

Designing a hand gesture recognition system for Indian Sign Language(ISL) is more challenging than other sign languages due to the following reasons.

[1] Unlike other sign languages (American Sign Language, German Sign language) Indian Sign Language uses both hands to make sign.
[2] Some signs involve overlapping of both the hands and complicated hand shapes.
[3] One hand moves faster than the other at times in dynamic hand gestures.

[10] Conclusion and Future Scope

In this paper, we present the current status of the research on computer assisted sign language recognition systems. After thorough analysis, the following are conclusions for future research in sign language recognition:
[1] Current systems are mainly focused on static signs/ manual signs/ alphabets/ numerals.
[2] Standard dataset not available for all countries/sub continents / languages.
[3] A need for large vocabulary database is the demand for current scenario.
[4] Focus should be on continuous or dynamic signs and nonverbal type of communication.
[5] Sign language recognition systems should adopt data acquisition in any situation (not restricted to laboratory data).
[6] Systems should be able to distinguish face, hand (right/left) and other parts of body simultaneously.
[7] Systems should perform recognition task in a convenient and faster manner.

References

[1] A. Bonnaccorsi, "On the Relationship between Firm Size and Export Intensity," Journal of International Business Studies, XXIII (4), pp. 605-635, 1992. (journal style)
[2] R. Caves, Multinational Enterprise and Economic Analysis, Cambridge University Press, Cambridge, 1982. (book style)
[3] M. Clerc, "The Swarm and the Queen: Towards a Deterministic and Adaptive Particle Swarm Optimization," In Proceedings of the IEEE Congress on Evolutionary Computation (CEC), pp. 1951-1957, 1999. (conference style)
[4] H.H. Crokell, "Specialization and International Competitiveness," in Managing the Multinational Subsidiary, H. Etemad and L. S, Sulude (eds.), Croom-Helm, London, 1986. (book chapter style)
[5] K. Deb, S. Agrawal, A. Pratab, T. Meyarivan, "A Fast Elitist Non-dominated Sorting Genetic Algorithms for Multiobjective Optimization: NSGA II," KanGAL report 200001, Indian Institute of Technology, Kanpur, India, 2000. (technical report style)
[6] J. Geralds, "Sega Ends Production of Dreamcast," vnunet.com, para. 2, Jan. 31, 2001. [Online]. Available: http://nl1.vnunet.com/news/1116995. [Accessed: Sept. 12, 2004]. (General Internet site)

Author Profile

Taro Denshi received the B.S. and M.S. degrees in Electrical Engineering from Shibaura Institute of Technology in 1997 and 1999, respectively. During 1997-1999, he stayed in Communications Research Laboratory (CRL), Ministry of Posts and Telecommunications of Japan to study digital beam forming antennas, mobile satellite communication systems, and wireless access network using stratospheric platforms. He now with DDI Tokyo Pocket Telephone, Inc.