HAND GESTURE DETECTION USING CNN

KODATI SASI
PG Scholar, Department of Computer Science,
SVKP & Dr K S Raju Arts & Science College,
Penugonda, W.G.Dt., A.P, India
sasikodati123@gmail.com

K. LAKSHMANA REDDY
PG Scholar, Department of Computer Science,
SVKP & Dr K S Raju Arts & Science College,
Penugonda, W.G.Dt., A.P, India
klreddy29@yahoo.co.in

ABSTRACT

One of the major drawbacks of our society is the barrier that is created between disabled or handicapped persons and the normal person. Communication is the onlymedium by which we can share our thoughts or convey the message but for a personwith disability (deaf and dumb) faces difficulty in communication with normal person. For many deaf and dumb people, sign language is the basic means of communication. Sign language recognition (SLR) aims to interpret sign languages automatically by a computer in order to help the deaf communicate with hearing society conveniently. Our aim is to design a system to help the person who trained the hearing impaired to communicate with the rest of the world using sign language or hand gesture recognition techniques. In this system, feature detection and featureextraction of hand gesture is done with the help of CNN and OpenCV.

INTRODUCTION

1.1 Introduction

Image processing is a rapidly growing area in diverse applications, such as multimedia computing, secured data communication, biomedical, biometrics, remote sensing, texture understanding, pattern recognition, content-based retrieval, compression, and many more. This is all about how a computer can sense pictorial data after processing an image. Among the set of gestures intuitively performed by humans when communicating with each other, pointing gestures are especially interesting for communication and is perhaps the most intuitive interface for selection. They open up the possibility of intuitively indicating objects and locations, e.g., to make a robot change moving direction or simply mark some object. This is particularly useful in combination with speech recognition as pointing gestures can be used to specify parameters of location in verbal statements. This technology can be a boon for disable people who are not able to speak hence can't communicate. Also if the person has different language than receiver, then also, it can be used to as translator. There has been always considered a challenge the development of a natural interaction interface, where people interact with technology as they are used to interact with the real world. A hand free interface, based only on human gestures, where no devices are attached to the user, will naturally immerse the user from the real world to the virtual environment.

Hands are human organs which are used to manipulate physical objects. For this very reason

hands are used most frequently by human beings to communicate and interact with machines. Mouse and Keyboard are the basic input/output to computers and the use of both of these devices require the use of hands. Most important and immediate information exchange between man and machine is through visual and actual aid, but this communication is one sided. Computers of this age provide humans with 1024 * 768 pixels at a rate of 15 frames per second and compared to it a good typist can write 60 words per minute with each word on average containing 6 letters. To help somewhat mouse remedies this problem, but there are limitations in this as well. Although hands are most commonly used for day to day physical manipulation related tasks, but in some cases they are also used for communication. Hand gestures support us in our daily communications to convey out- messages clearly. Hands are most important for mute and deaf people, who depends their hands and gestures to communicate, so hand gestures are vital for communication in sign language. If computer had the ability to translate and understand hand gestures, it would be a leap forward in the field of human computer interaction. The dilemma, faced with this is that the images these days are information tick and in-order to achieve this task extensive processing is required.

2. LITERATURE SURVEY

[1] Title: Using Multiple Sensors for Mobile Sign Language Recognition

Author: Helene Brashear, Thad Starner, Paul Lukowicz & Holger Junker

The authors built a constrained, lab-based Sign Language recognition system with the goal of making it a mobile assistive technology. They examine using multiple sensors for disambiguation of noisy data to improve recognition accuracy. The experiment compares the results of training a small gesture vocabulary using noisy vision data, accelerometer data and both data sets combined. The authors chose to use a rule–based grammar for sentence structure in the training and testing process. Speech recognition often uses statistical grammars for increased accuracy. These grammars are built by tying together phonemes (the simplest unit of speech) and training on the transition between the phonemes. The sets are usually done with bigrams (two phonemes tied together) or trigrams (three phonemes). Training using bigrams or trigrams requires considerably more data because representations of each transition of each word are now needed. In our case, the bigrams and trigrams would be built by tying together gestures. The current data set is too small to effectively train using bigrams or trigrams, but we intend to continue collecting data with the goal of implementing these techniques.

[2] Title: A Vision Based Dynamic Gesture Recognition of Indian Sign Language on Kinect based Depth Images

Author: Geetha M, Manjusha C, Unnikrishnan P and Hari Krishnan R

Indian Sign Language (ISL) is a visual-spatial language which provides linguistic information using hands, arms, facial expressions, and head/body postures. The proposed work aims at recognizing 3D dynamic signs corresponding to ISL words. With the advent of 3D sensors like Microsoft Kinect Cameras, 3D geometric processing of images has received much attention in recent researches. The authors have captured 3D dynamic gestures of ISL words using Kinect

camera and has proposed a novel method for feature extraction of dynamic gestures of ISL words. While languages like the American Sign Language (ASL) are of huge popularity in the field of research and development, Indian Sign Language on the other hand has been standardized recently and hence its (ISLs) recognition is less explored. The method extracts features from the signs and converts it to the intended textual form. The proposed method integrates both local as well as global information of the dynamic sign. A new trajectory-based feature extraction method using the concept of Axis of Least Inertia (ALI) is proposed for global feature extraction. An Eigen distance-based method using the seven3D key points- (five corresponding to each finger tips, one corresponding to center of the palm and another corresponding to lower part of palm), extracted using Kinect is proposed for local feature extraction. Integrating 3D local feature has improved the performance of the system as shown in the result. Apart from serving as an aid to the disabled people, other applications of the system also include serving as a sign language tutor, interpreter and also be of use in electronic systems that take gesture input from the users.

3. OVERVIEW OF THE SYSTEM

3.1 Existing System

Existing proposed a technique of first computing the similarity of different gestures and then assign probabilities to them using Bayesian Interface Rule. These classes consist of Hu moments with geometrical attributes like rotation, transformation and scale in variation which were used as features for classification. Performance of this technique was very well and it was giving 95 °/o accurate results.

Existing system also proposed a similar technique which also uses H U-moments along with modified KNN (K-Nearest Neighbor) algorithm for classification called as Locally Weighted Naive Bayes Classifier.

Classification results where this technique were 93°/o accurate under different lighting conditions with different users.

Disadvantages:

- Irrelevant object might overlap with the hand. Wrong object extraction appeared if the objects larger than the hand.
- Performance recognition algorithm decreases when the distance is greater than 1.5 meters between the user and the camera.
- System limitations restrict the applications such as the arm must be vertical, the palm is facing the camera and the finger color must be basic color such as either red or green or blue.
- Ambient light affects the color detection threshold.

3.2 Proposed System:

The first step for our proposed system is the capturing of the video using webcam where different alphabets were taken into consideration. Skin Filtering was performed to the input video frames for detection of hand gestures. It was done so that the required hand could be extracted from the background. Skin Filtering is a technique used

for separating the skin coloured regions from the non-skin coloured regions. In our proposed system there are

5 modules: real time Input image from webcam, pre-processing and segmentation, feature extraction, classification and Results analysis (gesture recognition). For gesture recognition is real time recognized in live camera

.The proposed system are used in SVM (Support Vector Machine), K- Neighbours-Classifier, Logistic-Regression, MLP-Classifier, Naive Bayes, Random-Forest-Classifier algorithms. We propose an easy-to-use and inexpensive approach to recognize single handed as well as double handed gestures accurately. This system can definitely help millions of deaf people to communicate with other normal people. A fast, novel and robust system was proposed for recognition of different alphabets of Indian Sign Language for video sequences. The proposed system is a real time video processing that is based on a real time application system.

Advantages:

- There are no moving parts, so device wear is not an issue.
- We have proposed a system which is able to recognize the various alphabets of Indian Sign Language for Human-Computer interaction giving more accurate results at least possible time.
- Accuracy rate obtained was 98% but it lacks proper Skin filtering with changes in illumination.
- CNN Algorithm were used to recognize the gestures.

3.3 System Modules

Dataset collection:

In this stage data set is collecting by running camera and pictures are captured for each gesture and folder is saved with one alphabet for each gesture.

pre-processing:

In this stage data set is pre-processing by converting images to required size and images are covered to required back ground corrections.

Splitting Dataset:

data set is divided in to folders features and labels and where features are images signs and labels are alphabets.

Initialize algorithm:

In this stage CNN algorithm is used and training data is fit in to algorithm.

Save Model:

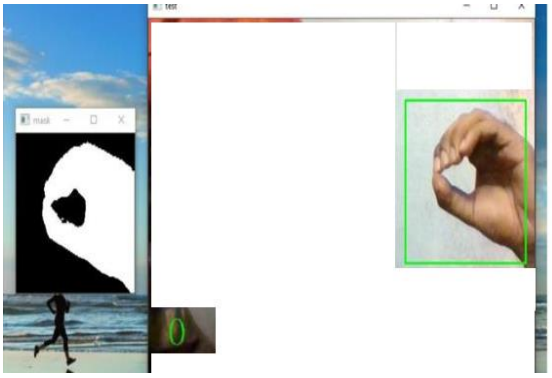
After training data is saved as .h5 model which is used for predicting purpose in future.

Predict:

In this stage live camera will open and signs are given by user based on signs alphabets will be displayed on camera.

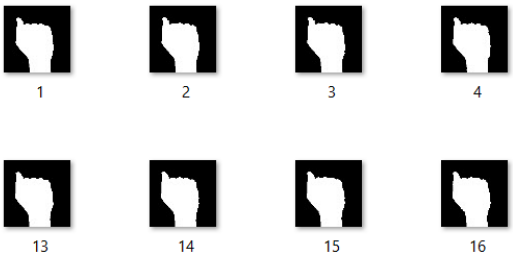
4. RESULTS

Main page

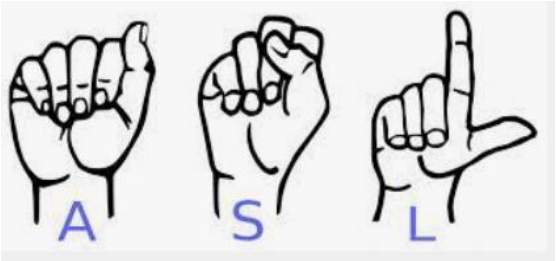


Detected as 0

Example gestures for training



Training set sample for Alphabet A.



Gestures for Prediction

5. CONCLUSION

we have proposed a system for recognizing a dynamic hand word gesture of Indian sign language and conversion of recognized gesture into text and speech and vice versa i.e. dual way communication. In this system skin color filtering technique has used for segmentation. Eigen vectors and Eigen values technique has used for feature extraction. For classification, Eigen value weighted Euclidean Distance based classifier has used. Prediction of words sign using one or both hands, working with Indian Sign language dynamic hand gesture words dataset and dual way communication has proposed in this system.

Future Enhancements:

we can use hand gesture applications in real time applications for handling various applications by connecting to hardware devices.

- For future work, there's still so many possibilities of improvement, like noise reduction, dataset collection, better feature extraction, or better model to be used.
- In future, we are looking at developing a system for Indian sign language words that works in real-time. And, we will make efforts to extend our work towards more words and sentences of Indian sign language.

REFERENCES

[1] A. Chaudhary, J.L. Raheja, K. Das, S. Raheja, "A Survey on Hand Gesture Recognition in context of Soft Computing", Published as Book Chapter in "Advanced Computing" CCIS, Springer Berlin Heidelberg, Vol. 133, 2011, pp. 46-55

[2] A. Chaudhary, J.L. Raheja, K. Das, "A Vision based Real Time System to Control Remote Robotic hand Fingers", In proceedings of the IEEE International Conference on Computer Control and Automation, South Korea, 1-3 May, 2011, pp. 118-122

[3] J.L. Raheja, A. Chaudhary, S. Maheshwari, "Automatic Gesture Pointing Location Detection", Optik: International Journal for Light and Electron Optics, Elsevier, Vol. 125, Issue 3, 2014, pp. 993-996.

[4] A. Chaudhary, K. Vatwani, T. Agrawal, J.L. Raheja, "A Vision-Based Method to Find Fingertips in a Closed Hand", Journal of Information Processing Systems, Korea, Vol. 8, No. 3, Sep 2012, pp. 399-408.

[5] A. Chaudhary, J.L. Raheja, K. Das, S. Raheja, "Fingers' Angle Calculation using Level-Set Method", Published as Book Chapter in "Computer Vision and Image Processing in Intelligent Systems and Multimedia Technologies", IGI,

USA, April 2014, pp.191-202

[6] D. Sturman, D. Zeltzer, "A survey of glove-based input", IEEE Transactions on Computer Graphics and Applications, Vol. 14, No. 1, Jan. 1994, pp. 30-39.

[7] T.S. Huang, and V.I. Pavlovic, "Hand gesture modeling, analysis and synthesis", Proceedings of international workshop on automatic face and gesture recognition, 1995, pp.73-79.

[8] O.Y. Cho, and et al. "A hand gestue recognition system for interactive virtual environment", IEEK, 36-s(4), 1999, pp. 70-82.

[9] M. Sonka, V. Hlavac, R. Boyle, "Image processing, analysis, and machine vision"2014, Cengage Learning.

ABOUT AUTHORS:

KODATI SASI is currently pursuing MCA in SVKP & Dr K S Raju Arts & Science College, affiliated, to Adikavi Nannaya University, Rajamahendravaram. His research interests include Data Structures Web Technologies, Operating Systems and Artificial Intelligent.

K. Lakshmana Reddy is working as an Associate in the Department of Computer Science in SVKP & Dr K S Raju Arts &Science College, Penugonda, A.P. He received MCA from Andhra University, 'C' level from DOEACC, New Delhi and M.Tech from Acharya Nagarjuna University, A.P. He attended and presented papers in conferences and seminars. He has done online certifications in several courses from NPTEL. His areas of interests includes Computer Networks ,Network Security and Cryptography, Formal Languages and Automata Theory and Object Oriented programming languages.