Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 48 (2015) 673 – 678

Procedia Computer Science

International Conference on Intelligent Computing, Communication & Convergence

(ICCC-2015)

Conference Organized by Interscience Institute of Management and Technology,

Bhubaneswar, Odisha, India

Sign Language Unification: The Need for Next Generation Deaf Education

Vinay Kumar K*, R.H.Goudar*, V T Desai**

*Dept. Of CNE, Visvesvaraya Technological University, Belagavi-590018, Karnataka

**Dept. of MCA, KLE Dr. M S Sheshgiri College of Engineering and Technology, Belagavi.

vinu.vinay036@gmail.com, *rhgoudar@gmail.com*

Abstract

Speech plays a basic role in communication between people, in education, in sharing ideas and in maintaining social contacts. The hearing-impaired have to challenge communication barriers in a mostly hearing-capable society. Availability of real-time speech transcription anywhere, anytime, represents a potentially life-changing opportunity for the hearing impaired to improve their communication capability. Sign Language is the most important communication way between hearing impaired community and normal persons. There are approximately three hundred Sign Languages that are used around the world today and it is considered to be the sixth largest language used worldwide. The Sign Languages are developed depending on the nation and locality of the deaf people. Since there are many different Sign Languages, it poses a difficulty for a deaf person to communicate with other deaf person from different region. The aim is to develop a Sign Language translator that eases the communication of the deaf people.

Keywords: American Standard Langauage;Chineese Sign Language;Indian Sign Langauge;

Peer-review under responsibility of scientific committee of International Conference on Computer, Communication and Convergence (ICCC 2015)
doi:10.1016/j.procs.2015.04.151

1. Introduction

There are differently able/ challenged people who have some obvious problems that make them feel that they are different and which makes them feel very difficult when they want to communicate with the rest of the normal world. Deaf people cause serious problems when sharing their feelings to the people who cannot understand their Sign Languages. The only way deaf people feel easy to communicate with others is using Sign Languages and hence there is a need of unified Sign Language that becomes common for entire world.

A Sign Language is a mode of communication which is used by deaf people. Sign languages are widely used by deaf people around the globe and it is the most primary mode of communication for them. Most of the deaf people use Sign Languages of their own country and each Sign Language is influenced by their own nativity and culture. There are more than three hundred Sign Languages across the world and hence it throws a great difficulty if a deaf person wants to communicate with his counterpart from other countries. Hence there is a need to translate one Sign Language to other so that the communication and other things can be easily understood.

Objective

- The objective is to develop a Sign Language translator that eases the communication of the deaf people.
- It consists of a camera with sensors that capture's the motion of the deaf as they sign and the sensors in that identifies its trajectory.
- Once the trajectory is identified, then it is converted to other sign language by matching the performed sign with the sign which is archived.
- The converted sign is also displayed on a device and a text message is also displayed to ease communication with a person who can hear.

2. Literature review

The lack of early Sign Language acquisition pose a difficulty for deaf children to learn basic concepts of daily life. The learning of such basic daily life concepts should occur when the child recognizes patterns in the world and start to identify linguistic labels for them. Later, the child learns to ask questions in order to clarify doubts, and starts to form relations that alter their cognitive structures: Children begin to combine, compare, infer, deduce and extrapolate old and new knowledge, in a mental process that is mediated by social experiences and language. This lack of mother tongue acquisition is detrimental to the intellectual development of Deaf children, and brings about severe consequences, such as the inability to perform daily tasks for the development of intelligent action; the inability to learn and plan; the inability to overcome impulsive behavior, the dependency on concrete, visual situations; difficulties to control themselves and to socialize, among others. Thus deprived of intellectual development and use of language, members of the Deaf community face intellectual and social barriers of information access and knowledge creation, needed for identity formation and full citizenship. Additionally, the Deaf grows into a reality where there is little written material in Sign Language [1].

Gesture recognition is possible based on 3D data from swiss Ranger SR-2 camera which produces both depth as well as intensity image of a scene. The gesture recognition is based on the finding motion of the primitives in the 3D data. The primitives are represented compactly and view invariant using harmonic shape content. Many detecting devices are used and one which is used here is the CSEM swiss ranger camera. Each pixel in this camera provides a depth value. Even though the resolution of the data is less, they are used in typical computer vision applications like face detection, face tracking, shape analysis and robot navigation. The camera also provides a amplitude value corresponding to an intensity value for each pixel. The data acquisition here is through the capture of these using the camera. The camera is based on the time of flight and emits a audio frequency modulated light in the near infrared spectrum. The movements are detected using a 3D version of 2D double differencing. This is done by subtracting the depth value pixel wise in two pairs of depth images, threshold and finally ANDING two binary images[2].

Plan view projection of a real-time depth imaginary can improve the statistics of its intrinsic 3D data and allows for cleaner separation of occluding and closely interacting people. Objective is to build a probabilistic, real time multi person tracking system upon a plan view image substrate that well preserves both shape and size information of

foreground objects. The tracking robustness derives in part from its plan view template person models, which capture people's detailed body configurations. The cost of the camera are inexpensive and can be used to produce dense depth or disparity imaginary, obtained from stereo camera. Person tracking and activity recognition into a single real time framework based on Plan view template model is obtained. Then these models are allowed to adapt over time during tracking, to reflect changes in the people's body configuration and visibility to the camera. SUPPORT VECTOR MACHINES(SVM) are used to quickly and reliably discriminate between poses such as sitting, standing and reaching or facing left rather than right. Conversions of the SVM outputs to possibilities and integration of these probabilities over time, allows for principled, stable detection of human activities and events such as sitting down and turning face left[3].

Sign language (SL) recognition, although has been explored for many years, is still a challenging problem for real practice. The complex background and illumination conditions affect the hand tracking and make the SL recognition very difficult. Fortunately, Kinect is able to provide depth and color data simultaneously, based on which the hand and body action can be tracked more accurately and easily. Therefore, 3D motion trajectory of each sign language vocabulary is aligned and matched between probe and gallery to get the recognized result. Based on proposed 3D trajectory matching algorithm, a Sign Language recognition and translation system is built to connect hearing impaired community and normal persons. System consists of two modes: Translation Mode, in which it translates sign language into text or speech and Communication Mode, in which a normal person can communicate with the signer through an avatar. Translation mode includes isolated word recognition and sentence recognition. In current system, raising and putting down hands are defined as the start and end gesture of each SL word to be recognized. For sentence recognition, all the words can be input continuously and then the system will give the results by integrating both the SL matching score and the probability given by SL language model. Also, the signer can adjust or confirm the results by manual interaction[4].

Depth imaging technology has advanced dramatically over the last few years, and has finally reached a consumer price point. Pixels in a depth image indicate the calibrated distance in meters of 3D points in the world from the imaging plane, rather than a measure of intensity or color. The Kinect depth camera is used to provide our input data. Kinect uses structured infrared light and can infer depth images with high spatial and depth resolution at 30 frames per second. Using a depth camera gives several advantages for human pose estimation. Depth cameras work in low light conditions (even in the dark), help remove ambiguity in scale, are largely color and texture invariant, and resolve silhouette ambiguities. They also greatly simplify the task of background subtraction[5].

3. Proposed Work and Methodology

Method 1: To make library of archive's from different sign language and to teach very basic sign language of other language to deaf people.

For this approach, it needs extensive work of analyzing and collecting the different kinds of sign language and to make them archive of that. The library constructed should contain all the possible as well as available signs. There are almost 300 sign languages available and it needs a great effort in collecting all the sign's of different language and use them appropriately. To achieve this it needs that there should be an expertise way so that care should be taken that no sign is missed from a particular sign language. It should also be noted that at least some emergency as well as some basic sign languages should be made common so that all the deaf people across the world use the same sign languages in case of an emergency.

Advantages: The main advantage of this approach is that it reduces the amount of difficulty in conveying the things when deaf people travel to some other parts of the world.

Method 2: To have a tool or application that can convert one Sign Language to another.

This is the most important problem that has to be addressed. There is a need to build an application which enables the deaf to convert one Sign Language to other. The application should contain an efficient algorithm that can be used to convert Sign Language to another. The algorithm should take any Sign Language as input and should be able

to convert it to desired Sign Language. This conversion of Sign Language from one another is very challenging as it needs real time capture of actions and to show necessary sign using some animation.

Advantages: There will be no need of interpreters. The communication becomes easy and it creates a platform that eases the deaf to express their ideas.

Method 3: To create a Sign Language that becomes universal.

This needs the researchers from different sign language have to work together. Identify the most common signs that are used and to build a unique Sign Language that becomes universal. It should cover all signs and some more signs need to be added so that it can cover all the sign which can be used in that particular native of the deaf people. If this becomes reality then there is no need of some interpreters or some other assistance to communicate with other counterparts of the deaf.

Method 4: To have a real time gesture recognition system.

This approach aims in building a gesture recognition system so that it captures the gestures or Sign Language and it may be used to convert it to some other word. It should contain real time sensors which capture the real time gestures of Sign Language and should provide the equivalent sentence or a sign as an output.

Fundamentally how it is done :

- To develop an algorithm which is capable of transforming one Sign Language to another.
- As soon as the application is started we should be able to choose between the two Sign Languages for which the conversation is needed.
- Application makes use of the camera to capture the sign performed.
- Once after finishing with the sign, the algorithm starts the process of converting the Sign Language.
- The main need here is to correctly identify the trajectory of the sign and to convert it to exactly what it means.
- Sensors correctly identify the trajectory and they match with the predefined vocabulary and convert it to necessary Sign Language.
- Some animations may be used which shows the Sign Language in other form.
- It should also provide voice output to ease communication with the hearing enabled person.
- The application should be portable to any smart phones and system.

1: Image acquisition

The signing by a user is obtained by using a camera. The camera with sensors captures all sign and provides an input to the next stage.

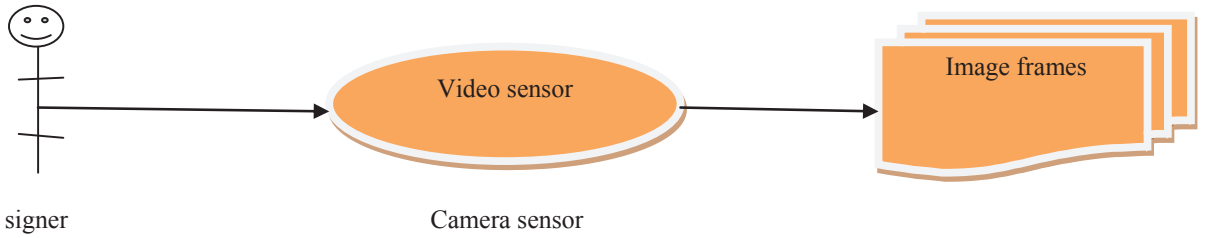

Fig 1: image acquisition

2: Pre-processing

Various factors like illumination, background, camera parameters and viewpoint are used to address scene complexity. These scene conditions affect images of camera object dramatically. A moving average or median filter is used to remove noise from image scenes. Gaussian average method is used to obtain background subtraction.

3: Feature extraction

Hand shape and movement of the hand are major one's in guessing word or sentences. Features are divided into hand shape and hand movement and the state of the hand are given by parameters called point of interest (POI).

4: Combinational Neural Network Systems

The Combinational Neural Networks are specifically developed for sign language recognition and it is based on gestures of signs that are stored in a database. It can be defined as a multiprocessor system network with high degree of interconnection and adaptive interaction between elements. A neural network image processor can free imaging applications from various constraints in terms of video acquisition, lighting and hardware settings. Object recognition using neural networks is based on the learning of the surveillance network using a database set of objects.It consists of three layers stage1, stage 2 and stage3. Each layer or stage takes elements of feature vector as input and outputs the class of object.

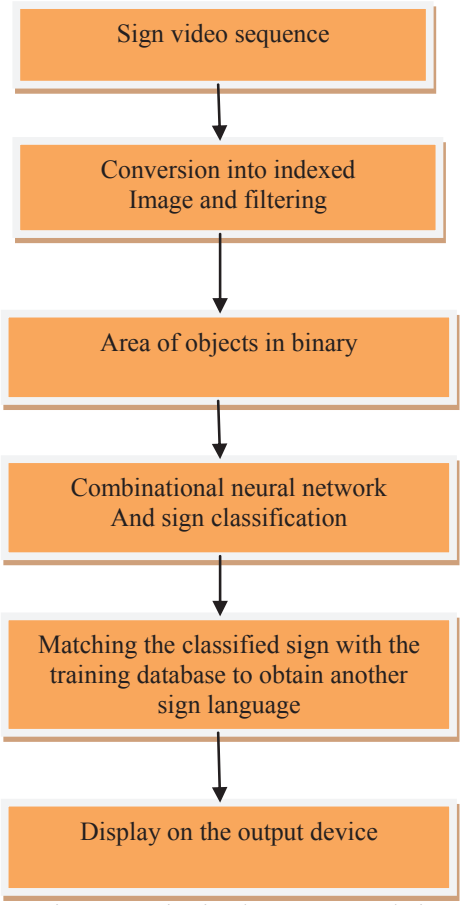

Fig 2: steps in sign language translation

Conclusion

In this paper we have investigated the difficulties faced by deaf people in various aspects such as communication with others as well as their counterparts from different parts of the world and the problem they face in education, understanding technology and conveying their feelings to others. There is no tool existing that can help in betterment of deaf in understanding various aspects and to express their views, thoughts and ideas. Hence in this paper it is proposed here a unified Sign Language or a Sign Language translator that helps in reducing the problem the deaf are facing today**.**

References

1. Cayley Guimarães, Jeferson F. Guardezi, Luis Eduardo Oliveira, Sueli Fernandes "Deaf Culture and Sign Language Writing System – a Database for a New Approach to Writing System Recognition Technology" Federal University of Parana – UFPR {cayleyg, jfguardezi, lesoliveria
2. M.B. Holte, T.B. Moeslund and P. Fihl "Fusion of Range and Intensity Information for View Invariant Gesture Recognition" Computer Vision and Media Technology Laboratory Aalborg University, Denmark
3. Michael Harville Dalong Li, Palo Alto "Fast, Integrated Person Tracking and Activity Recognition with Plan-View Templates from a Single Stereo Camera" Hewlett-Packard Laboratories Center for Signal and Image Processing., CA 94304 Georgia Inst. of Technology, Atlanta.
4. Xiujuan Chai, Guang Li, Yushun Lin, Zhihao Xu, Yili Tang "Sign Language Recognition and Translation with Kinect" , Xilin Chen Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS Beijing, China {xiujuan.chai, guang.li, yushun.lin, zhihao.xu, yili.tang, xilin.chen}vipl.ict.ac.cn
5. Jamie Shotton, Ross Girshick, Andrew Fitzgibbon,Mat Cook, Mark Finocchio, Richard Moore, "Efficient Human Pose Estimation from Single Depth Images", Senior Member, IEEE.
6. Priyanka Mekala[1], Senor Zafer Erdogan[2], Jeffrey Fan "Automatic Object Recognition Using Combinational Neural Networks in Surveillance Networks" Dept. of Electrical and Computer Engineering, Florida International University, Miami, Florida, U.S.A Dept. of Software Engineering, Maltepe University, Istanbul, Turkey
7. Priyanka Mekala1, Ying Gao2 , Jeffrey Fan1 , Asad Davari3 "Real-time Sign Language Recognition based on Neural Network Architecture" 1Dept. Of Electrical and Computer Eng., Florida International University, FL, U.S.A 2Electrical Eng. Dept., University of Wisconsin, Platteville, WI, U.S.A 3Dept. Of Electrical and Computer Eng., West Virginia University Institute of Tech, WV, U.S.A.