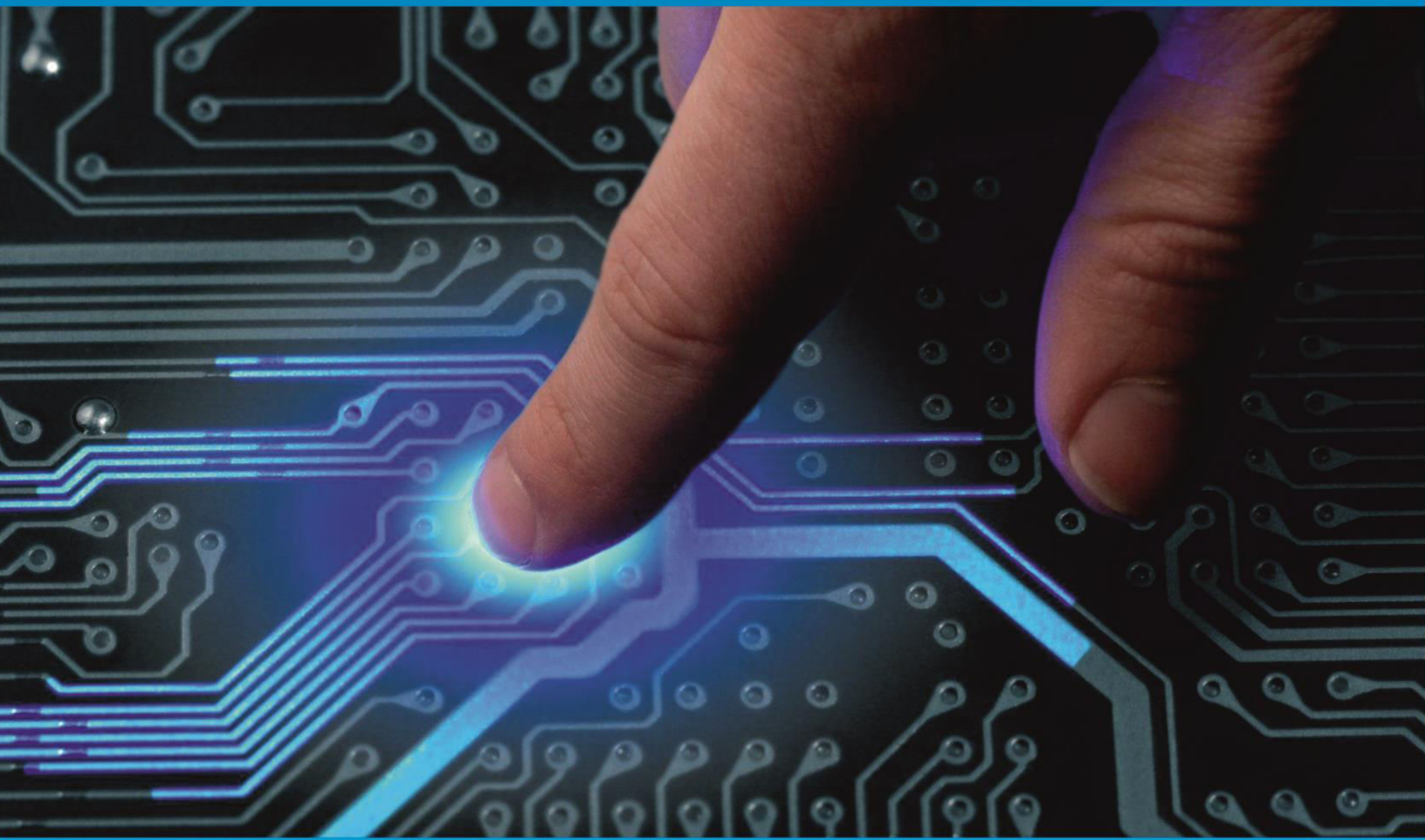
e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 3, March 2021

Impact Factor: 7.488

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021. 0903114|

Machine Learning Algorithm for Sign Language Recognition through Image Processing

Aarthi A, Akila A, Aparnaa S, Dharani K, Dr. Deepa S

U.G. Scholar, Department of ECE, Panimalar Engineering College, Chennai, India

U.G. Scholar, Department of ECE, Panimalar Engineering College, Chennai, India

U.G. Scholar, Department of ECE, Panimalar Engineering College, Chennai, India

U.G. Scholar, Department of ECE, Panimalar Engineering College, Chennai, India

Professor, Department of ECE, Panimalar Engineering College, Chennai, India

ABSTRACT: The Sign Language Recognition (SLR) Problem is a mostessential research topic, to increase the ability interaction between those people who are hearing- impaired in speech. However, there are several limitations of the existing methods. Most applicationsneed different necessities like making the user wear multi-coloured or sensor-based gloves or usage of a specific camera. We propose a simple but robust system that can be used without the need of any specific accessories. The proposed system consists of three main steps. First, we are going to apply segmentation to the face and hand region by using Fuzzy C-Means Clustering (FCM) and Thresholding. FCM is a clustering technique which is used to employs fuzzy partitioning, in an iterative algorithm. After the face and hands are segmented, the feature vectors are extracted. The feature vectors are chosen among the low-level features such as the bounding ellipse, bounding box, and centre of mass coordinates, since they are known to be more robust to segmentation errors due to low resolution images. After the process of Segmentation and Clustering is completed, we get the enhanced images and we convert that images into speech. In total there are 23 features for each hand, the system is shown to be working with 85.8% accuracy in the user independent case and 100% in user dependent case.

KEYWORDS: Sign Language Recognition, Machine Vision, Machine Learning, Discrete Hidden Markov Model, Fuzzy C- Means Clustering.

I. INTRODUCTION

People express time-varying motion patterns (Gestures) in order to transmit a message to a receiver. In this generation the most popular receivers are the computers. Speech and Gesture are the naturally used methods for communication between humans and these techniquesare imitated by the Human-Machine Interface (HMI) systems. The research on this speech recognition is more mature than gesture recognition. As a result of the recent development in the quality and availability of the phone and web cameras, gesture recognition has started to attract attention of many researches and shown to be growing potential. Probably the most explicit example of human gestures is sign language, which has a well-defined vocabulary and grammar. Sign language gestures (SLG) composed of many elements such as facial expression, body shape and hand movements. However, the most important information about the performed gesture is conveyed through the hands symbol. This is the reasonwhy we concentrate on this paper mainly on the hand gestures. Most of the researches in the literature are based on the SLR which requires the multi-coloured are sensor-based gloves or usage of a specific camera. This reduces the user-friendliness of the system. In this paper work, a generic method to recognise the isolated sign language gestures in a signer independent way is presented.

II. LITERATURE SURVEY

Most of the SLR systems can be divided into the 3 main steps. First, data acquisition, and tracking of the hand or the body parts must be utilised. In the Second step, features describing the manual or non-manual information must be defined and extracted from the input data. In the Last step by using the data gathered from the feature descriptors, the performed sign is classified. This paper covers the manual features and non-manual features are not in the scope of this study [1].

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021. 0903114|

Varun Tiwari, Vijay Anand, A.G. Keskar, V.R. Satpute, 2015. This paper makes the interaction easier between the normal people and people with hearing disabilities. By using Kinetic depth camera and neural network through we can capture and crop the image by removing the unwanted region. R.K. Shangeetha, V. Valliammai, S. Padmavathi, 2012. Hand gestures corresponding to Indian Sign Language English alphabets are captured through a webcam. In the captured frames the hand is segmented and the state of fingers is used to recognize the alphabet words. The number of fingers is used to recognize the alphabet.

Beatriz Tomazela Teodora, Joao Bernardes, Luciano Antonio Digiampietri, 2017.Skin language automatic recognition mitigate the obstacles of deaf people. This language recognizes those image processing, human skin segmentation, object tracking through videos for the skin colour-based segmentation stage [9]. In this we useBSL sign (Brazilian Sign Language). Geethu G Nath, V.S. Anu, 2017. Deaf and Dumb conversation is attached on sign language which make words through hand and finger symbols [2]. Main objective of the project is building a device that help deaf and dumbpeople to convey their messages to normal people. Sign language is the only medium through which disabled people can connect to rest of the world through different hand gestures [6]. In current time, the building was developed for dumb person using flex sensor, there user hand is attached with the flex sensors. In this technique the flex sensor reacts on bend of each finger individually. By taking that value controller starts to react with speech, each flex sensor otter unique voice stored in APR kit and for each sign it will play unique voice. And in other existing system, the work is done only for some alphabets and not for the words are sentences, and accuracy obtained is very low. The system is shown to be working with 85.8% accuracy in the user independent case and 100% in dependent case. The limitations of existing systems are it's restricted to only 10 voice announcements it may reduce product capacity. The controller may think that the user is giving command and finally it may result in unwanted results and less hardware lifetime.

III. PROPOSED SYSTEM

The system will be implemented through a desktop with a Full-HD web camera. The camera will capture the images of the hand that will be feed in the system. Note that the signer will adjust to remove to the size of the frame so that the system will be able to capture the orientation of the signer's hand. The camera has already captured the gesture from the user, then segmentation and clustering will be processed to get enhanced features. Then feature extractions and classification algorithms are used to translate the sign language into English text. This translation is converted to speech using text to speech Application Programming Interface (API). In the vision-based method, web camera used to capture images. After that, image segmentation has done [10]. Feature like palm, finger extracted from input image. Different hand motion is there that is half closed, fully closed, semi closed was detected. Data is saved in vector and that vector is used for sign recognition of alphabets. For implementation, hand gesture was already captured and stored in the file. Histogram is used for feature extraction. Feature extraction was used to extract motions and shape of hands. For further processing neural network is used.

3.1 Sign Language Recognition

Recognition is the main step in an SLR system. Previously gathered feature vectors are transformed into the meaningful signs in this step. The first studies in SLR started with ANN, HMM based methods are more widely used lately. After the HMMs are applied to the speech recognition successfully, they become widely used for gesture recognition due to the similar nature of two pattern recognition problems. Although the most widely used methods are HMM and ANN, there exist other methods used to classify the sign gestures [8]. In this thesis work, to recognize the gestures HMMs are used.

3.2 Hidden Markov Model Application

After the segmentation and feature extraction steps feature vectors are ready for recognition phase. There are 5 different feature vectors set for every gesture and for each performer. For testing one performer's data is used and 5 different test feature vectors could be used for testing in user independent case. For user dependent classification 4 feature vector set of every gesture and for each performer is used and the remaining 1 feature vector sets used for classification. While constructing the HMMs, parallel training approach is accepted [4].

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0903114|

3.3 Hand Segmentation

Hand segmentation is the beginning stage in many of the SLRs. In order to gather the valuable information and analyse the acquired video correctly, one has to extract the desired data in the entire set of pixels from the sampled image of the video. Since hands do not have a strict shape, and change in size for different gestures, segmentation could be a very challenging task in an SLR [5].

3.4 Clustering

Clustering can be considered as the most essential not done learning problem; so, as every other problem of this kind, it deals with finding a structure in a group of things unlabelled data. In the first case, the data is grouped in an limited way, so that if a certain datum belongs to a definite bunch then it could not be included in another cluster. On the country, the second type, the overlapping bunching, uses appearance sets to cluster data, so that each point may belong to two or more bundles with different degrees of membership. The third type, the hierarchical bunching algorithm is based on the union between the two nearest bunches. After a few repetitions it reaches the final bunches wanted. Finally, the last kind of bunching uses a completely probabilistic approach.

Proposed solution is SLR. SLR can be a desirable interpreter which can help both the community general and deaf. In this work, we proposed an idea for feasible communication between hearing and normal person with the help of Machine learning approach. The proposed HRF employs a hierarchical recurrent architecture to encode the visual semantics with different visual granularities. The core steps of the Hierarchical deep Recurrent Fusion(HRF) are based on an encoder-decoder framework, the proposed HRF translates a video into neural languages after encoding the entire visual content. This framework can solve the disordered gloss label issue and we employ Adaptive Clip Summarization(ACS) to explore sign action patterns in SLT [3]. Differing from previous studies that extract key frames are clips with a fixed time interval or fixed clip number, we propose an adaptive temporal segmentation scheme, i.e., ACS. The proposed scheme automatically obtains variable-sized key frames/clips and implements dynamic temporal pooling on less important frames/clips.Figure 1 shows the Block Diagram of the Proposed system.

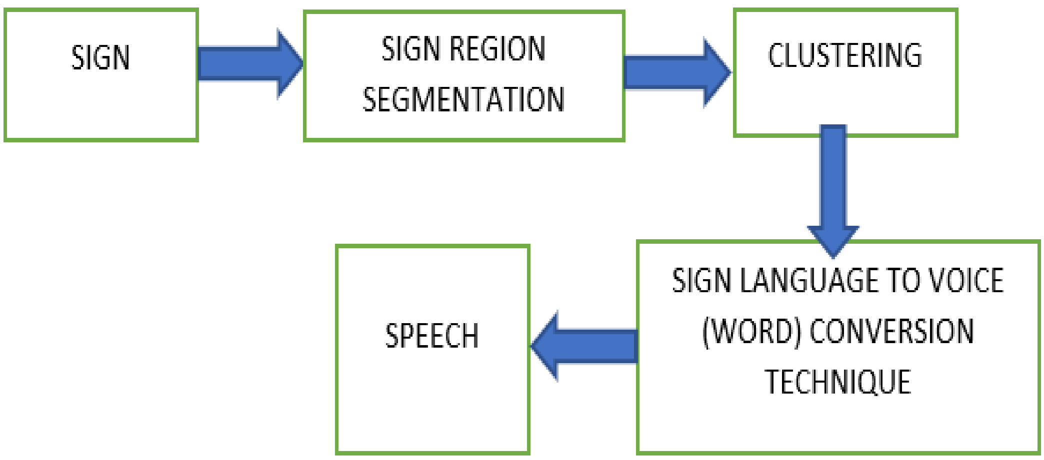


Figure 1: Block Diagram

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE2021. 0903114|

IV. EXPERIMENTAL RESULTS

Firstly, we capture the hand image frame using a webcam. Figure 2 and 3 shows that theframes are segmented to get enhanced features.

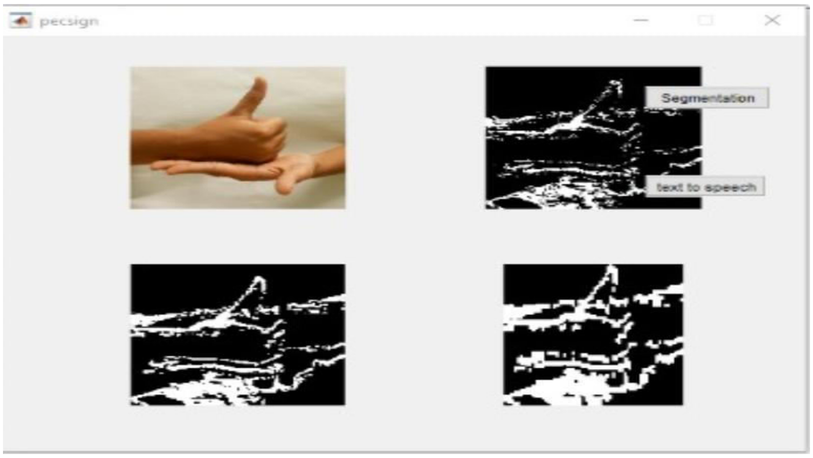


Figure 2: Segmentation

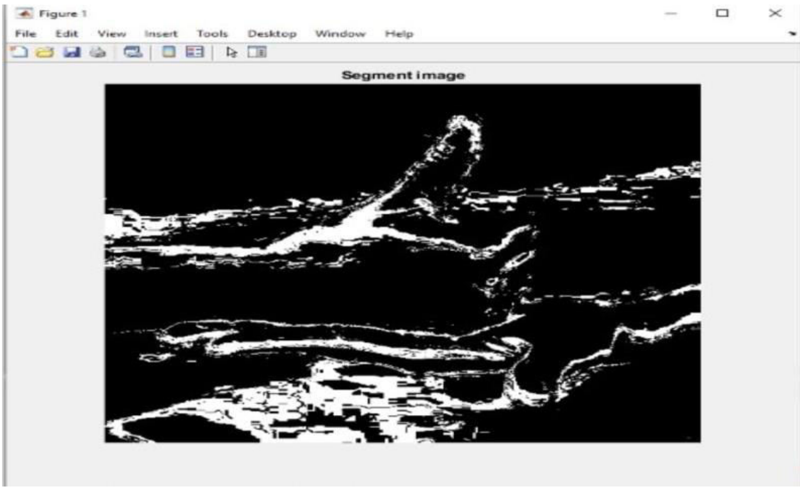


Figure3: Segmented image

After segmenting the image, the next step we have do is clustering. Figure 4 represents that the images are clustered to get enhanced features. The clustering step is done according to the method proposed by [7].

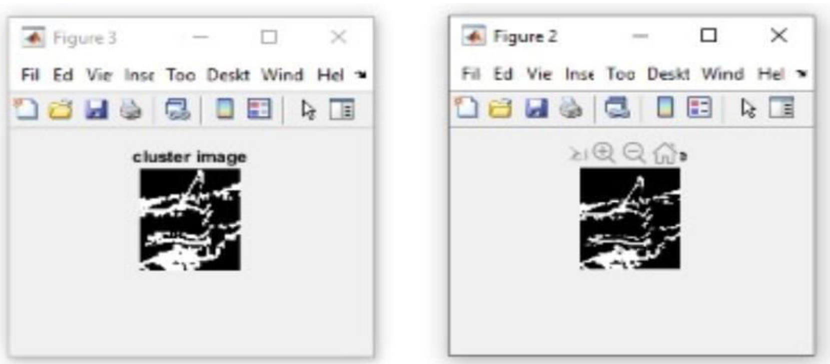


Figure 4: Clustering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE2021. 0903114|

After clustering,Figure 5 represents that voice is identified by the system and compare with histogram and we get audio signal as the output.

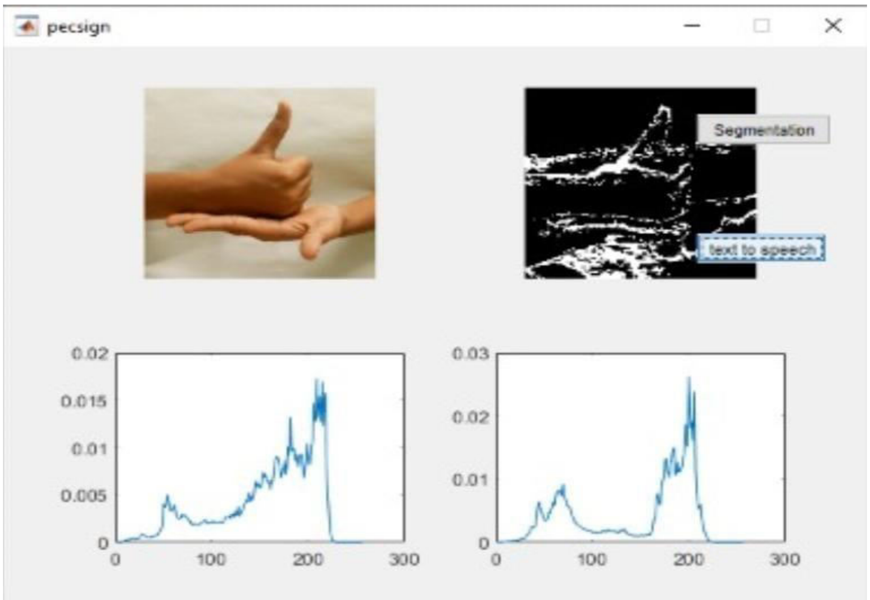


Figure 5: Histogram analysis

TABLE I

SIGN IMAGES AND IT'S CORRESPONDING WORDS

The table shows that the captured sign images and its corresponding words.

S.NO	SIGN IMAGES	OUTPUT
1		Help
2		Yes

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE2021. 0903114|

3		No
4		Play
5		House
6		Correct
7		Friend

V. CONCLUSION

In this study, SLR system which recognizes the hand-based signs are proposed. The system mentioned here does not need coloured are sensor-based gloves or specialized camera system. It is aimed to work with the videos recorded by the user's phone. And also, in this review paper, we have different techniques of sign language recognition are reviewed on the basis of sign acquiring methods and sign identification methods. In this techniquefor sign acquiring methods, vision-based methods and for sign identification methods, artificial neuron network proves a strong candidature. As a future work the proposed system could be transferred to a mobile platform since the system is planned to work with phone cameras. By doing this system could reach much more people since nearly everyone has a phone with a decent camera lately. The gesture set could be expanded and to include more gestures.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 3, March 2021 ||

| DOI: 10.15680/IJIRCCE.2021. 0903114|

REFERENCES

[1] O. Aran, "Vision based sign language recognition: modelling and recognizing isolated signs with manual and non-manual components", Ph.D. dissertation, Dept. Comput. Eng., Bogazici University, 2008.

[2] G. Awad, J. Han, and A. Sutherland, "A Unified System for Segmentation and Tracking of Face and Hands in Sign Language Recognition", 18th International Conference on Pattern Recognition, 2006 (ICPR), Vol. 1, pp.239-242, IEEE,2006.

[3] N. D.Binh, E. Shuichi and T. Ejima, "Real-Time Hand Tracking and Gesture Recognition System", Proceedings of International Conference on Graphics, Vision and Image Processing (GVIP-05), Vol., No., pp. 362-368, 2005.

[4] M. Brand, "Coupled Hidden Markov Models for Modelling Interacting Processes", Technical report, MIT Media Lab Perceptual Computing, June 1997.

[5] H. Cooper, B. Holt and R. Bowden, "Sign language recognition using linguistically derived sub-units", Workshop on the Representation and Processing of Sign Languages: Corpora and Sign Languages Technologies, Valetta, Malta, May 2010.

[6] H. Cooper, B. Holt and R. Bowden, "Sign language recognition", in Visual Analysis of Humans, London, Springer, 2007, pp. 88-97.

[7] Y. Chahir and A. Elmoataz, "Skin- color detection using fuzzy clustering", IEEE- EURASIP International Symposium on Control, Communications, and Signal Processing, March 2006.

[8] G. Fang and W. Gao, "An SRN/HMM System for Signer- independent Continuous Sign Language Recognition", Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition, 2002.

[9] M. W. Kadous, "Machine recognition of Auslan Signs using power gloves: Towards large-lexicon recognition of sign language", Proceedings of the Workshop on the Integration of Gesture in Language and Speech, pp.165-174,1996.

[10] J. Zieren and K. -F. Kraiss, "Robust person-independent visual sign language recognition", in Pattern recognition and image analysis, London, Springer, 2005, pp. 520-528.