IMAGE BASED SIGN LANGUAGE RECOGNITION SYSTEM FOR SINHALA SIGN LANGUAGE

H.C.M. Herath, W.A.L.V.Kumari, W.A.P.B Senevirathne and M.B Dissanayake

Department of Electrical and Electronic Engineering, Faculty of Engineering, University of Peradeniya, Sri Lanka. Email: maheshid@ee.pdn.ac.lk

ABSTRACT

A novel method for hearing impaired people to communicate with others effectively by means of technology is presented. The goal of this research is to achieve a tool that will help a hearing impaired person to communicate with a person who is not aware of sign languages. This paper presents a low cost approach to develop an image processing based Sinhala sign language recognition application for real time applications.
New concept of mapping the gesture using centroid finding method, which is capable of mapping the input gesture with the data base independent from hand size and position, is explored.

***Key words*:** hearing impaired, Sinhala sign language, mapping

1. INTRODUCTION

Deaf or inarticulate person is a one who suffers from hearing impairment or difficulty of hearing. They communicate with each other using sign languages. Hence, it is considered as the top of the gesture hierarchical taxonomy. The importance such a language establishes the automatic recognition of sign languages as a research challenge. There are several factors that impede the task of automatic sign language recognition. For instance, it is the least standardized, can be modified according to grammatical function, largely symbolic and referential, and most gestures are both sequential and simultaneous.

Research on sign language recognition has been done around the world, using many sign languages, including American Sign Language [3, 5, 2], Korean Sign Language [7], Taiwanese Sign Language [8], Chinese Sign Language [4, 6], Japanese Sign Language [9], and German Sign Language [1]. Many sign language recognition systems use Hidden Markov Models (HMMs) for their abilities to train useful models from limited and potentially noisy sensor data [6, 5, 3]. To help people with disabilities, Newby worked on the recognition of the letters and numbers of the ASL manual alphabet based upon statistical [10]. Starner and Pentlands' American Sign Language system [11, 12, 13] could recognize short sentences of American Sign Language (ASL) with 40 vocabularies, each was attached with its part of speech, which greatly reduced the computational complexity. The feature vector was fed to a hidden Markov model (HMM) for recognition of the signed words. Furthermore, Nam's system [14] tried to recognize hand movement patterns and Liang and Ouhyoung's proposal [15] aims at recognizing a large set of vocabularies in a sign language by recognizing constructive postures and context information.

Although a large number of approaches have been proposed, robust automatic sign language recognition still remains an open problem. This research will help to fill the gap created by the non-existence of proper software to convert dynamic gestures made by the user of a computer to a predefined word/phrase or command for Sinhala Sign Language. The research presented is of high importance to community and it can aid the country's development immensely.

2. METHODOLOGY

Figure 1 illustrates the high level architecture of the system. In the prototype developed for this project a green background is used to capture the image for the simplicity of the implementation. First, the RGB image captured from the web camera is separated into the three matrices, red (R), green (G) and blue (B). Next G matrix is subtracted from the R matrix. This is done because it was experimentally found that red is the most dominant color of the skin as shown in the figure 2 and the background used is in green color as well. However, the algorithm presented can be fine tuned to be used in a background with any constant color. Shadows are removed in this process as in figure 3. In figure 3 (b) shadow effect remove by subtracting G matrix from R matrix and it is convert to binary image as in figure 3 (c)

Then resulted image is converted to binary image by defining a threshold. This is generated to facilitate faster mapping. The resulted binary image accuracy is depended on lighting condition at which the image is captured. If the

lighting intensity is sufficient to capture the image with its natural colors or closer to natural colors then the binary image is noise free as illustrate in figure 4 Next boundaries of the hand are identified by drawing smallest possible rectangle around the hand and the image is cropped to extract the region to interest. Then the cropped image is equally divided in to four parts as shown in Figure 5(b).

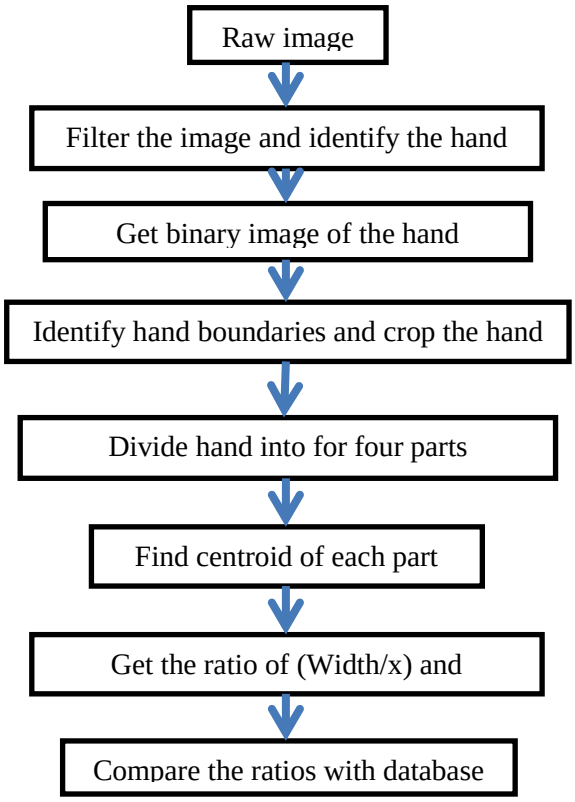

Figure 1: High level architecture of the approach

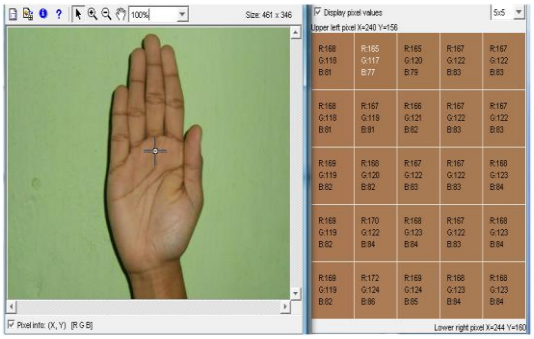

Figure 2: R, G, B value of a point in the hand

Figure 3: Shadow effect removal

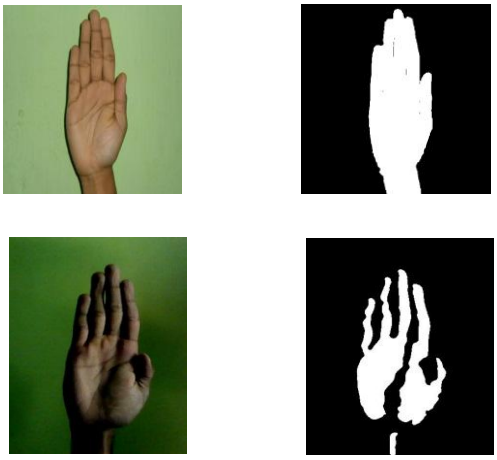

Figure 4: real time image captured in different light intensities and the converted binary images

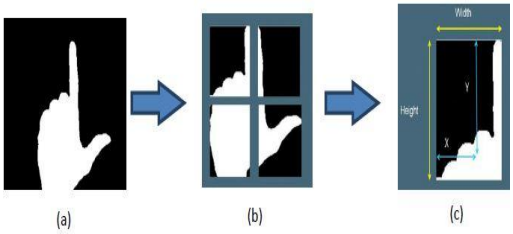

Figure 5: Segmented hand and ratio calculation

Next centroid of each segment is calculated as shown in Figure 5 (c). The (Height/y) and

(Width/x) ratios are calculated for each segment and then they are compared against pre calculated values that are in the database. Errors of the ratios are also calculated using eq. (01).

Error = (ratio in the database) – (ratio calculated for the real time image) (01)

Finally, the image with minimum error is selected as the matched image. The prototype of the system is developed using Matlab simulation package, a portable camera (Intex Model No IT-309WC, 16MP) and green back ground.

3. RESULTS AND DISCUSSION

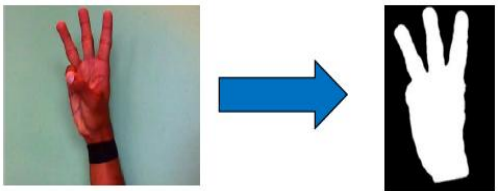

Figure 6: Real and matched images

Figure 6 shows a possible outcome of the system. The proposed prototype was tested in real time with 5 random participants, against a database of 15 sign gestures that is shown in Figure 7. According to the results presented in Figure 8, the system identified 10 numbers of gestures with 100% accuracy, 4 numbers of gestures with 80% accuracy and one gesture with 60% accuracy. In other words, the prototype correctly recognized 92% of gestures. Therefore the proposed algorithm shows good adaptability and acceptable level of performance for a random selection of users.

The results for gestures 4, 6, 9, 12 and 14 have shown less accuracy. That is because of the high level of correlation present between the signed gesture and other gesture in the date base. For example, binary image of the gestures 4, 8 are similar to the gesture 14. To avoid these issues finger detection algorithms can be introduced to the system as the next step

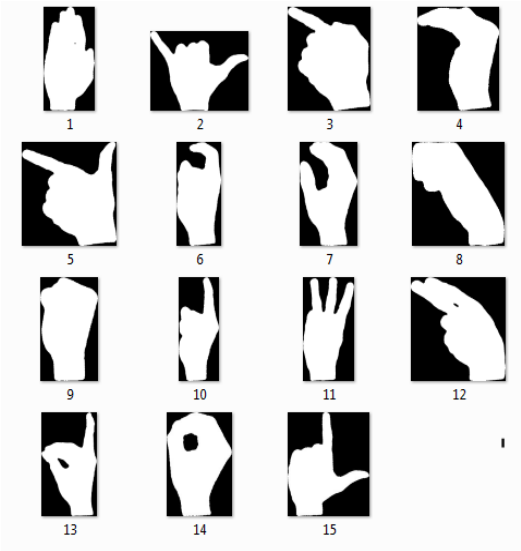

Figure 7: Database gestures

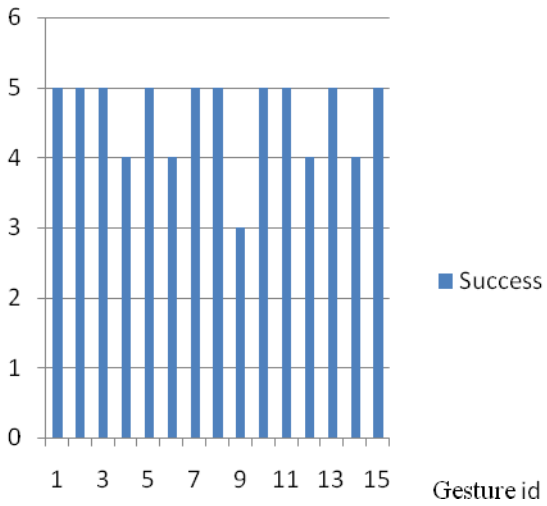

Figure 8: Successful matches vs. gesture id

4. CONCLUSIONS

The paper presents a prototype for image based Sinhala sign language recognition mechanism. The goal of this research is to produce a tool that will help a hearing impaired person to communicate with a person who is not aware of sign languages with the help of image processing technology. This research will help to fill the gap created by the non-existence of proper user application to convert dynamic gestures made by the user of a computer to a predefined word/phrase or command in Sinhala sign language. Since the research idea is fresh, there are fewer resources available to refer to. Further, the research team hopes the novelty of the finding presented in this paper will lead to many new research challenges.

5. REFERENCES

[1] B. Bauer, H. Hienz, and K. Kraiss, "*Video-based continuous sign language recognition using statistical methods,"* in: Proceedings of the 15th International Conference on Pattern Recognition, vol. 2, pp 463–466, September 2000.

[2] H. Brashear, T. Starner, P. Lukowicz, and H. Junker, "*Using multiple sensors for mobile sign language recognition,"* in: Proceedings of the Seventh IEEE International Symposium on Wearable Computers, pp. 45–52, 2003.

[3] C. Vogler and D. Metaxas, "*Hand shapes and movements: Multiple-channel American sign language recognition,"* in: Springer Lecture notes in Artificial Intelligence, vol 2915, pp. 247–258, January 2004.

[4] G. Fang, W. Gao, and D. Zhao "*Large vocabulary sign language recognition based on hierarchical decision trees*" in: International Conference on Multimodal Interfaces, pp. 125–131, 2003..

[5] T. Starner and A. Pentland, "*Visual recognition of American sign language using hidden markov models,"* in: Proceedings of the International Workshop on Automatic Face and Gesture Recognition, 1995.

[6] W. Gao, G. Fang, D. Zhao, and Y. Chen, "*Transition movement models for large vocabulary continuous sign language recognition (csl),"* in: Sixth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 553–558, 2004.

[7] J. S. Kim, W. Jang, and Z. Bien, "*A dynamic gesture recognition system for the Korean sign language KSL,"* in: IEEE Transactions on Systems, Man and Cybernetics, 354–359, 1996.

[8] R. Liang and M. Ouhyoung, "*A real-time continuous gesture recognition system for sign language,"* in: Third International Conference on Automatic Face and Gesture Recognition, pp. 558–565, 1998.

[9] H. Sagawa and M. Takeuchi, " *A method for recognizing a sequence of sign language words represented in a Japanese sign language sentence,"* in: Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 434–439, March 2000.

[10] G.B. Newby, "*Gesture Recognition Based upon Statistical Similarity*," MIT Press, vol. 3 pp. 236-243, 1994.

[11] T. Starner, "*Visual Recognition of American Sign Language Using Hidden Markov Models,"* Master Thesis, MIT Media Lab, 1995.

[12] T. Starner. and A. Pentland, "*Visual Recognition of American Sign Language Using Hidden Markov Models,"* Technical Report TR306, Media Lab, MIT, 1995.

[13] T. Starner and A. Pentland, "*Real-time American Sign Language Recognition from Video Using Hidden Markov Models*," Technical Report TR375, Media Lab, MIT, 1996.

[14] Y. Nam and K-Y. Wohn. "*Recognition of Space-Time Hand-Gestures using Hidden Markov Model*," Proc. of the ACM Symposium on Virtual Reality Software and Technology, Hong Kong, pp. 51-58, July 1996.

[15] R-H. Liang and M. Ouhyoung, "*A Sign Language Recognition System Using Hidden Markov Model and Context Sensitive Search*" Proc. of the ACM Symposium on Virtual Reality Software and Technology, Hong Kong, pp. 59-66, July 1996.