HAND DETECTION FOR SIGN LANGUAGE RECOGNITION APPLYING HIDDEN MARKOV MODEL

[1]MOHD NOAH A. RAHMAN, [2]ARMANADURNI ABDUL RAHMAN, [3]AFZAAL H. SEYAL, [4]NORHAFILAH BARA

[1]School of Computing & Informatics, INSTITUT TEKNOLOGI BRUNEI, Brunei Darussalam, [2]Royal Brunei Airlines, Brunei Darussalam
E-mail: [1]noah.rahman@itb.edu.bn, [2]norhafilah.bara@rba.com.bn

Abstract- The study of developing interface with computers using gestures of human body such as hand movements keeps growing in interest. This is to provide more natural, innovative and sophisticated way of non-verbal communication, such as sign language, in human computer interaction. A review on this hand gesture detection and recognition based on computer vision technique remains a very challenging task. Nevertheless, this paper explores hand detection and hand gesture recognition. The hand detection uses skin color spaces such as HSV and YCrCb are applied. However, there are limitations that are needed to be considered. Almost all of skin color space models are sensitive to quickly changing or mixed lighting circumstances. The Hidden Markov Model (HMM) is proposed to be implemented in hand recognition for alphabetical sign language. There are certain restrictions in order for the hand recognition to give better results such as the distance of user's hand to the computer camera and the posture of the hand shown by user.

Keywords- Hand Detection, Hand Gesture, Sign Language, Hidden Markov Model.

I. INTRODUCTION

A gesture is basically the physical action perform by the individual to convey meaningful information [1]. Gesture (sign language) recognition is a computerized process of mathematical interpretation of human actions utilizing the power of a computer. Gestures commonly originate from hand can be recognized by computer which enables humans to interact with machines. It allows human-computer-interaction (HCI) that benefits the speech and hearing impaired people. Using this technology, it is possible to move a cursor on the screen just by pointing a finger. In this project, hand gesture recognition is focused using computer vision technique instead of image processing.

Most visual hand gesture recognition programs work only under circumscribed environment due to some effects such as lighting and complex background. There is still ongoing literature in this computer science and language technology field with the objective of developing a program that produces improved and most accurate result.

Sign Language (SL) is used as a non-verbal communication medium for people with speech and hearing impaired around the world. Hand gesture is an important part of SL such as finger-spelling. When a word does not have a sign, it is spelt out using a manual alphabet. It uses hand movement as the primary element of communication. Each country has its own native SL but some may share similarities to signs used in another country. Some of the notable existing SL around the world are Japanese SL [1]-[5].

In Brunei Darussalam, she borrows SL from Malaysia and USA because Brunei Darussalam does not has her own SL. In an interview with The Brunei Times which is a local newspaper, it was mentioned that they wanted the community to be involved, not just to focus on parents and teachers, noting the difficulty that speech and hearing impaired people face in communication outside their group [6]

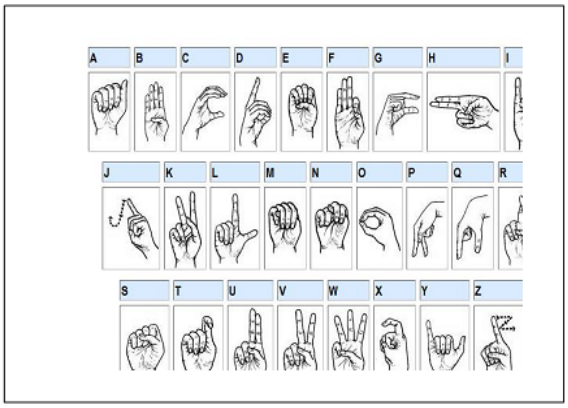


Fig.1. Alphabet sign language used in Brunei Darussalam

SL is not commonly used in society and recognizing them is a challenging task for people who are not familiar with SL. Instead of finding a skilled teachers or translator for SL, HCI system for SL can be installed anywhere possible. SL recognition research faces difficulty on how to make recognition approaches scale with large vocabularies.

This study utilizes the development of visual interface using hand gestures and captures the images using a webcam which is one of the most challenging part to program the computer to recognize the hand gestures.

The organization of this paper is sequenced as follow: Firstly, related studies are compiled and analyzed. Secondly, a Hidden Markov Model (HMM) is introduced. Thirdly, an experiment is conducted to test all SL activities and subsequently leads to discussions and a conclusion is made.

II. LITERATURE SURVEY

SL is one of the communication media for people who are having speech and hearing impaired to interact and communicate with others. There exist diverse hand gesture detection and recognition techniques ranging from neural networks [7]-[8], self-organizing features maps [9], self-growing and self-organized neural gas network [10], genetic algorithm [11] and support vector machines [12] among others.
A widely used statistical method introduced in the 70s known as HMM was originally used in the implementation for speech processing applications [13]. Subsequently, various studies were conducted on HMM to convert a series of motions into a description of activities such as SL [14]-[18].

III. METHODOLOGY

This section reports a method for developing a gesture interface using HMM. It is a well-known and widely used statistical technique. Instead of using geometric features, gestures are transformed into sequential symbols. The gesture signal can be well characterized as a parametric random process, and that the parameters of the stochastic process can be determined in a precise, well-defined manner.

3.1. Hidden Markov Model

HMM consists of three sets of states which are known as the initial state set, output symbols state set and transitions state set [19]. Each state of the hand gesture represents a set of possible hand positions.
Mathematically, it can be observed based on the equation [20] which states that the probability distribution for each state j determines the probability of generating observation ot at time t. This is illustrated by the equation:

$$b_{ij}(o_t) = \sum_{i=1}^{m} \omega_i N(o_t; \mu_i, \sigma_i) \quad (1)$$

where m is the mixture component in state j. □i is the weightage of m-th component. N is a Gaussian multivariate with vector mean μi and covariance matrix of σi.

3.2. Posture Recognition

The approach of a posture (static hand gestures) recognition is to assume a simple probabilistic model of posture input data. The model of each specific posture P produces an observation Y with probability P(P,Y). The goal is to decode the posture data with the observation so that the decoded posture has the maximum a posteriori (MAP) probability [21].

$$P(P|Y) = P(Y|P)P(P)/P(Y) \quad (2)$$

Since P(Y) is independent of P,

$$\hat{P} = \arg\max P(Y|P)P(P) \quad (3)$$

P(Y|P) estimates the probability of a observation conditioned on a specific posture. P(P), is called the language model since it describes the probability associated with P in the specific set of SL.

3.3. Posture Recognition

Gestures are usually understood as hand and body movement which can pass information from one to another. In Hand Gesture Detection for Alphabet SL system, there are two non-static hand gestures for Alphabet Sign Language such as J and Z. HMM of a gesture can be represented as Figure 2. Each state among S2, S3, and S4 represents an evaluation of a posture.

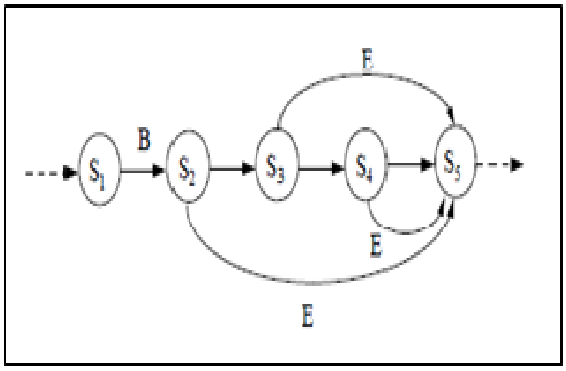


Fig.1. HMM representation of a dynamic data

In order to simplify the learning process, consider only the first order Markov Chain, which is based on the assumptions:

i. Markov assumption: a new state is entered based only on the current state.

ii. Output-independent assumption: the output probability distribution function depends only on the state at the particular time despite of when and how the state is entered.

IV. EXPERIMENTAL RESULT AND ANALYSIS

This section discusses the efficiency and performance of hand detection for numerical recognition based on some experiments conducted. This stage is necessary in order to measure the project's success.
In this experiment, we used a PC with a 4th generation Intel Core i5 with a 3.96 GHz processor with 4 GB RAM and 1TB hard drive. Notepad is used as the editor for the coding of HTML, PHP and MySQL as the database management systems (DBMS). The DBMS is used because it has profound benefits and ease of use [22] to store data in the database

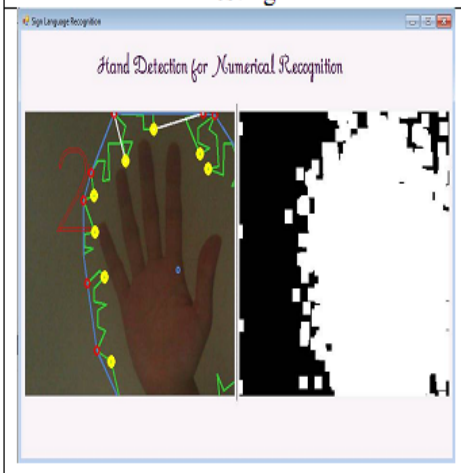
From the experiment testing above, it can be concluded that hand detection for numerical recognition can detect different shapes or sizes of hand. This testing is conducted to check the efficiency of the system since different user has different shapes or sizes.

Table 1: Different shapes or sizes of hands

Person	Hand shown (Testing)	Detect	Recognize
1		True	True
2		True	True
3		True	True

The experiment was conducted to test the performance of detection under darker lightings. It was also conducted to test the performance of Hand Detection for Numerical Recognition (HDfNR) when there are objects at the background

Table 2: Result under less lighting sources condition

Testing	Result
	False

There are three experiments conducted to test the performance and efficiency of HDfNR in this project. From the testing, it can be analyzed that this system has several conditions that need to be followed in order to get better and accurate results.

Table 3: Result when there are objects at the background

Testing	Result
	True

From the first experiment, it can be concluded that HDfNR can detect different shapes or sizes of hand.
From the testing in the second experiment conducted earlier, it can be concluded that HDfNR is sensitive to lighting conditions. As discussed earlier, almost all approaches in hand detection are sensitive to quickly changing or mixed lighting circumstances, hence, it was proven.
From the third experiment conducted, it showed that HDfNR can detect hand regions and eliminate the unwanted background image.
There are other factors, other than those three experiments, that can be conducted to test the performance and efficiency of the HDfNR.

CONCLUSIONS

This study reports a few related studies on application of SL using HMM have been analyzed and reviewed. There are a few demonstrations and source code examples available from the internet on hand gesture detection and recognition, unfortunately, they did not explain explicitly which method they were using to recognize the hand gesture applications.
For this study, it is expected that the system can detect hand gesture and recognize it as an alphabet in SL. However, due to crucial, critical and limited period of time, only numerical signs can be recognized. A vision-based approach, using built-in computer webcam, is deployed as an input device. HSV skin detector and YCrCb skin detector are utilized to detect hand and convert the image to binary image. Convex hull and convexity defects approaches are implemented for recognition purposes. The system can count the number of fingers through the convexity defects. Subsequently, the system will show the result to users.
There are many characteristics to be considered too for examples data background, lightings, and the size of user's hand. With limited time and skills, we managed to finish Hand Detection for Numerical Recognition even though it is not the proposed system initially. There are several experiments conducted to test the performance and efficiency of

HDfNR. Nevertheless, HDfNR can be enhanced for in the future to achieve our enthusiastic objectives.

This research has equipped us on the understanding of how hand detections and recognition processes of different approaches work. With limited theoretical knowledge on this field and scarce skill on C# programming, it has been a challenging and tough journey but a good experience for us. We have gained something positive in many aspects through this project such as improved self-learning skill and time management.

REFERENCES

[1]. M. Panwar, "Hand gesture based interface for aiding visually impaired," Proc. Of International Conference on Recent Advances in Computing and Software Systems (RACSS), April 25-27 2012, pp. 80-85, Chennai, India.

[2]. K. Murakami and H. Taguchi, "Gesture recognition using recurrent neural networks," Proc. Of International Conference of Human Factors in Computing Systems, pp. 237-242, 1991.

[3]. C. Oz and M.C. Leu, "Recognition of finger spelling of American Sign Language with artificial neural network using position/orientation sensors and data glove," Proc. International Symposium on Neural Network (ISNN), pp. 157-164 , 2005.

[4]. T. Shanableh and K. Assaleh, "Arabic sign language recognition in user independent mode," Proc. International Conference on Intelligent and Advanced Systems (ICIAS), pp. 597-600, 2007.

[5]. Y.H. Lee and C.Y. Tsai, "Taiwan sign language (TSL) recognition based on 3D data and neural networks," Expert Systems with Applications, Elsevier Ltd, pp. 1-6, 2007.

[6]. A. Adnan, Learning Sign Language at UBD. The Brunei Times, Retrieved September 20, 2013 www.bt.com/happenings/2013/07/03/learning -sign-language-ubd.

[7]. Y. Zhang and J. Yuan, "Gesture recognition for human computer interaction using neural networks," Proc. 8th International Conference on Neural Information Processing (ICNIP), Novmber 14-18, pp. 735-740, Shanghai China,

[8]. C.L. Huang and W.Y. Huang, "Sign language recognition using model-based tracking and a 3D Hopfield neural network," Machine Vision and Application, vol. 10, pp. 292-307, 1998.

[9]. G.L. Fang, W. Gao and D.B. Zhao, "Large vocabulary sign language recognition based on fuzzy decision trees," IEEE Transactions vol. 34, no. 3, pp. 305-314, 2004.

[10]. E. Stergiopoulou and N. Papamarkos, "A new technique for hand gesture recognition," Proc. Of IEEE I International Conference on Image Processing,, pp. 2657-2660, 2006.

[11]. M.A. Abed, A.N. Ismail and Z.M. Hazi, "Pattern recognition using genetic algorithm," International Journal of Computer and Electrical Engineering (IJCEE), vol. 2, no. 3, pp. 1793-1863, June 2010.

[12]. Y. Yuan and K. Barner, "Hybrid feature selection for gesture recognition using support vector machines," Proc. Of International Conference on Accoustics, Speech and Signal Processing (ICASSP), pp. 1941-1944, 2008.

[13]. L. Rabiner and B.W. Huang, Fundamentals of Speech Recognition, Prentice Hall International Inc., pp. 321-481, 1993.

[14]. H. Buxton, "Learning and understanding dynamic scene activity: A review," Image and Vision Computing, vol. 21, no. 1, pp. 125-136, January 2003.

[15]. J.M. Siskind and Q. Morris, "A maximum likelihood approach to visual event classification," Proc. 4th European Conference on Computer Vision (ECCV), pp. 347-360, 1996.

[16]. Y.A. Ivanov and A.F. Bobick, "Recognition of visual activities and interactions by stochastic parsing," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 8, pp. 852-872, August 2000.

[17]. J. Lou, Q. Liu, T. Tan and W. Hu, "Semantic interpretation of object activities in a surveillance system," Proc. International Conference on Pattern Recognition (ICPR), pp. 777-780, 2002.

[18]. N.M. Oliver, B. Rosario and A.P. Pentland, "A Bayesian computer vision system for modeling human interactions," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 8, pp. 831-843, August 2000.

[19]. C. Eugene, Statistical Language Learning, MIT Press, Cambridge, 1993.

[20]. M. Maebatake, I. Suzuki, M. Nishida, Y. Horiuchi and S. Kuroiwa, "Sign language recognition based on position and movement using multi-stream HMM," Proc. Of International Symposium on Universal Communication, pp. 478-481, 2008.

[21]. N.S. Salleh, J. Jais, L. Mazalan, R. Ismail, S. Yussof, A. Ahmad, A. Anuar and D. Mohamad, "Sign language to voice recognition: Hand detection techniques for vision-based approach,". Proc. of Current Development in Technology Assisted Education, pp. 967-972, 2006.

[22]. M.N. Rahman, A.H. Seyal and A.Y. Mohamed, "An empirical framework of DBMS usage in Brunei Darussalam," Proc. Of the 5th Annual Global Information Technology Management World Conference (GITM), pp. 189-192, June 2004.

★★★