STATIC HAND GESTURE RECOGNITION USING HAAR-LIKE FEATURES

WONG KAI SIN

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF SCIENCE WITH HONOURS

MATHEMATICS WITH COMPUTER GRAPHICS
FACULTY OF SCIENCE AND NATURAL RESOURCES
UNIVERSITY OF MALAYSIA SABAH

2015

ACKNOWLEDGEMENT

I would like to express sincere gratitude to my final year project supervisor, Madam Zaturrawiah binti Ali Omar, who guides me in accomplishing this project. This dissertation would be impossible without her advice and guidance.

I wish to express the deepest appreciation to my family and friends for helping in data collection and testing experiment.

In addition, a thank you to every lecturer who gives me a lot of knowledge for completing this project.

CERTIFIED BY

Signature

Supervisor

(Madam Zaturrawiah binti Ali Omar) _____

DECLARATION

I declare that this dissertation is based on my original work, except for quotations and summaries, each of which has been fully acknowledged.

_____

WONG KAI SIN

(BS12110705)

17 May 2015

ABSTRACT

Hand gesture recognition plays a crucial role in communication between human and computer or robot. It is used to improve Human-Computer Interaction (HCI) for the sake of making the communication more natural and much easier. Static hand gesture or posture recognition using Haar-like features is being presented in this paper. Two static hand gestures which are index finger and fist are trained using Haar-like features algorithm. Index finger represents left click mouse event while fist represents right click mouse event. AdaBoost algorithm is applied in the training phase to increase accuracy and robustness of the system. Since this is a real-time system, built-in webcam is used to capture the image of the gesture. Brightness and distance are tested for evaluation of this system. Some static imported images are also tested. The experimental results show that both static hand gestures achieve the highest accuracy under a high degree (80%-100%) of brightness. Index finger and fist achieve 90.4% and 91.2% accuracy respectively under a high degree of brightness. The best distance is 80cm from the screen. Index finger achieves 92% accuracy for 80cm distance while the fist achieves 95.2% for both 80cm and 100cm distances.

PENGIKTIRAFAN ISYARAT TANGAN STATIK MENGGUNAKAN CIRI-CIRI HAAR

ABSTRAK

Pengiktirafan isyarat tangan memainkan peranan yang signifikan antara komunikasi manusia dengan komputer atau robot. Ia digunakan untuk meningkatkan interaksi manusia-komputer. Justeru, komunikasi manusia antara komputer lebih mudah dan bersifat. Pengiktirafan isyarat tangan statik atau postur menggunakan ciri-ciri Haar disampaikan dalam kajian ini. Dua isyarat tangan statik iaitu jari telunjuk dan penumbuk diiktiraf dengan menggunakan ciri-ciri Haar. Jari telunjuk adalah klik kiri manakala penumbuk adalah klik kanan. Algoritma AdaBoost juga digunakan untuk meningkatkan ketepatan. Sistem ini adalah sistem semasa. Oleh itu, webcam digunakan untuk menangkap gambar isyarat tangan. Kecerahan dan jarak diuji untuk penilaian sistem ini. Beberapa gambar juga diimport untuk tujuan penilaian. Hasil kajian ini menunjukkan kedua-dua isyarat tangan statik mencapai ketepatan yang paling tinggi dalam kecerahan yang tinggi (80%-100%). Jari telunjuk mencapai 90.4% manakala penumbuk mencapai 91.2% ketepatan dalam keadaan yang sangat cerah. Jarak yang paling baik adalah 80cm dari skrin. Jari telunjuk mencapai 92% ketepatan untuk 80cm manakala penumbuk mencapai 95.2% untuk 80cm dan 100cm.

CONTENTS

CHAPTER 2 LITERATURE REVIEW

CHAPTER 3 METHODOLOGY

CHAPTER 4 DESIGN AND IMPLEMENTATION

CHAPTER 5 RESULT AND DISCUSSION

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

HCI	Human-computer interaction
OpenCV	Open Computer Vision Library
AdaBoost	Adaptive boosting
ASL	American Sign Language
HMM	Hidden Markov model
GMM	Gaussian mixture model
EM	Expectation maximization
SIFT	Scale Invariant Feature Transform
FSM	Finite State Machine
PCA	Principal Component Analysis
IDE	Integrated development environment

LIST OF SYMBOLS

ii(x,y)	Integral image
s(x,y)	Cumulative row sum
F(x)	Haar-like feature
$h_j(x)$	Haar classifier

CHAPTER 1

INTRODUCTION

1.1 Overview

This project is about to improve Human-Computer Interaction (HCI) by developing a system using static hand gesture to communicate with computer. Haar-like feature is used as the algorithm to train the hand gestures.

1.2 Background of the Project

In this era of science and technology, in order to communicate with robots or machines more efficiently, hand gesture recognition always plays its role in this field. Hand gesture recognition makes the communication between human and machine more natural and much easier. The example that closest to us in daily life is using a touch-screen hand phone. Samsung Galaxy Note 3 is equipped with the function that can browse gallery using air gestures. There are two types of hand gesture, which are dynamic and static hand gesture. Dynamic hand gesture is the motion originates from hand and it may refer to certain movement which conveys certain messages. Static hand gesture is the posture of hand that may convey some messages, we may call it as hand posture. Static hand gesture will be the focus in this study. In the field of human-computer interaction, static hand gesture recognition can be used to improve Human-Computer Interaction (HCI) in order to make a more user-friendly system. Wang & Wang (2008) presented a research of static hand gesture recognition for human robot interaction. These postures are tested for the experiment. Figure 1.1 illustrate a person interacts with the robot.

Figure 1.1 Illustration of a person interacts with robot.

There are many technology gadgets in the market nowadays to aid in communicating with the computer, such as keyboard, mouse and touch-screen. These gadgets may be redundant when we have a system without any mechanical devices, such as keyboard and mouse. We can use our hand to be an input device instead of using mechanical device (Chen *et al.* , 2007). Static hand gesture recognition can be used to detect and recognize certain gesture or posture of user's hand and therefore executes certain command. For example, system detects a posture of showing thumb and then starts to run a program. This will save the time of searching program under certain conditions. Many researchers have used data glove in the research to track the hand. Davis & Mubarak (1994) presented a model-based method for hand gesture recognition which uses Computer Vision. In this work, marked glove is worn by user for the recognition of 7 gestures: top, down, left, right, stop, rotate and grab. These gestures are used to communicate with the computer. Kumar *et al.* (2012) used hand data glove for gesture recognition in their research of real-time HCI. It is found that the interaction is more accurate and natural when using data glove compare to conventional keyboard and mouse. There are no any limitations by using data glove compare to conventional keyboard and mouse. Accelerometer and Flex sensors are setup on the glove in the research which is done by Kataware & Bombale (2014). These sensors are used to track the movements and so the gesture presented by user. User can use the glove freely as the glove is wireless and simple gesture is easy to present. Figure 1.2 shows the wearable wireless device used in the study.

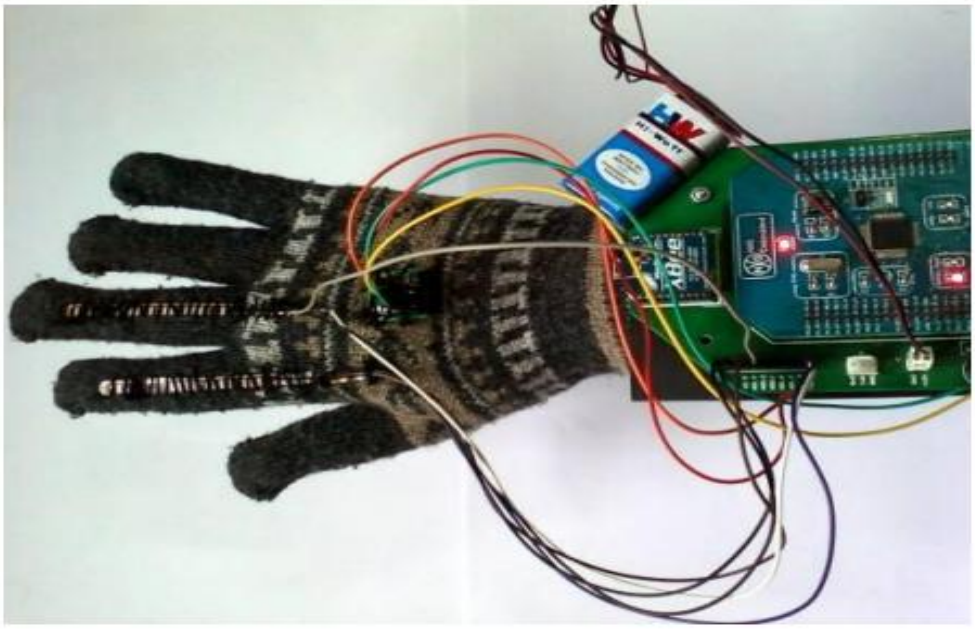

Figure 1.2 Wearable wireless device.

However, data glove costs a lot and it is cumbersome (Chen *et al.* , 2007). Figure 1.3 shows the cumbersome data glove.

Figure 1.3 Cumbersome data glove.

User may not afford to buy an expensive and unpleasant-looked glove in order to control their computers. Therefore, more research of vision-based of hand

tracking is needed for the sake of making a user-friendly system without any mechanical devices or glove. In the vision-based hand posture detection system, 2D video input is connected to the computer to capture and track the hand. User does not need to wear any device to be tracked by the system. By using several algorithms, the gesture will be determined. Figure 1.4 shows the vision based gesture recognition system without using glove.

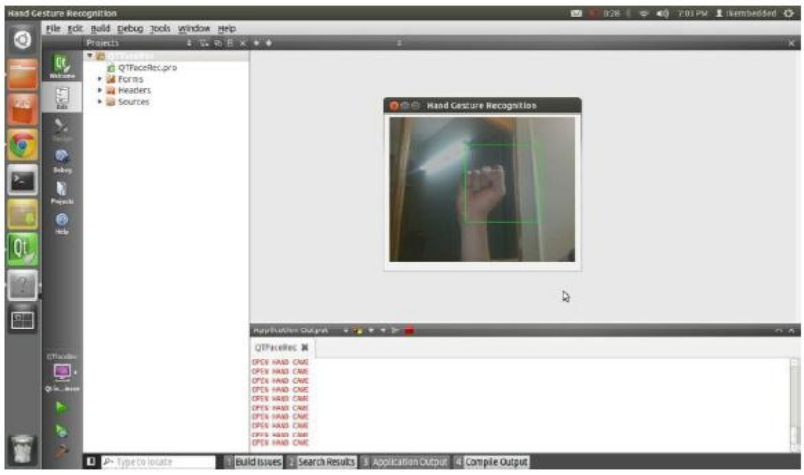

Figure 1.4 Vision based posture recognition without using glove.

To recognize the static hand gestures, Haar-like feature and related techniques are applied. Haar-like feature is a technique that uses feature for recognition rather than pixels. Haar-like feature is well-known in face detection. For example, Viola & Jones (2004) used this feature for face detection and it achieves high rate for the detection in real time.

1.3 PROBLEM STATEMENT

In this era of modernization, technology development is in rapid growth. There are many computer gadgets to allow us to communicate or control our computers more easily. However, sometimes we may weary to use those devices in order to communicate with our computers. The mechanical devices could be a redundancy when we travel somewhere. For example, it is difficult to use a mouse in an airplane due to its confined space. Other than that, prolonged use of mouse leads to some healthy problems. For instance, it may cause repetitive strain injury. Thus, a more

user-friendly system in Human-Computer Interaction (HCI) without mechanical device is needed. Moreover, HCI without mechanical device leads us to a more decent life. In order to make a more user-friendly HCI system, it is good to implement static hand gesture recognition without mechanical device. By using this system, user can command their computers with certain static hand gesture. Static hand gestures are used to make left click and right click events instead of using mouse or touchpad. Therefore, there is no any mechanical device is needed to communicate with the computer.

1.4 Aim

The aim of this project is to build a system which can recognize static hand gesture using Haar-like features.

1.5 Objective

There are two main objectives in this project:

i. To build a static hand gesture recognition system using Haar-like features.
ii. To measure the accuracy of the system for static hand gesture recognition.

1.6 Scope

This project is carried out to recognize static hand gesture. Haar-like feature is used as the algorithm in training the static hand gesture. Two different types of static hand gestures are trained: index finger and fist. Each static hand gesture represents one operation. Thus, we have two operations which are right click and left click. Index finger represents left click while fist represents right click. This project focuses on medium light to light skin datasets of static hand gestures. Other than that, this project recognizes hand gestures in a black background. Some other related algorithms are applied to transform input color space and detect hand. To implement this algorithm, language C++ is written in Microsoft Visual Studio 2012 with the aid of Open Computer Vision library (OpenCV).

A laptop with built-in webcam is used in this project. Built-in webcam with resolution is used to capture and detect the hand.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

Several methods used for the phases in static hand gesture recognition system will be discussed in this section. Hand detection, classification and training using Haar-like features are the main phases. Some prominent methods such as Gaussian mixture model, colour transformation and AdaBoost algorithm are also discussed.

2.2 Static Hand Gesture Recognition

Gloved-based system is the starting point for hand gesture recognition. Meaningful gesture can be used to command a computer. It is like a deaf mute trying to communicate with each other. Back in the 19th century, despite the use of mechanical device, the invention of data glove provides a new way to communicate with computer. Premaratne (2014) defined data glove, as the wired part on the fingers of the glove that will be detected by the computer when user moves their fingers. There is no preprocessing needed as compared to the vision-based hand gesture recognition. Unlike the glove-based system, vision-based system needs to consider some external factors during implementation. The major factor is the illumination changes (lighting). However, the first vision-based system came in 1980s. In the early development of vision-based gesture recognition system, it is very similar with a marker system. User wears a glove of distinct color on the fingers and sits in front of the camera connected to computer. The camera will detect the color of fingers and the motion.

Generally, there are two types of hand gesture, namely dynamic and static hand gesture. Dynamic hand gesture is motion gesture while static hand gesture is posture of hand. Static hand gesture is the main focus in this paper. According to Messer (2009), a static hand gesture recognition can be referred as a person performs posture in front of a machine, the posture will be captured and analysed in defining the meaning of the posture. American Sign Language (ASL) recognition and electronic devices control using static hand gesture are two famous example in static hand gesture recognition (Messer, 2009). Figure 2.1 shows the typical flow of gesture recognition process (Premaratne, 2014). The image is captured by input video stream. After that, performed gesture will be determined in gesture isolation and feature extraction. Tracking or classification will be used for the purpose of recognition.

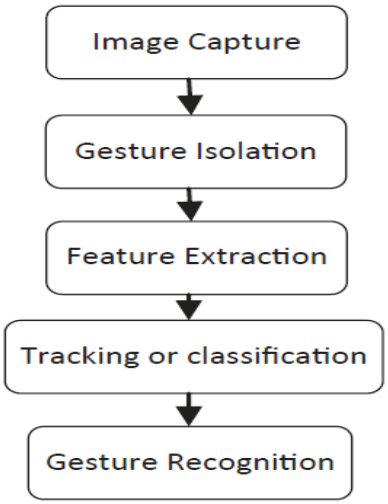

Figure 2.1 Typical flow of gesture recognition.

Pansare *et al.* (2012) presented an experiment of real-time static hand gesture recognition for ASL in complex background. 26 hand gestures represent A to Z alphabets are tested in this experiment. This recognition system achieves 90.19% success rate in complex background. Figure 2.2 shows the ASL symbol.

Figure 2.2 ASL symbol.

Dewi *et al.* (2010) used hand gestures and postures recognition to control television. Two model structures, one with computer, camera and television, another with television equipped with camera, are used in the experiment. The first structure will be performed by computer and the second structure will be performed by television. This system are used to change channel and volume by hand gestures and postures without using remote control.

Hasan & Abdul-Kareem (2012) presented static hand gesture recognition using neural network. There are six static hand gestures to be recognized: Open, Close, Cut, Paste, Maximize, Minimize. These postures are used for HCI. Neural network with hand contour and with hand complex moments are used for the recognition. Neural network with hand contour has 70.83% recognition rate while neural network with hand complex moments has 86.38% recognition rate.

2.2.1 Hand Detection

Hand detection is the first step of vision-based hand gesture recognition (Zabulis *et al.*, 2009). Hand detection is the process of discovering the shown hand in front of the camera which is the image capture in Figure 2.1. During this process, shown hand is segmented from the background which is the gesture isolation in Figure 2.1.

REFERENCES

Acharya, T. & Ray, A. K. 2005. *Image Processing: Principles and Applications.* John Wiley & Son, New Jersey.

Bhuyan, M. K., Kumar, D. A., MacDorman, K. F. & Iwahori, Y. 2014. A novel set of features for continuous hand gesture recognition. *Journal on Multimodal User Interfaces*, **8** (4): 333-343.

Chen, Q., Georganas, N. D. & Petriu, E. M. 2007. Real-time vision-based hand gesture recognition using Haar-like features. *Instrumentation and Measurement Technology Conference Proceedings*, 1-3 May 2007, University of Ottawa, Canada.

Davis, J. & Shah, M. 1994. Recognizing hand gestures. *Third European Conference Proceedings on Computer Vision*, 2-6 May 1994, Stockholm, Sweden, pp. 331-340.

Dewi, A.R., Haryanto, Adriani Y. K., Ryan, P. & Yohanes, M. 2010. Hand Gestures and Posture Recognition to Control Television Using Haar-Like Features. *Proceedings of The Second International Workshop on Open Source and Open Content WOSOC 2010*, 24 November 2010, University of Gunadarma, Indonesia.

Elmezain, M., Al-Hamadi, Ayoub, Appenrodt, J. & Michaelis, B. 2008. A hidden Markov model-based continuous gesture recognition system for hand motion trajectory. In: Borgefors, G. & Flynn, P. (eds) *Proceedings of the 19th International Conference on Pattern Recognition*, 8-11 December 2008, New Jersey: 1-4.

Fang, Y., Wang, K., Cheng, J. & Lu, H. 2007. A real-time hand gesture recognition method. *2007 IEEE International Conference on Multimedia and Expo Proceedings*, 2-5 July 2007, Beijing, China, pp. 995-998.

Fernandes, B. & Fernandez J. 2010. Using Haar-Like Feature Classifiers for Hand Tracking in Tabletop Augmented Reality. *XII Symposium on Virtual and Augmented Reality*, Brazil, pp. 6-13.

Ghotkar, A. S. & Kharate, G. K. 2013. Vision based real time hand gesture recognition techniques for human computer interaction. *International Journal of Computer Applications*, **70** (16): 1-6.

Greggio, N., Bernardino, A., Laschi, C., Dario, P. & Santos-Victor, J. 2012. Fast estimation of Gaussian mixture model for image segmentation. *Machine Vision and Applications*, **23** (4): 773-789.

REFERENCES

Acharya, T. & Ray, A. K. 2005. *Image Processing: Principles and Applications.* John Wiley & Son, New Jersey.

Bhuyan, M. K., Kumar, D. A., MacDorman, K. F. & Iwahori, Y. 2014. A novel set of features for continuous hand gesture recognition. *Journal on Multimodal User Interfaces*, **8** (4): 333-343.

Chen, Q., Georganas, N. D. & Petriu, E. M. 2007. Real-time vision-based hand gesture recognition using Haar-like features. *Instrumentation and Measurement Technology Conference Proceedings*, 1-3 May 2007, University of Ottawa, Canada.

Davis, J. & Shah, M. 1994. Recognizing hand gestures. *Third European Conference Proceedings on Computer Vision*, 2-6 May 1994, Stockholm, Sweden, pp. 331-340.

Dewi, A.R., Haryanto, Adriani Y. K., Ryan, P. & Yohanes, M. 2010. Hand Gestures and Posture Recognition to Control Television Using Haar-Like Features. *Proceedings of The Second International Workshop on Open Source and Open Content WOSOC 2010*, 24 November 2010, University of Gunadarma, Indonesia.

Elmezain, M., Al-Hamadi, Ayoub, Appenrodt, J. & Michaelis, B. 2008. A hidden Markov model-based continuous gesture recognition system for hand motion trajectory. In: Borgefors, G. & Flynn, P. (eds) *Proceedings of the 19th International Conference on Pattern Recognition*, 8-11 December 2008, New Jersey: 1-4.

Fang, Y., Wang, K., Cheng, J. & Lu, H. 2007. A real-time hand gesture recognition method. *2007 IEEE International Conference on Multimedia and Expo Proceedings*, 2-5 July 2007, Beijing, China, pp. 995-998.

Fernandes, B. & Fernandez J. 2010. Using Haar-Like Feature Classifiers for Hand Tracking in Tabletop Augmented Reality. *XII Symposium on Virtual and Augmented Reality*, Brazil, pp. 6-13.

Ghotkar, A. S. & Kharate, G. K. 2013. Vision based real time hand gesture recognition techniques for human computer interaction. *International Journal of Computer Applications*, **70** (16): 1-6.

Greggio, N., Bernardino, A., Laschi, C., Dario, P. & Santos-Victor, J. 2012. Fast estimation of Gaussian mixture model for image segmentation. *Machine Vision and Applications*, **23** (4): 773-789.

Gurav, R. M. & Kadbe, P. K. 2015. Vision Based Hand Gesture Recognition with Haar Classifier and AdaBoost Algorithm. *International Journal of Latest Trends in Engineering and Technology,* **5**(2): 155-160.

Hasan, H. & Abdul-Kareem, S. 2012. Static Hand Gesture Recognition Using Neural Networks. *Artificial Intelligence Review*, **41**(2): 147-181.

Jalilian, B. & Chalechale, A. 2013. Face and hand shape segmentation using statistical skin detection for sign language recognition. *Computer Science and Information Technology*, **1** (3): 196-201.

Kataware, Y. D. & Bombale, U. L. 2014. A wearable wireless device for effective human computer interaction. *International Journal of Computer Applications*, **99** (9): 9-14.

Kawulok, M., Grzejszczak, T., Nalepa, J. & Knyc M. 2013. Database for Hand Gesture Recognition. http://sun.aei.polsl.pl/~mkawulok/gestures.

Khan, R. Z. & Ibraheem, N. A. 2012. Hand gesture recognition: A literature review. *International Journal of Artificial Intelligence & Applications*, **3** (4): 161-174.

Kolsch, M. & Turk, M. 2004. Fast 2D hand tracking with flocks of features and multi-cue integration. *Proceedings of the 2004 Conference on Computer Vision and Pattern Recognition Workshop*, 27 June-2 July 2004, Washington, D.C.

Kumar, P., Prasad, S. & Verma, J. 2012. Hand data glove: A wearable real-time device for human-computer interaction. *International Journal of Advanced Science and Technology*, **43**: 15-25.

Lienhart, R. & Maydt, J. 2002. An Extended Set of Haar-Like features for Rapid Object Detection. *Proceedings International Conference on Image Processing*, 22-25 September 2002, IEEE, pp. 900-903.

Madhuri, K. & Kumar, L. P. 2013. Cursor movements controlled by real time hand gestures. *International Journal of Science and Research*, **2** (2): 283-286.

Meethien, A. & Imran, B. M. 2013. Real time gesture recognition using Gaussian mixture model. *International Journal of Scientific & Engineering Research*, **4**(8): 1582-1586.

Messer, T. 2009. *Static Hand Gesture Recognition.* University of Fribourg, Switzerland (Unpublished).

Noor, A. I., Mokhtar, M. H., Rafiqul, Z. K. & Pramod K. M. 2012. Understanding Color Models: A Review. *ARPN Journal of Science and Technology*, **2** (3): 265-275.

Pansare, J. R., Gawande, S. H. & Ingle, M. 2012. Real-Time Static Hand Gesture Recognition for American Sign Language (ASL) in Complex Background. *Journal of Signal and Information Processing*, **3**: 364-367.

Park, H. 2010. *A method for controlling mouse movement using a real-time camera.* Master. Thesis. Brown University, United States.

Premaratne, P. 2014. *Human computer interaction using hand gestures.* Springer, Singapore.

Rekha, J., Bhattacharya, J. & Majumder, S. 2011. Hand gesture recognition for sign language: A new hybrid approach. In: Arabnia, H. R., Deligiannidis, L. & Schaefer, G. (eds) *Proceedings of the International Conference on Image Processing, Computer Vision and Pattern Recognition*, 18-21 July 2011, Las Vegas: 80-86.

Reynolds, D. A. 2009. Gaussian mixture models. In: Li, S. Z. & Jain, A. (eds) *Encyclopedia of Biometrics*. Springer, New York, pp. 659-663.

Rezaei, M. 2014. Haar Training Tools and Dataset. https://www.cs.auckland.ac.nz/~m.rezaei/Downloads.html

Safaei, A. & Jahed, M. 2013. 3D hand motion evaluation using HMM. *Journal of Electrical and Computer Engineering Innovations*, **1** (1): 11-18.

Schapire, R. E. 2013. Explaining AdaBoost. In: Scholkopf, B., Luo, Z. Y. & Vovk, V. (eds) *Empirical Inference: Festschrift in Honor of Vladimir*. Springer, Berlin Heidelberg, pp. 37-52.

Viola, P. & Jones, M. J. 2004. Robust Real-Time Face Detection. *International Journal of Computer Vision*, **57** (2): 137-154.

Wang, C. C. & Wang, K. C. 2008. Hand Posture Recognition Using Adaboost with SIFT for Human Robot Interaction. In: Lee, S. H., Suh, I. H. & Kim M. S. (eds) *Recent Progress in Roboctics: Viable Robotic Service to Human.* Springer, Berlin Heidelberg, pp. 317-329.

Wilson, P. I. & Fernandez, J. 2006. Facial Feature Detection Using Haar Classifiers. *Journal of Computing Sciences in Colleges*, **21** (4): 127-133.

Yin, Y. 2014. *Real-time continuous gesture recognition for natural multimodal interaction.* PhD. Thesis. Massachusetts Institute of Technology, Massachusetts.

Zabulis, X., Baltzakis, H. & Argyros, A. 2009. Vision-based hand gesture recognition for human-computer interaction. In: Stephanidis, C. (eds). *The Universal Access Handbook*. CRC Press, Florida.

Zaher Hamid Al-Tairi, Rahmita Wirza Rahmat, M. Iqbal Saripan & Puteri Suhaiza Sulaiman. 2014. Skin segmentation using YUV and RGB colour spaces. *Journal of Information Processing Systems*, **10** (2): 283-299.

Zhu, J., Zou, H., Rosset, S. & Hastie, T. 2009. Multi-Class AdaBoost. *Statistics and Its Interface*, **2**: 349-360.