AN INTELLIGENT GESTURE RECOGNITION SYSTEM

By

WAN MOHD RIDZUAN BIN WAN AB MAJID

(0730610169)

A thesis submitted

in fulfillment of the requirements for the degree of

Master of Science (Mechatronic Engineering)

School of Mechatronic Engineering

UNIVERSITI MALAYSIA PERLIS

2012

© This item is protected by original copyright

UNIVERSITI MALAYSIA PERLIS

DECLARATION OF THESIS

Author's full name : WAN MOHD RIDZUAN BIN WAN AB MAJID

Date of birth : 5 JUNE 1984

Title : AN INTELLIGENT GESTURE RECOGNITION SYSTEM

Academic Session : 2011/2012

© This item is protected by original copyright

I hereby declare that the thesis becomes the property of Universiti Malaysia Perlis (UniMAP) and to be placed at the library of UniMAP. This thesis is classified as :

☐ **CONFIDENTIAL** (Contains confidential information under the Official Secret Act 1972)

☐ **RESTICTED** (Contains restricted information as specified by the organization where research was done)

☒ **OPEN ACCESS** I agree that my thesis is to be made immediately available as hard copy or on-line open access (full text)

I, the author, give permission to the UniMAP to reproduce this thesis in whole or in part for the purpose of research or academic exchange only (except during a period of _____ years, if so requested above).

Certified by:

_____ _____

SIGNATURE **SIGNATURE OF SUPERVISOR**

840605-11-5639 Assoc. Prof. Dr Paulraj MP

(NEW IC NO. / PASSPORT NO.) **NAME OF SUPERVISOR**

Date : 21 March 2012 Date : 21 March 2012

ACKNOWLEDGEMENT

First of all, I would like to say my gratitude to the one and only almighty God, Allah S.W.T. for giving me the strength, good health and talent to complete this research work as one of the requirement for the conferment of the Degree.

I would like to express my sincere and profound gratitude to the Vice Chancellor of Universiti Malaysia Perlis, Y. Bhg. Brigedier Jeneral Dato' Prof. Dr. Kamarudin b. Hussin for granting me permission to study in this university.

I would like to express my thanks to the Dean of School of Mechatronic Engineering, Universiti Malaysia Perlis, Associate Prof. Dr. Abdul Hamid Adom for providing support during my research work.

I would like to extend my highest gratitude to my supervisor, Associate Prof. Dr. Paulraj Murugesa Pandiyan for his willingness to spend his precious time in giving me guidance and assistance throughout my research work. This thesis would not have been possibly finished without his encouragement and support.

© This item is protected by original copyright

In addition to that, my gratitude goes to Dr. Hazry Desa, who is the co-supervisor of my research project and Prof. Dr. Sazali Yaacob, Head of Acoustic Applications Research Cluster for their kindness in sharing their knowledge.

I also would like to thank my family, my wife and friends for giving me all support from the beginning of the research work. Especially, to all members of Acoustic Applications Research Cluster who have contributed directly and indirectly towards the completion of this research.

Also, to all staffs of Sekolah Rendah Pendidikan Khas Perlis, Kangar, for give me a lot of information regarding the Sign Language. Furthermore, I would like to dedicate thousands of appreciations to all my beloved lecturers who had given me the motivation and inspiration to make this research work a dream come true. May Allah bless you all with love and happiness.

© This item is protected by original copyright

TABLE OF CONTENT

ACKNOWLEDGEMENT III

TABLE OF CONTENT V

LIST OF FIGURES X

LIST OF ABBREVIATION XII

ABSTRAK XIV

ABSTRACT XV

Chapter 1 1

Introduction 1

1.1 Overview 1

1.2 Problem Statement 2

1.3 Need of the Research work 3

1.4 Research Objectives 4

1.5 Thesis Organization 6

Chapter 2 9

Literature Review 9

2.1 Introduction 9

2.2 Hearing impaired and Hearing Impairment 10

2.3 Hearing Impairment in Malaysia 12

2.4 Term of Kod Tangan Bahasa Melayu 13

2.5 Sign Language 15

2.6 Sign Language Recognition 16

2.6.1 Sign Language Recognition Based on Image 16

2.6.2 Sign Language Recognition Based on Video 21

2.7 Summary 25

© This item is protected by original copyright

Chapter 3 27

Image Processing and Features Extraction Methods 27

3.1 Introduction 27

3.2 Data Collection 29

3.2.1 Selection of Image Acquisition Device 29

3.2.2 Data Acquisition Setup 30

3.2.3 Requirement for the Subject during Data Acquisition 32

3.2.4 MATLAB software interface 33

3.3 Sign Language Image Processing 34

3.4 Feature Extraction 41

3.4.1 Image Feature Extraction Algorithm-Centroid 41

3.4.2 Image Feature Extraction Algorithm-Centroid-Magnitude 44

3.4.3 Image Feature Extraction Algorithm-Angle 46

3.4.4 Image Feature Extraction Algorithm-Area of Triangle 48

3.4.5 Image Feature Extraction Algorithm-Distance 52

3.5 Summary 54

Chapter 4 55

Applications and Implementation of Neural Networks Algorithms 55

4.1 Introduction 55

4.2 Artificial Neural Network 55

4.3 Back Propagation Network 57

4.3.1 Training Preparation 59

4.3.2 Back Propagation Training Algorithm 60

4.4 Implementation of Neural Network for Centroid Features 62

4.5 Implementation of Neural Network for Centroid-Magnitude Features 64

4.6 Implementation of Neural Network for Angle Features 66

© This item is protected by original copyright

4.7 Implementation of Neural Network for Area of Triangle Features 68

4.8 Implementation of Neural Network for Distance Features 70

4.9 Summary 72

Chapter 5 73

Result and Discussion 73

5.1 Introduction 73

5.2 Results of Centroid Features 74

5.3 Results of Centroid-Magnitude Features 76

5.4 Results of Angle Features 78

5.5 Results Area of Triangle Features 80

5.6 Results of Distance Features 82

5.7 Comparison of Feature Extraction 85

5.8 Development of Graphical User Interface (GUI) for an Intelligent Gesture Recognition System 86

5.8.1 Development of GUI 87

5.8.2 Software Description 88

5.9 Summary and Conclusion 92

Chapter 6 94

Conclusion and Recommendation 94

6.1 Contributions 97

6.2 Suggestions for Future Work 98

REFERENCES 100

LIST OF PUBLICATIONS 105

AWARDS AND MEDALS 107

APPENDIX A 108

© This item is protected by original copyright

© This item is protected by original copyright

LIST OF TABLES

Table 3.1 The List of Gesture Used in Project 37

Table 5.1 Neural Network Training Results for Centroid Features 74

Table 5.2 Confusion Matrix result for Centroid Features 75

Table 5.3 Neural Network Training Results for Centroid-Magnitude Features 77

Table 5.4 Confusion Matrix result for Centroid-Magnitude Features 78

Table 5.5 Neural Network Training Results for Angle Features 79

Table 5.6 Confusion Matrix result for Angle Features 80

Table 5.7 Neural Network Training Results for Area of Triangle Features 81

Table 5.8 Confusion Matrix result for Area of Triangle Features 82

Table 5.9 Neural Network Training Results for Distance Features 83

Table 5.10 Confusion Matrix result for Distance Features 84

Table 5.11 Comparison of Classification Accuracy for Various Feature Extraction Method 85

Table 5.12 Comparison of Computational Time of Training for Various Feature Extraction Method 85

© This item is protected by original copyright

LIST OF FIGURES

Figure 2.1 Conductive Hearing Impairment(*TheFreeDictionary.com*., 2010a) 10

Figure 2.2 Sensorineural Hearing Impairment(*TheFreeDictionary.com.*, 2010b) 11

Figure 2.3 Gesture/sign for the word Many (Banyak) 14

Figure 2.4 Gesture/sign for the word Animal (Binatang) 14

Figure 3.1 Block Diagram of an Intelligent Gesture Recognition System 28

Figure 3.2 Image Acquisition Device 30

Figure 3.3 The Complete Experimental Setup for Data Acquisition 31

Figure 3.4 The layout of the MATLAB Interface 34

Figure 3.5 Image Acquisition Device 36

Figure 3.6 Gesture Video Sequence for the word Big (Besar) 38

Figure 3.7 Gesture Video Sequence for the Word Small (Kecil) 39

Figure 3.8 Binary Image of Gesture for the word Shirt (Baju) 40

Figure 3.9 Binary Image of Gesture for the word Guitar (Gitar) 40

Figure 3.10 Centroid Feature for Guitar (Gitar) sign 42

Figure 3.11 Centroid Feature Plot for Five Different Signs 44

Figure 3.12 Centroid-magnitude Feature for Guitar (Gitar) sign 45

Figure 3.13 Centroid-Magnitude Features Plot for Five Different Signs 46

Figure 3.14 AngleFeature of Image 47

Figure 3.15 Angle Feature Plot for Five Different Signs 48

Figure 3.16 Formation of Area of Triangle for Word Table (Meja) 48

Figure 3.17 Formation of Area of Triangle for Word Guitar (Gitar) 49

Figure 3.18 Area of triangle Feature Plot for Five Different Signs 49

Figure 3.19 Distance between centroid of the current frame and the previous frame 52

© This item is protected by original copyright

Figure 3.20 Distance Feature Plot for Five Different Signs 53

Figure 4.1 Architecture of Back Propagation Network (Sivanandam & Paulraj, 2003) 58

Figure 4.2 Neural Network Architecture for Centroid Feature 63

Figure 4.3 Neural Network Architecture for Centroid-Magnitude Feature 65

Figure 4.4 The Architecture of Neural Network Model for Angle Feature 67

Figure 4.5 Neural Network Architecture for Area of triangle Feature 69

Figure 4.6 Neural Network Architecture for Distance Feature 71

Figure 5.1 GUI Layout 89

Figure 5.2 Open Dialogs to Load the Sign Language File 90

Figure 5.3 GUI output for the Gesture of Guitar (Gitar) word 90

Figure 5.4 GUI output for the Gesture of Many (Banyak) word 91

Figure 5.5 GUI output for the Gesture of Bicycle (Basikal) word 91

© This item is protected by original copyright

LIST OF ABBREVIATION

ANN	Artificial Neural Network
ASL	American Sign Language
AVI	Audio Video Interleave
BIM	Bahasa Isyarat Malaysia
BP	Back Propagation
CCD	Charge-Coupled Device
CIF	Common Intermediate Format
CMOS	Complementary Metal–Oxide–Semiconductor
DOS	Disk Operating System
GUI	Graphical User Interface
HCI	Human Computer Interface
HMM	Hidden Markov Model
KTBM	Kod Tangan Bahasa Melayu
NNM	Neural Network Model
OKU	Orang Kurang Upaya
PC	Personal Computer
PNP	Plug And Play
PVBG	Pure Visual Based Gesture
RGB	Red Green Blue
SWDM	Sosial Welfare Department Malaysia
TSL	Taiwanese Sign Language
USB	Universal Serial Bus

© This item is protected by original copyright

WHO World Health Organization

© This item is protected by original copyright

ABSTRAK

SISTEM PENGECAMAN KOD TANGAN PINTAR

Pada masa kini, maklumat dan pengetahuan berkembang dengan banyak dan pesat. Walaubagaimana pun, orang kurang upaya (OKU), seperti orang cacat pendengaran, sering tersisih kerana terdapat jurang komunikasi yang besar di antara mereka dengan orang yang normal. Salah satu kaedah komunikasi yang dapat mengurangkan jurang komunikasi ini ialah dengan menggunakan bahasa isyarat. Kebiasaannya bila orang normal ingin berkomunikasi dengan orang cacat pendengaran, penterjemah diperlukan untuk menterjemahkan bahasa isyarat kepada isyarat bunyi. Oleh sebab itu, satu kaedah mudah telah dibangunkan untuk menukar bahasa isyarat kepada isyarat bunyi dengan menggunakan ciri-ciri yang diperolehi dari gerakan tangan. Dengan menggunakan kamera, system ini menerima video bahasa isyarat daripada orang cacat pendengaran dalam bentuk rangkaian video Merah Hijau Biru (Red Green Blue: RGB) dengan kedalaman warna layar 24-bit dan resolusi 320 x 240 piksel. Untuk setiap rangka gamba, dua bahagian tangan diasingkan dan kemudian diubah menjadi imej binari. Model pengekstrakan ciri-ciri kemudian diterapkan pada setiap gambar yang diasingkan untuk mendapatkan ciri-ciri yang paling penting dari gambar. Rangkaian Saraf Tiruan (Artificial Neural Network: ANN) memberikan alternatif bentuk pengkomputeran yang berkeupayaan untuk meniru fungsi otak manusia. Sebuah model rangkaian neural dibangunkan untuk pengecaman bahasa isyarat dari rangkaian video. Sistem audio dipasang untuk memainkan maksud bahasa isyarat itu kepada orang biasa.

© This item is protected by original copyright

ABSTRACT

AN INTELLIGENT GESTURE RECOGNITION SYSTEM

Information and knowledge are expanding in quantity and accessibility. However, people with functional limitations, such as hearing impaired, often left out of conversation where there are wide communication gaps between them with the ordinary people. The sign language is the fundamental communication method between people who suffer from hearing defects. In order for an ordinary people to communicate with hearing impaired community, a translator is usually needed to translate the sign language into natural language. This project presents a simple method for converting sign language into voice signal using features obtained from the hand gestures. Using a camera, the system receives sign language video from the hearing impaired subject in the form of video streams in RGB (red-green-blue) colour with a screen bit depth of 24-bits and a resolution of 320 x 240 pixels. For each frame of images, two hand regions are segmented and then converted into binary image. Feature extraction model is then applied on each of segmented image to get the most important feature from the image. Artificial Neural Network (ANN) provides alternative form of computing that attempts to mimic the functionality of the brain. A simple neural network model is developed for sign recognition directly from the video stream. An audio system is installed to play the particular word for the communication between the ordinary people and hearing impaired community.

© This item is protected by original copyright

Chapter 1

Introduction

1.1 Overview

Communication is the most important means that people in a society use to convey or exchange messages, news, information, thoughts and feelings. Information and knowledge are expanding in quantity and accessibility. People with functional limitations, such as the hearing impaired people, often experience wide communication gaps, even though, most of them have normal intellectual capacity. Due to the hearing limitation, hearing impaired people have developed their own culture and methods for communicating among themselves and also with hearing groups.

© This item is protected by original copyright

Sign language is one of the most important and natural communication modality. A sign language is a language which, instead of acoustically express sound patterns, uses visually transmitted sign patterns (manual communication, body language and lip patterns) to expresses one's thought. A speaker's expression can be represented by simultaneously combining the hand shapes, orientation and its movements. It is a static expressing system that is composed of signs by using hand motion aided by facial expressions. Sign language is mainly employed by hearing impaired people to communicate with each other. However, communication with normal people is a big handicap for them since the normal people do not understand their sign language. Communication between people using different sign

languages is easier than different spoken languages meet. Sign language provides access to an international hearing impaired community. Sign language recognition system is needed in providing an easy, efficient and accurate mechanism to transform sign language into speech or word.

The aim of this research work is to develop a sign language recognition system based on neural networks. The work consists of the data collection, video and image processing, feature extraction, training Neural Network Models (NNMs), and recognizing sign language with the NNMs are proposed in recognizing sign language. The main motivation and objectives of this research work have been discussed in this chapter.

© This item is protected by original copyright

1.2 Problem Statement

Hearing impaired community and hard of hearing people often have problems to understand and communicate with other people. Often hearing impaired community and hard of hearing people feel left out of conversation and feel ignored by other people. This is very difficult for hearing impaired people to deal with, as they want to join in but may feel embarrassed to ask people what they are talking about and to ask for things to be repeated again and again. The sign language is the fundamental communication method between those who suffers from hearing defects. In order for an ordinary person to communicate with the hearing impaired

people, a translator is usually needed to translate the sign language into natural language. In order to help the hearing impaired community, an intelligent system that can convert the symbolic signs of the sign language into its equivalent voice signal has been proposed.

1.3 Need of the Research work

Sign language is an important language used daily by the hearing impaired community as a means of communication. In order to communicate with family members, friends and the public, sign language has to be used. Unfortunately, communication with normal people is a quite difficult task for them since the normal people do not understand the sign language. Sign language recognition is needed for realizing a human oriented interactive system which can perform a normal interaction communication. The first motivating factor is the possibility of reducing the communications barrier which exists between the hearing impaired and hearing communities.

© This item is protected by original copyright

The developed system is user-friendly and can be accessed easily. This system will be helpful for hearing impaired and hard of hearing people to solve their communication problems and give them the same whole life experience to which their hearing colleagues are accessing.

1.4 Research Objectives

The main objective of this research work is to develop an intelligent gesture recognition system for hearing impaired community. In order to achieve this, four objectives are presented.

© This item is protected by original copyright

a) To develop intelligent algorithms that will extract the features from the head and hand gesture images

Feature extraction plays an important role in sign language recognition system. There are many types of feature extraction methods available in the literature. In this research work, the sign language code from "Kod Tangan Bahasa Melayu"(KTBM) is considered. Simple feature extraction methods are to be developed to extract the features from the sign language code represented in the form of an image. The performance of the various feature extraction methods are need to be evaluated and to be compared.

b) To develope Matlab software interfaces to interface a camera as well as a voice generating system to a computer.

A simple software system is to be developed to interface the camera to the computer. Further, a suitable software interface is to be developed for generating and displaying the equivalent words.

c) To design and develop neural network model embedded using Matlab in the computer system for image recognition and voice association .

Designing a decision making system is the important part in the automatic recognition tool. Development of classifiers for automatic detection of sign language recognition is important. There are various classifiers have been proposed such as neural network. Each classifier has its own advantages and disadvantages. The selection of suitable classifiers for a particular problem requires good knowledge. Neural Network provides alternative form of computing that attempts to mimic the functionality of the brain. As neural network classifier is one of the best classifiers, it has been proposed to develop a suitable neural network model.

© This item is protected by original copyright

d) To develop software interfaces using GUI Interface

Development of PC based automatic recognition system for sign language recognition is an important part of this research. In this research work, an effort has

been go through for development of PC based on sign language recognition system using a commercially available platform such as MATLAB. MATLAB® is a high-level technical computing language and interactive environment for algorithm development, data visualization, data analysis, and numeric computation. In this research work, MATLAB GUI is used to design the software interface which is to standardize and simplify the use of the computer programs.

© This item is protected by original copyright

1.5 Thesis Organization

This thesis explores the topic of sign language recognition system using image and video processing and artificial intelligence techniques. The research works that were carried out are presented in six chapters in this thesis.

In Chapter 1, the introduction to the topic of the interest is explained. This chapter presents an overview of the research work that describes the problem statement that occurs to the hearing impaired and hard of hearing community and needs of the research work to overcome this problem. This chapter also deals with the objective of the proposed research work and organization of the thesis.

The literature reviews on hearing impaired, the causes of hearing impaired, sign language and sign language recognition is presented in Chapter 2. A detailed review of the early works performed by other researchers has been discussed.

In Chapter 3, the development of image processing and features extraction methods in identifying the sign language is described. The method used to develop an intelligent gesture recognition system is discussed and the way of preprocessing of image and video processing is proposed. The experimental setup for the research work and the configurations for the experiment have also been discussed.

The development and implementation of Neural Network is explained in Chapter 4. The proposed Neural Networks based sign language recognition is suggested and its performance is discussed. The neural networks are trained using the Back Propagation Neural Network model structure. Experiments were conducted to evaluate the performance characteristics of the proposed sign language recognition system.

The result and discussion of the proposed feature and suggested classifier together with the development of the graphical user interface for the sign language recognition system using MATLAB GUI is presented in Chapter 5. Each proposed

© This item is protected by original copyright

feature results and the GUI interface for this research work is incorporated in this chapter.

In Chapter 6, the problems encountered and the conclusion of the research work done is discussed. Few suggestions for future research works were given.

© This item is protected by original copyright

Chapter 2

Literature Review

2.1 Introduction

Sign language is the fundamental communication modality among the hearing impaired. Normal people generally require a translator to communicate with the hearing impaired. The hearing impaired people have developed their own culture and methods for communicating among themselves as well as with ordinary people by using sign language. Sign language are a non-verbal visual language, different from the spoken language, but serving the same function. The types of sign language can be classified according to the various regions around the world. For example, Algerian and Ghana sign languages are from the continent of Africa; American and Argentinean sign languages are from the United States of America; Australian and Malaysian sign languages are from the Asia/Pacific region; and Austrian and British sign languages are from the continent of Europe(Flodin, 2004).

© This item is protected by original copyright

In this research work, literature reviews were conducted on the issues related to the development of an intelligent gesture recognition system. These include reviews on hearing impaired and hearing impairment, the hearing impaired communities in Malaysia, sign language and the previous research work of sign language recognition.