DESIGN AND CONSTRUCTION OF ROBOT HEAD SYSTEM USING SPEECH RECOGNITION TO EXPRESS EMOTIONS

LIEW QUEK SHIUNG

BACHELOR OF MECHATRONICS ENGINEERING WITH HONOURS
UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2019

DESIGN AND CONSTRUCTION OF ROBOT HEAD SYSTEM USING SPEECH RECOGNITION TO EXPRESS EMOTIONS

LIEW QUEK SHIUNG

A report submitted
in partial fulfillment of the requirements for the degree of
Bachelor of Mechatronics Engineering with Honours

Faculty of Electrical Engineering

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2019

DECLARATION

I declare that this thesis entitled "DESIGN AND CONSTRUCTION OF ROBOT HEAD SYSTEM USING SPEECH RECOGNITION TO EXPRESS EMOTIONS is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature	:	
Name	:	LIEW QUEK SHIUNG
Date	:	24 MAY 2019

APPROVAL

I hereby declare that I have checked this report entitled "DESIGN AND CONSTRUCTION OF ROBOT HEAD SYSTEM USING SPEECH RECOGNITION TO EXPRESS EMOTIONS" and in my opinion, this thesis it complies the partial fulfillment for awarding the award of the degree of Bachelor of Mechatronics Engineering with Honours

Signature : .....

Supervisor Name : .....

Date : .....

DEDICATIONS

To my beloved mother and father

ACKNOWLEDGEMENTS

My final year project would not have been possible without the contribution and collaboration of others. First of all, I would like to thank the Almighty God who granted me health and long life, without which I could not have finished this project.

Next, I want to thank my supervisor, Dr. Fariz bin Ali @ Ibrahim for his support and constant supervision which contributed immensely to my personal development. I also want to thank him for his guidance which was a remarkable force that enabled me to successfully complete the final year project.

After that, I would also like to thank the rest of the lecturers for their supports and guidance which helped me to overcome the challenges I faced during my final year project.

Last but not least, special thanks to all my friends and family for sharing their experiences, time and commitment during my final year project. I am grateful because I have a lot of friends and family who were always there for me to help and support me throughout the course of completing it.

ABSTRACT

Humanoid robot is a robot that has the characteristics of a human. It is designed to improve the quality of human life. Humanoid robot is able to have human-like behaviours and engage in effective robot-human synergy. Facial expressions are very important in communication between human. The objectives of this project are to develop a robot head system that is able to express five basic expressions which are happiness, sadness, anger, surprise and fear, to control the robot's expressions using voice commands and to analyse the performance of the robot head system in terms of position, the recognition rate of each facial expression and the speech recognition rate. The control points of the robot head are proposed by combining the action units. The robot head that is able to express various emotions have been provided with 5 DOFs. A survey has been conducted on the design after finished drawing using SolidWorks. Evaluation on the conceptual design of the robot head is done, later, to be implemented on the final product of the robot head. Hardware experiment has been conducted and Arduino Mega is used as the controller for the system. Besides, a microphone with voice recognition module is designed to control the expressions of the robot using speech recognition. This project also focused on the robot head system's ability to function in terms of position of the actuators, recognition rate for facial expressions and speech recognition. The result of the relative position error for the chosen actuator is less than 5%. This increases the ability of the robot head to express the emotions more effectively. Besides, a survey related to the recognition rate for the five facial expressions has been done on 30 subjects. The result obtained for most of the recognition rate is more than 50% except for fear which is 23.34%. The speech recognition rate is repeated for 30 times. The recognition rate for most of the emotions are above 90%. The objectives of this project have been achieved and a robot head system which is able to express emotions using speech recognition has been successfully designed and constructed. However, in the future work, the facial expression recognition rate is expected to be improved by adding more DOFs or improve the hardware integration. The robot head system can also be integrated with other robotic bodies and addition of sensors.

ABSTRAK

Humanoid robot merupakan sejenis robot yang mempunyai ciri-ciri seorang manusia. Tujuan ia direka adalah untuk meningkatkan kualiti kehidupan manusia. Humanoid robot mampu bertingkah laku sebagai seorang manusia dan manjalani interaksi robot-manusia dengan efektif. Ekspresi wajah memainkan peranan yang penting dalam kalangan manusia terutama komunikasi. Objektif projek ini adalah untuk membina satu sistem kepala robot yang dapat mengekspreskan lima emosi asas, iaitu kegembiraan, kesedihan, kemarahan, kejutan, dan ketakutan, mengawasi ekspresi wajah melalui perintah suara, dan menganalisis prestasi sistem kepala robot terutama di atas kedudukan aktuator, kadar pengecaman untuk ekspresi wajah dan suara . Titik kawalan kepala robot adalah dicadangkan dengan menggabungkan unit aksi. Kepala robot yang terbaru ini adalah diberikan 5 darjah kebebasan untuk melaksanakan pelbagai emosi. Satu kajian berkaitan dengan rekabentuk telah dijalankan selepas dilukis menggunakan SolidWorks. Penilaian rekabentuk konsep kepala robot juga telah dijalankan, kemudian, digunakan pada produk terakhir kepala robot. Uji kaji perkakasan telah dijalankan dan Arduino Mega digunakan sebagai kawalan untuk sistem kepala robot. Tambahan pula, satu mikrofon dengan modul pengecaman suara telah direka untuk mengawal emosi robot melalui pengecaman suara. Projek ini memberi tumpuan kepada ujian prestasi sistem kepala robot terutama di atas kedudukan penggerak, kadar pengecaman untuk ekspresi wajah dan suara. Hasil ralat kedudukan relatif bagi penggerak yang terpilih adalah kurang daripada 5%. Oleh hal yang demikian, kepala robot tersebut dapat melakukan emosi dengan efektif. Selain itu, satu tinjauan di atas kadar pengecaman ekspresi wajah telah dijalankan dengan 30 orang responden. Kebanyakan hasil kadar pengesanan emosi yang didapati adalah melebihi 50% kecuali ketakutan, iaitu sebanyak 23.34%. Tinjauan untuk pengecaman suara juga telah diulangi sebanyak 30 kali. Kebanyakan hasil kadar pengesanan bagi pengecaman suara adalah melebihi 90%. Objektif projek ini telah dicapai dan satu kepala robot yang mampu mengekspreskan emosi melalui pengecaman suara telah direka dan dibina. Walau bagaimanapun, pada masa depan, kadar pengesanan emosi dijangka akan bertambah baik dengan menambah darjah kebebasan atau memperbaiki integrasi perkakasan. Kepala robot juga boleh berintegrasi dengan badan robotik lain dan menambah pengesan.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

AUs	-	action units
DOFs	-	degree of freedoms
mm	-	millimeter
VCC	-	voltage collector to collector
IDCU		Computer and Intelligent Drive & Control Unit

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Motivation

Autism spectrum disorder (ASD) is a complex developmental condition where it will affect a person's ability to communicate and interact with the others, involve issue with non-verbal communication, and behaviours. The effects of ASD and its severity vary in each individual. Usually, ASD is diagnosed when the individual is around 2 – 3 years with many of the most obvious signs presenting. However, some of the individual with ASD may develop normally until they stop learning or lose previously acquired skills [1].

In [1], it explains that ASD varies from person to person in severity and combination of symptoms. However, the common characteristics of ASD are difficulty in interacting and communicating, having issues relating to things and limited behaviours. Individuals with ASD have issues in a continuous conversation, lessen sharing of interests, challenges in responding to interaction such as eye contact and facial expressions, trouble keeping a relationship, and others.

According to UNICEF's record, there are around 30,000 children registered with disabilities in 2012, and 19,150 of them were having learning disabilities. It is approximated that 1 out of every 600 children in Malaysia is born with ASD based on BERNAMA's report in 2014. The latest statistics also showed that there are about 47,000 Malaysians that are autistic, with roughly 4 out of every 10,000 suffering from severe autism [2]

A child with physical or mental disabilities needs more attention and involvement to assist, and all of these cost money. In 2010, it is estimated that raising a child with disabilities required a total cost of RM96,700 each year. The cost covered diagnosis, therapies, special needs education, and others [2].

Therefore, the end goal here is to build a robot system that is able to help children with autism and special needs. The robot system will be started by constructing its head first. The robot head is designed with the ability to express

emotions. Besides, the robot head will be able to communicate with the children so that their social interactions can be improved. Children with autism have issues staring into one's eyes and the ability to remember faces that are unfamiliar. Thus, a robot head system with facial expressions is able to help them because it can help them stay focused and interact with it.

1.2 Problem Statement

There is a theory supported by the research conducted by electric engineering and computer science department at Vanderbilt University and issued in the journal Neural Systems and Rehabilitation Engineering saying that robotic technology works with children with ASD. The research was carried out by testing on six children with ASD and six children in stages of normal development. The result showed that a child with ASD spent more time staring at the robot compared to the normally-developing child. Thus, this study shows that a robot head system could actually help the children with autism in staying focused better [3].

In [4], it mentions that robot-assisted therapies are showing potential in becoming assessment and therapeutic tools. More researches have shown that child with ASD engage better with robots rather than human, because they are simpler and more predictable. In [5], Assoc. Prof. Dr. Hasnah Toran – a leading authority on autism in the country supports the idea of robot-assisted therapies. It is because there is a shortage of healthcare professionals in the field. There are around 300 psychiatrists in the country and only 20 of them are child psychiatrists.

From the above statement, it is believed that a robot capable of showing facial expressions and communicating with the autistic children is able to help them. Autistic children are able to interact more with a robot than a human. This is because the facial expressions, the voice tone and the body movement of the human prevent the autistic child to concentrate more, thus, they are not able to interact well with the human.

There are a few important criteria that have to be considered in designing and constructing the robot head system so that it is effective in helping the autistic children learn to socially adapt:-

1) The design concept and structure of a human-like robot head
2) The degree of freedom of the robot head movement
3) The expression of emotion of the robot head

4) The recognition rate of each emotion expressed by the robot head

5) The ability of the robot head to recognize the speech given by the user

6) The speech recognition rate based on the success rate on expressing correct facial expression

The hardware for the robot head system has been chosen carefully because the selection of hardware components is capable of influencing the performance of the system. Arduino Mega is used as the controller for the system because it has 54 digital I/O pins. The model of the actuators selected are based on its performance in terms of angular position, torque and velocity. The actuators chosen must be capable of high precision for the robot head movement.

1.3 Objectives

The objectives of the research study are:

1) To design and construct a robot head system with five basic expressions.
2) To control the robot's expressions using voice commands.
3) To analyse the performance of the robot head system in terms of position, the recognition rate of each facial expression and the speech recognition rate.

1.4 Scopes

The scopes of the research study are:

1) The construction of the robot head focuses on the human-like movement which consists of the eyebrows, eyelid and mouth. Each part is able to move independently to produce rich facial expressions. The motion of two eyebrows, two eyelids which linked as one, and upper and lower lips are taken into considerations to increase the effectiveness of developing a lifelike robotic face.
2) The robot head system consists of 5 degree of freedoms (DOFs) with 2 DOFs for the eyebrows, 1 DOF for the eyelid, 1 DOF for the upper lip and 1 DOF for the lower lip.
3) The 5 basic expressions are happiness, sadness, anger, fear and surprise.
4) This project covers the position repeatability test by repeating each of the expression for 15 times as it has been done before in [6].

5) This project also covers the facial expression recognition test on 30 subjects by using the final product of the robot head system. To justify using 30 subjects, it has been showed in [11], the recognition test used 27 subjects; in [12], it is repeated with 20 subjects; and in [16], the test is repeated with 32 subjects.
6) The motion of the robot head is being controlled by Arduino Mega which consists of 54 digital I/O pins.
7) A microphone with a voice recognition module is installed in the system to process the speech given by the user. The speech must be the keywords that have been programmed in the system.
8) Since facial expression recognition test is done with 30 subjects, the speech recognition, too is repeated 30 times. The recognition is done with the keywords provided to the subjects. It is counted as success if the robot head is able to express the correct facial expression with the correct keyword.

CHAPTER 2

LITERATURE REVIEW

2.1 Theory

2.1.1 Design and Construction of Robot Head System

The study of humanoid robots has been conducted for many years and the robot head system becomes the main factor in determining the effectiveness of human-robot interaction. Different type of design of the robot head will have distinct interpretation to the user. Therefore, there are a few considerations to be taken into when designing the robot head [7]:

1) The design of a face with understandable expressions, and
2) The design of a friendly face and not react unfavourably to it.

In [8], there is a discussion in if the robot head should be more human-like or more technical optimized. The benefit of a technical head is that, there is no restriction on the design parameters such as size and shape. This particular detail reduces the effort for mechanical construction. Contrarily, if the robot head is used for human-robot interaction, an anthropomorphic head could help increase the performance of the system.

The technical complexity is measured by the number of sensors or actuators in constructing the robot head. The main objective of this project is to construct a robot head with 5 DOFs so that it is able to express the 5 basic expressions in the best recognition rate. The construction of the robot head included the eyebrows, eyelids, and lips.

2.1.2 Facial Action Coding System (FACS)

The Facial Action Coding System (FACS) refers to a set of facial muscle movements that match with a displayed emotion. It is created by Carl-Herman Hjortsjö with 23 facial motion units in 1970. It is later further developed and published by Paul Ekman and Wallace Friesen in 1978. Then, it was substantially updated and published in 2002 by Ekman, Friesen, and Joseph C. Hager [9].

Action units (AUs) are the basic actions used for the expression of emotions by means of facial expression. The AUs are formed by a group of muscles to produce various facial expression. The five basic expressions introduced are happiness, sadness, surprise, anger and fear.

In this project, 4 AUs are used to produce the five facial expressions as shown in Table 2.1. 10 control points and their movements on the facial skin are selected corresponded to the action units as shown in Figure 2.1.

Table 2.1 Action Units and Control Points for Five Facial Expressions

Action Units	Motion Changes	Control Point	
		Right	Left
1	Eyebrow Motion	1, 3	2, 4
2	Eyelid Motion	5, 6	
3	Upper Lip Motion	7, 8	
4	Lower Lip Motion	9, 10	

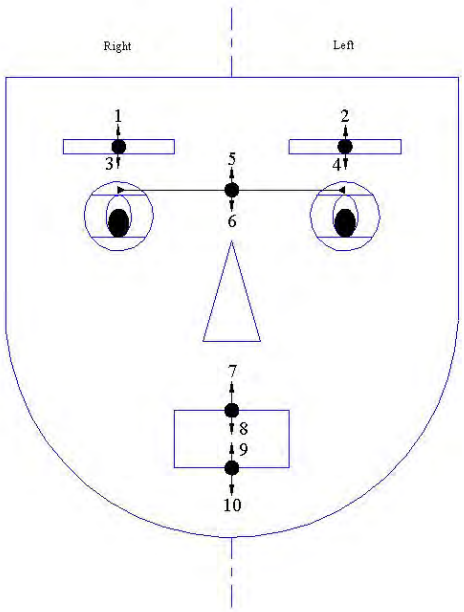

Figure 2.1 Position of Control Points and Their Direction of Movement

2.2 Research and Study

2.2.1 Facial Expressions

Table 2.2 to 2.6 show the review for 6 research papers on facial expressions. The robot names of the 6 research papers are ExpressionBot, KOBIAN-Refined, SAYA, F & H Robot, Flobi and ROMAN respectively. There are a few elements that are compared to help design and construct a robot head system for this project. The elements compared are the type of head used, the propose method of the robot head system, the type of expression, the DOFs of the robot head system, the type of motor used, the type of mechanism used, the dimension of the robot head, the weight of the robot head, the recognition rate of the facial expressions and the design overview of the robot head system.

Table 2.2 Review for 6 Research Papers on Facial Expressions

No.	**Paper 1** [10]	**Paper 2** [11]
Name of Robot	ExpressionBot	KOBIAN-Refined
Type of Head	Human-like	Human-like
Propose Method	1. Neck system controls the projector and mask position. 2. Display system consists of a small projector. 3. Animation application displays a face animation along with speech and emotion.	1. The head should have high facial expression ability. 2. Facial color is suggested in creating computer graphics (CG) images to affect the impression (e.g. sadness – check becomes little blue-tinged, fear – blue forehead color expression).
Type of Expression	6 basic expressions (anger, disgust, fear, joy, sadness, surprise) and neutral	Ekman's 6 basic expressions (anger, sadness, happiness, surprise, disgust, and fear)
Degree of Freedoms (DOFs)	3 DOFs (150º of yaw, 30º of pitch, 30º of roll)	24 DOFs (3 for eyes, 5 for eyelids, 8 for eyebrows, 7 for lips, and 1 for jaw)
Type of Motor	N/A	DC motor and ultrasonic motor

Table 2.3 Review for 6 Research Papers on Facial Expressions (cont.)

Type of Mechanism	1. Dell DLP M110 Portable Projector. 2. Nikon Fisheye Converter FCE8. 3. A mold is designed using the 3D model of the neutral face in Autodesk Maya (as a mask).	1. Eyes are driven by a gimbal structure where metal ring turns around low friction PTFE sleeve. 2. Eyelids are driven by an ultrasonic motor TULA-70. 3. Eyebrows are driven by a magnet through the cover.
Type of Controller	Computer	Computer
Skin Material	N/A	Silicone-like (Dragonskin)
Dimension (mm)	N/A	150 x 214 x 181
Weight (kg)	N/A	1.7
Recognition Rate	Joy – 100% Anger – 85% Sadness – 100% Disgust – 40% Surprise – 100% Fear – 55%	Happiness – 71.5% Anger – 92.3% Sadness – 73.1% Surprise – 96.2% Fear – 19.2% Disgust – 57.7%
Design Overview		

Table 2.4 Review for 6 Research Papers on Facial Expressions (cont.)

No.	**Paper 3** [12]	**Paper 4** [13]
Name of Robot	SAYA	F & H Robot
Type of Head	Human-like	Human-like
Propose Method	1. Generate Action Units (AUs) necessary for producing the various facial expressions. 2. Use McKibben type pneumatic actuator to control displacement of control points on facial skin. 3. Facial skin made from soft urethane resin.	1. Anti-cracking, anti-aging and coloring properties of face film material must be treated. 2. Adopts motor plus guiding tube plus diving mechanism of non-metal rope.
Type of Expression	6 basic expressions (surprise, fear, disgust, anger, happiness, sadness) and calm	Ekman's 6 basic expressions (happy, anger, sad, surprise, fear, disgust), initial (nature), and solemnity
Degree of Freedoms (DOFs)	6 DOFs (2 for eyeballs, 1 for chin, 2 for mechanical frames, and 1 for eyelid)	8 DOFs (3 for neck, 2 for eyeballs, 1 for eyebrow, and 2 for lower jaw)
Type of Motor	DC motor	DC servo motor
Type of Mechanism	1. McKibben type actuators that are attached in a mechanical frame. 2. Oculomotor mechanism on mechanical frame. 3. Human-like motion using coil spring	1. Small-pneumatic-element mechanism 2. Liquid-drive mechanism. 3. Rope-drive mechanism
Type of Controller	Computer	Computer and IDCU
Skin Material	Soft urethan resin	Concave die and metal mould
Dimension (mm)	115 x 200 x 155	N/A
Weight (kg)	1.5	N/A