ADAPTIVE LANGUAGE PROCESSING UNIT FOR MALAYSIAN SIGN LANGUAGE SYNTHESIZER

BY

HARIS AL QODRI MAARIF

A thesis submitted in fulfillment of the
requirement for the degree of Doctor of Philosophy
(Engineering)

Kulliyyah of Engineering
International Islamic University Malaysia

JUNE 2021

ABSTRACT

Language Processing Unit (LPU) is a system built to process text-based data to comply with the rules of sign language grammar. This system was developed as an important part of the Sign Language Synthesizer system. Sign Language uses different grammatical rules from the spoken/verbal language, which only involves the important words that Hearing/Impaired Speech people can understand. It needs word classification by LPU to determine grammatically processed sentences for the sign language synthesizer. The existing language processing unit in SL synthesizers suffers time lagging and complexity problems, resulting in high processing time. The two features, i.e., the computational time and successful rate, become trade-offs which means the processing time becomes longer to achieve a higher success rate. To address this problem, this thesis proposes an adaptive Language Processing Unit (LPU) that allows processing the words from spoken words to Malaysian SL grammatical rule that results in relatively fast processing time and a good success rate. It involves n-grams, NLP, and Hidden Markov Models (HMM)/Bayesian Networks as the classifier to process the text-based input. As a result, the proposed LPU system has successfully provided an efficient (fast) processing time and a good success rate compared to LPU with other edit distances (Mahalanobis, Levenstein, and Soundex). The system has been tested on 130 text-input sentences with words ranging from 3 to 10 words. As a result, the proposed LPU could achieve around 1.449ms processing time with an average success rate of 84.49% for a maximum of ten-word sentences.

خلاصة البحث

وحدة معالجة اللغة (LPU) هي نظام تم إنشاؤه لمعالجة البيانات المستندة إلى النصوص لتتوافق مع قواعد لغة الإشارة. تم تطوير هذا النظام كجزء مهم من نظام مُرَكِّب لغة الإشارة. تستخدم لغة الإشارة في قواعد نحوية مختلفة عن اللغة المنطوقة / اللفظية، والتي تتضمن فقط الكلمات المهمة التي يمكن للمستمع / ضعاف الكلام فهمها. يحتاج إلى تصنيف الكلمات بواسطة LPU لتحديد معالجة الجمل نحويًا لمركب لغة الإشارة. تعاني وحدة معالجة اللغة الموجودة في أجهزة توليف اللغة الإنجليزية من مشاكل التأخر الزمني والتعقيد، مما يؤدي إلى رفع وقت المعالجة. تصبح السمتان، أي الوقت الحسابي والمعدل الناجح، مفاضلات مما يعني أن وقت المعالجة يصبح أطول لتحقيق معدل نجاح أعلى. لمعالجة هذه المشكلة، تقترح هذه الأطروحة وحدة معالجة اللغة التكيفية (LPU) التي تسمح بمعالجة الكلمات من الكلمات المنطوقة إلى القاعدة النحوية للغة الإنجليزية الى لغة المصدر الماليزي والتي تؤدي إلى وقت معالجة أسرع نسبيًا ومعدل نجاح جيد. يتضمن n-grams و NLP و Hidden Markov Models (HMM) / Bayesian Networks كمصنف لمعالجة المدخلات المستندة إلى النص. نتيجة لذلك، نجح نظام LPU المقترح في توفير وقت معالجة فعال (سريع) ومعدل نجاح جيد مقارنة بـ LPU مع مسافات التحرير الأخرى (Mahalanobis و Levenstein و Soundex) . تم اختبار النظام على 130 جملة إدخال نصية بكلمات تتراوح من 3 إلى 10 كلمات. نتيجة لذلك، يمكن أن تحقق LPU المقترحة حوالي 1.449 مللي ثانية من وقت المعالجة بمتوسط معدل نجاح 84.49٪ لجمل من عشر كلمات كحد أقصى.

.

APPROVAL PAGE

The thesis of Haris Al Qodri Maarif has been approved by the following:

_____
Rini Akmeliawati
Supervisor

_____
Teddy Surya Gunawan
Co-Supervisor

_____
Yasir Mohd Mustafah
Internal Examiner

_____
Kamarul Hawari bin Ghazali
External Examiner

_____
Ali Chekima
External Examiner

_____
Akram Zeki Khedher
Chairman

DECLARATION

I hereby declare that this thesis is the result of my own investigations, except where otherwise stated. I also declare that it has not been previously or concurrently submitted as a whole for any other degrees at IIUM or other institutions.

Haris Al Qodri Maarif

Signature :  Date : 11 June 2021

DEDICATION

To IIUM and My Lecturers,
(for Educated People Serving the Ummah)
My Parents, My Wife, Family, And Muslim Ummah

INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

DECLARATION OF COPYRIGHT AND AFFIRMATION OF FAIR USE OF UNPUBLISHED RESEARCH

ADAPTIVE LANGUAGE PROCESSING UNIT FOR MALAYSIAN SIGN LANGUAGE SYNTHESIZER

I declare that the copyright holders of this thesis are jointly owned by the student and IIUM.

By signing this form, I acknowledged that I have read and understood the IIUM Intellectual Property Right and Commercialization policy.

Affirmed by Haris Al Qodri Maarif

……..……..
Signature

Date : 11 June 2021

ACKNOWLEDGEMENTS

First of all, I would like to thank Allah SWT, who has allowed me to complete all steps in my thesis project. Without His mercy and blessing, it is impossible for me to work hard and reach all of the objectives of this project.

I gratefully acknowledge my supervisors, Prof. Dr. Rini Akmeliawati and Prof. Dr. Teddy Surya Gunawan, for their full support, assistantship, supervision, and giving me spirit in doing this project to complete my work.

Special thanks are especially given to my family, wife, and daughters for their endless support, encouragement, and love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ASL	American Sign Language
ATLAS	Automatic Translation into Sign Language
AEWLIS	ATLAS Extended Written Italian Sign Language
AUSLAN	Australian Sign Language
BER	Bit Error Rate
BLEU	Billinguaal Evaluation Understudy
BN	Bayesian Network
BSL	British Sign Language
DAG	Direct Acrylic Graph
DSP	Digital Signal Processing
HMM	Hidden Markov Models
HSI	Hearing/Speech Impaired
LGP	Língua Gestual Portuguesa
LPU	Language Processing Unit
MSL	Malaysian Sign Language
NEE	Name Entity Extraction
NLP	Natural Language Processing
POS	Part of Speech
RAM	Random Access Memory
SASL	South African Sign Language
SER	Sign Error Rate
SISI	Say It Sign It
SL	Sign Langauge
SLC	Sentence Length Calculator

LIST OF SYMBOLS

d	Euclidean distance
l	Length
s	String
max	Maximizing value
t	Transposition
d_ω	Jaro Winkler distance
μ	Mean feature vector
D_M	Mahalanobis distance
S	Covariance matrix
R_p	P Dimensional space
$\boldsymbol{S}$	HMM hidden states
Q_t	Number of distinct observation symbols per hidden state
A	State transition probability
B	Probability matrix
π	Pi value = 3.14159265359
O	Observation sequence
T	Number of observation in sequence
$\mathbf{Pa}_G$	Set of parent nodes
$p(w_i \mid w_{i-1})$	Conditional probability
$C(w_{n-1}w_n)$	Raw text
w_n	Word sequence

CHAPTER ONE

INTRODUCTION

1.1 BACKGROUND

The Hearing and Speech Impaired (HSI) refers to the person who cannot hear and/or speak the voice. The disability leads to a problem that HSI people cannot communicate with non-HSI people or vice-versa, who use spoken language to communicate daily. The impossibility of communicating between HSI and non-HSI people provides a communication problem for HSI people, as most people live as non-HSI.

The definition of HSI also refers to the deaf. The deaf can be divided into three categories. The categories are born as deaf or became deaf from the baby or considered as before language development, became deaf after language development, and become deaf in old age (Ladner, 2009). These categories need a specific language that is used as the communication tool

Sign Language (SL) is the primary language and can be considered as the *mother tongue*[1] for the HSI people. Sign language only needs some important words compared to spoken language(Islam et al., 2017). It only needs important words to form sentence. Generally, sign language uses Subject, Verb, Noun, and Adverb, as well as, there is no other suffixes, prefixes, and particles It is a non-verbal language that uses hand movement, hand orientation, face expression, head movement, posture, and body orientation (Ong & Ranganath, 2005). Since sign language is a non-verbal language,

[1] Many people who are born deaf learn sign language as their primary language, and it remains their preferred, or first, language. There is no written form of sign language, so deaf people have to rely on reading and writing in their second or less-preferred language. A significant proportion of deaf people therefore have a strong preference for accessing information in sign language rather than as written text.

the understanding of sign language has been compulsory for HSI people to communicate with each other (between HSI people).

The awareness of sign language for non-HSI people is little, or many do not know sign language. It provides an obstacle in communication in the community, especially if it needs interaction between non-HSI and HSI people. As obstacles arise in contact with the community, the communication bridge must fill the gap between them. The options are sign language translator and sign language synthesizer technology that translates spoken language to sign language (Joy & Balakrishnan, 2014a).

Using a sign language translator to communicate between non-HSI and HSI has been limited since sign language translators are limited in Malaysia. As early as 2017, there are only less than 100 certified SL translators to cater to more than 30,000 persons of HSI. While in the world, The World Federation of the Deaf reported that there are about 70 million HSI people (Joy & Balakrishnan, 2014a) and 138 living sign language, which is according to the Ethnologue catalog (Karpov et al., 2016).

Sign language synthesizer consists of three main modules, i.e., the voice recognition module, language processing unit module, and signing module. Each module has its components and algorithms which need a different approach to development. In this thesis, the main focus is on the language processing module, which transforms the input language.

The language processing module alters input language into output language that suitable for output sign language. The input and output language are in the sequence of words (text), in which some methodology is required to do the transformation process properly.

The development of language processing unit has been made and implemented in much different sign language, for example, American Sign Language (Wolfe, Cook, et al., 2011b), British Sign Language (Darren Murph, 2007), South African Sign Language (Zijl & Olivrin, 2008) and Australian Sign Language (Wong, 2004).

In Malaysia, the language processing unit has not been implemented as an integral part of the sign language synthesizer. Furthermore, the language processing unit for *Bahasa Isyarat Malaysia* has not been implemented. A comprehensive review of the existing work and proposed work on the language processing unit is presented in this thesis. Various methods such as Edit Distance, Natural Language Processing, HMM methods, and Bayesian Network are discussed.

1.2 PROBLEM STATEMENT

The interaction between the Hearing and Speech Impaired (HSI) people and the non-HSI people has faced challenges when there is a communication gap. Sign Language is the tool for HSI people to communicate with each other and with non-HSI people. The understanding of SL removes the communication gap between HSI and non-HSI. Unfortunately, the number of people who understand SL is relatively small, and the SL translator population is considered less.

The sign language synthesizer has been a tool to convert spoken language into sign language. There are three main sign language modules: the voice recognition module (capturing the speech input and process it into particular purposes), language processing module (processing the language data into suitable output for sign language), and signing module (signing in the form of animation, recorded video, or robot movement). The voice recognition module and signing module have been developed in

different languages and Bahasa Malaysia. Otherwise, the language processing unit has been implemented for many languages but not in *Bahasa Malaysia.*

This thesis aims to develop a language processing unit suitable for *Bahasa Malaysia.* The research develops a language processing unit that is fast processing time- and has an excellent success rate[2] for *Bahasa Isyarat Malaysia* using natural language processing and classifier.

1.3 RESEARCH PHILOSOPHY

The sign language (SL) synthesizer is a tool to address the communication gap between HSI to non-HSI people. It should have a robust speech recognition system to understand normal/spoken sentences from non-HSI people. A classifier must adjust grammatical rules in normal/spoken language to grammatical in sign language. It is also required to arrange processing input from the classifier to be understandable by HSI people. Finally, the SL synthesizer should have an animator agent to visualize signs to HSI people.

1.4 RESEARCH OBJECTIVES

This thesis aims to develop an efficient language processing unit for Malaysian Sign Language (MSL). The following steps are considered to achieve the objectives:

a. To evaluate and select appropriate edit distance for the language processing unit.
b. To implement selected edit distance with natural language processing method and hidden Markov model and Bayesian network for the language processing unit.

[2] Processing time and successful rate are key factors for this research. This research is aimed to contribute in real time to the overall sign language synthesizer system and should have good output which is understandable by HSI people.

c. To develop an adaptive system for the language processing unit

1.5 RESEARCH SCOPE

This research's scope is limited to the development of a language processing unit for Malaysia Sign Language. Input to the language processing unit is the sequence of spoken word (text) in the Malay language, and output is the sequence of words suitable to Malaysian Sign Langauge. The proposed system has been developed using MATLAB 2014 and running at Intel i5 5200U processor and 4GB RAM to process the output from speech data.

1.6 RESEARCH METHODOLOGY

The following works have been carried out to achieve the objective of the research:

1. The research starts with a literature review. The study covers the language processing unit in the existing sign language synthesizers.
2. A language processing unit is developed by selecting edit distance, natural language processing, the hidden Markov model, and the Bayesian network. The development of an adaptive method is also done.
3. The proposed language processing unit is evaluated. The performance is assessed in terms of its processing time and success rate.

1.7 THESIS OUTLINE

The thesis is organized as follows. Chapter 1 provides the thesis's introduction, where the problem statement, objective, research scope, and methodology are presented.

Chapter 2 presents the literature review. It provides the review context related to the definition of sign language and sign language synthesizer. The language

processing unit's description is elaborated on and the existing methods used for sign language synthesizer.

Chapter 3 discusses the basic theory for Sign Language Synthesizer. The basic theory discussed is an analysis of the Edit Distance, Hidden Markov Model, and Bayesian Network. At the end of this chapter, the advantages and disadvantages of using these three methods are presented.

Chapter 4 discusses the Proposed Sign Language Synthesizer System. In the proposed system, we discuss the general structure and discuss each part of the proposed system.

Experimental results are presented in Chapter 5. In this chapter, the experimental setup and results are shown. Results obtained from each experiment are discussed.

Finally, chapter 6 provides conclusions about the experimental results. Besides, the contribution is clearly explained, and the last one is future works, which are expected to be the basis for research related to the Sign Language Synthesizer.

CHAPTER TWO

LITERATURE REVIEW

2.1 INTRODUCTION

Many researchers have researched sign language recognizers and sign language synthesizers to reduce the communication gap between HSI and non-HSI people and their quality of life in the community. Sign language recognizer and sign language synthesizer are based on non-HSI people and HSI people, respectively.

Sign language is a language that is spoken by HSI people using hand movement and facial expression. It is available in many countries, and those are different from each other. For example, Malaysian sign language (MSL)[3] has different signs compared to British sign language (BSL) and American sign language (ASL). On the other hand, sign language has a common similarity that only uses important words in a sentence.

From non-HSI people's view, a sign language recognizer is an exact tool to understand HSI people as they speak by sign language. The sign language recognizer converts sign language into text or spoken words. The sign language recognizer captures hand movement and facial expression to get the meaning of the sign. The systems understand the hand movement and match it with the database where the output is given in text or voice.

On the contrary, sign language synthesizer benefits HSI people. In other words, HSI people can understand what non-HSI people speak in verbal language. The sign language synthesizer converts the voice input into the corresponding sign language.

[3] Malaysian Sign Language is based on American Sign Language. However, the language development of MSL has shown specificity in the sign used.

Figure 2.1 shows modules of the sign language synthesizer. The system has three main modules: voice recognition module, language processing unit module, and signing module.

The voice recognition module captures data from the speech input. The second module is the language processing unit, which alters input language from the first module into suitable output language with sign language. The third module is a signing module where sign movement appears in some possible method, such as avatar, recorded video, or hand robot movement.

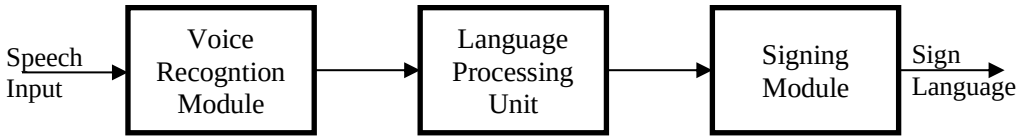

Figure 2.1 The Sign Language Synthesizer Module

The voice recognition module and signing module of the sign language synthesizer have been developed in many languages such as Arabic (Eljawad et al., 2019), English (Ullah & Min, 2016), Thai (Prajongjai et al., 2018), and Malaysian language. However, a language processing unit has not been developed for Malaysian Language (Mokhtar et al., 2017).

This literature review consists of techniques and an existing system for language processing unit for sign language synthesizer. The technique for the language processing unit provides a literature background for the language processing unit using natural language processing. The existing language processing unit works on available sign language synthesizers for six different countries.