Comparison Binary Search and Linear Algorithm for German-Indonesian Sign Language Using Markov Model

Fridy Mandita
Thai-German Graduate School of Engineering (TGGS) - KMUTNB
Bangkok, Thailand
Email: fridymandita@live.com

Toni Anwar
Universiti Teknologi Malaysia
Johor Bahru, Malaysia
Emaill: tonianwar@utm.my

Harnan Malik Abdullah
Telecommunication Engineering
Polytechnic of Malang City
Malang, Indonesia
Email: malik_pmpg@yahoo.com

Abstract—**Research in the area of sign language gives significant results in the decade of years. A lot of research has been done focusing on the signer-independent schemas that contain sign language and video. This paper contributes the statistical translation words from German-Indonesia sign language and vice versa. The process of translation used Markov model and parsing tree to translate words. A binary search and linear algorithms in collaboration with Markov model are used to implement the translation process. The execution time between 2 models of algorithms has been analyzed. A binary search and Markov model have given a better execution than sequential algorithm and Markov model when translates words.**

Keywords—Sign language; Markov model; binary search; linear search.

I. INTRODUCTION

Impaired hearing is the inability of ears to hear a sound at all which could be happened by; (1) complications at birth (2) infectious diseases, and (3) drugs [17]. Sign language is used by deaf or impaired hearing people to communicate each other. Sign language is a language that uses signs via hand gestures or other movements, such as facial expression, body posture, gaze, and lip patterns [16]. It is a self-contained language which has its own characteristics of the structure and grammar. On the other side, sign language is the most natural and expressive way for deaf people to communicate with their fellows. The developments of sign language show significant results within a decade.

Currently, sign language has spread around the world. Almost every countries in the world have their own sign language but not universal, such as American Sign Language (ASL), British Sign Language (BSL), Australian Sign Language (Auslan), German Sign Language (DGS), and Indonesia Sign Language (ISL) [5].

Generally, sign languages do not have a relation with their spoken language where they were developed. In contrast, from the pronunciation of the words, there is no standardization of sign language.

In order to communicate, deaf people can use different signs for the same word as long as they understand it [15]. Each sign language has its own sign to a word and does not relate between them. Different with spoken language that is used acoustic signal to give the words, sign language uses the visual communication to deliver the words [16]. The device of visual communication can be divided into two types: the manual components and the non-manual components.

I-CHAT (I Can Hear and Talk - Indonesian Sign Language Computer Application for the Deaf) is an application that is designed to help people, particularly children with impaired hearing problem to master Indonesian language as their spoken language [8]. A tool called SIGNUM Database was built with the purpose to develop a video corpus based on the basic vocabulary of signs in German language. The tools of SIGNUM Database contain isolated and continuous video copies of signers.

In this paper, translator of words from German Sign Language (DGS) to Indonesia Sign Language (ISL) and vice versa is proposed using Markov model combined with binary search and linear algorithm. The input words or sentences in sign language are inputted by users and the tools will translate and then display it as sign language video corpus.

II. RELATED WORKS

This section gives a short overview related works in sign language and machine translation. In the recent decades, a lot of research has been done in the field of sign language and machine translation. Framework for statistical machine translation based on sign language recognition was proposed by Bauer et al. [3]. The framework is used to translate recognized video based continuous sign language to spoken language and vice versa.

A rule-based is the good model of translation from English sentences to American Sign Language. The model is suitable for translation process but the larger corpora are difficult to be fulfilled when training and translating the words [9].

The adaptation method could be obtained a better result of Sign Language recognition by increasing the number of training signers [15]. As mentioned in [15], the vision based recognition of sign language can be quicker to adapt unknown signers.

In 2010 [1], I-CHAT tool was built for impaired hearing people. This tool was designed to assist the people to master their spoken language which is Indonesia Language. It's prepared with a dictionary module to improve vocabulary in learning sign language and speech readings or pronunciation of some word in Indonesia language.

Nur Afifah et al. has published a paper about electronic dictionary for translating words from Indonesia language to Javanese language. The combination of a binary search algorithm and Marko model were used for the translation schemas [11].

Chris Callison-Burch et al. have implemented binary search algorithm to arrange alphabetic data [4]. The data is stored as an array of integers where array length equal to the length of the corpus. The complexity of the computation is therefore limited by (2 log (n)), where n is the length of the corpus.

The analysis of binary search and linear algorithms was described by Asagba P.O et.al in 2010 [2]. The paper conducted experiments on the execution time of binary search and linear search algorithms.

III. SYSTEM ARCHITECTURE

The statistical machine translation has the function to translate DGS to ISL and vice versa as well as provides video corpus. Consider the block diagram in Figure 1. From the block diagram, it can be seen that the deaf or hear impaired input the data in sentences or phrases.

At this point, the input could be a simple word. The input is analyzed by using parsing methods. This process will read and split the input into string of character. The translation model would be done afterward.

There are basically two models used for the sign language translator: binary search and Markov model, and linear search and Markov model search. Markov model is used for the word pattern recognition by searching the value of each word which is entered by users.

Binary search algoritm works by comparing input value to the middle element of array and compare it with the the record that we look for. As the results the system will show the output of the translation of the words. Linear search algorithm works by checks the first item following second item until find the target that is looked for.

The databases of statistical machine translation are taken from SIGNUM databases and I-CHAT. SIGMUN library provide 450 isolated video and 780 continuous sentences, was performed once by 25 native signers of different sexes and ages. I-CHAT was prepared with a dictionary module to improve vocabulary in learning sign language and speech

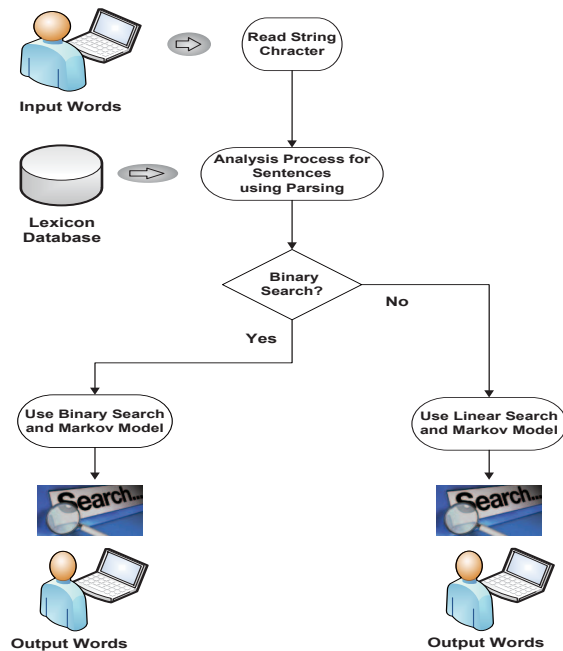readings or pronunciation of some word in Indonesian language.

Fig. 1. German-Indonesia Sign Language Translator Systems Architecture

IV. MATERIALS AND METHODS

A. Binary Search Algorithm

In computer science, binary search or half-interval search algorithm is more efficient than a sequential algorithm [11]. It due to this algorithm based on divide and conquer strategy, which divide array into two lists and search by comparing input value to the middle element of an array. The process of binary search is shown in figure 2.
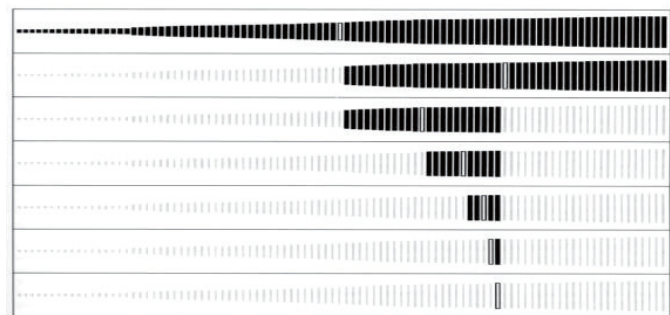

Fig. 2. Binary Search Illustrations (from [6])

Fig. 2 shows how a binary search algorithm works. For checking the value of middle (approximately) item in the list, it can be described as follows. If it is not the target search and the value of the target is smaller than the middle item, the target should be in the first half of the list. If the search of target value is greater than the middle item, the target should be in the last half of the list. By the way of comparison as mentioned before, it manages to reduce the number of items to check out the half [2].

112

The search continues by examining the middle item in the remaining half of the list. If it is not target of the search, the search process narrows to half of the remaining part of the target list could on. The separation process continues until the target is found or the rest of the list consists of only one item. If the last item is not the target, then it is not in the list [6].

B. Linear Search Algorithm

Linear search algorithm is a simple method to find the list of elements. It check the first item, second item, third item, and the next item until the item that is being sought has been found. The algorithm works by analyzing the each key of the data node. The search process is carried out by comparing each element in a data node one by one, starting from the first element to the next element until the sought element is found or all the elements have been examined. It doesn't able to go to the previous process, except the search process starts again from the beginning.

The nature of the data processing is first in first out, which is data reading from this file should be started at the earliest data. The nature of the data processing is first in first out, which is data reading from this file, should be started at the earliest data [13, 20]. The process of linear algorithm can be shown bellows:

Fig. 3 Linear search illustrations (from [9])

C. Markov Model

Markov model is stochastic model processes which have the Markov Property. This model has been introduced by Andrew A. Markov a mathematician from Russian (Alexander Volfovsky, 2007). The property describes the condition of given state where the state depends on the previous states. Markov property has form:

$$P(Xn = x|X_o,...,X_{n-1}) = P(X_n = x|X_{n-1}) \ V_n V_x. \qquad (1)$$

Markov chain is a structure that consists of stationary entities called state (Eric, 1998). The next state depends on the current state and it does not depend on the sequence of event that preceded it. If we have a *set of states*, $Y = \{y_1, y_2, ..., y_r\}$. The next output (Y_t) can be defined by:

$$Y_t = f \{y_{t-1}, y_{t-2}, ..., y_r\} \qquad (2)$$

Add the noise to make stochastic the formula is described by:

$$_P(Y_t \mid y_{t-1}, y_{t-2}, ..., y_r) \qquad (3)$$

If the chain is currently in the state y_i then moves into the state y_j in the next step with a probability can be denoted by p_{ij} and this probability does not depend which states in the chain were in before the current state. The probabilities of p_{ij} are called transition probabilities. The process can remain in the state, and this occurs with value probability p_{ii}.

An initial probability distribution, which is defined on *Y* specify the initial state. Usually this is done by specifying a particular state as the initial state. Markov chains example can be shown as follows:
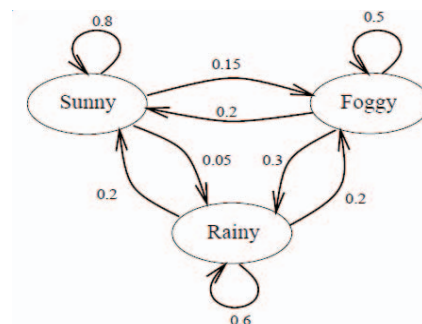

Fig. 4 Binary Search Illustrations (from [7])

V. RESULTS

In order to translate words from German Sign Language (DGS) to Indonesian Sign Language (ISL) and vice versa; a test case was designed to test the translation of words from DGS to ISL. The test case is divided in 2 models of sign language translator; 1) German-Indonesian Sign Language, 2) Indonesian-German Sign Language.

The execution time for the translation of sign language words or sentences were also been analyzed to compare the experimental results of the used method. As mentioned earlier, 2 different methods: 1) binary search algorithm and Markov mode, 2) linear search and Markov model would be implemented.

The test cases are designed to translate German Sign Language to Indonesian Sign Language and vice versa instead of word or sentences as much as five times from each examples which has been prepared.

The images shows a comparison of the time when translate German-Indonesia Sign Language and vice versa; one of the methods will provide a better performance when translated word and other methods will give a better performance when translated sentences. The test case is prepared for the simple sentences which consist of 3 words.

A. German-Indonesia Sign Language Translator

TABLE I. EXECUTION TIME GERMAN-INDONESIA TRANSLATION USING BINARY SEARCH AND MARKOV MODEL

No	Example of word	Testing (ms)					Mean
		1	2	3	4	5	
1	ich	14	50	36	22	46	33.6
2	halle	51	4	51	49	56	42.2
3	lieber	54	12	46	41	33	37.2

Table I shows the implementation of German-Indonesia sign language translation using binary search and Markov

model. The test model was prepared with 3 different example of words that are '*ich', 'halle', and 'lieber'.* The best execution time is 33.6ms when translating the word *'ich'*. The worst execution time is 42.2ms occured when it was used to translate the word *'halle'*. Another, the translation word *'lieber'* need 37.2ms execution time in average.

TABLE II. EXECUTION TIME GERMAN-INDONESIA TRANSLATION USING LINEAR SEARCH AND MARKOV MODEL

No	Example of word	Testing (ms)					Mean
		1	2	3	4	5	
1	Ich	31	42	28	37	32	34
2	Halle	54	35	31	32	40	38.4
3	Lieber	52	35	31	35	54	41.4

The results of the implementation of German-Indonesia sign language translation using linear search and Markov model is shown in Table II.The sample word were same with being used by binary search and Markov model test. The best execution time is 34ms for the word *'ich'* and the worst execution time was 41.4.6ms when translated the word *'lieber'*. The word *of 'halle'* gave results 80.8ms.

B.   Indonesian-German Sign Language Translator

TABLE III. EXECUTION TIME INDONESIA-GERMAN TRANSLATION USING BINARY SEARCH AND MARKOV MODEL

No	Example of word	Testing					Mean
		1	2	3	4	5	
1	Saya	59	34	36	6	48	36.6
2	Kamu	50	27	33	29	44	36.6
3	Saying	18	39	30	25	31	28.6

Table III shows the result of execution time for Indonesia-German sign language translation using binary search and Markov model. This test model provides 3 different words such as *'saya, kamu, and sayang'* is used to test the performance of the binary search and Markov model. For the best time execution is 28.6ms when translate the word *'sayang'*. The worst of time execution is 36.6ms where used to implement translation of the word *'saya'* and *'kamu'*.

TABLE IV. EXECUTION TIME INDONESIA-GERMAN TRANSLATION USING BINARY SEARCH AND MARKOV MODEL

No	Example of word	Testing					Mean
		1	2	3	4	5	
1	Saya	52	34	58	35	33	42.4
2	Kamu	21	48	52	25	38	36.8
3	Saying	45	15	53	52	2	33.4

The execution time result of Indonesia-German sign language translation using linear search and Markov model is shown in Table IV that used the same test word. The word *'sayang'* has the best time execution that is 28.6ms. The worst

of time execution is 42.4ms where used to implement translation of the word *'saya'*. For other execution time is 36.8ms when translated the word *'kamu'*.
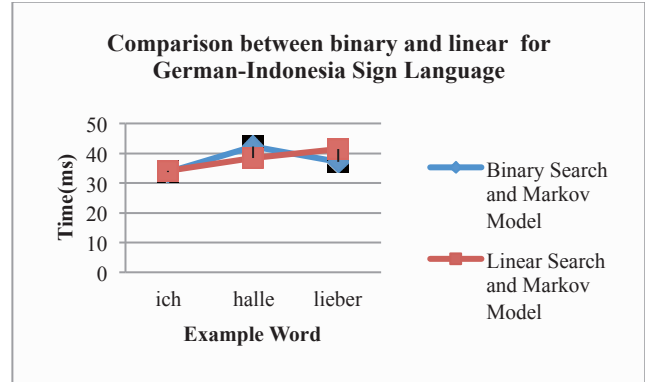
C.   Comparison For German-Indonesia Translator

Fig. 5. Comparison between binary and linear for German-Indonesia sign language translation time

The results of execution time for translate German-Indonesia sign language is depicted in Figure 5. The combination of Markov model and Binary Search gave the faster sign language translation rather than Linear Search and Markov Model. The time execution is 33.6ms when translate the word *'ich'* using Binary Search where it would need 34ms if Linear Search was used. On the other hand, the slowest time performance is 42.2ms when translating the word *'halle'* by using Binary Search.

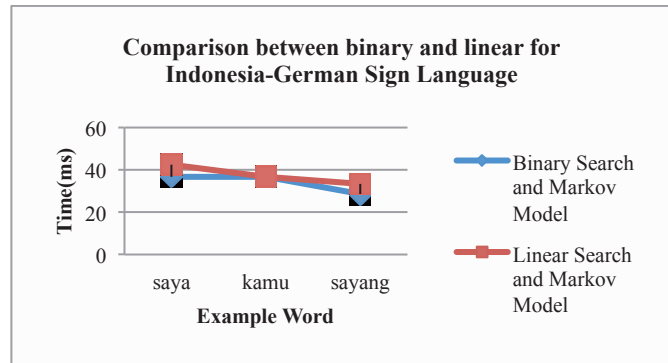*D.   Comparison For Indonesia-German Translator*

Fig. 6. Comparison between binary and linear for Indonesia-German sign language translation time

The comparison execution time for translate Indonesia-German sign language word can be seen in Figure 6. Implementation of Binary Search and Markov Model has given the fastest translation of sign language rather than linear search. It give the better performance than the linear search one as depicted in the figure.

The best execution time is 28.6ms when using binary search and Markov model for the word *'sayang'* where Linear Search and Markov Model need 33.4ms for the same word. In

114

contrast, the slowest time performance is 36.6ms when translating word *'saya' and 'kamu'.* The slowest execution time for linear search and Markov model is 42.4ms when used to translate the word *'saya'*.

VI.   CONCLUSIONS

In this paper, the test models and analysis of translating words from German Sign Language (DGS) to Indonesian Sign Language (ISL) and vice versa has been carried out. The test models were provided for 3 different words of sign language. From the overall test results, it can be summarized that combination of Markov Model and Binary Search give the better performance to translate the words of sign language to be compared with Linear Search and Markov Model. It's also proven by particular result for each test word.

REFERENCES

[1].   Andreas W. Yanuardi, Samudra Prasetio, and Johannes Adi Purnama P.: Indonesian Sign Language Computer Application for The Deaf.s *2010 nd International Conferece on Education Technology and Computer (ICETC)*, 2010.

[2].   Asagba P. O, Osaghae E. O. and Ogheneovo, E. E.: IS BINARY SEARCH TECHNIQUE FASTER THAN LINEAR SEARCH. TECHNIQUE?. *Scientia Africana, Vol. 9 (No. 2) Pp 83-92 Faculty of Science, University of Port Harcourt, Printed in Nigeria, December 2010.*

[3].   B. Bauer, S. Nießen, and H. Hienz.: Towards an automaticas Sign Language translation system. In *Proc. of the Int. Workshop on Physicality and Tangibility in Interaction*, Siena, Italy, 1999.

[4].   Chris Callison-Burch, Colin Bannard, and Josh Schroeder: Scaling Phrase-Based Statistical Machine Translation to Larger Corpora and Longer Phrases, 2004.

[5].   D. Rybach: Appearance-Based Features for Automatic Continuous Sign Language Recognition. Diplomarbeit im Fach Informatik Rheinisch-Westf˜alische Technische Hochschule Aachen, 2006.

[6].   Dewi Martina Andayani: Pembuatan Kamus Elektronik Kata-Kata Bahasa Indonesia-Jawa Menggunakan Metode Binary Search Berbasis Perangkat Lunak. Tugas Akhir D-4, PENS-ITS, Surabaya, 2009.

[7].   Eric F.L.: Markov Models and Hidden Markov Models A Brief Tutorial, INTERNATIONAL COMPUTER SCIENCE INSTITUTE, December 1998.

[8].   "I-CHAT". Retrieved December 20, 2013, http://app.i-chat.web.id/

[9].   M. Huenerfauth: A multi-path architecture for machine translation of English text into American Sign Language animation. In *Proc. Student Workshop at Human Language Technologies conference HLT-NAACL*, Boston, MA, USA, 2004.

[10]. "Linear   Search".   Retrieved   19   April   2014, http://dopey.cs.vt.edu/courses/cs1706/slides/searching.html

[11]. Nur Afifah: "Pembuatan Kamus Elektronik Kalimat Bahasa Indonesia dan Bahasa Jawa Menggunakan Markov ", Tugas Akhir D-4, PENS-ITS, Surabaya, 2012.

[12].   "Princeton". Retrieved December 20, 2013, https://www.princeton.edu/~achaney/tmve/wiki100k/docs/Binary_search_algorithm.html

[13]. Sapinder, Ritu, amd Arvinder Singh: Analysis of Linear and Binary Search Algorithms. International Journal of Computers & Distributed Systems. Volume 1, No.1, Jun 2012

[14]. "SIGNUM   Database".   Retrieved   December   20,   2013, http://www.phonetik.uni-muenchen.de/forschung/Bas/SIGNUM/

[15]. U. von Agris, D. Schneider, J. Zieren, and K.-F. Kraiss: Rapid signer adaptation for isolated sign language recognition. *24th IEEE International Conference on Computer Vision and Pattern Recognition*, June 2006.

[16]. U. von Agris, J. Zieren, U. Canzler, B. Bauer, and K.-F. Kraiss: Recent developments in visual sign language recognition. *In Universal Access in the Information Society, Special Issue Emerging Technologies for Deaf Accessibility in the Information Society*, 6(4):323{362, February 2008.

[17].   "WHO".   Retrieved   27   December,   2013   from http://www.who.int/topics/deafness/en/

[18]. W. Stokoe, D. Casterline, C. Croneberg: A Dictionary of American Sign Language on Linguistic Principles, Gallaudet College Press, Washington D.C., USA, 1965.

[19]. Yudi Wibisono: Penggunaan Hidden Markov Model untuk Kompresi Kalimat, Skripsi-S1, ITB, Bandung, 2008.

[20]. Yunia Dinarsih: "Pembuatan Kamus Digital Jawa-Indonesia Menggunakan Metode Sequential Searching ", Tugas Akhir D-3, PENS-ITS, Surabaya, 2009.