Real Time SIBI Sign Language Recognition based on K-Nearest Neighbor

Fitrah Maharani Humaira
Electrical Engineering
Politeknik Negeri Madura
Sampang, Indonesia
humaira1124@yahoo.com

Supria
Informatics Engineering
Politeknik Negeri Bengkalis
Riau, Indonesia
phiya1287@gmail.com

Darlis Herumurti
Department of Informatics
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
darlis@if.its.ac.id

Kukuh Widarsono
Electrical Engineering
Politeknik Negeri Madura
Sampang, Indonesia
kukuh@poltera.ac.id

***Abstract*— Persons with disabilities also have the right to communicate between each other, both with normal people and people with other disabilities. People with disabilities will be difficult to communicate with other people. They use 'sign language' to communicate. That's why other normal people will be difficult to communicate with them. Because there are not many normal people that can understand the 'sign language'. The system can help to communicate with disabilities people are needed. In this paper, we proposed sign language recognition for Sistem Isyarat Bahasa Indonesia (SIBI) using leap motion based on K-Nearest Neighbor. Technology of leap motion controller will generate the existence of coordinate points on each bone in hand. As an input, we used the value of distance between the coordinates of each bone distal to the position of the palm, which were measured using Euclidean Distance. This feature of distance will be used for training and testing data on K-Nearest Neighbor method. The experiment result shows that the best accuracy is 0,78 and error 0,22 with proposed parameter of K = 5.**

Keywords*— *dissability, communication, sign language recognition, leap motion controller, K-nearest neighbor

I. Introduction

People with special needs or disabilities are also part of society, where they have the same right to interact and socialize with the surrounding environment. Persons with disabilities, such as deaf and speech impaired sometimes looks like a normal person. However, problems arise when communication with others, the deaf cannot hear, whereas speech impaired cannot be answered conversation. Classically, this problem can be answered, in which the deaf using hearing aids. While the speech impaired using sign language, through hand gestures or body movements. While there are many different types of gestures, the most structured sets belong to the sign languages. In sign language, each gesture already has assigned meaning, and strong rules of context and grammar may be applied to make recognition tractable.

However, the constraints, if normal people who talk understands sign language, certainly not all. Therefore, this issue would have to find an alternative solution, need a system that can overcome these problem. In previous research, the technology used is Kinect camera or webcam-based image for sign language recognition. But has built research-based sign language recognition leap motion controller such as American Sign Language (ASL) [1], the introduction of sign language Arabic Sign Language (ArSL) [2], [3]. In this paper, we proposed an identification of hand gesture alphabet using leap motion controller (LMC). Hopefully the system is able to overcome the weakness of communication with deaf and speech impaired and easier for normal people to communicate with them.

The study proposed a sign language recognition system using leap motion controller and K-Nearest Neighbor (KNN). By using the leap motion controller, every coordinate point bone in the hand will be detected, so that the coordinate points can be used as input feature. The use of distance feature between palm position with type distal measured by Euclidean distance. These features will be used for training and testing data for the k-nearest neighbor classification method. Training data used are 10 samples for each letter, the number of 26 letters of the alphabet, so that the total amount of training data 260 for one person.

Section II of this paper describes the related works. Section III describes the proposed algorithm used for SIBI recognition. The experimental results and conclusions are given in sections IV and V respectively.

II. Related Works

A. Sign Language Alphabet

Sign language alphabet or finger spelled alphabet is the process of spelling out words by using signs that correspond to the letters of the word. Sign language is a technique to interpret the writing alphabet of writing a signal in the form of motion of the fingers, hand shape, orientation and hand gestures. There is a sign language used by the deaf and speech impaired. There has been no successful international signaling alphabet is applied, because in each country have different sign language. Even in Indonesia, there are two types of alphabet sign language, namely: Bisindo (Bahasa Isyarat Indonesia) and SIBI (Sistem Isyarat Bahasa Indonesia) [3].

American Sign Language (ASL) is the dominant sign language used by deaf communities in the United States and most of Canada. Figure 1 is a American Sign Language Alphabet.

B. Leap Motion Controller (LMC)

Leap Motion Controller (LMC) is a new technology developed by the Leap Motion Company. This USB-based technology is a sensor capable of detecting the movement of the hands and fingers, as well as small objects like a pen. This tool is made to replace the mouse and keyboard to control or play computer using motion [1]. Figure 2 shows the shape leap motion, Figure 3 is a part-section of the leap motion.

The device is claimed to be even more precise 100x than Kinect, where accuracy in detecting each fingertip position is about 0.01 mm, with a frame rate of up to 200-300 fps. And be able to track all interactions finger of 10 users simultaneously [2]. LMC uses two high-precision infrared camera sensor and 3 LEDs to capture information in the hands of the active range. LMC create 3-dimensional space with the three axes, namely X, Y, and Z as shown in Figure 4 that can extend the distance range from about 1 inch to 2 meters.

Hand gesture detection is the first step in the identification of gesture recognition as the value of the alphabet. This detection task is to extract lines hands formed a whole region of the captured image. LMC is able to capture the image of a hand that can automatically separates the background, then we can ignore the morphological operations and hand detection can be done easily.

Color segmentation generates binary form. Using a binary image with a black area shows the region of the hand and the white area indicates the region of template. Here we consider the image of the hand will always be placed in the middle of a fixed-size frame, so that long-distance image capture of close hand will have no effect. Each image of the alphabet will be trained by 10 different hands image frame based on the long finger; the width of a finger; the position of the x, y, z axis; palms, and so on.

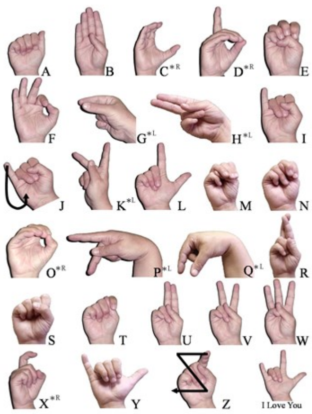

Fig.1. American Sign Language Alphabet

Fig. 2. Leap Motion Controller

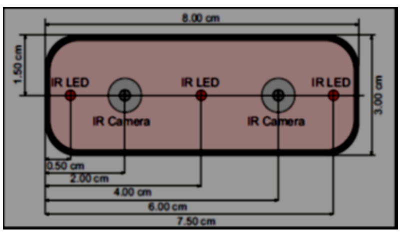

Fig. 3. Part-section of leap motion controller

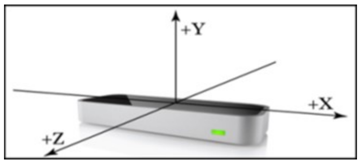

Fig. 4. 3-dimensional space of LMC

C. K-Nearest Neighbor

Methods K-Nearest Neighbor (K-NN) becoming one of the oldest and popular methods based NN. K values used in here stated number of nearest neighbors were involved in the determination of the class label prediction on the test data. K nearest neighbors were voted then do voting class of the K nearest neighbors. Class with the highest number votes of neighbor given as class labels prediction results on the test data [4].

There are several key issues that affect the performance of KNN, including the vote of the value of K [5]. If K is too small, have consequences of bias prediction result which sensitive to the presence of noise. On the other hand, if K is too large then the selected nearest neighbor might be too much of another class that was not relevant because the distance is too far. To estimate the value of K at its best, can be done by using cross validation. It is important to affirm that k = 1 may perform other K values, in particular for small data sets, including typically when used for research or practice in the classroom. However, if given sufficient samples, then the larger value of K will be more resistant to noise.

Another issue is the approach to combine classes. The simplest method is by majority vote as described above. However, it can be a problem if the nearest neighbor have varying distances, while neighbor closer more reliable shows the class label. More recent approach, which is usually much less sensitive to the choice of K, which is weighted by a nearest neighbor distance. Varying options that could be used e.g. weighting factor that is often taken from the inverse square of the distance, as in:

$$W_i = \frac{1}{d(x_i,\ z)^2} \tag{1}$$

Wi is weights for the data xi selected as nearest neighbors, whereas d(xi, z) is the distance (dissimilarity) between data xi of the test data z.

Another important issue on KNN weightlessness is the value of K is used, whether even or odd. For K odd with an even number of classes will facilitate voting because it is guaranteed there will be no two classes received the same voting power. However, if even, there will be the possibility of two classes received the same voting power. For the case when there are two or more classes with the same voting and the majority then the class label can be taken between classes earlier.

The algorithm to calculate the predicted outcome class voting or weighted class voting y' as in:

$$y' = \arg\max_{v \in C} \sum_{y_i \in D_z} w_i \; x \; I(v = y_i) \qquad (2)$$

III. Research Method

In this study, there are several stages in the process of introduction sign language using leap motion controller. The stages are collect distal phalanges coordinate data of each finger and palm position, measure the distance between each distal phalanges on each finger with the palm position and classify testing data towards training data using K-nearest Neighbor method. The general description of the proposed research methods can be seen in Figure 5.

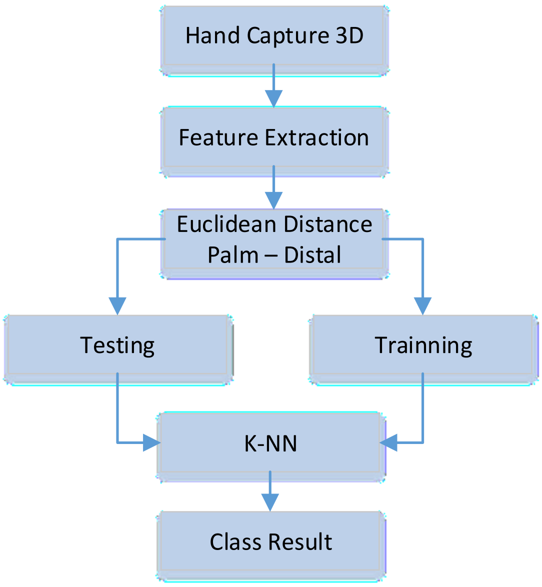

Fig. 5. SIBI recognition diagram

A. Preprocessing

Leap motion controller is used to capture hands to get the coordinate points on the hands. The coordinate points generated by a leap motion controller is palm position, the distal phalanges, intermediate phalanges, proximal phalanges, and metacarpals. In this study only used the coordinates of palm position and distal phalanges in each finger. The distance between the palm position and the distal phalanges measured by Euclidean distance, so that the distance is used as the input feature in this study.

B. Distance

To measure the distance d between the palm position and the distal phalanges used Euclidean distance, where measuring the distance from palm position with distal phalanges on every finger, namely the thumb, index finger, middle finger, ring finger and little finger. As for distance measurement using Euclidean distance can be seen in equation 3.

$$d(xyz) = \sqrt{(x_2 - x_1)^2 + (y_2 - xy_1)^2 + (z_2 - z_1)^2} \qquad (3)$$

C. Classification

K-nearest neighbor is the method used to predict the class of testing data which input based on class of training data that were located closest to the class. Learning data projected onto many-dimensional space, where each dimension represents the features of the data. The space is divided into sections based on the learning data classification. A point in this space marked class c if class C is the most common classification k nearest neighbors on that point. Near or far neighbors are usually calculated based on Euclidean distance.

In the learning phase, the algorithm is simply to store the vectors of features and classification of learning data. In the classification phase, the same features that are calculated to test data (which classification is not known). The distance of this new vector of all learning data vector is calculated, and taken the number of closest k. New classification point predicted included in the classification most of these points.

The best value of K for this algorithm depends on the data; n general, a high k value will reduce the effect of noise on the classification, but makes the boundaries between each classification becomes more blurred. The great value of K can be selected with parameter optimization, for example by using cross-validation.

K-NN algorithm accuracy is greatly influenced by the presence or absence of features that are not relevant, or if the weight of such features is not equivalent to its relevance to the classification.

IV. Experiment and Result

This study focused on the introduction of 26 alphabet on Sistem Isyarat Bahasa Indonesia (SIBI). The data taken using leap motion controller. It is taken from the distal coordinate data on each finger and the data inputted is 15 frames hand training data for each letter of the alphabet so that a total of 390 frames as training data. Then for testing the data needed one hand frame. For an introduction to the alphabet used KNN (k-nearest neighbor) method. The feature used in here is a distance feature.

To measure the success of the proposed system is then measured by calculating the accuracy and error rate. For accuracy by calculating the correctly predicted total data to the total amount of data that is predictable. To calculate the accuracy of the results can be seen in equation 4. As for calculating the error rate (prediction error) is shown in equation 5.

$$Accuration = \frac{\sum true\ data\, prediction}{\sum\ all\ data\ prediction}\ x\ 100\% \qquad (4)$$

$$Error\ rate = \frac{\sum false\ preddiction\ data}{\sum all\ data\, prediction}\ x\ 100\% \qquad (5)$$

The Results of SIBI sign language recognition accuracy by using leap motion controller and K-NN performed on the 26-letter alphabet. Testing is done in real-time towards training data. In this paper is determined 3 parameter value of K nearest neighbor namely 5, 10 and 15. The accuracy of the test results can be seen in Table 1, 2 and 3.

The results of SIBI sign language recognition by using leap motion controller based on KNN method has a very high accuracy. Still in a few letters have low accuracy due to leap motion controller cannot detect the coordinates when the area coordinate point is protected or covered by the toe or the other, and when the letters that have similar data is very high, the system is difficult to distinguish. For several letters like the letter NR, C, F, L, O, P, W and Y have a very high accuracy, it is because all the coordinates of the finger not covered by sections other fingers, and also does not much similar with another letter.

The test results indicate that the parameter value K = 5 has the highest accuracy compared with the K = 10 and K = 15. It shows that the parameter value K = 5 has the best average accuracy is 0,78 and error rate is 0,22.

V. CONCLUSION

The study proposed a sign language recognition system using leap motion controller based on K-Nearest Neighbor. This system allows normal people to understand sign language by hand of the deaf and speech impaired. LMC as the latest infrared sensor technology that can detect the coordinate points on the hands and can facilitate the developer in terms of hand recognition. The results of SIBI sign language recognition by using leap motion controller based on KNN method has a very high accuracy. In the future, this system is suitable to applied by people with disabilities, especially the deaf and speech impaired to communicate not only the alphabet, but has been able to detect word by word.

REFERENCES

[1] C. Chuan, E. Regina, C. Guardino, A. L. M. Controller, and I. Apis, "American Sign Language Recognition Using Leap Motion Sensor," pp. 541–544, 2014.

[2] M. Mohandes, S. Aliyu, and M. Deriche, "Arabic Sign Language Recognition using the Leap Motion Controller," 2014.

[3] A. S. Elons, M. Ahmed, H. Shedid, and M. F. Tolba, "Arabic Sign Language Recognition Using Leap Motion Sensor," pp. 368–373, 2014.

[4] S. Tan, "Neighbor-weighted K-nearest neighbor for unbalanced text corpus," Expert Syst. Appl., vol. 28, no. 4, pp. 667–671, 2005.

[5] X. Wu, V. Kumar, Q. J. Ross, J. Ghosh, Q. Yang, H. Motoda, G. J. McLachlan, A. Ng, B. Liu, P. S. Yu, Z.-H. Zhou, M. Steinbach, D. J. Hand, and D. Steinberg, Top 10 algorithms in data mining, vol. 14, no. 1. 2008.

TABLE 1. Accuracy Result with K = 5

Recognition	**Accuracy (%)**	**Error (%)**	**Recognition**	**Accuracy (%)**	**Error (%)**
NR	1,0	0			
A	0,6	G(0,2), O(0,2)	N	0,4	S(0,2), T(0,4)
B	0,8	W(0,2)	O	1,0	0
C	1,0	0	P	1,0	0
D	0,8	X(0,2)	Q	0,8	L(0,2)
E	0,4	G(0,2), M(0,2), T(0,2)	R	0,8	U(0,2)
F	1,0	0	S	0,4	E(0,2), T(0,4)
G	0,6	Q(0,4)	T	0.4	A(0,2), N(0,2), S(0,2)
H	0,4	K(0,2), R(0,4)	U	0.4	H(0,2), K(0,2), R(0,2)
I	0,8	Z(0,2)	V	0,6	K(0,2), U(0,2)
J	0,8	I(0,2)	W	1,0	0
K	0,6	W(0,4)	X	0,6	D(0,2), Z(0,2)
L	1,0	0	Y	1,0	0
M	0,6	E(0,1), S(0,2)	Z	0,8	D(0,2)
			Average Accuracy		**0.78**
			Average Error		**0.22**

TABLE 2. Accuracy Result with K = 10

Recognition	**Accuracy (%)**	**Error (%)**	**Recognition**	**Accuracy (%)**	**Error (%)**
NR	1,0	0			
A	0.6	G(0,2), O(0,2)	N	0,5	A(0,1), S(0,2), T(0,2)
B	0,5	NR(0,1), W(0,4)	O	0,6	A(0,1), E(0,3)
C	0,9	O(0,1)	P	0,6	K(0,4)
D	0,9	P(0,1)	Q	0,6	G(0,2), L(0,2)
E	0,6	M(0,2), T(0,2)	R	0,4	H(0,3), K(0,1), U(0,2)
F	1,0	0	S	0,4	A(0,3),E(0,1), N(0,1), T(0,1)
G	0,6	L(0,1), Q(0,2), X(0,1)	T	0,3	A(0,2), M(0,2), N(0,2), S(0,1)
H	0,3	K(0,2), R(0,3), U(0,2)	U	0,6	H(0,3), K(0,1)
I	0,4	J(0,2), Y(0,4)	V	0,7	K(0,2), U(0,1)
J	0,5	I(0,4), Y(0,1)	W	0,9	B(0,1)
K	0,6	P(0,4)	X	0,7	P(0,1), Z(0,2)
L	0,6	Q(0,4)	Y	1,0	0
M	0,3	E(0,1), N(0,1), S(0,2), T(0,3)	Z	0,8	D(0,2)

Average Accuracy	**0.63**
Average Error	**0.37**

TABLE 3

Accuracy Result with K = 15

Recognition	**Accuracy (%)**	**Error (%)**	**Recognition**	**Accuracy (%)**	**Error (%)**
NR	1,0	0			
A	0,6	E(0,07), G(0,07), M(0,07), N(0,2)	N	0,27	A(0,07), E(0,2), M(0,13), S(0,2), T(0,13)
B	1,0	0	O	0,47	C(0,13), E(0,2), G(0,07), N(0,13
C	0,7	N(0.13)	P	0,4	H(0,2), K(0,27), V(0,07), Z(0,07)
D	0,73	P(0,07), X(0,2)	Q	0,53	G(0,27), L(0,2)
E	0,53	M(0,27), S(0,07), T(0,13)	R	0,33	H(0.13), K(0,07), U(0,13), V(0,33)
F	0,93	NR(0,07)	S	0,4	A(0,2), E(0,07), M(0,07), O(0,07), T(0,2)
G	0,73	L(0,07), Q(0,2)	T	0,33	E(0,27), M(0,27), N(0,07), S(0,07)
H	0,33	K(0,2), R(0,27), U(0,2)	U	0,4	H(0,27), K(0,2), R(0,13)
I	0,6	J(0,4)	V	0,4	H(0,13), K(0,27), R(0,13), U(0,07)
J	0,47	I(0,27), Y(0,27)	W	0,67	NR(0,07), B(0,27)
K	0,47	H(0,07), U(0,07), V(0,4)	X	0,53	Z(0,33), D(0,07), P(0,07)
L	0,67	Q(0,33)	Y	0,93	J(0,07)
M	0,4	E(0,27), N(0,07), S(0,13), T(0,13)	Z	0,53	D(0,33), X(0,13)
				Average Accuracy	**0.59**
				Average Error	**0.41**