Tone Classification Matches Kodàly Handsign with the K-Nearest Neighbor Method at Leap Motion Controller

Muhammad Croassacipto[1], Muhammad Ichwan[1], Dina Budhi Utami[1*]

[1]*Institut Teknologi Nasional*
Bandung, Indonesia

*muh.croassacipto@gmail.com

Abstract

Hands can produce a variety of poses in which each pose can have a meaning or purpose that can be used as a form of communication determined according to a general agreement or who communicate. Hand pose can be used as human interaction with the computer is faster, intuitive, and in line with the natural function of the human body called Handsign. One of them is Kodàly Handsign, made by a Hungarian composer named Zoltán Kodály, which is a concept in music education in Hungary. This hand sign is used in interactive angklung performances in determining the tone that will be played by the K-Nearest Neighbor (KNN) algorithm classification process based on hand poses. This classification process is performed on the extracted data from Leap Motion Controller, which takes Pitch, Roll, and Yaw values based on basic aircraft principle. The results of the research were conducted five times with the value of k periodically 1,3,5,7,9 with test data consisting pose of 874 Do', 702 Si, 913 La, 612 Sol, 661 Fa, 526 Mi, 891 Re, and 1004 Do punctuation on 21099 training data. The test results can recognize hand poses with the optimal k value k=1 with an accuracy level of 94.87%.

Keywords: Classification, K-Nearest Neighbor, Kodàly Handsign, Leap Motion Controller

I. INTRODUCTION

Angklung is one of the traditional musical instruments from the West Java region that is in great demand from various ages. This Angklung can be played individually or in a crowd directly without training first or can be called the Interactive Angklung performance. Because it is played in tandem, it is necessary to use the method used by the conductor. The method makes it easier to lead the Interactive Angklung performance in telling the tone that will be played to the angklung player, namely with a score or with a hand sign (hand pose). Generally, it uses a hand sign method because this method is easy to understand. Where the conductor will move the hand, which means a tone, then the angklung player follows the hand sign of the conductor. The set of tones will produce a song. So this method is applied to determine the tone in the Interactive Angklung performance.

Using some special equipment (one of them is Leap Motion Controller), hand poses can be used as human interaction with computers (Human-Computer Interaction) to play certain music. Hands can produce a variety of poses where each pose can have a meaning or purpose that can be used as a form of communication according to agreement. One method of communication using hands is Kodály Handsign [1].

CROASSACIPTO ET AL.:
TONE CLASSIFICATION MATCHES KODÀLY HANDSIGN WITH THE K-NEAREST NEIGHBOR METHOD AT LEAP MOTION
CONTROLLER

41

There are various examples of technological innovation in human and computer interactions such as gesture-tracking from Leap Motion Controller and Depth-sensing from Microsoft Kinect. This technology provides opportunities for users to interact with computers using hand or body movements and even hand poses. In addition to these technologies, research on the use of hands in human and computer interaction and research on the introduction of hand poses have been widely applied in several scientific journals, including research on the introduction of hand poses using the HuMoment method [2]. In addition, there is research on the introduction of sign language, namely the introduction of American Sign Language (ASL) based on the estimated value of hand pose from sensor depth in real-time [3], and there is also research on the use of hands as a natural interaction with the Augmented-Reality Interface [4].

In an interactive angklung, it is played in several people, and if there are obstacles, it can be replaced by a machine where the machine can be controlled by one person. So to make it easier to play the angklung, it is necessary to determine the hand sign tone from the visualization of the tone with the Leap Motion Controller. It is used to get the hand pose data carried by the conductor, followed by the classification process using the K-Nearest Neighbor (KNN) method to classify the tones with hand pose, according to Kodály Handsign.

This research is focused on building hand pose recognition systems based on gesture-tracking obtained from the Leap Motion Controller. The difference with previous research is the hand movement recognition system based on the overall skeleton position of the body rather than on the skeleton position of the hand. In this study, the hand pose recognition system uses the K-Nearest Neighbor (KNN) algorithm. KNN is done by looking for groups of objects in the training data that are closest (similar) to objects in the new data or test data. The use of KNN has been used in various studies in the area of object recognition, one of which is the use of KNN in the development of the introduction of Arabic Sign Language [5].

II.   RELATED WORK & LITERATURE

The input to the finger pose recognition system is the gesture-tracking obtained from the skeleton-tracking on the Leap Motion Controller sensor, which is in each position in accordance with Kodály Handsign. From the gesture-tracking data obtained, the value of Pitch, Roll, and Yaw then classified with the K-Nearest Neighbor algorithm, where there are two process streams, namely the training process and the introduction process shown at Fig.1.
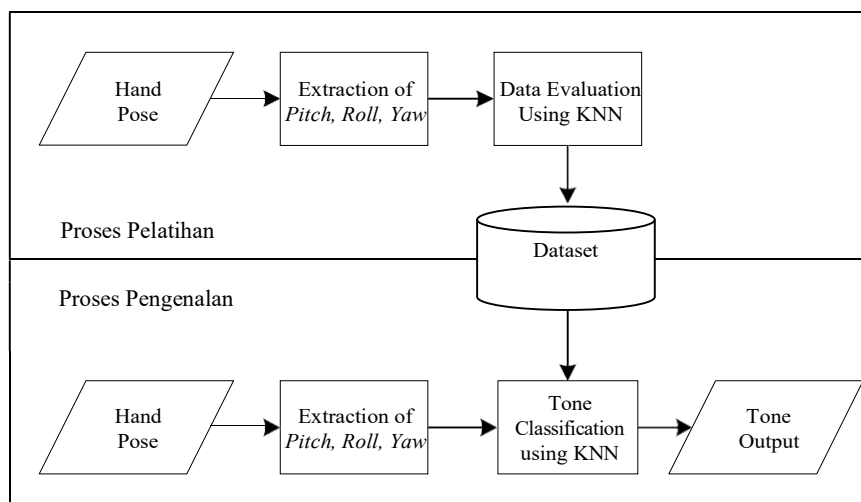

Fig. 1. Block Diagram

The training process is a process of collecting data on samples of hand poses captured with Leap Motion Controller, then extracting the value of Pitch, Roll, Yaw. This training data is then evaluated using the KNN algorithm to validate the training data if the evaluation results state the data is valid then the data is stored as a reference for the recognition process.

The introduction process begins with taking test data. The process of taking test data on the hand pose is captured with Leap Motion Controller, then extracted the value of Pitch, Roll, Yaw. Using the KNN algorithm, the data is classified into tones based on the training data that was stored previously. The result of this process is in the form of a tone that has a value that is closed to the training data.

A. Kodály Handsign

Kodály Hand-sign made by a Hungarian composer named Zoltán Kodály is a concept in music education in Hungary. This method was adopted from Curwen's Solfege hand signs which were proposed by John Curwen [6].

B. Leap Motion Controller

The Leap Motion Company developed a device called the Leap Motion Controller that detects hands and fingers according to their position and movement in real-time [7]. Extraction of data to be classified from Leap Motion Controller [8], which is the visualization of detecting the right hand [9] and get data for tone determination consisting of Pitch, Roll, Yaw [10] obtained from the hand orientation tracking based on the normal palm.

Pitch, Roll, Yaw values are taken based on the basic motion principle of the plane. Pitch is a movement up and down the nose of a plane that moves rotation on the x-axis. Roll is a rolling motion that moves rotation on the z-axis while Yaw is a movement to the right or to the left of the nose of the plane that moves rotation on the y-axis. The same thing with the hand, the principle of aircraft movement can be applied to see the direction and slope of the hand.

C. K-Nearest Neighbor Classification

K-Nearest Neighbor (KNN) is an object classification method based on training data that has the closest distance to the object [11]. It has K values. With the k values, it can be predicted the value of the accuracy of the test data classification with training data based on the object.

Near or far distance to neighbors can be computed based on Euclidean Distance presented in equation 1 as follows.

$$D(a,b) = \sum_{k=1}^{d} (a_k - b_k)^2 \qquad (1)$$

where the matrix D (a, b) denotes the scalar distance (Euclidean Distance) of both vectors a and b of the d dimensional matrix. a and b are training and testing data (classification). k denotes the index of the feature, and d is dimensions (number of k).

CROASSACIPTO ET AL.:
TONE CLASSIFICATION MATCHES KODÀLY HANDSIGN WITH THE K-NEAREST NEIGHBOR METHOD AT LEAP MOTION
CONTROLLER

43

Next is the flowchart of the KNN algorithm for the tone classification process, as shown in Fig.2.

```
                    ┌──────────┐
                    │  Start   │
                    └────┬─────┘
                         │
                         ▼
              ┌─────────────────────┐
              │ Determine the value │
              │        of k         │
              └──────────┬──────────┘
                         │
                         ▼
              ┌─────────────────────┐
              │ Calculate the       │
              │ Euclidean Distance  │
              │ value between       │
              │ training data and   │
              │ test data           │
              └──────────┬──────────┘
                         │
                         ▼
              ┌─────────────────────┐
              │ Sort data that has  │
              │ the smallest        │
              │ distance            │
              └──────────┬──────────┘
                         │
                         ▼
              ┌─────────────────────┐
              │ Determine the group │
              │ of test results     │
              │ based on the        │
              │ majority class from │
              │ the k-nearest       │
              │ neighbor            │
              └──────────┬──────────┘
                         │
                         ▼
                    ┌──────────┐
                    │   End    │
                    └──────────┘
```
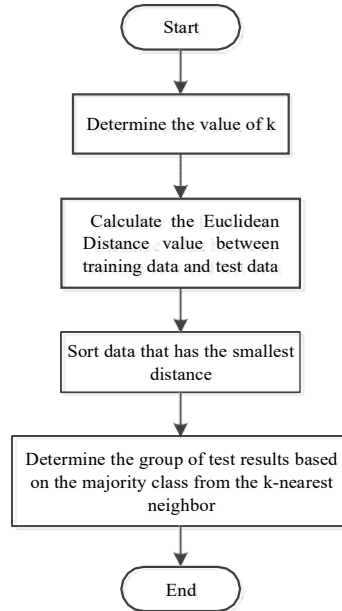
Fig. 2. Classification Process

In using the KNN algorithm, there are two processes, namely the preparation process of training data and test data (classification process).

I. Training Data Preparation Process

The process of preparing training data (Fig.3) is used as a parameter to determine the tone according to the Kodàly hand-sign, where a sample of Pitch, Roll, and Yaw value data will be stored in the dataset according to the tone.

II. Classification Process

Test data or classification process (Fig. 4) is done by comparing the value of Pitch, Roll, and Yaw received with the value of the previous training data stored in the dataset.
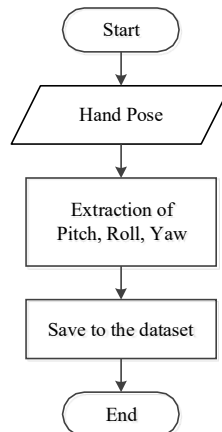
```
                    ┌──────────┐
                    │  Start   │
                    └────┬─────┘
                         │
                         ▼
                   ╱─────────────╲
                  ╱  Hand Pose    ╲
                  ╲───────────────╱
                         │
                         ▼
              ┌─────────────────────┐
              │ Extraction of       │
              │ Pitch, Roll, Yaw    │
              └──────────┬──────────┘
                         │
                         ▼
              ┌─────────────────────┐
              │ Save to the dataset │
              └──────────┬──────────┘
                         │
                         ▼
                    ┌──────────┐
                    │   End    │
                    └──────────┘
```
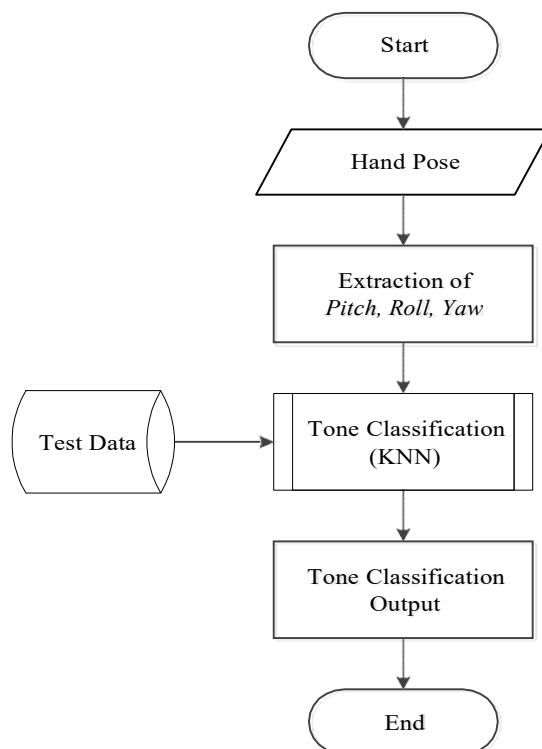
Fig. 3. Training Data Preparation Process

Fig. 4. Tone Classification Process using the KNN Algorithm

III. RESULTS AND DISCUSSION

From the test data obtained consisted of 1004 Do tone poses, 891 Re tone poses, 526 Mi tone poses, 661 Fa tone poses, 612 Sol tone poses, 913 La tone poses, 702 Si tone poses and 874 Do 'tone poses against 21099 data train which has been classified based on tone. The data were tested using the K-Nearest Neighbor algorithm with a k value determined periodically, namely 1, 3, 5, 7, 9. Here are the results of the tests carried out on the classification of the tone of the k that has been determined.

TABLE I
Result Summary

k	Accuracy	Error
1	94.87%	5.13%
3	94.49%	5.51%
5	92.83%	7.17%
7	92.75%	7.25%
9	92.59%	7.41%

Based on the test results in Table I, according to the k that has been determined, it is found that the optimal k value is at k = 1 with an accuracy of 94.87% as visualized in Fig.5 for each tone. Where for the Do tone, there are 831 data that are suitable, and 173 data is not suitable (similar to the Mi tone) from 1004 data while the Re tone has 679 data that is appropriate and 212 data that is not suitable (similar to La tone).
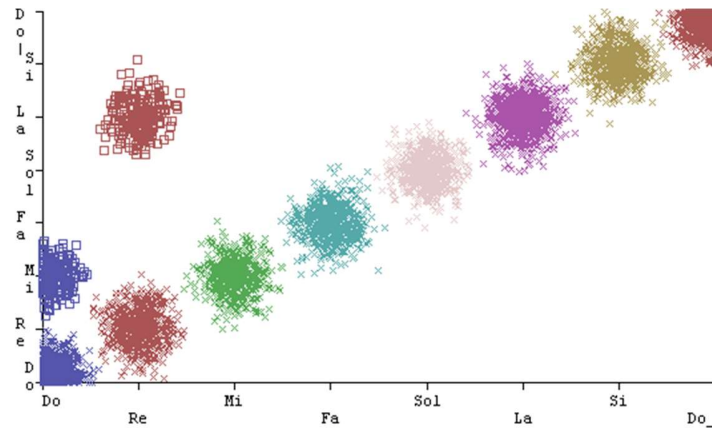
CROASSACIPTO ET AL.:
TONE CLASSIFICATION MATCHES KODÀLY HANDSIGN WITH THE K-NEAREST NEIGHBOR METHOD AT LEAP MOTION
CONTROLLER

45

Fig. 5. Visualization of Tone Classification with *K-Nearest Neighbor*

IV. Conclusion

Based on the results of the study, the hand pose recognition system using the K-Nearest Neighbor algorithm can recognize hand poses with an optimal k value of k = 1 with an accuracy of 94.87%. The system can recognize the hand pose in determining the tone according to the Kodály Handsign based on Pitch, Roll, Yaw. In addition, the system can recognize hand poses despite changes in position changes.

REFERENCES

[1]   M. Croassacipto, M. Ichwan, and D. B. Utami, "Klasifikasi Nada Sesuai Kodàly Handsign Dengan Metode K-Nearest Neighbor Pada Leap Motion Controller," *Indones. J. Comput.*, 2019, doi: 10.21108/indojc.2019.4.1.269.

[2]   D. B. Utami and M. Ichwan, "Pengenalan Pose Tangan Menggunakan HuMoment," *J. INFOTEL*, 2017, doi: 10.20895/infotel.v9i1.177.

[3]   S. J. Reyadh Naoum, Hussein H. Owaied, "Development of a new Arabic Sign Language Recognition Using K-Nearest Neighbor Algorithm," *J. Emerg. Trends Comput. Inf. Sci.*, 2012.

[4]   M. Billinghurst, T. Piumsomboon, and H. Bai, "Hands in space: Gesture interaction with augmented-reality interfaces," *IEEE Comput. Graph. Appl.*, 2014, doi: 10.1109/MCG.2014.8.

[5]   C. Keskin, F. Kiraç, Y. E. Kara, and L. Akarun, "Real time hand pose estimation using depth sensors," in *Proceedings of the IEEE International Conference on Computer Vision*, 2011, doi: 10.1109/ICCVW.2011.6130391.

[6]   L. Chosky, *The Kodaly method I: comprehensive music education*. New Jersey: Upper Saddle River, NJ: Prentice, 1999.

[7]   F. Weichert, D. Bachmann, B. Rudak, and D. Fisseler, "Analysis of the accuracy and robustness of the Leap Motion Controller," *Sensors (Switzerland)*, 2013, doi: 10.3390/s130506380.

[8]   "Leap Motion Controller." [Online]. Available: https://developer.leapmotion.com/documentation.

[9]   "Leap Motion Forum." [Online]. Available: https://forums.leapmotion.com.

[10]   M. Mohandes, S. Aliyu, and M. Deriche, "Arabic sign language recognition using the leap motion controller," in *IEEE International Symposium on Industrial Electronics*, 2014, doi: 10.1109/ISIE.2014.6864742.

[11]   D. Yuristiawan, "Identifikasi Kualitas Daging Sapi Berbasis Android Dengan Ekstraksi Fitur Warna Dan Klasifikasi KNN," Universitas Dian Nuswantoro., 2015.