Alphabet Sign Language Recognition Using K-Nearest Neighbor Optimization

Fitri Utaminingrum¹*, I Komang Somawirata², Gilbert Dany Naviri ¹
¹Brawijaya University, 8 Veteran Road Malang, Indonesia.
²National Institute of Technology (ITN Malang), KM 2 Karanglo Malang, Indonesia.

* Corresponding author. Email: f3_ningrum@ub.ac.id
Manuscript submitted November 16, 2018; accepted December 29, 2018.
doi: 10.17706/jcp.14.1 63-70

Abstract: Hand gesture recognition is an interesting and challenging study, especially for human interaction with computer. There are many method can used for hand gesture recognizing such as Random K-NN, Tree K-NN, Fuzzy K-NN. K-Nearest Neighbor method is interesting to be examined. While the weighting method Simple Multi Attribute Rating Technique (SMART) can be used to improve classify accuracy result. That method known very simple to use start from determine criteria of weighting, normalization, utility value and recommendation. Thus, in our research, we combine SMART weighting and K-NN classification for alphabet sign language recognition to help disabled to communicate with each other. The simulation result, we have the average accuracy 94%, 95% and 96% for rather dark, normal and rather bright respectively.

Key words: K-NN, recognition, SMART.

1. Introduction

Being able to communicate well is important. The hearing and speech are the primary points of communicating with the people. However, a lot of people are unfortunate and lack, they can’t hear and speech well, they call the mute and deaf people. Currently, research in the field of deaf and mute are growing rapidly, especially on the alphabet sign recognition. This paper, we will focus on the research for normal people who want to have communication with the deaf and mute using alphabet sign language. Simple method, high accuracy and easy to implemented in low cost device are needed. Many approaches method have been used for image classification. One of which is K-Nearest Neighbor (K-NN) [1]. K-NN is a famous method for image classification, which is because it’s simple and easy to be implemented. However, standalone K-NN doesn’t have a good in accuracy of image classification. Determination of the weighting value in KNN is greatly affect for accuracy value in the image classification, especially in the introduction of alphabet sign language. So that in this research, we will optimize the weight value for obtaining the optimum accuracy value.

Simple Multi Attribute Rating Technique (SMART) is now being used for weighting, which is due simple and easy to used. Weighting using the SMART method can be used to help make decisions, which is based on the theory that each alternative consists of a number of criteria and each criteria have values and weights. This weighting is used to assess each alternative in order to obtain the best alternative. The SMART weighting method has five steps as follow: define number of criteria; calculate weight of criteria; obtain relative weight; calculate utility value, and the last is calculate final value. The research with SMART
is comparing to Analytical Hierarchy Process (AHP) and SMART, where each indicator before AHP, examined with SMART. The result of SMART weighting is proved that it will give a better recommendation of obtaining conclusion while combined with AHP [2]. Based on the background, we will combining SMART and K-NN method for hand gesture recognition.

Many researcher are focus in the field of hand gesture recognitions, such as in [3] that built alphabet sign language (ASL) recognition using Back propagation Neural Network. 390 samples data images have been used to train the Network and 208 sample images have been used as the test set. After they were got 4000 epochs, the Mean Square Error (MSE) goes to an acceptable level of 0.01, than 5 input neurons are used to take the input as 1 5 feature vector, 26 output neurons to classify 26 individual signs. They have 80% accuracy [3]. The recognition using Centroid, Roundness and Scan line features for recognition and make classification is proposed by [4]. The tested image must be converted into gray scale and then by using OTSU’s algorithm to obtain binary image. The boundary of desired object for shape details is based on edge of image, and then uses Moore Neighbor Contour Tracing Algorithm for feature extraction. As the result, about 81% accuracy reached [4]. Another research was used difference of Gaussian and Scale Invariant Feature Transform (SIFT) algorithm for recognition [5]. The key-points derived from the image are placed in an array. The match performance based on similarity measures is not made for every point, instead a dimensionality reduction is done. There are 2 test conditions: First, within 26 alphabets and 10 alphabets repeat entries with difference entries and orientation. Second, live capture image for testing the vulnerability and performance. The result show that the system can recognize image in different background, lighting, scale, and illumination variance, different orientation, and different size hand [5].

2. System Overview

This study proposes a method of alphabet sign recognition that is combining the method of Simple Multi Attribute Rating Technique (SMART) and K-Nearest Neighbor (KNN). Those methods have simple formulas. K-NN has a better classify if being combined and SMART, the weighting will support to improve the accuracy. The initial process of our proposed method is capturing the hand image from camera. The size of image is resized to 300×400 pixels by using Adaptive Manipulation Interpolation Kernel [6], then continue to detect hand area using skin detection analysis. It is used to distinguish the hand from the background. Skin detection can be obtained by converting color image (RGB) to HSV ((Hue, Saturation, Value) and YCbCr color. SMART weighting uses binary image who obtained from segmentation process. Then, we divided the group recommendation into three (3) criteria based on shape of hand (Holding Hand, Straight Finger and Leaning Finger), these used for identifying the data training that will be used. Feature extraction is implemented to get a statistic data from distance between center image and edge of hand. The last process is an image classification using K-Nearest Neighbor (KNN). Our research was used Euclidean-Distance to get smallest data process, that is used to choose the smallest point as the selected category. The steps of our proposed method could be illustrated in Fig. 1.

2.1. Hand Detection

This step combines two color-spaces to improve skin color detection. These are HSV and YCbCr models [6]. HSV (Hue, Saturation, Value) model is manipulated from RGB color space into nonlinear transformation. Hue means a specific color form, it’s green, red or yellow. The brightness color corresponds to Value. Equation (1) is express to convert from RGB to HSV.

\[ I_{HSV} = C(I_{RGB}) \]  

(1)

where \( C[.] \) is conversion operator to convert RGB color space to HSV color space.
Then, Equation (2) is used to convert RGB color space to YCbCr:

\[ I_{YCbCr} = E[I_{RGB}] \]  

(2)

where \( E[.] \) is conversion operator to convert RGB color space to YCbCr color space.

In this research, skin color is obtained from thresholding process that combine between HSV and YCbCr color space as in equation (3).

\[ 130 \leq C_r \leq 180 \quad \& \quad 130 \leq C_b \leq 180 \quad \& \quad 0.01 \leq H \]  

(3)

Fig. 1 is Hand detection process. The capturing image of the hand area uses camera is illustrated in Fig. 3(a). Fig. 3(b) shows the skin detection based on equation (3). Fig. 3(c) is binary image that resulted from Fig. 3(b). Fig. 3(b) is hand segmentation by skin color detection. Then, the segmented image is converted into binary image. The binary image is obtained by following steps: first, convert segmented image from skin detection into grayscale image using luminance algorithm and then use Otsu Algorithm to create binary image. Furthermore Fig. 3(d) is cropped the binary image base on the top-end \( x \) and \( y \) coordinate (bounding box process).
2.2. Alphabet Grouped

26 alphabets are grouped into three criteria (K1, K2 and K3) based on shape of hand such as holding hand, straight finger and leaning finger. We use Simple Multi Attribute Rating Technique (SMART) to obtain the Weight (W) based on the group alphabet of data training. The steps of Simple Multi Attribute Rating Technique (SMART) to obtain weighting recommendation as follow: first step is obtain the number of criteria. In this paper we divide the criteria into K1, K2 and K3 which is based on the hand shape. Criteria one (K1) is hand shape based on holding hand which have seven alphabet members ("A,E,M,N,O,S,T"). Criteria two is based on straight finger shape, there have thirteen alphabet members ("B,C,D,F,I,K,L,R,U,V,W,X,Y") and the others alphabet is criteria three (K3) which have six members ("G,H,J,P,Q,Z"). The second step is obtain the weight of each criteria that calculated by dividing the number of criteria by total number of alphabet (26) as shown in Table 1.

<table>
<thead>
<tr>
<th>Criteria (K)</th>
<th>Hand Shape Base</th>
<th>Alphabet</th>
<th>Amount Criteria Member (N)</th>
<th>Weight (W)</th>
<th>Normalization (R)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K1</td>
<td>Holding Hand</td>
<td>A,E,M,N,O,S,T</td>
<td>7</td>
<td>7</td>
<td>( \frac{7}{26} )</td>
</tr>
<tr>
<td>K2</td>
<td>Straight Finger</td>
<td>B,C,D,F,I,K,L,R,U,V,W,X,Y</td>
<td>13</td>
<td>13</td>
<td>( \frac{13}{26} )</td>
</tr>
<tr>
<td>K3</td>
<td>Leaning Finger</td>
<td>G,H,J,P,Q,Z</td>
<td>6</td>
<td>6</td>
<td>( \frac{6}{26} )</td>
</tr>
</tbody>
</table>

Table 2. Weight Criteria and Normalization

<table>
<thead>
<tr>
<th>Number</th>
<th>Conclusion Alternative</th>
<th>Final Value Na</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>G1</td>
<td>27</td>
</tr>
<tr>
<td>2</td>
<td>G2</td>
<td>50</td>
</tr>
<tr>
<td>3</td>
<td>G3</td>
<td>23</td>
</tr>
</tbody>
</table>

The last steps is obtain the utility value and final value. Utility value \( U_{ki} \) is obtained by equation (4) that only have two values (0 and 100). Utility value for criteria one (K1) has 100 value, if the number of \( g_y \) is greater than 90 and \( g_x \) is greater than 200, while zero for other. The number of \( g_y \) and \( g_x \) are counted from binary image. Another criteria values are expressed in equation (5). Utility value for criteria two (K2) has 100 value, if the number of \( g_y \) is greater than 90 and \( g_x \) is lower than 200, while zero for other. The last is utility value for criteria three (K3) illustrated in equation (6). It has 100 value, if the number of \( g_y \) is lower than 90, while zero for other. Table 1 shows three criteria for 26 alphabet which is based on hand shape.

\[
U_{k1} = \begin{cases} 
100 & \text{if } g_y > 90 \text{ and } g_x > 200 \\
0 & \text{otherwise}
\end{cases} \quad (4)
\]

\[
U_{k2} = \begin{cases} 
100 & \text{if } g_y > 90 \text{ and } g_x < 200 \\
0 & \text{otherwise}
\end{cases} \quad (5)
\]

\[
U_{k3} = \begin{cases} 
100 & \text{if } g_y < 90 \\
0 & \text{otherwise}
\end{cases} \quad (6)
\]

The total number of \( g_y \) and \( g_x \) values were obtained after the experiment process on all data training alphabets. The result shows that every alphabet members of K3 have \( g_y \) less than 90, every alphabet members K2 have \( g_y \) more than 90 and \( g_x \) less than 200, and every alphabet members K1 have \( g_y \) more.
than 90 and \( g_x \) more than 200.

The conclusion alternative is based on the final value which is calculated by equation (7), and then the binary tested image will be directed recommendation of group based on Table 2.

\[
N_a = \sum_{i=0}^{3} R_1 \times U_{ki}
\] (7)

Every conclusion alternative is a group alphabets \( G_j \) and every group alphabet have alphabet members similar with the data training. Members of \( G_1 \), \( G_2 \) and \( G_3 \) are also similar with \( K_1 \), \( K_2 \), and \( K_3 \) respectively.

### 2.3. Feature Extraction

We use distance between center and edge of coordinate from the hand binary image for data training. The equation (8) is used to get distance \( d \) between center point coordinate \( (x, y) \) and every edge of hand coordinate \( (e_x, e_y) \) as coordinate of hand edge. The center coordinate \( (x, y) \) is obtained by divided high and width of cropped binary image by two.

\[
d = \sqrt{(center_x - e_x)^2 + (center_y - e_y)^2}
\] (8)

The pattern to get the distance \( d \) will be applied into all alphabets that have been set as data training and it is used to get the data test feature. Fig. 4 is an illustration of equation (8) for binary image.

### 2.4. K-Nearest Neighbor (KNN)

K-Nearest Neighbor (KNN) classification is famous method that is used for image classification [7]. We use \( K=1 \) and Euclidean-Distance to get smallest data process which is used for ranking. Data training \( (L) \) sets is “\( d \)”, it is resulted from every group members of binary image. Data test \( (U) \) sets by “\( d \)” was obtained from binary image that has been processed from captured image. Equation (7) is used for calculate the Euclidean-Distance.

\[
D = \sqrt{\sum_{i=1}^{1100} (U_i - L_i)^2}
\] (7)

The absolute decline after the training of data \( (L) \) has been risen sequence by all data testing. We get always positive ratings points that caused by \( K = 1 \), so we have to choose the smallest point as the selected category.

### 3. Evaluate and Analysis

The method is implemented on Matlab and the camera is shoot focused to hand of user that use alphabet sign language. Fig. 5 shows the result of hand detection with alphabet sign and center point detection. Fig. 5(a) up to (c) are sample of alphabet detection for “L”, “M” and “C”. The center point detection is shown in a binary image with blue point. The images in Fig. 5 are binary image that have been cropped from actual size. We use auto cropping image that coordinates start from the first on the left in the top area of white pixel on an image until the last of right in the bottom side.

We evaluate the performance of proposed method using accuracy [8]-[10]. Fig. 6 shows the accuracy result of the proposed method. We experiment in different source light condition, there are rather dark, normal and rather bright. In each experiment, we repeat until ten times. The experiment result shows the lighting will be affect to accuracy of alphabet sign recognition. The accuracy of alphabet sign recognition is good if the lighting is rather bright, and the accuracy will be decrease if the lighting down to dark.
average accuracy are 94%, 95% and 96% for rather dark, normal and rather bright respectively.

![Fig. 4. Illustration of distance “d” from center point to one edge.](image)

Fig. 4. Illustration of distance “d” from center point to one edge.

![Fig. 5. Alphabet sign detection and center point detection.](image)

Fig. 5. Alphabet sign detection and center point detection.

![Fig. 6. Average of accuracy testing with different lighting. Accuracy 1 for rather dark, Accuracy 2 for normal and Accuracy 3 for rather bright.](image)

Fig. 6. Average of accuracy testing with different lighting. Accuracy 1 for rather dark, Accuracy 2 for normal and Accuracy 3 for rather bright.

4. Conclusion

This paper proposed alphabet sign recognition by combining SMART weighting and K-Nearest Neighbor
(K-NN) classification. The SMART weighting is optimal to improve the accuracy of K-NN classification. However, accuracy is also affected by lighting condition. If the lighting is down to dark then the accuracy will be decrease and vice versa. For different lighting condition we have the average accuracy 94%, 95% and 96% for rather dark, normal and rather bright respectively.

References


Fitri Utaminingrum received the bachelor degree in electrical engineering from National Institute of Technology Malang Indonesia (2000-2004), and received the master degree in the same major from Brawijaya University, Indonesia (2004-2007) and Ph.D degree from Kumamoto University Japan (2011-2014). She has been a full time lecturer in Informatics Department of Brawijaya University, Malang-Indonesia from 2008. Her research interests are image processing and computer vision.

I Komang Somawirata received the bachelor degree in electrical engineering from National Institute of Technology Malang Indonesia, and received the master degree in the same major from Brawijaya University, Indonesia in 2004-2007 and the Ph.D degree from Kumamoto University Japan. He has been a full time lecturer in National Institute of Technology (ITN
Gilbert Dany Naviri received the bachelor degree in computer science in Informatics Department of Brawijaya University Malang Indonesia. His research interests are image processing and computer vision.