Journal of Physics: Conference Series

Discrete Wavelet Transform on static hand gesture recognition

To cite this article: Erizka Banuwati Candrasari *et al* 2019 *J. Phys.: Conf. Ser.* **1367** 012022

Discrete Wavelet Transform on static hand gesture recognition

Erizka Banuwati Candrasari[1], **Ledya Novamizanti**[2], **Suci Aulia**[3]

[1,2]School of Electrical Engineering, Telkom University, Indonesia
[3]School of Applied Science, Telkom University, Indonesia

E-mail: [1]erizkabanuwatic@gmail.com, [2]ledyaldn@telkomuniversity.ac.id, [3]suciaulia@telkomuniversity.ac.id

Abstract. The technology in present-day is very useful for human-physical recognition and the hand gesture recognition is one of them. Nevertheless, that technology still had various weakness, for example in the image brightness, contrast, recognition time, and accuracy rate. The objective of this paper was to construct hand gesture recognition system using RGB images. Pre-processing is wrought by resizing the image, separate the hand area, and pick the specific layer. This experiment used the YCbCr because its derived directly from RGB and had a higher contrast compared to other layers. The feature value was gathered from feature extraction on Discrete Wavelet Transform (DWT) using Low-Low sub-band and 2nd level decomposition. The current sub-band had smoothest contours in comparison with other sub-band. The final process was gesture classification using Hidden Markov Models (HMM) and K-Nearest Neighbor (KNN). The amount of training and testing data used were 150 and 100 images respectively, divided into five gestures with accuracy using HMM and KNN consecutively was 58% and 100%. The research novelty was that the classification impacted positively on accuracy level.

1. Introduction

A newer technology is able to recognizing a human-physical-body and a hand gesture recognition is one member of them. Hand recognition would be implemented in sign languange. The movement can be widely used, even for the deaf. In addition hand gestures are used for general nonverbal communication in the community

The objective of this research is to contrive a hand gesture recognition system based on the digital image, utilizing DWT and HMM extraction feature as classification algorithm. Subsequently, testing the result and analysing the system performance of DWT method and HMM classification. Research problems of this paper were as follows: how to create a construction and simulation of the hand gesture recognition against the dataset processed by DWT method and HMM classification, how value changes of the input parameter would affect system performance, and to determine of how were the comparable accuracy and system computing time of DWT and HMM classification.

Preceding research was conducted by Rajat Agarwal et al. [1] using *Discrete Wavelet Transform* (DWT) extraction feature and *Support Vector Machine* (SVM) classification method. In this research, 94% accuracy was acquired. Test was also conducted by utilizing an image of 256x256 pixels with level 5 decomposition that would generate 93,14% accuracy. DWT could deliver time and frequency information simultaneously and *wavelet* was adjustable and adaptable in accordance with purposes [2]. HMM had vantage point to solve evaluation, interference, and learning issues [3]. HMM was frequently used in various application, such as voice recognition [4], [5]; speech recognition [6], [7]; musical genre classification [8]; pattern recognition [9]; handwriting recognition [10], [11]; human action recognition

[12]; face recognition [13]; hand tracking [14]; and sign language recognition [15] was an effective learning algorithm and able to handle variation in record structure [16]. Refering to previous research conducted by Radu-Laurentiu Vieriu et al. [17] introduction static hand gesture with HMM method would deliver 93,38% of average accuracy.

Research method employed in this research was literature study that aimed to enhance the basic theories regarding DWT and HMM classification. Subsequently, data collection was conducted by utilizing data set Sebastien Marcel Static Hand Posture Database comprising a collection of hand gesture images. The dataset that applied in this study is a set of static hand gesture images. Yet, the dataset used in this study and in Pradyumna Narayana et al [18] research is different because the last used a dynamic gesture. The planning of the system would employ DWT as a hand recognition methodology and HMM classification to classify hand gesture. System testing was utilized to analyse the performance result of implementation that had been conducted by using DWT method and HMM classification and conclusion.

2. Research Method

This section would explain the fundamental theory of DWT method and HMM classification.

2.1. Discrete Wavelet Transform

The scalar and wavelet functions are shaped differently in 2D form, in this case are stated respectively as $\boldsymbol{\phi(x,y)}$ and $\boldsymbol{\psi(x,y)}$. The scalar base function and translations defined as [19]

$$\phi_{j,m,n}(x,y) = 2^{j/2}\phi\left(2^j x - m, 2^j y - n\right), \tag{1}$$

$$\psi^i_{j,m,n}(x,y) = 2^{j/2}\psi^i\left(2^j x - m, 2^j y - n\right), \qquad i = \{H,V,D\}. \tag{2}$$

wherein the index *i* refering to the direction of the wavelet. There are three different wavelet functions, $\boldsymbol{\psi^H(x,y)}, \boldsymbol{\psi^V(x,y)}$, and $\boldsymbol{\psi^D(x,y)}$. The wavelet function can be separated to $\boldsymbol{f(x,y) = f_1(x) f_2(y)}$. This equation then written as:

$$\phi(x,y) = \phi(x)\phi(y) \tag{3}$$

$$\psi^H(x,y) = \psi(x)\phi(y) \tag{4}$$

$$\psi^V(x,y) = \phi(x)\psi(y) \tag{5}$$

$$\psi^D(x,y) = \psi(x)\psi(y). \tag{6}$$

These (3), (4), (5), and (6) would enumerate the intensity or functional variations for pictures throughout different directions*:* $\boldsymbol{\psi^H(x,y)}$ scales the variations along the columns, $\boldsymbol{\psi^V(x,y)}$ reacts to rows variations, and $\boldsymbol{\psi^D(x,y)}$ corresponds with diagonals variations.

If the above functions defined as a separate functions, it easier to analyze 2D functions and focus on 1D wavelets design and scalar functions. Both analysis and equations changed to [20]

$$W_\phi(j_0,m,n) = \frac{1}{\sqrt{MN}} + \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\,\phi_{j_0,m,n}(x,y) \tag{7}$$

$$W^i_\psi(j,m,n) = \frac{1}{\sqrt{MN}} + \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\,\phi^i_{j,m,n}(x,y), \qquad i = \{H,V,D\} \tag{8}$$

$$\begin{aligned} f(x,y) = {} & \frac{1}{\sqrt{MN}}\sum_m\sum_n W_\phi(j_0,m,n)\,\phi_{j_0,m,n}(x,y) \\ & + \frac{1}{\sqrt{MN}}\sum_{i=H,V,D}\sum_{j=j_0}^{\infty}\sum_m\sum_n W^i_\psi(j,m,n)\,\phi^i_{j,m,n}(x,y), \qquad i = \{H,V,D\} \end{aligned} \tag{9}$$

where M is a row, N is a column, $\boldsymbol{j_0}$ is a mutable outset scale and the coeefficient of $\boldsymbol{W_\phi(j_0,m,n)}$ eludicate the $\boldsymbol{f(x,y)}$ approximation at scale $\boldsymbol{j_0}$. The $\boldsymbol{W^i_\psi(j,m,n)}$ coefficients append the horizontal, vertical, and diagonal details for scales $\boldsymbol{j \geq j_0}$. Allow $\boldsymbol{j_0 = 0}$ and choose $\boldsymbol{N = M = 2^j}$ so that $\boldsymbol{j = 0, 1, 2, \ldots, j-1}$ and $\boldsymbol{m, n = 0, 1, 2, \ldots, 2^j - 1}$. Supposed the $\boldsymbol{W_\phi}$ and $\boldsymbol{W^i_\psi}$ of (7), (8), and (9), $\boldsymbol{f\,(x,y)}$ is acquired from Inverse Discrete Wavelet Transform [21].

2.2. Hidden Markov Models

Each HMM is prescribed by a state, the probability of state, transition, emission and initial. To delineate it, five elements must be specified [22],[23]:

- N is the state of the model, potrayed as:

$$S = \{S_1, \dots, S_N\}. \quad (10)$$

- The M is the observation symbols per state $V = \{v_1, \dots, v_M\}$. If the observations are continuous, then M would valued infinite
- The state transition probability distribution $A = \{a_{ij}\}$, where a_{ij} is the probability that the state at time $t+1$ is S_j, is given when the state at time t is S_i. The connection architecture for the model defined by the cofiguration of this stochastic matrix. The coefficient a_{ij} will abide even through the training process if it valued zero, so noway a transition from the state S_i to S_j is occured.

$$a_{ij} = p\{q_{t+1} = j | q_t = i\}, 1 \leq i,\ j \leq N \quad (11)$$

where q_t shows the ongoing state. The transition probabilities shall meet the normal stochastic constraints, $a_{ij} \geq 0, 1 \leq i, j \leq N$ and $\sum_{j=1}^{N} a_{ij} = 1,\ 1 \leq i \leq N$.

- The Observation symbol probability distribution in every state sketched as, $B = \{b_j(k)\}$ where $b_j(k)$ is probability that symbol v_k is casted in state S_j.

$$b_j(k) = p\{o_t = v_k | q_t = j\}, 1 \leq j \leq N, 1 \leq k \leq M \quad (12)$$

where v_k shows the k^{th} observation symbol in the alphabet and o_t is the present parameter vector. The ensuing stochastic constraints must be met: $b_j(k) \geq 0, 1 \leq j \leq N$, $1 \leq k \leq M$ and $\sum_{k=1}^{M} b_j(k) = 1,\ 1 \leq j \leq N$.

- The HMM lead off state distribution $\pi = \{\pi_i\}$, where π_i is the probability that the model is in state S_i at the time $t = 0$ with

$$\pi_i = p\{q_1 = i\} \text{ and } 1 \leq i \leq N. \quad (13)$$

In order to analyze further, we have to solve the two basic problems with HMM, that is [24].

- The evaluation and forward algorithm problems, There are other calculation methods with very low complexity that using additional variables.

$$\alpha_t(i) = p\{o_1, o_2, \dots, o_T, q_t = i | \lambda\} \quad (14)$$

α_t is called **forward variable**, and $o_1, o_2, \dots, o_T$ is the partial observation sequence. Out of this, the recursive relationship

$$\alpha_{t+1}(j) = b_j(o_{t+1}) \sum_{i=1}^{N} \alpha_t(i)\, o_{ij}, 1 \leq j \leq N, 1 \leq t \leq T-1 \quad (15)$$

with $\alpha_1(j) = \pi_j b_j(o_1)$ where $1 \leq j \leq N$. $\alpha_T(i),\ 1 \leq i \leq N$ can be calculated using this recursion. So the required probability is given by

$$p\{O|\lambda\} = \sum_{i=1}^{N} \alpha_T(i) \quad (16)$$

This method is well known as the **forward algorithm**.

The backward variable $\beta_t(i)$ is specified at (17)

$$\beta_t(i) = p\{o_1, o_2, \dots, o_T | q_t = i, \lambda\}. \quad (17)$$

Given the present state i, $\beta_t(i)$ is the partial observation sequence probability, $o_{t+1}, o_{t+2}, \dots, o_{T.}$ $\beta_t(i)$ could be figured effectively by applying a recursive

$$\beta_t(i) = \sum_{j=1}^{N} \beta_{t+1}(j) \alpha_{ij} b_j(o_{t+1}), 1 \leq i \leq N, 1 \leq t \leq T-1 \quad (18)$$

where $\beta_t(i) = 1, 1 \leq i \leq N$. Further we can see that in equation (19)

$$\alpha_t(i)\beta_t(i) = p\{O, q_t = i|\ \lambda\}, \qquad 1 \le i \le N, 1 \le t \le T. \tag{19}$$

There are two ways to calculate $p\{O|\lambda\}$, we can either using forward or backward variable

$$p\{O|\lambda\} = \sum_{i=1}^{N} p\{O, q_t = i|\lambda\} = \sum_{i=1}^{N} \alpha_t(i)\beta_t(i) \tag{20}$$

- Learning problem, the next problem with HMM is how to set the model parameters $\lambda = (A, B, \pi)$ if given an maximum observation line $O = o_1, o_2, \dots, o_T, p\{O|\lambda\}$. To solve this problem, the Baum Welch algorithm is used

$$Q(\lambda, \bar{\lambda}) = \sum_{q} p\{q|O, \lambda\}\, log\big[p\{O, q, \bar{\lambda}\}\big]. \tag{21}$$

We need to interpret two more supporting variables to forward and backward variables. The first one of these variables is:

$$\xi_t(i,j) = \frac{p\{q_t = i, q_{t+1} = j|O, \lambda\}}{p\{O|\lambda\}} \tag{22}$$

The second variable is the posteriori probability,

$$\gamma_t(i) = \ p\{q_t = i|O, \lambda\} \tag{23}$$

The next function known as reestimation formulas and is used to update the parameters

$$\bar{\pi}_i = \gamma_1\,(i), 1 \le i \le N \tag{24}$$

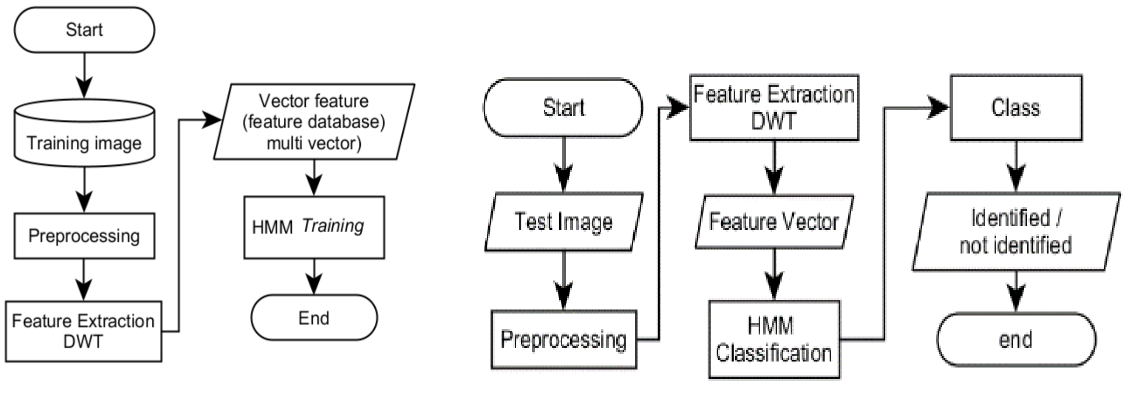
$$\bar{a}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)}, 1 \le i \le N, 1 \le j \le N \tag{25}$$

$$\bar{b}_j(k) = \frac{\sum_{t=1}^{T} \gamma_t(j)}{\sum_{t=1}^{T} \gamma_t(j)}, O_t = \ v_k, 1 \le j \le N, 1 \le k \le M. \tag{26}$$

These reestimation formulas can be altered well to tackle the continuous density case [8].

3. System Design

A system that enable to recognize hand gesture through image had been designed in this research. Generally, system design in this research illustrated in Figure 1.



(a) (b)

Figure 1. Flowchart of system design. (a) Training process, (b) Testing process.

General scheme exhibited in Figure 1 would be the foundation of this research. Explanation of flowchart diagram in training process in Figure 1.b of the system design:

- Inputs on the system consist of training image database from dataset Sebastien that has RGB (Red Green Blue) layer.
- Second step that had to be carried out was the processing of the first image that comprised of image resizing with size of 76x66 pixels and skin color segmentation. Then, produced conture images.

- Third sequent of the process, extraction carried out using DWT method. Later, it produced vector features.
- The final process was to train foward and backward parameter of the image in every class using HMM along with characteristic vector and training image as the inputs.

Explanation of flowchart in testing process in Figure 1.b of the system design:

- Inputs consists of training image that had layer RGB (*Red Green Blue*).
- Second sequence of the process, processing the early image by resizing the image and skin color segmentation. It produced conture image.
- Third sequence or the process, feature extraction was conducted by using DWT methods.

Last sequence, of the process for test image, classification using HMM. The process consisted of computing the forward parameter and determine the class with highest probability.

3.1. Preprocessing

Preprocessing of the image illustrated in Figure 1. Initial processing of the image consisted of separating the background process. Object in this research was a right hand as exhibited in Figure 2.

(a) (b) (c) (d) (e)

Figure 2. Hand image gesture. (a) Letter A, (b) Letter B, (c) Letter C, (d) Gesture Point, (e) Number 5.

Subsequently, convert it in the required layer. Explanation about the initial processing of the images is explained as follow:

- First step, resize the image. In this process the image was resized into 76×66 pixels.
- Second step, convert the image with RGB layer into YCbCr layer version. In this process, inputs consists of hand gesture image with RGB layer.
- Third step, segmenting the skin by determining the pixel value threshold of YCbCr (Cr).
- Fourth step is denoising. Process of erasing noises on the signal whilst keeping the characteristic of the signal.
- Fifth step, filling up the unerasable noises from earlier process.
- Sixth step is dilation process. Dilation is process of thickening the segmented image border from fifth step in order to make the required pixel detectable.
- Seventh step is erotion process. Erotion process is an exfoliation on the segmented image border from sixth step. It is carried out in order to dispose unsed pixel.
- Outputs from all sequences is a hand image conture with layer YCbCr (Cr).

3.2. Feature Extraction

Process of finding hand feature in this research was carried out using *Discrete Wavelet Transform* method. *Discrete Wavelet Transform* method utilized in order to create characteristic matrix of an image to represent the value of the related matrix and image [25].

As illustrated in Figure 1, following is the explanation of feature extraction process sequences:

- First step is average value computation of each pair of hand gesture conture pixel row..
- Second is average value computation from each pixel column pair by inputing the previous computation result. It will produce and output of image conture value subband LL, LH, HL, and HH.
- Third step, choose the required subband as characteristic matrix. The process will produce an output of characteristic matrix and it will be combined with previous conture image characteristic.

- Final step is repeating the DWT characteristic extraction. It will be finalized if each conture image data are extracted.

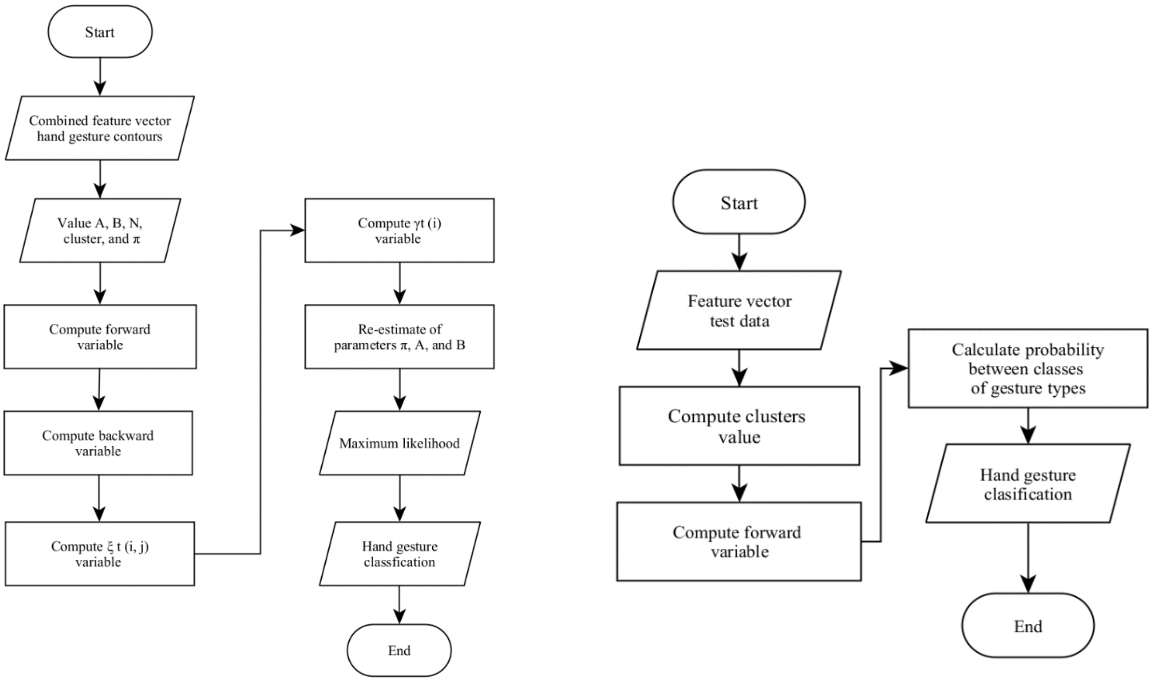
3.3. Classification

Classification process is to classify characteristic value that refer to matrix value of data obtained from feature extraction. Classification process using HMM method illustrated in Figure 3. Explanation of the procedure illustrated in Figure 3(a) classification training process using HMM method is as follow*:*

- Inputs consisted of matrix combination of image characteristic matrix obtained from feature extraction process using DWT method. Furthermore, HMM required value of A, B, π, state, and cluster. So it is necessary to determine the required state value and cluster value computation as value of observation by finding k-means value.
- Next process is the computation of forward variable that consist of initialization [9], [11], recursion and termination [4]. Prior to forward algorithm calculation process, scaling function calculation process is required.
- After pass through forward algorithm calculation, the next step must be backward algorithm. The process consists of two step namely: Installation and recursion step. Prior to backward algorithm calcutaion scalling function calculation process is required [4].
- Calculation variable ξ_t(i, j) and variable γ_t(i) based on the defined variable from the previous forward and backward procedure.
- Once four of the variable are obtained, reestimate the Parameter A, B, and π.

After reestimating parameter A, B, and π, next process is testing the data from classification process. Explanation of the procedure illustrated in Figure 3.b of classification testing process using HMM method is as follow:

- Inputs is single test characteristic matrix. Next process is computing the cluster value with k-means algorithm.
- Recalculate the forward algorithm as in training process. After termination, determine the test image probability for every hand gesture classes.
- Once the test image probability value obtained for every hand gesture classes. Final step is taking the highest test image probability value as the final value of hand gesture classification.



(a) (b)

Figure 3. Classification flowchart design. (a) training process, (b) testing process.

4. Results and Analysis

System testing was carried out on dataset Sebastien Marcel Static Hand Posture Database [26] by the image that had been through resizing process into 76×66 pixels. The objective of the system testing was to find the accuracy and system computing time comparisons and to find the best performance of hand gesture introduction system. In total there were 250 image data used in this research and they were gathered from dataset Sebastien Marcel Static Hand Posture Database. Hand gesture consisted of 5 classes. Every word classes consisted of 50 images.

4.1. System Parameter Testing

The objective of parameter parameter testing was to obtain parameter result with optimum performance that could be seen from its accuracy and computing time.

4.1.1. Influence of Layer Type Parameter. The test carried out by using one type of layer that would broght to the test as the DWT characteristic extraction would be carried out and classified by using HMM as exhibited in Table 1. It could be concluded that the optimum parameter lays on layer type YCbCr (Cr).

Table 1. Layer Type Parameter Performances

Layer Type	Total Testing Data	Total correct data	Accuracy (%)	Computing Time (second)
Red	100	32	32	1,5
Green	100	40	40	1,4
Blue	100	45	45	1,6
Grayscale	100	45	45	1,2
Binary	100	40	40	1,2
YCbCr (Cr)	100	58	58	1,2
HSV (V)	100	40	40	1,4

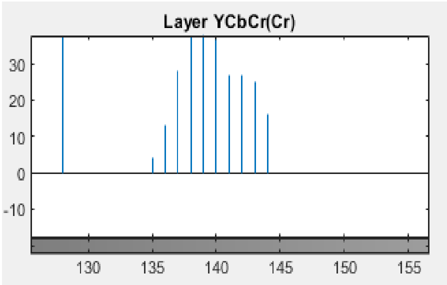


Figure 4. Histogram of Image with Layer YCbCr (Cr).

It could be concluded from Table 1 that layer YCbCr (Cr) has optimum accuracy. Inasmuch as the value of pixel frequence lays between 0 to 43, thus more the high intensity pixel tend to appear the most compared to other layer type as shown in Figure 4.

4.1.2. Influence of Sub-Band Type Parameter. The test was carried out using the layer that has optimum performance in prior test, namely layer YCbCr (Cr) and DWT parameter, there were type 4 sub-band consisted of Low-Low (LL), Low-High (LH), High-Low (HL), and High-High (HH). The result of sub-band type parameter performance is listed in Table 2. Table 2 shows that sub-band LL is the best parameter. Sub-band LL has smooth conture compared to other type of Sub-Band as shown in Figure 5.

Table 2. Sub-band Type Parameter Performances.

Sub-band Type	Total Testing Data	Total Correct data	Accuracy (%)	Computing Time (Second)
LL	100	58	58	1,2
LH	100	20	20	2,1
HL	100	20	20	2,2
HH	100	22	22	2,1

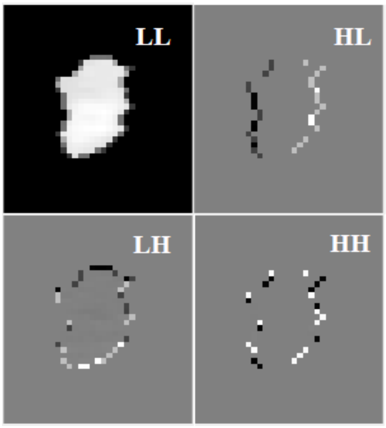


Figure 5. Image Illustration in Sub-Band type

4.1.3. Influence of Decomposition Level Parameter Previous test carried out by using layer with best performance namely, layer YCbCr (Cr) and DWT parameter. The test carried out with YCbCr (Cr) dan sub-band LL. Results of parameter performance listed in Table 3. As can be inferred from Table 3, Decomposition level 2 shows the best parameter. Characteristic graphic affected by the decomposition level shown in Figure 6. The decomposition level had the characteristic remained with least characteristic. It would affect the accuracy and computing time. The fewer decomposition level value would affect the corresponding computing time to be briefer.The same effect did not work on accuracy.

Certain decomposition level parameter value hold the clear characteristic value to distinguish between classes.

Table 3. Decomposition Level Parameter Performances.

Level	Total Testing Data	Total Correct Data	Accuracy (%)	Computing Time (second)
1	100	40	40	2,4
2	100	48	48	1,2
3	100	37	37	1,2
4	100	22	22	1,2

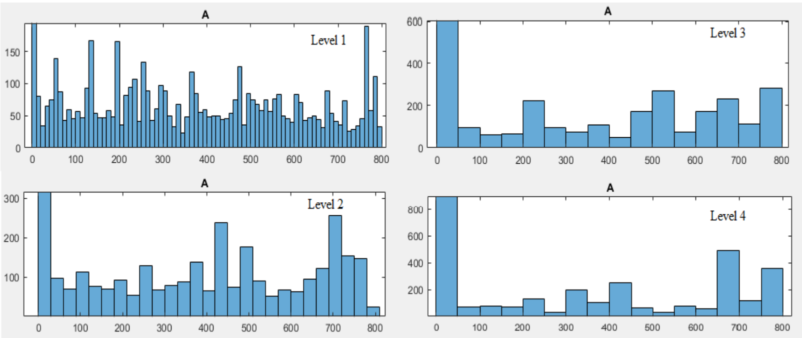


Figure 6. Characteristic Chart to Decomposition Level Value Changes

4.1.4. Influence of Mother Wavelet Type Performance Previous testing carried out by involving four types of mother wavelet parameter namely, Haar, db3, db5, dan db7. Previous scenario showed an optimum parameter combination consisted of layer YCbCr (Cr), sub-band LL, and decomposition level 2. Performance test resulted from mother wavelet type listed in Table 4. According to the test shown in Table 4, optimum accuracy and computing time in Mother Wavelet db5 type were obtained.

Table 4. Mother Wavelet Parameter Performances

Mother Wavelet Type	Total Testing Data	Total Correct Data	Accuracy (%)	Computing Time (second)
Haar	100	39	39	1,6
db3	100	40	40	1,1
db5	100	55	55	1,2
db7	100	20	20	1,7

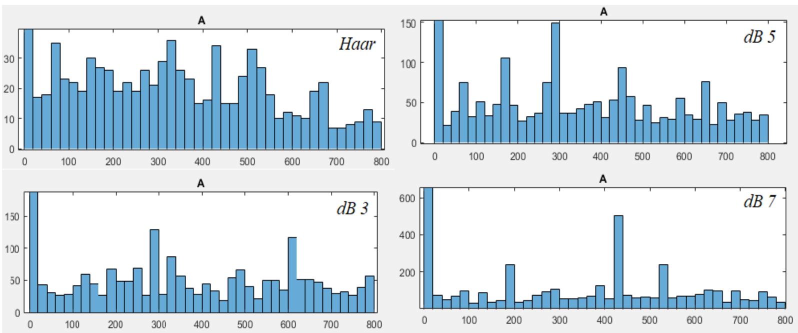


Figure 7. Characteristic to change in type of mother wavelet type chart.

As shown in Figure 7, characteristic chart displays the difference in mother wavelet type that leads to a difference in form of characteristic in among class.

4.1.5. Influence of Cluster Amount Parameter. Previous test involved cluster parameter test conducted in HMM classification method. Clusters that would be taken into the analysis namely, 50, 100, 200,400, 800, and 1000. Test that result in the optimum parameter had been conducted in previous test scenario. Cluster amount parameter performance result shown as in Table 5. Table 5 exhibit the optimum cluster amount parameter as 800. Figure 8 shows that the decision of selecting characteristic produced by cluster amount of 1000 had led the characteristic of each gesture to have the similar characteristic value, therefore gaining less accuracy compared to cluster amount 800.

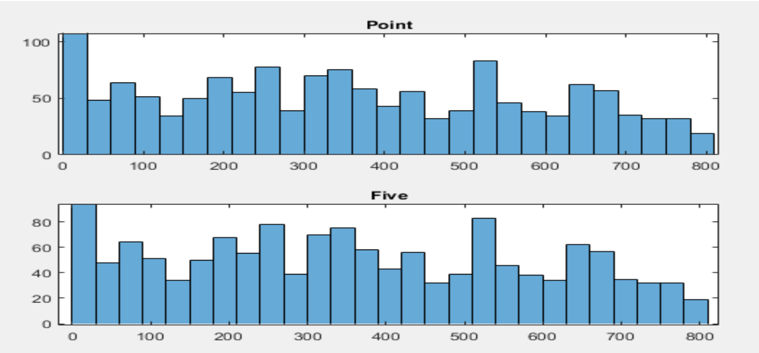


Figure 8. Characteristic chart of cluster amount 1000.

Table 5. Cluster Amount Parameter Performances.

Number of Clusters	Total Testing Data	Total Correct Data	Accuracy (%)	Computing Time (Second)
50	100	38	38	2
100	100	21	21	1,2
200	100	20	20	1,2
400	100	43	43	1,2
800	100	58	58	1,2
1000	100	34	34	1,2

4.1.6. Influence of State Amount Parameter. Next step would be state parameter testing that would make use for HMM classification to accuracy and system computing. State 4, 5, 25, 50, 100, and 150 would be brought to the analysis. The best performance of state amount parameter testing occurred in 5 state listed in Table 6.

Table 6. State amount parameter performances.

Number of State	Total Testing Data	Total Correct Data	Accuracy (%)	Computing Time (second)
4	100	46	46	2,7
5	100	58	58	2,4
25	100	20	20	11,1
50	100	20	20	11,1
100	100	20	20	11,1
150	100	35	35	11,1

According to Table 6, the best parameter achieved by state amount 5. It occurred due to the concept of the HMM that basicly split the data to the amount of desired state. Hence, once the selected state value false, testing data identification would be arduous.

4.2. Data Amount Testing

Data amount testing comprised of two tests namely, training image and testing image sequentialy as many as 40 training and 10 testing, 30 training and 20 testing, 25 training and 25 training using the optimum parameter shown in Table 7.

Table 7. Data amount testing performances.

Number of Data (Training Data – Testing Data)	Accuracy (%)	Computing Time (detik)
40 – 10	20	1,1
30 – 20	58	2,4
25 – 25	39	3

It could be inferred from Table 7 that introduction system using DWT extraction methods and HMM classification enable to thoroughly identifying, providing the percentage between training and testing data is 60% and 40% from all data of each class.

4.3. Classification Testing
Classification testing carried out by comparing accuracy and computing time against two classification methods namely, K-NN (K-Nearest Neighbor) and HMM. Classification testing performed by testing the training data to training data and training data to testing data as exhibited in Table 8.

Table 8. Classification testing performances to classification methods.

Dataset	Classification	Accuracy (%)		Computing Time (second)	
		Training Data	Testing Data	Training Data	Testing Data
Dataset *Sebastien*	K-NN	100%	100%	3,78	2,7
	HMM	38%	58%	2,6	2,01
Dataset Erizka	K-NN	100%	100%	58	40
	HMM	55%	72%	68	53

It can be infered from Table 7, HMM has low accuracy value when testing training data to training data, compared to training data to testing data. It occurred due to the data amount percentage of 50%-50% when testing training data to training data. In the other hand, when testing training data to testing data, data amount percentage lays consequentively in 60%-40%.

Table 8 shows that the testing also carried out with other data set namely, Erizka's dataset with better accuracy and computing time compared to performances of input data set Sebastien Marcel Static Hand Posture Dataset. Erizka's data set has resize image of 128×128 pixels.

5. Conclusion

Hand gesture recognition using Discrete Wavelet Transform and Hidden Markov Models has been designed. There are 5 type of gestures namely, letter A, letter B, letter C, point and number 5 (five). The data amount was 30 training image and 20 testing image. High accuracy layer was those with exceptional contrast and brightness value. Low quality images had shorter pixel intensity range that span between 128 to 144. Image brightness value range came to low caused by pixel frequency value with high intensity span between 0 to 43. Those case led to low accuracy value of 58% for Sebastien Marcel Static Hand Posture dataset. While, dataset Erizka with resize image of 128×128 pixels, achieved accuracy 72%. DWT sub-band type parameter meant processing the image to obtain smooth image characteristic and it occurred well in sub-band LL. Decomposition level parameter is the process of image conversion into simplified form in order to obtain good feature characteristic. Mother wavelet type had caused the difference between feature characteristic. Cluster parameter on HMM affect the selection of characteristic that would be used. The determination of cluster value should be correct as the characteristic value of each class would be similar. The finest classification was KNN by cause of the image was compressed into 38x33 pixel when using DWT. Contradictively, HMM compressed the image into five separate segments causing the system to classify the image with a much tinier image. Because of the differences in image size caused by compression, the accuracy value with HMM was smaller than it was with KNN classifier.

Reference

[1] Agarwal R, Raman B, Mittal A. 2015. Hand Gesture Recognition using Discrete Wavelet Transform and Support Vector Machine. In: 2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN). Noida: IEEE; pp 489–93

[2] Sifuzzaman M, Islam MR, Ali MZ. 2009, Application of Wavelet Transform and its Advantages Compared to Fourier Transform. J Phys Sci.

[3] Prasetyo MEB. 2011. Teori Dasar Hidden Markov Model. Makal Probab Stat STEI.

[4] Buono A, Ramdhan A, Ruvinna. 2008. Pengenalan Kata Berbahasa Indoensia dengan Hidden Markov Model (HMM) Menggunakan Algoritme Baum-Welch. J Ilmu Ilmu Komputer;6(2): pp 32–40

[5] Jamaludin A, Huda AF, Sahyandari R. 2016. Pengenalan Lafal Hukum Nun Mati Menggunakan Hidden Markov Model. LOG!K@;6(1):1–10.

[6] Daniel J, Martin JH. 2018. Speech and Language Processing an introduction to natural language processing, computational linguistics, and speech recognition (Second Edition). PEARSON.

[7] Gopi ES. Digital Speech Processing Using Matlab [Internet]. 2014. Available from: http://link.springer.com/10.1007/978-81-322-1677-3

[8] I. Ikhsan, L. Novamizanti, I. N. A. Ramatryana. 2014. Automatic musical genre classification of audio using Hidden Markov Model. 2nd International Conference on Information and Communication Technology, ICoICT. Bandung. 28-30 May 2014

[9] Fink GA. 2014. Markov Models for Pattern Recognition. Markov Models for Pattern Recognition. Springer London.

[10] Krishna N, Andrew S, Jayashree S. 1996. Intialization of hidden markov models for unconstrained on-line handwriting recognition. In: 1996 IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings. Atlanta. p. 3502–5.

[11] Mardhiyya A, Hidayat B, Aulia S. 2015. Deteksi Tulisan Tangan Menggunakan Metode Segmentasi Adaptif Dan Hidden Markov Model. CITEE 2015.p. 196–200.

[12] Moghaddam Z, Piccardi M. 2014. Training initialization of hidden markov models in human action recognition. In: IEEE Transactions on Automation Science and Engineering. p. 394–408.

[13] Syakhala AR, Puspitaningrum D, Purwandari EP. 2015. Perbandingan Metode Principal Component Analysis (PCA) dengan Metode Hidden Markov Model (HMM) dalam Pengenalan Identitas. J Rekursif [Internet]. ;3(2):68–81. Available from: https://ejournal.unib.ac.id/index.php/rekursif/issue/archive

[14] Keskin C, Erkan A, Akarun L. 2003. Real time hand tracking and 3D gesture recognition for interactive interfaces using HMM. Icann/Iconipp;26–9.

[15] Trigueiros P, Ribeiro F, Reis LP. 2015. Generic System for Human-Computer Gesture Interaction: Applications on Sign Language Recognition and Robotic Soccer Refereeing. J Intell Robot Syst Theory Appl.;80(3–4):573–94.

[16] Khiatani D, Ghose U. Weather forecasting using Hidden Markov Model. 2017 Int Conf Comput Commun Technol Smart Nation, IC3TSN 2017. 2018;2017-Octob:220–5.

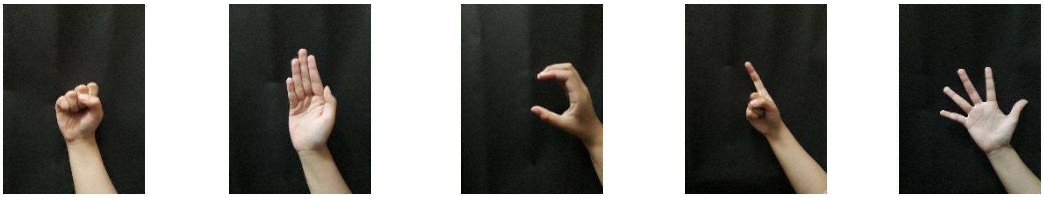
[17] Vieriu RL, Mironică I, Goras BT. 2013. Background invariant static hand gesture recognition based on Hidden Markov Models. ISSCS 2013 - Int Symp Signals, Circuits Syst.;1–4.

[18] Gonzalez RC, Woods RE. 2006. Digital Image Processing (3rd Edition). Prentice-Hall, Inc. Upper Saddle River, NJ, USA ©2006.

[19] Narayana P, Beveridge JR, Draper BA. 2018. Gesture Recognition: Focus on the Hands. Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit.;5235–44.

[20] Liu C-L. 2010. A Tutorial of the Wavelet Transform. NTUEE.
[21] Sundararajan D. 2015. Discrete Wavelet Transform: A Signal Processing Approach. Discrete Wavelet Transform: A Signal Processing Approach. WILEY.
[22] Dymarski P. 2011. Hidden Markov Models, Theory And Applications. InTech Web.ORG. InTech. 314 p.
[23] Giron-sierra JM. 2017. Digital Signal Processing with Matlab Examples. Vol. 1.
[24] Petrov Y. 2012. Hidden Markov Models, Theory and Applications. PLoS One;
[25] Novamizanti L, 2015. Kurnia A. Analisis Perbandingan Kompresi Haar Wavelet Transform dengan Embedded Zerotree Wavelet pada Citra. Elkomnika.;3(2):161–76.
[26] Marcel S. 1999. Hand posture recognition in a body-face centered space. In: Proceedings of the Conference on Human Factors in Computer Systems (CHI),. Pittsburgh.

Appendix



(a) (b) (c) (d) (e)

Erizka's hand gesture dataset. (a) Leter A, (b) Leter B, (c) Leter C, (d) Gesture Point, (e) Number 5