

# Similarity Measurement Of Segmented Image Using Euclidean Distance Method

1<sup>st</sup>Meira Parma Dewi  
Mathematics Department  
Universitas Negeri Padang  
Padang, Indonesia  
meiradaud@gmail.com

2<sup>nd</sup>Silfinia  
Mathematics Department  
Universitas Negeri Padang  
Padang, Indonesia

3<sup>rd</sup>Dewi Murni  
Mathematics and Science Faculty  
Universitas Negeri Padang  
Padang, Indonesia  
dewimunp@gmail.com

**Abstract**—This research aim to measure the similarity of two image. The similarity measurement is used to get the original image of the segmented image. The query image is an image that suspected to get undergone a change. The query image must be segmented first before measuring process. The similarity measurement process use Euclidean distance method. The process undertaken to identify the image of the segmentation result is to segment the image of the database, determine the co-occurrence matrix of each image, perform the texture feature analysis of each image, then calculate the distance of the texture feature of each image. The image identification stage of the query image in the program begins by inputting the color image to be segmented, then the image of the database is segmented, followed by texture feature value analysis for image and image imagery in the database, and calculating the value of similarity texture feature using Euclidean Distance method.

**Keywords**— Euclidean Distance Method, Image Segmentation, Similarity Measure.

## I. INTRODUCTION

Image is one of the multimedia component that plays an important role as a form of social information. Image needs to be done an image processing so that the required information can be met and the image of the processing can be used for certain purposes. Tools needed in image processing is a computer, so the image used is a digital image[2].

Digital imagery is a clear picture of objects that can be processed with a computer. The larger the image size will require a large storage area as well. There are several types of digital imagery that is color image (RGB), grayscale scale image and binary image. One type of digital image processing is image segmentation.

Segmentation is a process to separate an object from its background. Image segmentation is a process to retrieve information contained in the image and understand the full image features or used to process to a further level. After the segmentation is done then the original image will not be known, if you want to know the original image form of segmentation results need to be done to the image of the image segmentation by calculating the similarity of texture features between the image of the results of segmentation

with the image contained in the database. The problem raised in this research is how to find the original image form from the image of the result of segmentation.

The identification process is done by comparing the texture feature value of the image of the segmentation result with the texture feature value of the image in the database. The regularity texture of certain patterns formed from the arrangement of the pixels in a digital image [1]. Calculation of equality value of an image is one of the essential components of the process of an algorithm or method to match between the image of the query with a new image as a comparison. The accuracy of the similarity / similarity value calculation determines whether the query image has similarity or similarity characteristics to the image in the database.

The process of matching similarity or similarity characteristics of an image is one of them by using the method of distance measurement. One method of measuring the distance in measuring the similarity or similarity of an image is using the euclidean distance method, so it can be known which image texture is most similar to the image of the query.

## II. METHODS

This research is a basic research. The method used in this research is literature study. In the settlement of the problem, the steps taken are as follows:

- 1) Collect relevant theories and literatures with euclidean distance method popping in texture pattern recognition techniques.
- 2) Study and examine more in depth theories relating to the application of texture pattern recognition techniques to the euclidean distance method.
- 3) Create stages in identifying the image of segmentation results (query image).
- 4) Create algorithms for each stage.
- 5) Apply the algorithm in the form of computer program assisted matlab application program.
- 6) Conducting trial and evaluation of the application to identify the image of the result of segmentation that has been made.
- 7) Drawing conclusions from the results and discussions that have been done.

III. RESULTS AND DISCUSSION

a. Identify The Image Of The Segmentation Result Through The Introduction Of The Texture Pattern By The Euclidean Distance Method.

The image of the segmentation result is a binary image. Binary imagery is an image that each pixel is only black and white. In this research will be done image identification process of segmentation result, which with this research we will know how to form original image before done segmentation process. To find out how to form the original image of the image of the result of segmentation, the identification of the image of the segmentation result through the introduction of texture pattern with euclidean distance method. Each identification process is described in the following steps:

1) Perform Image Segmentation

The following are the steps of image segmentation which is to change the image in the form of color image into gray image using matlab application, determine the gray level of each pixel represented in matrix form, convert gray image into binary image with *thresholding* concept. The development concept is to classify the gray level of each pixel into 2 classes are 0 (black) and 1 (white). The following is the mining function [3] :

$$f(i, j) = \begin{cases} 1 \text{ (white)}, & f_g(i, j) \leq T \\ 0 \text{ (black)}, & \text{else} \end{cases}$$

$$T = \frac{\text{with } maxv + minv}{2}$$

which in this case,  $f_g(i, j)$  is the pixel and T is the thresholding where  $maxv$  is the highest gray level and  $minv$  is the lowest gray level, changing the gray level of the image with 0 (black) and 1 (white) according to section c, so the result of segmentation is obtained.

Example :

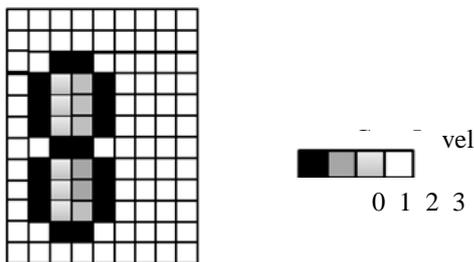


Fig. 1. Figure Number Eight

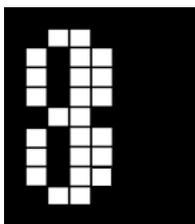


Fig. 2. Result Of Image Segmentation

2) Specifies The Co-Occurrence Matrix For The Query Image With The Database Image

Co-occurrence matrix is a matrix that describes the frequency of occurrence of pairs of two pixels with certain intensity within a certain distance and direction in the image[4].The following steps to make the co-occurrence matrix of an image is to determine the distance between two points (pixels) in the vertical and horizontal direction( $d = (dx, dy)$ ), where in this study the author uses  $d = (1,1)$ , , calculate the pixel pixels that have the value of the  $i$  and  $j$  and spaced  $d$  pixels in the image, and put the calculation result of each pair of intensity values in the matrix according to their coordinates, where the abscissa value for the intensity  $i$  and ordinate for the intensity value  $j$ .

Example :

Find the co- occurrence matrix 2 pieces of image below:

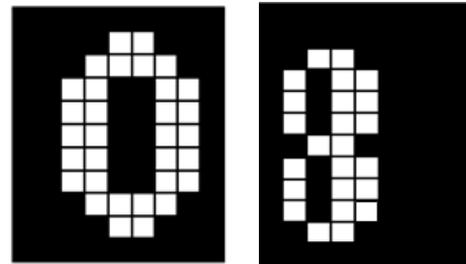


Fig 3. Image Number 0 (Query Image) and Image Number Eight (Database Image)

Completion :

In this research the authors use the distance between two pixels is  $d = (1,1)$ , which means one pixel to the right and one pixel down.

To determine the matrix co-occurrence image number 0 first pixels contained in the image represented in the form of a matrix namely:

$$f(x, y) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Get the matrix co-occurrence :

		j		
		0	1	
i	0	32	15	× $\frac{1}{79}$
	1	16	16	

Of the co-occurrence matrix for the image number 0 is:

$$p(0,0) = \frac{32}{79}, \quad p(0,1) = \frac{16}{79}, \quad p(1,0) = \frac{15}{79}, \quad \text{and} \\ p(1,1) = \frac{16}{79}$$

To determine the matrix co-occurrence image eight first the pixels contained in the image are represented in matrix form as:

$$f(x,y) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Get the matrix co-occurrence:

		i		
		0	1	
j	0	50	14	× $\frac{1}{84}$
	1	14	10	

From matrix co-occurrence for the image number eight as :

$$p(0,0) = \frac{50}{84}, \quad p(0,1) = \frac{14}{84}, \quad p(1,0) = \frac{14}{84}, \quad \text{dan} \\ p(1,1) = \frac{10}{84}$$

### 3) Calculated The Value Of Texture Fiture From Query Image and Database Image

The formula to calculate the value of texture feature in image is [1]:

Entropy is useful for measuring randomness from intensity. The value of entropy is maximum if all elements  $p(i,j)$  are equal, is the matrix corresponding with the image which there is no particular arrangement in the intensity pair. The value of entropy can be defined as follows:

$$e = - \sum_i \sum_j p(i,j) \log p(i,j) \\ \text{which} \\ e = \text{Entropy value}$$

$p(i,j)$  = the probability from pairs of two pixels with intensity i and j

Energy refers to the homogeneity of the texture, if the value of energy is high then the number of homogeneous areas are large and if the value of energy is low then the number of homogeneous areas are less.

$$E = \sum_i \sum_j [p(i,j)]^2 \\ \text{which} \\ E = \text{Energy value}$$

$p(i,j)$  = the probability from pairs of two pixels with intensity i and j

Contrast is useful for measuring the homogeneity of intensity variations in images. The value of contrast is increases if the intensity variation in the image is high, and decreases if the intensity variation in the image is small. Contrast can be defined as follows:

$$c = \sum_i \sum_j (i-j)^2 p(i,j) \\ \text{which} \\ c = \text{Contrast value}$$

$p(i,j)$  = the probability from pairs of two pixels with intensity i and j

The homogeneous texture is to contain a repeating and uniform structure that produces a particular pattern.

Homogeneity is useful for measuring the homogeneity of intensity variations in images. Homogeneity can be defined by:

$$H = \sum_i \sum_j \frac{p(i,j)}{1 + |i-j|} \\ \text{which} \\ H = \text{homogeneity value}$$

$p(i,j)$  = the probability from pairs of two pixels with intensity i and j

The value of homogeneity is large when the intensity of variation in the image is small.

The average of gray intensity is the disperse size of an image. The average of gray intensity shows the average of a distribution of gray image intensity values.

$$\mu = \frac{1}{N} \sum_{k=0}^{N-1} x_k$$

which

- $\mu$  = value of average gray intensity
- $x_k$  = the value of gray intensity from pixel-k, k = 0, 1, ..., N-1
- N = many pikxel in image

Standard deviation is a reflection of the mean deviation of data from the mean of gray intensity. The standard deviation shows the distribution of pixels in the image field.

$$\sigma = \sqrt{\frac{1}{N} \sum_{k=0}^{N-1} (x_k - \mu)^2}$$

which

- $\sigma$  = standar deviation value
- N = sum pixel in the image
- $x_k$  = gray intensity from pixel-k, k= 0, 1, ..., N-1
- $\mu$  = value of average gray intensity

The texture features of the query image and database image is :

In this research can be query image is image number 0. The calculation of texture features for image number 0 is that after the values contained in the co-occurrence matrix of

image number 0 are substituted into the formula, it is obtained:

$$\begin{aligned} e_1 &= -(p(0,0)\log p(0,0) + p(0,1)\log p(0,1) + p(1,0) \\ &\quad \log p(1,0) + p(1,1)\log p(1,1)) \\ &= -\left(\left(\frac{32}{79} \times \log \frac{32}{79}\right) + \left(\frac{15}{79} \times \log \frac{15}{79}\right) + \left(\frac{16}{79} \times \log \frac{16}{79}\right)\right. \\ &\quad \left.+ \left(\frac{16}{79} \times \log \frac{16}{79}\right)\right) \\ &= -((0,405 \times \log 0,405) + (0,189 \times \log 0,189) \\ &\quad + (0,203 \times \log 0,203) + (0,203 \times \log 0,203)) \\ &= 1,3283 \end{aligned}$$

Obtained the value entropy for *query* image (image number 0) as 1,3283.

$$\begin{aligned} E_1 &= [p(0,0)]^2 + [p(0,1)]^2 + [p(1,0)]^2 + [p(1,1)]^2 \\ &= \left[\frac{32}{79}\right]^2 + \left[\frac{15}{79}\right]^2 + \left[\frac{16}{79}\right]^2 + \left[\frac{16}{79}\right]^2 \\ &= 0,2822 \end{aligned}$$

Get energy value for *query* image as (image number 0) 0,2822.

$$\begin{aligned} c_1 &= (0-0)^2 p(0,0) + (0-1)^2 p(0,1) + (1-0)^2 p(1,0) \\ &\quad + (1-1)^2 p(1,1) \\ &= \left(0 \times \frac{32}{79}\right) + \left(1 \times \frac{15}{79}\right) + \left(1 \times \frac{16}{79}\right) + \left(0 \times \frac{16}{79}\right) \\ &= 0,3924 \end{aligned}$$

Obtained the contrast value for *query* image (image number 0) as 0,3924.

$$\begin{aligned} H_1 &= \frac{p(0,0)}{1+|0-0|} + \frac{p(0,1)}{1+|0-1|} + \frac{p(1,0)}{1+|1-0|} + \frac{p(1,1)}{1+|1-1|} \\ &= \frac{\frac{32}{79}}{1+0} + \frac{\frac{15}{79}}{1+1} + \frac{\frac{16}{79}}{1+1} + \frac{\frac{16}{79}}{1+0} \\ &= 0,8038 \end{aligned}$$

Obtained the homogeneity value for *query* image as 0,8038.

Get the intensity average gray as :

$$\mu_1 = \frac{1}{99} \times ((67 \times 0) + (32 \times 1)) = 0,3232$$

Get the intensity average gray for *query* image as 0,3232.

Obtained the standart deviation value as:

$$\begin{aligned} \sigma_1 &= \sqrt{\frac{1}{99} \times [(0-0,3232)^2 \times 67 + (1-0,3232)^2 \times 32]} \\ &= \sqrt{\frac{1}{99} \times [(-0,3232)^2 \times 67 + (0,6768)^2 \times 32]} \\ &= \sqrt{\frac{1}{99} \times [0,1044 \times 67 + 0,458058 \times 32]} \\ &= \sqrt{\frac{1}{99} \times [6,9948 + 14,6578]} \\ &= \sqrt{\frac{1}{99} \times [21,6526]} \\ &= \sqrt{0,2187} \\ &= 0,4676 \end{aligned}$$

Obtained the standart deviation from gray intensity for *query* image (image number 0) as 0,4676.

After the value contained in the *co-occurrence* matrix from image number eighth, substituted in to formula, as:

$$\begin{aligned} e_2 &= -(p(0,0)\log p(0,0) + p(0,1)\log p(0,1) + p(1,0) \\ &\quad \log p(1,0) + p(1,1)\log p(1,1)) \\ &= -\left(\left(\frac{50}{84} \times \log \frac{50}{84}\right) + \left(\frac{14}{84} \times \log \frac{14}{84}\right) + \left(\frac{14}{84} \times \log \frac{14}{84}\right)\right. \\ &\quad \left.+ \left(\frac{10}{84} \times \log \frac{10}{84}\right)\right) \\ &= -((0,596 \times \log 0,596) + (0,167 \times \log 0,167) \\ &\quad + (0,167 \times \log 0,167) + (0,214 \times \log 0,214)) \\ &= 0,2822 \end{aligned}$$

Obtained we get the entropy value for *database* image (image number eight) as 0,2822.

$$\begin{aligned} E_2 &= [p(0,0)]^2 + [p(0,1)]^2 + [p(1,0)]^2 + [p(1,1)]^2 \\ &= \left[\frac{50}{84}\right]^2 + \left[\frac{14}{84}\right]^2 + \left[\frac{14}{84}\right]^2 + \left[\frac{10}{84}\right]^2 \\ &= 0,4240 \end{aligned}$$

Obtained we get the energy value from *database* image (image number eight) as 0,4240

$$\begin{aligned} c_2 &= (0-0)^2 p(0,0) + (0-1)^2 p(0,1) + (1-0)^2 p(1,0) \\ &\quad + (1-1)^2 p(1,1) \\ &= \left(0 \times \frac{50}{84}\right) + \left(1 \times \frac{14}{84}\right) + \left(1 \times \frac{14}{84}\right) + \left(0 \times \frac{10}{84}\right) \\ &= 0,333 \end{aligned}$$

Obtained the contrast value for *database* image (image number eight) as 0,333.

$$H_2 = \frac{p(0,0)}{1+|0-0|} + \frac{p(0,1)}{1+|0-1|} + \frac{p(1,0)}{1+|1-0|} + \frac{p(1,1)}{1+|1-1|}$$

$$= \frac{\frac{50}{84}}{1+0} + \frac{\frac{14}{84}}{1+1} + \frac{\frac{14}{84}}{1+1} + \frac{\frac{10}{84}}{1+0}$$

$$= 0,8810$$

Obtained the homogeneity value for for *database* image (image number eight) as 0,8810.

Get the intensity average gray as :

$$\mu_2 = \frac{1}{108} \times ((84 \times 0) + (24 \times 1)) = 0,2222$$

Get the intensity average gray for *database* image as 0,2222

Obtained the standart deviation value as :

$$\sigma_2 = \sqrt{\frac{1}{108} \times [(0 - 0,2222)^2 \times 84 + (1 - 0,2222)^2 \times 24]}$$

$$= \sqrt{\frac{1}{108} \times [(-0,2222)^2 \times 84 + (0,7778)^2 \times 24]}$$

$$= \sqrt{\frac{1}{108} \times [(0,04937 \times 84) + (0,6049 \times 24)]}$$

$$= \sqrt{\frac{1}{108} \times [4,147 + 14,5176]}$$

$$= \sqrt{\frac{1}{108} \times [18,6646]}$$

$$= \sqrt{0,1728}$$

$$= 0,4157$$

Obtained the standart deviation from gray intensity for *database* image (image number eight) as 0,4157.

#### 4) Calculated The Distance Of Texture Feature Between Query Image and Database Image with Euclidean Distance Method

The formula to measured the distance of texture features between the *query* image with the *database* image are:

$$d_{ab} = \|ab\| = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2 + \dots + (a_n - b_n)^2}$$

Subtited the value of each texture feature of the *query* image (image number 0) and the *database* image (image number eight) where the vector of feature for the *query* image and the *database* image are:

The vector featurefor *query*image(image number 0) denoted by A where,

$$A = [e_1, E_1, c_1, H_1, \mu_1, \sigma_1]$$

$$A = [1,3283, 0,2822, 0,3924, 0,8038, 0,3232, 0,4676]$$

The vector feature for *database*image(image number eight) denoted by B where,

$$B = [e_2, E_2, c_2, H_2, \mu_2, \sigma_2]$$

$$B = [2,2301, 0,4240, 0,333, 0,8810, 0,2222, 0,4157]$$

Substituted the values on vectors feature A and B to euclidean distance equations to measure the level of texture feature similarity between the *query* image (image number 0) and the *database* image (image number eight) is obtained:

$$d_{AB} = \sqrt{(e_1 - e_2)^2 + (E_1 - E_2)^2 + (c_1 - c_2)^2 + (H_1 - H_2)^2 + (\mu_1 - \mu_2)^2 + (\sigma_1 - \sigma_2)^2}$$

$$d_{AB} = \sqrt{(1,3283 - 2,2301)^2 + (0,2822 - 0,424)^2 + (0,3924 - 0,333)^2 + (0,8038 - 0,8810)^2 + (0,3232 - 0,2222)^2 + (0,4676 - 0,4157)^2}$$

$$= \sqrt{(1,3283 - 2,2301)^2 + (0,2822 - 0,424)^2 + (0,3924 - 0,333)^2 + (0,8038 - 0,881)^2 + (0,3232 - 0,2222)^2 + (0,4676 - 0,4157)^2}$$

$$= 0,925058533$$

So the distance of texture features between the query image (image number 0) with the database image (image number eight) is 0,925058533.

#### 5) The Process of Design the Application for Identify The Image Of The Segmentation Result Through The Introduction Of The Texture Pattern By The Euclidean Distance Method



Fig. 4. Application program measure the similarity of segmentes image

#### IV. CONCLUSION

In this study, the conclusions are : (1) The similarity of two image measure by introducing texture pattern with Euclidean distance method. The image must be segmented first using minimum spanning tree method; (2) The process to measure the similarity of two image are : first process is to segment the image of the database. The second process specifies the co-occurrence matrix for the database image with the query image. So we get a co-occurrence matrix, where the entry value contained in the matrix will be used to determine the values of each texture feature. The third process is to perform texture feature analysis for image database. The fourth is to perform texture feature analysis

for query image. And the last process is to calculate the distance of each texture feature in the database image and the query image with the Euclidean distance method.

#### REFERENCES

- [1] Ahmad. Usman, *Pengolahan Citra Digital dan Teknik Pengolahan*, Penerbit Graha Ilmu: Bogor, 2005.
- [2] Alvini. Silvia, “Penerapan Pohon Rentang Minimum Pada Graph Dalam Segmentasi Citra”, Skripsi, 85 Hal., Universitas Negeri Padang, Padang, Indonesia, September 2016.
- [3] Janakiraman,T. N & Chandra, Mouli. 2010. “Image Segmentation Using Graph Euler.” *Int. J. of Computers, Communications & Control*.
- [4] Prasetyo. Eko, *Pengolahan Citra Digital Dan Aplikasinya Menggunakan Matlab*, Gresik: Penerbit ANDI Yogyakarta, 2015.
- [5] Selfinia, “Identifikasi Citra Hasil Segmentasi Melalui Pengenalan Pola Tekstur Dengan Metode Jarak Euclidean”, Skripsi, 78 Hal., Universitas Negeri Padang. Padang, Indonesia, September 2017.