International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 1 (2018) pp. 449-455

Cotton Texture Segmentation Based On Image Texture Analysis Using Gray Level Co-occurrence Matrix (GLCM) And Euclidean Distance

Farell Dwi Aferi [1], Tito Waluyo Purboyo [2] and Randy Erfa Saputra [3]

[1]College Student, Faculty of Electrical Engineering, Telkom University, Bandung, Indonesia.

[2,3] Lecturer, Faculty of Electrical Engineering, Telkom University, Bandung, Indonesia.

[1]Orcid: 0000-0003-2983-4861, [2]Orcid: 0000-0001-98173185, [3]Orcid: 0000-0002-8537-2086

Abstract

A fabric is one of the staple materials that everyone owns. The fabric continues to evolve over time from both the functionality and style and texture. This fabric has a wide range of materials and types. One of them is a kind of cotton called combed cotton and viscose (CVC). Out of the two types of fabric it could be analyzed through the process of texture image segmentation to obtain patterns that indicate the type of cotton fabric. This research explains the cotton texture through the texture image segmentation process can give the feature value and can be used for classification, so the process can distinguish the type of cotton fabric. The process of image texture segmentation includes the gray-scaling process, image normalization and extracting feature. The Gray Level Co-occurrence matrix (GLCM) method is used in feature extraction to get the feature value later to be used for the classification process. In the classification method used is the Euclidean distance method using 4 test images and 2 training images by converting the original image into 4x4 and 6x6 pixels yielding 100% accuracy. This result shows that accuracy using Euclidean distance method produces high accuracy.

Keywords: Image texture Segmentation, Texture Analysis, Gray Level Co-Occurrence matrix(GLCM), Euclidean Distance

INTRODUCTION

Segmentation is the fundamental profession performed on images that divide the image into segments or homogeneous regions to facilitate the analysis of the images. The digital image is essentially an array into two-dimensional rectangular area known as pixels, In this gray-scale image case, the intensity of each pixel is represented by a numerical value. It is also described as a set of fixed values in digital form an element called pixels. [1] The fabric is a basic requirement which is used by many people. The fabric has a variety of material and type. [2] Among these types of cotton combed cotton and viscose (CVC). In the garment industry, cotton cloth combed has several types, among other things, 24s 20s combed cotton, 30s and 40s. Cotton is cotton fabric 40s combed which has the best quality, while cotton fabric which became the standard for the garment industry was cotton combed 30's. Because cotton combed 30's is suitable as a basic ingredient of making t-shirts. Cotton viscose (CVC) included in this type of carded cotton fabric which is very different from the combed cotton. The fundamental difference of cotton and cotton carded combed is in the process of finishing. not combed cotton fibers so remain. CVC cotton itself is not purely the result of carded cotton fabrics, because cotton to make CVC cotton combines combed with CVC cotton with a percentage of 55% cotton and 45% cotton combed CVC. Although the quality of cotton under the combed CVC cotton but the cotton has a cheaper price than the combed cotton. On cotton fabrics cotton and CVC combed their texture that can be used to perform segmentation of texture image based on image texture analysis using the method of Gray Level Co-occurrence matrix (GLCM) and classification of Euclidean distance. [3]

The method of Gray Level Co-occurrence Matrix (GLCM) used this method because it has a function for removing the image texture features of the value of cotton cloth that will be tested and then classified using Euclidean distance method to generate the value of accuracy and matching between image and imagery training test. [4]

RESEARCH METHOD

This section describes an explanation of previous research and related features of the Gray Level Co-occurrence Matrix (GLCM) feature and Euclidean classification method.

Texture Analysis of Image

Texture analysis aims to identify the exact parameters with characteristic or characteristics of objects in the image or the image. The parameters are extracted from the image or image is characteristic or characteristics of objects in the image are presented or texture in the picture or the image. [5]

The Texture is characterized by the spatial distribution of the gray degree within a set of adjacent pixels. Texture can be classified into 2 groups, namely:

1. The existence of primitive patterns of one or more pixels. These primitive patterns can be a point, a straight line, a curve, an area and others that are the basic elements of a form.
2. Primitive patterns appear repeatedly with a certain distance and direction so that it can be predicted or found characteristic repetition.

Image Texture Segmentation

Image recording against real objects in nature does not always indicate the region of image intensity or color uniform or homogeneous. For example, the image of the surface of cotton cloth is not uniform, but contain variations of the intensity of the colors that form a pattern repeat. Such a thing could be categorized as a texture. Therefore, analysis of image segmentation based on texture and texture always refers to two things: analysis of the degree of hardness of the surface and analysis of the orientation of the pattern and structure. These two things have become the basis for the development of the theory about image texture analysis. Statistical analysis approaches have been developed to know the surface roughness of texture while knowing the shape or texture pattern can be done through a structured approach. Image texture segmentation process includes three activities: acquisition image Pre-processing value, and search for gray scaling the features of the method of Grey Level Co-occurrence Matrix (GLCM). Later classification using Euclidean distance.

Gray scaling

Conversion of RGB colored image to grays image, which each pixel is a single sample of his, namely information intensity. This type of image is only made up of shades of gray that has different intensity. [6]

Normalization

The image that has been in grayscaling will get the matrix value according to the gray level. However, the matrix value that appears displays the value from 0 - 255, with the normalization process will get a new matrix value making it easier to find the value of each feature. The normalization formula as follows. [7]

$$IN = (I - Min)\frac{newMax - newMin}{Max - Min} + newMin$$

IN = Image Normalization Formula

I = Pixels to be normalized

newMax = The largest new value that has been determined

newMin = The smallest new value that has been detrmined

Max = Previous largest value level

Min = Previous smallest value level

Gray Level Co-occurrence Matrix (GLCM)

Gray level Co-occurrence matrix (GLCM) is a statistical method that can be used to analyze the textures. Co-occurrence matrix formed from an image by looking at the pixels-pixel pairs that have a certain intensity. Matrix describing the frequency of the appearance of a couple of pixels at d direction with the orientation and angle θ in the image that is used to calculate the GLCM features. Gray level Co-occurrence matrix (GLCM) is a method used to acquire an image from a trait that would result from the search features of an image that can be used as input to the image classified or classes. [8] There are four angle or direction that is used on a method of GLCM to determine relationships between pixels that have a pattern of neighborliness within a digital image. $0^{\circ}, 45^{\circ}, 90^{\circ}, 135^{\circ}$. The determination of adjacency relationship between pixels of the co-occurrence matrix is generated is illustrated in figure 1. the co-occurrence matrices obtained from spatial distance parameter d and angles θ denoted Md, θ (i,j). [9]

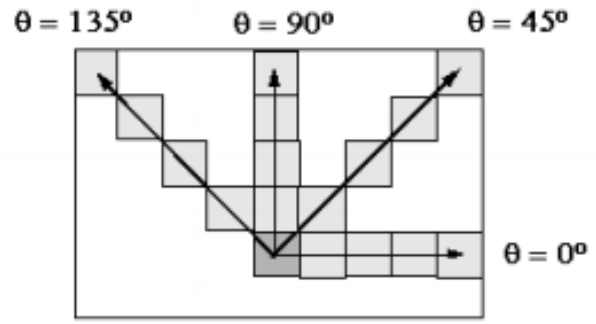

Figure 1: The direction angles that determines the adjacency to obtain ordered pairs of GLCM

As for the distance between pixels is usually defined by 1 pixel, 2 pixels, 3 pixels and so on. There are several variables used to obtain feature values from the GLCM method: [9]

i = value of row

j = value of column

In figure 2 there is an example of the co-occurrence matrix produced from the image.

1	1	5	6	8
2	3	5	7	1
4	5	7	1	2
8	5	1	2	5

Figure 2: Co-occurrence matrix construction

The above variable is used to find the character value of istik extraction features as follows:

1. Energy

It is a feature to measure the concentration of intensity pairs in matrix co-occurrence. The formula used to calculate energy is: [3] [8]

$$\text{Energy} = \sum_{i,j=0}^{n-1} p^2\ (i,j)$$

Where i and j are gray from the 2-pixel resolution that is close to each other, p (i, j) is the normalized symmetrical co-occurrence matrix.

2. Contrast

Contrast is a feature used to measure the strength of intensity differences in imagery. The contrast value enlarges if the image intensity varies high and decreases when the variation is low. The formula used to calculate contrast is: [3] [8]

$$\text{Contrast} = \sum_{i,j=0}^{n-1} pi,j\ (i-j)^2$$

Where i and j are gray from the 2-pixel resolution that is close to each other, p (i, j) is the normalized symmetrical co-occurrence matrix.

3. Homogeneity

Homogeneity is used to measure the homogeneity of variations in image intensity. Homogeneity's value will be more enlarge when the intensity variation in the image decreases. The formula used to calculate the homogeneity is: [3] [4]

$$\text{Homogeneity} = \sum_{i,j=0}^{n-1} \frac{p(i,j)}{1+|i-j|}$$

Where i and j are gray from the 2-pixel resolution that is close to each other, p (i, j) is the normalized symmetrical co-occurrence matrix.

Classification of Euclidean Distance

The classification process using Euclidean distance is to obtain accuracy and matching values between the testing image and training image. The formula in the Euclidean distance as follows. [10]

$$D = \sqrt{\sum_{i=0}^{N}(x_1 - x_2)}$$

D = Distance Value

x = Testing image Pattern

x = Training Image Pattern

Euclidean distance has a resemblance to the naïve Bayes method but this method underwent some sophistication. The simplification is due to the fact that all existing classes have equiprobable probabilities and also the same covariance matrix and are diagonal matrices.

RESULT AND ANALYSIS

In this stage, will analyze the methods will be applied to this research. the following stages are done in image texture segmentation based on texture analysis using the method of Gray Level Co-occurrence Matrix and Euclidian Distance Classification. in figure 3 is the flow of research process in form of flowchart.

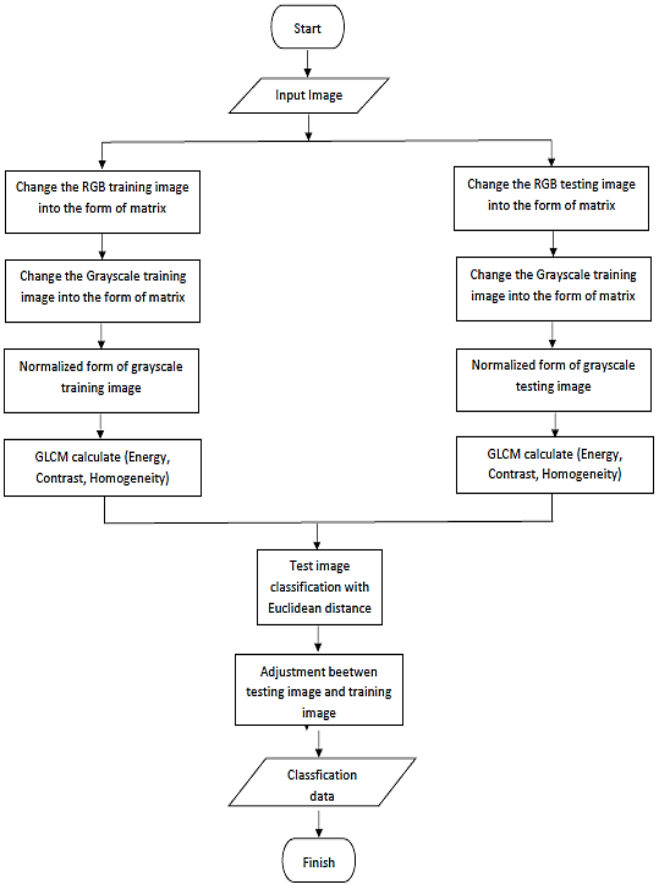

Figure 3: Flowchart system design

Image Input

Is the stage to take the image of the input media. The image is a cotton fabric that measures 150x150 pixels and is formatted .png

Figure 4: RGB Image

Change RGB image to Grayscale

The inserted image is converted to grayscale image while reducing the pixel format to 4x4 to make it easier for the next step and getting the matrix value from the grayscale image.

Resize

After the conversion from RGB mode to greyscale, it will get the value of gray scale with a range of 0-255 gray values, then the Resize process will be executed. The image that has gone through the grayscale phase by changing the range of values in the image. This process aims to save processing time.

GLCM Calculate Feature Value (Energy, Contrast, Homogeneity)

In GLCM there is a calculation of feature value. Calculate the feature value according to the direction of the designated shift to get the feature value. Feature value is used for the classification process.

Image Classification with Euclidean Distance Method

After getting the value of the feature, then the next process is the classification using the Euclidean distance method. Where the feature value is processed to get the results of the classification for the next process.

Match Between Testing image and Training Image

At this stage, the testing image and training image are matched to determine that the value of the training image has a similarity from the testing image or not.

Classification Data

After the training image and test image are matched then we get the classification data that will be used to calculate the accuracy value. Accuracy values are used to determine whether or not the classification results with the Euclidean distance method.

Method Analysis

In the analysis method, there are steps of research systems series. Where this analysis aims to analyze the method used in this study. The next step is done on image texture segmentation based on texture analysis using matrix, Gray Level Co-occurrence matrix and Euclidean Distance classification.

Image Input

Is the stage to take the image of the input media. The image is a cotton fabric that measures 150x150 pixels and is formatted .png

Analysis of Grayscaling Stages

Grayscaling stage analysis of a stage to convert RGB image into Grayscale image and get the matrix value of the image.

After the image is inserted then the image of RGB that was previously worth 150x150 pixels changed to 4x4 pixels and grayscaling process. So it removes the image matrix value in the following figure.

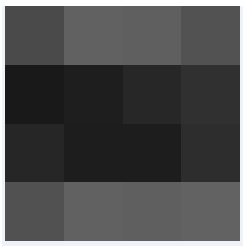

Figure 5: Grayscale Image

Image RGB in figure 5 has been resized to 4x4 pixels and became image grayscale. In figure 6, matrix values from image grayscale have been obtained.

74	96	98	82
26	28	38	49
36	29	31	43
81	97	96	99

Figure 6: A grayscale image in matrix form

Normalization

The value of the matrix in the gray scaling process is normalized to obtain new matrix values to make it easier to find the value of features present in the GLCM method and to save processing time. The following is the result of normalizing the matrix values in the grayscale image. In Figure 7, the known matrix values have been normalized.

3	2	2	4
0	0	0	0
0	0	1	0
4	2	2	4

Figure 7: Normalization Matrix

Analysis of GLCM Method Stages

The main step in the GLCM method is to take the normalized grayscale matrix value. Extraction featured image data using the GLCM cotton fabric that has been revamped into grayscale. Grayscale pixel value of each pair are counted based from 4 different angles, namely the angle $0^{\circ}, 45^{\circ}, 90^{\circ}, 135^{\circ}$. Of the 4 directions, calculated the statistical characteristics of the representative observed image. Types of Hallmark textural extracted in co-occurrence matrices here, there are three characteristics, namely: Energy, Contrast, and Homogeneity.

Table1: *Matrix Co-occurrence* 0°

GLCM (i)	Matrix Co-occurrence (j)				
	0	1	2	3	4
0	2	2	0	0	0
1	1	1	0	0	0
2	0	0	0	0	0
3	0	0	0	0	2
4	0	0	0	1	2

Table 1 is the work area matrix with an angle of 0 °

Then the matrix is summed with the matrix transpose to get the value of matrix in Table 2.

Table 2: Matrix Co-occurrence probabilities

GLCM (i)	Matrix Co-occurrence (j)				
	0	1	2	3	4
0	4	3	0	0	0
1	3	2	0	0	0
2	0	0	0	0	0
3	0	0	0	0	3
4	0	0	0	3	3

The matrix in table 2 is changed to its probability, the probability is obtained by dividing the pixel value by the total pixel value in the matrix.

The probability values obtained are shown in table 3. Where the value is later used to calculate the extraction features in the GLCM method of Energy, Contrast, and Homogeneity.

Table 3: Probabilitiy Value Co-occurrence Matrix

GLCM (i)	Matrix Co-occurrence (j)				
	0	1	2	3	4
0	0.1904	0.1428	0	0	0
1	0.1428	0.0952	0	0	0
2	0	0	0	0	0
3	0	0	0	0	0.1428
4	0	0	0	0.1428	0.1428

Then the extraction feature can be obtained as follows:

- The energy value obtained when angle 0 is as follows:

  Energy = $(0.1904)^2 + (0.1428)^2 + \cdots + (0.1428)^2 = 0.1464$

- The value of contrast obtained when angle 0 is as follows:

  Contrast = $(0-0)x(0.1904)^2 + (0-1)x(0.1428)^2 + \cdots + (4-4)x(0.1428)^2 = 0.57144$

- The Homogeneity value obtained when angle 0 is as follows:

  Homogeneity = $\frac{0.1904}{1+(0-0)^2} + \frac{0.1904}{1+(0-0)^2} + \frac{0.1904}{1+(0-0)^2} + \cdots + \frac{0.1904}{1+(0-0)^2} = 0.7143$

Using the same steps for a 45 °, 90 ° and 135 ° angle shift direction, then the value will be obtained as follows.

Table 4: Feature Values of the Input Image

	Energy	Contrast	Homogeneity
0°	0.1474	0.5714	0.7143
45°	0.124	7.389	0.126
90°	0.090	7.083	0.324
135°	0.080	6.810	0.341

Training Analysis

At this stage, the feature values obtained in the GLCM method will be used as a reference to classify the feature values contained in the image.

Table 5: Value Feature of Cotton Combed 4x4

	Energy	Contrast	Homogeneity
0°	0.1464	0.5714	0.7143
45°	0.124	7.389	0.126
90°	0.090	7.083	0.324
135°	0.080	6.810	0.341

Table 6: Value Feature of Cotton CVC 4x4

	Energy	Contrast	Homogeneity
0°	0.346	0.141	0.523
45°	0.224	3.481	0.215
90°	0.227	3.385	0.478
135°	0.290	3.379	0.246

With that results, there is two training image for classification.

Classification

Classification is done after the value in the image training has a feature value where the value of the feature will be used to classify. Here is an illustration of the classification phase using the Euclidean distance method.

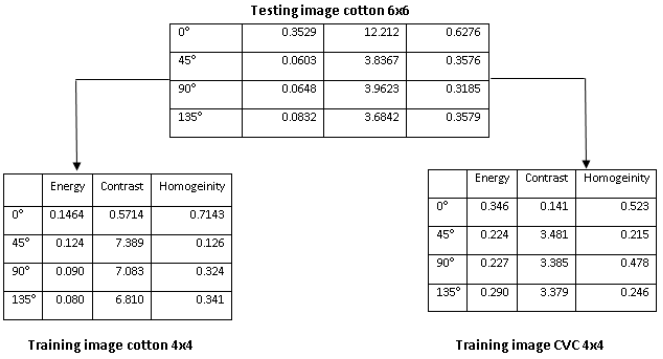

Figure 7: Illustration of Classification Euclidean Distance

D= $\sqrt{\sum_{i=1}^{n}(x1-x2)^2}$ = 10.90726

D= $\sqrt{\sum_{i=1}^{n}(x1-x2)^2}$ = 3.7611

x_1 = Testing image

x_2 = Training image cotton combed

$x_{2.1}$ = Training image cotton CVC

The data in figure 8 the classification result of Euclidean distance, the most minimum is the CVC cotton, therefore the test images are expressed as a cotton fabric with 3.7611.

Testing Analysis

In this stage will be tested by using a new texture image and matched with the existing training image on the dataset. After that, the image of the test results gets the classification data showing that the matching succeeded and issued a value of accuracy. Here are the results of a matching experiment between training image and testing image.

Table 8: Results of a matching experiment using Euclidean Distance

	Cotton	CVC
Cotton 4x4	0	4.1123
Cotton 6x6	10.908	6.8047
CVC 4x4	4.2212	0
CVC 6x6	7.8651	3.7611

In the table above it is known that the test result using Euclidean distance method resulted in the distance of CVC 4x4 is 0. Because the image is actually a training image be also being used as the test images, there by producing an exact resem balance.

The recognition similarity is obtained by calculating the Euclidean distance on each test image, the smaller the distance value the more likely it will be. So that image texture is sorted by type and the minimum distance value. The following table sorting the least distance value according type.

Table 9: Sorting of minimum value Euclidean Distance

Types of Input Image	Cotton	Type of Input Image	CVC
Cotton 4x4	0	CVC 4x4	0
CVC 6x6	7.8651	Cotton 6x6	6.8047
Cotton 6x6	10.908	CVC 6x6	3.7611
CVC 4x4	4.2212	Cotton 4x4	4.1123

Then classified to know the recognizable pattern of the test image in accordance with the known minimum distance value. The results of classification and accuracy value using the Euclidean distance method as follows.

Table 10: Classification and Accuracy

	Recognized Cotton	Recognized CVC
Cotton 4x4	1	0
CVC 4x4	0	1
Accuracy	100%	
Cotton 6x6	0	1
CVC 6x6	1	0
Accuracy	100%	

CONTRIBUTION

Based on previews research [11], the classification on texture that using Euclidean method for measuring acidity levels texture without determine accurate of texture values. For the research based on reference [11], features value using 5 steps, whereas in my own research is using different features, that is i only using 3 features and have accurate successfully 100%. In table 1, there is a comparison of this research.

Our research	**Research in reference** [11]
Used 2 object	Used 7 object
Using the grayscale method	Using the (SVD) method
using 3 feature values	Using 5 feature values
Got 100% accuracy from classification using Euclidean Distance	Didn’t get 100% accuracy

CONCLUSION

Based on the results of tests that have been done then got the conclusion that texture image segmentation method based on texture analysis using Gray Level Co-occurrence matrix(GLCM) can make pattern recognition which there is a value of feature used to do classification by Euclidean distance method. From the results of all tests, the Euclidean distance method generates a level of accuracy 100% with 4 testing images and 2 training images.

REFERENCES

[1] M. Gao, H. Chen, S. Zheng, B. Fang and L. Zhang, "Texture Image Segmentation Using Fused Features and Active Contour," in *International Conference on Pattern Recognition* , Mexico, 2016.

[2] R. Han and L. Zhang, "Fabric Defect Detection Method Based on Gabor Filter Mask," *Global Congres on Intelligent Systems,* pp. 1-5, 2 December 2013.

[3] B. Y. Salem and S. Nasri, "Texture Classification of Woven Fabric Based on A GLCM," in *International Multi-Conference on Systems, Signal and Devices*, Tunisia, 2009.

[4] R. K. S. H. Singh, "A Comprasion of Gray-Level Run Length and Gray-Level Co-Occurrence Matrix Toward Cereal Grain Classification," *International Journal of Computer Engineering & Technology (IJCET),* vol. 7, pp. 9-17, 2016.

[5] J. Malik, S. Belongie, T. Leung and J. Shi, "Contour and Texture Analysis for Image Segmentation," *International Journal of Computer Vision,* vol. 1, no. 43, pp. 7-27, 2001.

[6] K. Nurzynska, M. Kubo and K.-I. Muramoto, "Grey Scale Texture Classification Method Comparison

Considering Object and Lighting Rotation," *International Journal of Computer Theory and Engineering,,* vol. 5, no. 1, pp. 1-5, 2013.

[7] Wikipedia, "Wikipedia," Wikipedia, [Online]. Available: https://en.wikipedia.org/wiki/Normalization_(image_processing). [Accessed 21 11 2017].

[8] B.-h. Wang, H.-j. Wang and H.-n. QI, "Wood Recognition Based on Grey-Level Co-Occurrence Matrix," in *International Conference on Computer Application and System Modeling*, China, 2010.

[9] K. Meshkini and H. Ghasemmian, "Texture Classification Using Shearlet Transform and," in *2017 25th Iranian Conference on Electrical Engineering (ICEE)*, iran, 2017.

[10] T. M. Cole and J. T. Richtsmeier, "A Simple Method for Visualization of Influential Landmarks When Using Euclidean Distance Matrix Analysis," *AMERICAN JOURNAL OF PHYSICAL ANTHROPOLOGY,* vol. 107, pp. 273-283, 1998.

[11] Z. Xie, G. Liu, C. He and Y. Wen, "Texture Image Retrieval Based on gray level cooccurrence Matrix and Singular Value Decomposition," in *Multimedia Technology (ICMT), 2010 International Conference on*, Ningbo, 2010.