Batik Image Retrieval Based on Enhanced Micro-Structure Descriptor

Agus Eko Minarno*, Yuda Munarko*, Fitri Bimantoro†, Arrie Kurniawardhani† and Nanik Suciati†
*Jurusan Teknik Informatika, Universitas Muhammadiyah Malang, Malang, Indonesia
Email: agoes.minarno@gmail.com and yuda.munarko@gmail.com
†Institut Teknologi Sepuluh November, Surabaya, Indonesia
Email: beema.yichi@gmail.com, emailkuwardhani@gmail.com and nanik@if.its.ac.id

***Abstract*—This paper describes a novel method for extracting features of batik images. This method is called enhanced micro-structure descriptor (EMSD). EMSD is the enhanced model of micro-structure descriptor (MSD) which proposed by Guang-Hai Liu. Different with MSD that uses only edge orientation similarity for creating micro-structure map and then utilises this map along with color values; EMSD adds a new micro-structure map that is based on color similarity and then utilises this map along with edge orientation values. The combination of MSD and the additional micro-structure descriptor is used as feature extractor in EMSD. This method is tested on 300 batik images, Corel datasets with 5,000 images and 10,000 images. We also compared EMSD to MSD and multi-textons histogram (MTH), which EMSD performance is superior than the other two.**

***Index Terms*—image retrieval, batik, micro-structure descriptor**

I. Introduction

Several studies have been proposed for the classification of batik images. Nurhaida et al.[1] performed comparison of three feature extraction methods, those are Gray Level Co-occurrence Matrices (GLCM), Canny Edge Detection, and Gabor. The result shows that GLCM is superior to the others. Cheong and Loke [2] used Tchebichef Orthogonal Polynomials to get shape information from Multispectral Co-occurrence Matrix. The information is used to recognize and to classify the texture of colorful batik images and songket images. They indicated that the proposed method has higher accuracy compared to GLCM in terms of colorful texture classification. Furthermore, Cheong and Loke [2] also utilized Principal Component Analysis (PCA) in order to reduce 2% of features. This reduction can speed up the process of recognition and classification on batik images and songket images. Rangkuti et al. [3] developed Content-based Image Retrieval (CBIR), a system for batik image retrieval using the canny edge detection method and shape invariant moment method for texture and shape feature extraction. They also used Canberra distance for similarity measurement. However, this method cannot work optimally for a batik image with complicated motifs. In fact, the batik image may have several textures; therefore, a different class may have a similar pattern [1]. Nevertheless, there is a need for a better method for features extraction for a batik CBIR system.

CBIR is a retrieval system to find similar images using a matching process that is based on similarity measurement. There are different kinds of feature extraction methods applied to CBIR for color, texture and shape features which utilise co-occurrence matrices, for example, texton co-occurrence matrix [4], multi-textons histogram [5], and micro-structure descriptor (MSD) [6]. MSD is one method that reliable, simple and efficient for features extraction. MSD describes color, texture and shape features simultaneously based on color micro-structure and edge-orientation similarity. Furthermore, in this paper we propose a method for features extraction named enhance micro-structure descriptor (EMSD). This method is based on MSD by adding edge orientation features for the matching process. We found that our method is in a position to improve precision and recall.

II. Micro-structure Descriptor (MSD)

MSD features extraction method is inspired by the structural approach called texel (texture element). An example of texel is Julesz's texton which describes the structure of texture using texton [7]. However, texton theory emphasizes only certain natural texture and only works on grayscale images. Based on this problem, MSD was developed for extracting natural image features universally and utilizing color images. The main concept of MSD is the effective extraction of images content which consist of semantic meaning. Such features are then defined in a map structure. The goal is that the image can be differentiated or matched with other images. MSD describes the spatial correlation of images.

The first step of MSD is changing an image in RGB color space to image in HSV color space. Then, the changed image is quantized into 72 bins. The bins is composed by quantizing HSV color space to H=8, S=3, and V=3. The second step is detecting edge orientation using Sobel operator on each pixel, and then quantizing the theta angel to 6 bins of 180 degrees. So each bin has 30 degrees range. The third step is defining micro-structure from quantized edge. The results of this step are stored as micro-structure map. The final step is taking values of quantized color using micro-structure map as a mask, resulting histogram with 72 features. These features then are used to distinguish one image from the others. The MSD illustration can be seen on Fig. 1. Fig. 1(a) is the result of edge detection that has been quantized. Fig. 1(b) is the micro structure detection process, by looking for a value equal to the value of the center using a 3x3 kernel. Fig. 1(c) is the result of

978-1-4799-4568-9/14/$31.00 ©2014 IEEE

micro-structure detection. Fig. 1(d) is a micro-structure map. Fig. 1(e) is the combination of quantized color and micro-structure map as a mask. Fig. 1(f) is color features based on a micro-structure map.

III. Enhanced Micro-structure Descriptor (EMSD)

When MSD is only using 72 color features, EMSD is using 78 features with 6 new features. These new features are the value of edge orientation which has been quantized from 180 degrees to 6 bins. Steps being taken to EMSD are following Liu et al. [6] work.

The first step of EMSD is the conversion of RGB color into HSV color, and then quantized the result to H=8 bins, S=3 bins, and V=3 bins, so there are 72 bins in total which called color features.

The second step is the detection of edge orientation using Sobel operator, and then quantized the angle to 6 bins of 180 degrees. Furthermore, these 6 bins are used as edge orientation features and then added to 72 color features. So, in total, there are 78 features.

The third step is defining edge micro-structure maps using a 3x3 kernel. All kernel positions which have a same edge orientation value to the centre position are marked as map pattern. A similar process is also conducted to define color micro-structure maps. However, these maps are based on color histogram values rather than edge orientation value.

The final step is the acquisition of color intensity using edge micro-structure map and the acquisition of edge orientation using color micro-structure map. The results of this step are 72 color features and 6 edge orientation features. A detail description of each ESDM steps are described on the following section.

A. HSV color space and color quantization

Color in an image provides highly reliable spatial information for image retrieval and object recognition. Color histogram is widely used as features for image retrieval. The use of RGB color space is easy to use in practice, but does not describe the visual perception of the human eye. HSV color space is composed of a hue component that describes the type of color; by a saturation component that describes the purity of color or how big a color mixed with white; and by component value that describes color brightness.

The hue component (H) has a range of values between 0-360 degrees form a cylinder, the Saturation component (S) has a range value of 0-1 and the Value component (V) has a range of values 0-1. In this paper we quantize HSV color space into 72 bins consisting of 8 bins for hue component, 3 bins for saturation component, and 3 bins for component values. So the image which has been quantized has 72 features of colors.

B. Edge orientation detection in HSV color space

There are many methods for edge detection, such as the Robert operator, the Prewit operator, the Canny operator and the LOG operator. In this paper edge detection is performed on each component of HSV color space using the Sobel operator. In Cartesian space, a vector multiplication of *a(x1,y1,z1)* and *b(x2,y2,z2)* is defined as:

$$ab = x_1x_2 + y_1y_2 + z_1z_2 \tag{1}$$

$$cos(a,b) = \frac{ab}{|a||b|} = \frac{x_1x_2 + y_1y_2 + z_1z_2}{\sqrt{x_1^2 + y_1^2 + z_1^2}\sqrt{x_2^2 + y_2^2 + y_2^2}} \tag{2}$$

HSV color space is based on cylindrical coordinates, therefore it should be converted into Cartesian color space. If *(H,S,V)* is a cylindrical coordinate system and *(H',S',V')* is a Cartesian coordinate, so *H' = S.cos(H), S'=S.sin(H)* and *V'=V*. The Sobel operator is used to detect an edge in *H', S'* and *V'*. The result of the Sobel operator is a gradient x and y which is symbolized by *a(H'x, S'x, V'x)* and *b(H'y, S'y, V'y)* where *H'x* is horizontal gradient and *H'y* is vertical gradient.

$$|a| = \sqrt{(H'x)^2 + (S'x)^2 + (V'x)^2} \tag{3}$$

$$|a| = \sqrt{(H'y)^2 + (S'y)^2 + (V'y)^2} \tag{4}$$

$$ab = H'xH'y + S'xS'y + V'xV'y \tag{5}$$

Edge between a and b is

$$cos(a,b) = \frac{ab}{|a||b|} \tag{6}$$

$$\theta = arccos(ab) = arccos\left[\frac{ab}{|a||b|}\right] \tag{7}$$

After calculating the angel (θ) of edge orientation, the next step is quantizing this value into *m* bins where *m*={*6,12,18,24,30,36*}. In this case we used 6 bins with 30 degrees interval for each bin.

C. Micro-structure definition and map extraction

In order to define the micro-structure of an image, a 3x3 kernel is used. The kernel is used to localize a particular block of image. A micro-structure pattern is defined by looking for values under the kernel which equal to the midpoint value. For a micro-structure pattern detection on the edge of an image, there is an additional padding, so the detection process still can be performed. For all parts of the image, the kernel is moved from right to left and from top to bottom by two pixels each time shifted. Finally, all of the micro-structure patterns are combined to form a micro-structure map.

This process, then, is repeated from different starting points to address the possibility of overlapping on the detection of the final micro-structure map. Actually, there are four micro-structure maps, those are M1, M2, M3, and M4. The detection of M1 is started at (0,0), M2 is started at (1,0), M3 is started at (0,1), and M4 is started at (1,1). The combination of these four maps form the final micro-structure map. For this purpose, the following equation is used:

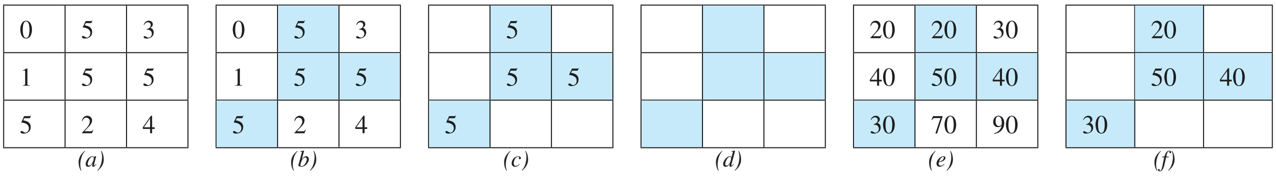


Fig. 1: An example of micro-structure detection. (a) result of edge quantization, (b) detection of micro-structure, (c) detection result, (d) a micro-structure map, (e) quantized color with a micro-structure map as mask, (f) color features based on a micro-structure map.

$$M(x,y) = Max(M1(x,y), M2(x,y), M3(x,y,), M4(x,y)) \tag{8}$$

In our approach, there are two types of final micro-structure map, those are the edge micro-structure map and the color micro-structure map. The edge micro-structure map is based on quantized edge orientation values and the color micro-structure map is based on quantized HSV color values. The illustration of the edge micro-structure map detection is shown at Fig. 2(a-g) and the illustration of the color micro-structure map detection is shown at Fig. 2(j-k).

D. Micro-structure feature representation

By utilizing the final micro-structure maps, color features and edge features are detected. The process of getting color features is performed by masking the edge micro-structure map to quantized HSV color values, producing 72 color features. Then, the process of getting edge features is performed by masking the color micro-structure map to the quantized edge orientation values, producing 7 edge features. Therefore, totally there are 78 features. The illustration of the detection of edge features and color features are described in Fig. 2(h-i,l).

IV. Experimental result

In our experiment, we compared the performance of EMSD to MSD and MTH. We tested the retrieval performance on batik dataset and Corel datasets. The experiment was conducted using the best quantization size for each method. Specifically, we used 72+6=78 bins in EMSD, 72 bins in MSD and 82 bins in MTH.

Our initial experiment is aimed to determine the best quantization size of color intensity and edge orientation. In this experiment, we found that the best quantization is 72 for color intensity and 6 for edge orientation. The next experiment is the comparison of recall and precision between EMSD and MSD on batik dataset. We were also comparing recall and precision of EMSD, MSD and MTH on Corel dataset with 5,000 images (Corel-5000)and Corel dataset with 10,000 images (Corel-10000). Finally, the last experiment is to contrast the efficiency of L1 distance matrix and the Modified Canberra distance matrix.

Batik dataset that was used in these experiments are composed by 4 data train and 2 data test from each class. These data were randomly chosen from 50 classes, so consequently there were 200 data train and 100 data test. For Corel datasets, we were randomly chosen 50 data test from each class. The Corel-10000 dataset contains 100 classes and the Corel-5000 dataset contains 50 classes, therefore, there were 5,000 data test and 2,500 data test for each Corel dataset respectively. In all experiments, for batik dataset we retrieved the best 4 similar images and for Corel datasets we retrieved the best 12 similar images. For all experiments, we utilised two types of distance matrix, those are L1 and the Modified Canberra.

A. Dataset

We used two types of datasets in experiments, those are batik dataset and Corel datasets. The first Corel dataset consists of 5,000 images (Corel-5000), and the second Corel dataset consists of 10,000 images (Corel-10000).

Batik image dataset is derived from 50 captured images of batik cloth. Then, each captured images are cut randomly to get 6 batik images. So, there are 300 batik images in dataset, with 50 classes, where each class has 6 batik images. Examples of batik images are shown in Fig. 3.

Corel dataset is a collection of Corel images which are already classified into 100 classes, such as stained glass, racing car, fruit, texture, bus, flower, and air plane. Each class is consisted of 100 images in JPEG format. Corel-5000 dataset is a subset of Corel-10000 dataset. However, the nature of this dataset is different with the Corel dataset that used by Liu et al. [6] which consisting various type of classes and having various type of images in each class.

B. Distance metric

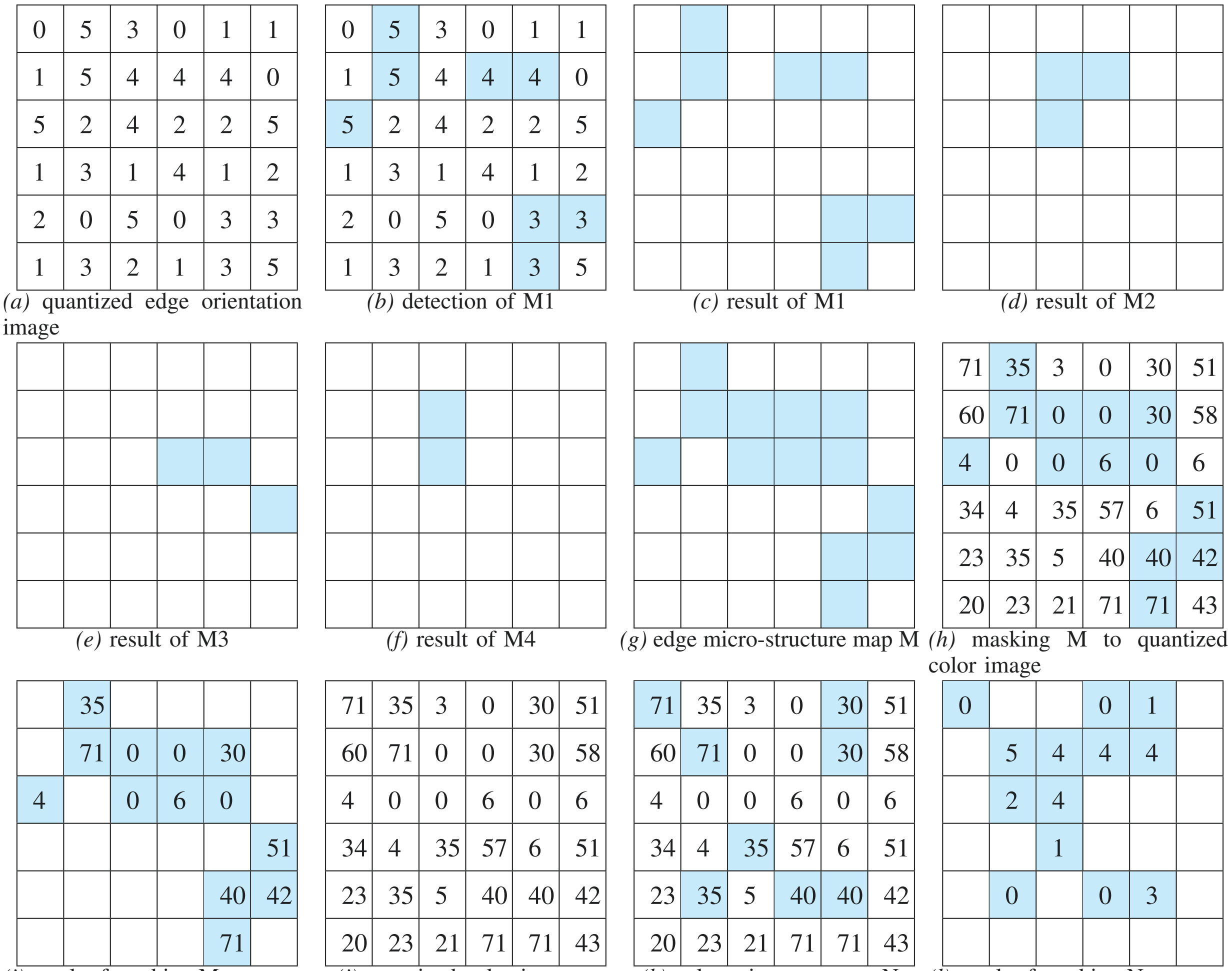
Similarity measurement in this experiment uses L1 and the Modified Canberra. For each data train *T=[T1,T2,...,Tm]* in dataset, if there are data test *Q=[Q1,Q2,...Qm]*, distance measurements using L1 and the Modified Canberra in order are:

$$D(T,Q) = \sum_{i=1}^{M} |T_i - Q_i| \tag{9}$$

$$D(T,Q) = \sum_{i=1}^{M} \frac{|T_i - Q_i|}{|T_i + \mu T| + |Q_i + \mu Q|} \tag{10}$$

C. Performance evaluation

In this experiment, performance of EMSD is measured using precision and recall which are defined as follow:



(a) quantized edge orientation image *(b)* detection of M1 *(c)* result of M1 *(d)* result of M2 *(e)* result of M3 *(f)* result of M4 *(g)* edge micro-structure map M *(h)* masking M to quantized color image *(i)* result of masking M to quantizer color image *(j)* quantized color images *(k)* color micro-structure N *(l)* result of masking N to quantized edge orientation image

Fig. 2: Example of features extraction using EMSD. (a-i) the detection of edge micro-structure map and color features, (j-l) the detection of color micro-structure map and edge features[6].

Fig. 3: An example of batik image dataset.

Fig. 4: An example of image retrieval. (the most left) image query, (the rest) results of four image retrievals.

$$P(N) = I_N/N \tag{11}$$

$$R(N) = I_N/M \tag{12}$$

Where I_N is the number of similar retrieved images, *N* is the number of relevant images, and *M* is the number of all relevant data in dataset.

D. Retrieval result

We conducted three types of experiments. The first experiment if for discovering the best quantization size of edge orientation. The second experiment is comparing MSD and EMSD based on their best quantization size. The final experiment is exploring performance EMSD, contrasted with MSD and MTH on Corel-5000 and Corel-10000.

In the first experiment, for color quantization, we used 72 bins of color features as proposed by Liu et al. [6] and we were only concerning on the quantization size of edge orientation. This experiment's result shown that the best quantization size of edge orientation is 6 bins. Furthermore, those bins become additional features in an image retrieval process. So, totally there are 72 + 6 = 78 features. This result is just the same as Liu et al. [6] result.

The main difference between MSD and EMSD is the representation of micro-structure. MSD is only used color intensity that identified by edge micro-structure map as image features. Enhancing MSD, EMSD is not only used color intensity, but also used edge orientation values that identified by color micro-structure map.

The result of the first experiment is shown at Table I. We can see that the best quantization size is 6 and 36. However, we may agree that a smaller size will contribute to the fastest computation process, therefore we choose 6 as the best bins. The measurement process was performed using L1 and the Modified Canbera distance matrix, which can be seen that L1 performs better.

The second experiment is the comparison between MSD and the proposed method (EMSD) on batik dataset. We can see the result in Table II that shows that EMSD performance is higher than MSD. Using the Modified Canberra distance matrix, the rate of EMSD precision and recall are 0.13 and 0.14 higher than MSD respectively. When using L1 distance matrix, the rate of EMSD precision is 0.02 higher than MSD and the rate of EMSD recall is 0.16 higher than MSD.

To demonstrate a fair comparison, we also compared EMSD to MSD and MTH on Corel-5000 and Corel-10000. The result of the experiment is presented at Table III. The table shows that the rate of EMSD precision and recall are 0.04 and 0.01 higher than MSD respectively; and 0.03 and 0.01 higher than MTH respectively.

TABLE I: Comparison of edge orientation bin size combined with 72 color features for retrieving batik image

Canberra	Edge orientation						
	6	8	12	18	24	30	36
Precision	0,62	0,61	0,60	0,61	0,60	0,57	0,62
Recall	0,62	0,61	0,60	0,61	0,60	0,57	0,62

L1	Edge orientation						
	6	8	12	18	24	30	36
Precision	0,85	0,82	0,79	0,79	0,74	0,75	0,85
Recall	0,85	0,82	0,79	0,79	0,74	0,75	0,85

TABLE II: Comparison between MSD and EMSD on batik dataset.

Canberra	MSD		EMSD	
Retrieval	Precision	Recall	Precision	Recall
4	0.60	0.60	0.62	0.62
6	0.47	0.70	0.48	0.72
Average	0.54	0.65	0.67	0.79

L1	MSD		EMSD	
Retrieval	Precision	Recall	Precision	Recall
4	0.84	0.84	0.85	0.85
6	0.60	0.62	0.62	0.93
Average	0.72	0.73	0.74	0.89

TABLE III: Comparison between MSD, MTH and EMSD on Corel-5000 and Corel 10000

	MSD		MTH		EMSD	
Dataset	Precision	Recall	Precision	Recall	Precision	Recall
Corel-5000	0.35	0.08	0.36	0.08	0.40	0.09
Corel-10000	0.18	0.03	0.17	0.03	0.21	0.04
Average	0.26	0.06	0.27	0.06	0.30	0.07

V. Conclusion

In this paper we propose an improvement of MSD to describe micro-structure. The improvement is simple, yet efficient by adding edge features based on color micro-structure map as additional features. The color micro-structure map is derived from co-occurence of color quantization on a 3x3 kernel. From conducted experiments, we conclude that the use of edge orientation as additional feature on micro-structure may improve image retrieval's recall and precision. We also found that EMSD performance is higher than MSD and MTH.

References

[1] I. Nurhaida, R. Manurung, and A. M. Arymurthy, "Performance comparison analysis features extraction methods for batik recognition," in *Advanced Computer Science and Information Systems (ICACSIS), 2012 International Conference on*. IEEE, 2012, pp. 207–212.

[2] M. Cheong and K.-S. Loke, "An approach to texture-based image recognition by deconstructing multispectral co-occurrence matrices using tchebichef orthogonal polynomials," in *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*. IEEE, 2008, pp. 1–4.

[3] A. H. Rangkuti, R. B. Bahaweres, and A. Harjoko, "Batik image retrieval based on similarity of shape and texture characteristics," in *Advanced Computer Science and Information Systems (ICACSIS), 2012 International Conference on*. IEEE, 2012, pp. 267–273.

[4] G.-H. Liu and J.-Y. Yang, "Image retrieval based on the texton co-occurrence matrix," *Pattern Recognition*, vol. 41, no. 12, pp. 3521–3527, 2008.

[5] G.-H. Liu, L. Zhang, Y.-K. Hou, Z.-Y. Li, and J.-Y. Yang, "Image retrieval based on multi-texton histogram," *Pattern Recognition*, vol. 43, no. 7, pp. 2380–2389, 2010.

[6] G.-H. Liu, Z.-Y. Li, L. Zhang, and Y. Xu, "Image retrieval based on micro-structure descriptor," *Pattern Recognition*, vol. 44, no. 9, pp. 2123–2133, 2011.

[7] B. Julesz, "Textons, the elements of texture perception, and their interactions." *Nature*, 1981.