Re-Ranking Image Retrieval on Multi Texton Co-Occurrence Descriptor Using K-Nearest Neighbor

1st Yufis Azhar
Informatics
Universitas Muhammadiyah Malang
Malang, Indonesia
yufis@umm.ac.id

2nd Agus Eko Minarno
Informatics
Universitas Muhammadiyah Malang
Malang, Indonesia
aguseko@umm.ac.id

3rd Yuda Munarko
Informatics
Universitas Muhammadiyah Malang
Malang, Indonesia
yuda@umm.ac.id

***Abstract*— Some features commonly used to conduct image retrieval are color, texture and edge. Multi Texton Co-Occurrence Descriptor (MTCD) is a method which uses all three features to perform image retrieval. This method has a high precision when doing retrieval on a patterned image such as Batik images. However, for images focusing on object detection like corel images, its precision decreases. This study proposes the use of KNN method to improve the precision of MTCD method by re-ranking the retrieval results from MTCD. The results show that the method is able to increase the precision by 0.8% for Batik images and 9% for corel images.**

Keywords—*MTCD, KNN*, Image Retrieval

I. INTRODUCTION

Image retrieval is one of the research topics in the identification of computer pattern and vision. There are numerous previous researches have been conducted on image retrieval system. However, it is still unresolved research topic so far. The retrieval of content-based imagery began in the 1970s when text-based image search was no longer effective [1].

In general, the image retrieval system is built using color, texture, and shape features either separately or in combination have proposed by Minarno [2], [3]. Indexing on CBIR for fast retrieved also have proposed by Munarko [4]. Color features are the most dominant image feature and the easiest to distinguish. J. Huang has proposed image indexing employing a three-dimensional table based on color and distance between pixels. The table describes the spatial relationships in image color alterations. The purpose of color indexing is that the system can distinguish an image with an image stored in the database [5]. Research on texture has also been proposed, both in the field of recognition of clustering patterns, classification and object detection. Julesz conducted, in his research, an analysis of texton interaction to texture discrimination. Moreover, a texton can consist of several pixels. The results of his research revealed that texton only using simple statistical methods was able to provide significant visual perception to distinguish texture [6]. Research on texture has also been conducted by Haralick using the Gray Co-Occurrence Matrix (GLCM). GLCM uses statistical methods of order one and two to produce 14 features, such as mean, variance, correlation, energy, homogeneity and so forth [7].

Jain et al. proposed image retrieval using feature shapes combined with color features to retrieve logo image. Jain's experiments used 400 logo images as data to test the performance of the proposed method. The image retrieval results on two most similar data obtained 99% accuracy [8]. Pradnyana also proposed classification of Endek (balinese fabric) Image using K-Nearest Neighbor obtained accuracy 91% with k value 3,4,7,8 [9]. Edge detection method one of powerful and simple technique to extract image. Syahrian et al proposed the method to detection cracks using Canny. The results show that the best value for smoothing is 10 and 5 for thresholding in getting not too blurred or to sharp result [10].

Guang Hai Liu developed a research on image retrieval using texton proposed by Julesz. The proposed method is known as Texton Co-Occurrence Matrix (TCM). TCM uses 5 types of texton as kernels to perform feature extraction which generates microstructure maps. There were 2,000 images used as data generated from Corel Dataset and Vistex MIT. The results obtained by TCM in image retrieval ranged from 41% to 43% [11].

Multi Texton Histogram (MTH) has been proposed by Liu to improve TCM performance. The difference between TCM and MTH lies in the type of texton used and the texton shift. MTH uses 4 different types of texton from TCM and switches per two pixels. Meanwhile, TCM shifts its texton per pixel. The MTH test uses Corel 5000 dataset and 10,000 images. MTH testing using Corel 5000 MTH was able to achieve 49.98% precision and 6% recall compared to TCM precision and recall achieved 27.36% and 3.28% respectively. Furthermore, the testing employing Corel 10000 MTH reaches precision of 40.87% and recall of 4.91%, and TCM reaches 20.42% precision and recall of 2.45% [12].

Guang Hai Liu et al. once again proposed an improved image retrieval system using a different approach than MTH. The proposed method is called Micro Structure Descriptor (MSD). MSD uses a 3x3 kernel variation to perform feature extraction. MSD compares the middle value of the kernel with its eight neighbors. Only the neighbor value is equal to the accumulated middle kernel value into a histogram. The MSD test uses Corel 5000 dataset and 10,000 images. At the time of using Corel 5000, MSD was able to produce 55.92% precision and 6.71% recall while testing using Corel-10.000, MSD reaches precision of 45.62% and recall of 5.48% [13].

Human visuals can basically easily distinguish the image edge orientation and color difference in the image. This is behind Guang hai Liu proposed Color difference histogram (CDH) to develop image retrieval system using psychology theory based on human visual perception. CDH is an improvement of MTH and uses the same data as MTH for testing. The results obtained by CDH were better at delivering retrieval results. The precision difference from MTH was 7.39%, and the recall was 0.89% [14].

MSD improvement has been proposed by Minarno et al. in his research using Enhanced Micro Structure Descriptor

(EMSD). EMSD adds an edge orientation feature to the MSD histogram. The performance retrieval test used 300 Batik images as its data, Corel 5000 and Corel 10000. The results obtained by EMSD provided better performance than MSD with 5% precision difference and 1% recall when using Corel 5000. While on the test using Corel 10000, it was found 3% precision difference and 1% recall [15].

MTH repair proposed again by Minarno by adding Gray Level Co-Occurrence Descriptor feature called MTCD. Minarno employed as many as 16 features using Entropy, Energy, Contrast and Corelation using 4 angles of 0, 45, 90 and 135 degree. The study was able to increase the precision by 2.8% for the Batik dataset and 3.4% for the Corel dataset.

This study was proposed to improve MTCD to improve precision by re-ranking index resulted from image retrieval using the KNN algorithm. The results from MTCD testing with KNN increased precision up to 9% for Corel dataset and 0.8% for Batik dataset.

II. RESEARCH METHOD

Multi Texton Co-Occurrence Descriptor with KNN

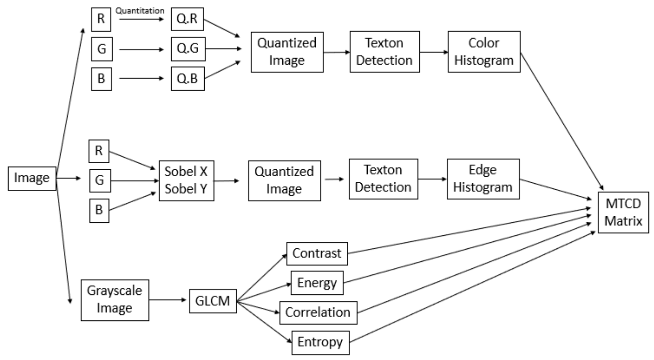

Figure 1. Flowchart of MTCD method

a. Extraction of Color Feature

According to Figure 1, it can be seen that an image will go through 3 extraction stages to acquire its features. The first stage is the color feature extraction. To do so, the first image will be broken down into 3 color channels, namely R (red), G (green) and B (blue). Afterwards, each channel will be quantized and then reunited. Following these processes, the texton detection process is completed by using 6 different texton types. A color histogram will then be arranged based on the matrix result of the texton detection process.

b. Extraction of Edge Feature

The second stage is edge extraction. The steps are almost identical as color extraction. The image will be split into 3 color channels. Consecutively, by using the sobel, each channel will be converted into grayscale image which will then be quantized and conducted texton detection as in the previous stage. This section will also be optimized in this research. Finally, a border histogram will be constructed.

c. Extraction of GLCM Feature

The third stage detects texture using GLCM method. The values which will be extracted through GLCM are energy, contrast, correlation and entropy. Finally, these three features will be incorporated into a 2-dimensional matrix. This matrix will be used for the preparation of table indexing in image retrieval.

d. Re-Ranking KNN

In this study, the proximity between images is calculated using the same method as MTCD, that is modified canberra [16]. The main difference between this study and MTCD is that once the retrieval results are obtained, the KNN method will be applied to re-ranking so that it increases the precision value. Taking the first image of the retrieval result and rearranging the image based on the number of occurrences of most of the images having the same class is the first procedure. Afterwards, the precision value will be calculated by looking at r-value of the first image, r ≤ K. The proposed redesign model is presented in Figure 2 (a). The first 2nd rank utilizes a B-categorized image.

If an A-class image is used as a query, so the precision value of the retrieval result is 0.67. With KNN, where, the image, with the most emerging class i.e. A image appearing 4 times, will be grouped into one, and placed in the top ranking. Therefore, this process will automatically shift the position of B image to the 5th position (Figure 2 (b)). In this way, the precision value will increase to 1.

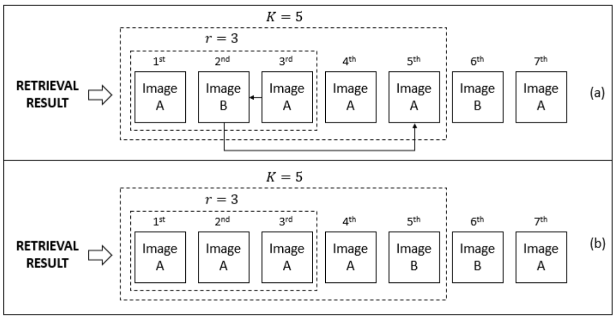

Figure 2. Re-ranking method on the retrieval results using KNN

III. RESEARCH RESULT AND DISCUSSION

A. Dataset

In this study, the dataset used consist of 2 types. The first dataset is 300 Batik images divided into 50 categories (each category has 6 pieces of image). This dataset is chosen because the image of Batik has a fairly complex and repeatable detail so that it is suitable for the application of the proposed method. The second dataset used in this study is Corel dataset. This dataset contains 10,000 images divided into 100 categories (each category contains 100 images). Unlike the first dataset which prioritizes the repetition of motifs, the categorization of corel images is based on the appearance of the same objects in an image. The example of categorization in corel dataset is categories of elephant, lion, church, fruit, vegetable, and so forth. It is a challenge to see how well the proposed method identifies the similarity of the objects which appear on two different images. Figures 1 and 2 show an example of the dataset image employed in this study.

Figure 1. Examples of Batik image dataset

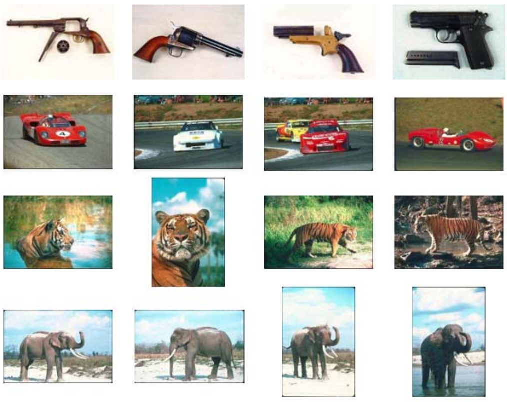

Figure 2. Examples of corel image dataset

In addition to the previously presented discrepancies, the other discrepancies between the two datasets are their dimensions. Batik datasets have the same dimension for all of its image, which is 128 x 128 pixels. As for corel dataset, there are image dimension of 126 x 187 (portrait) and also image dimension of 187 x 126 (landscape). The portrait or landscape image form actually does not affect texton search process, but because there is one dimension of odd value (187), so it needs to have special mechanism before texton search is implemented. This is because texton search is conducted by shifting in every 2 pixels, so if there is one dimension having odd value, there will be 1 row or 1 column that cannot be detected. Figure 3 (a) presents its example result, an image with dimension of 5 x 4. The 5th row cannot be detected by the system because the 6th row does not exist. Therefore, in this study, for any image whose number of rows /columns is odd, 1 additional row/column will be added with 0 for each element as presented in Figure 3 (b).

4	5	4	2
4	2	1	4
5	4	3	4
4	1	4	2
2	4	2	2

(a)

4	5	4	2
4	2	1	4
5	4	3	4
4	1	4	2
2	4	2	2
0	0	0	0

(b)

Figure 3. (a) Image having odd dimension; and

(b) Additional row for image dimension completion

B. Testing Results

The proposed method is built using the python programming language. Python is chosen because of the availability of the library to perform image retrieval. In this research, opencv library is used to perform image processing. As for calculating GLCM value, the skimage library is used.

This research uses MTCD method for feature extraction process of an image. This method will extract 3 main features, namely color, edge and texture. To extract color and edge features, Multi Texton Detection (MTD) method is employed. Meanwhile, to extract texture features, it uses GLCM method with specified parameters, energy, contrast, correlation and entropy.

Before applying color and edge extraction, an image will be quantized first. The quantization process is grouping the color ranges in an image into the n section (bit). In this study, the number of bits used for the color feature is 4 bits for each color channel. As for the edge feature, the number of bits used for is 18. The number of texton used is 6 pieces texton. This is in accordance with the previous research conducted by Minarno [16].

The contribution in this research is the application of KNN algorithm to increase precision value from MTCD method. The classification accuracy of the KNN method depends heavily on the value of K; therefore, this study utilizes different K values for each test scenario. This study uses only precision value because in image retrieval, recall value can not be used as a reference to assess quality of a method. We can just get a high recall value by taking all the image in the database. Another reason is that the image retrieved is small (when compared to the total image overall), the recall value must be very low, so this value is not significant enough to be calculated.

Table 1. Results of precision testing on Batik images

r	K	Precision	
		MTCD	MTCD + KNN
3	5	0.990	**0.997**
	7	0.990	0.983
	10	0.990	0.977
	12	0.990	0.989
	15	0.990	0.989
4	5	0.980	**0.986**
	7	0.980	**0.983**
	10	0.980	0.972
	12	0.980	**0.982**
	15	0.980	0.980
5	5	0.964	0.964
	7	0.964	**0.972**
	10	0.964	0.964
	12	0.964	**0.970**
	15	0.964	**0.969**

In the first test, Batik image dataset is used as a test object. From each category in the dataset, one image is taken for data testing, and the rest is used as training data. Hence, it produces 50 images for data testing, and 250 images of training data. Each test image is used as a query and calculated precisely from the r image successfully retrieved. The mean value of the precision value for each subsequent query will be obtained. This step is repeated as much as six times, so that each image which exist in a category has once become data testing. The precision value of the six tests will then be re-averaged. Table 1 shows that KNN method can increase the precision value of the MTCD method with the proper K value (this study uses 5 different K values). Although the increase recorded in the image of Batik is not very significant, the precision value of MTCD method itself is quite high.

From the test results, when r = 3, the proposed method can outperform MTCD method for one K-value (i.e. when K = 5). Meanwhile, the other four K-values are lower than MTCD method. When the r-value is increased to 4, then there are three K-values which get a higher precision values (i.e. when K = 5, K = 7 and K = 12). As for K = 10, the value is lower. Moreover, when K = 15, the value is the same as MTCD method. When the r-value is increased again to 5, the KNN can increase the MTCD precision value when K = 7, K = 12 and K = 15. As for K = 5 and K = 7, the precision value is identical to those of MTCD, and none of the K-value results in lower precision values than those of the MTCD method. From this test, it can be concluded that the KNN algorithm can effectively increase the precision value if the value of r is higher. Unfortunately, for the dataset of Batik image, each category has only six pieces of image, if one image is functioned as query, then the remaining images in the database are only 5 images. Therefore, the maximum r-value can be tested only to r = 5.

Furthermore, in the second test, the image of the corel dataset is employed. In this test, 2 different scenarios are performed. The first scenario, from each category, one random image will be used as data testing, and the rest will be used as data training. Thus, there are 100 data testing and 9,900 training data are obtained. This step will be repeated as much as ten times, then the average precision value of each step will be calculated. Table 2 presents the test results obtained.

Table 2. Results of precision testing for corel dataset (100 data test)

r	K	Precision		Increment (%)
		MTCD	MTCD + KNN	
5	20	0.465	0.572	10
	50	0.465	**0.623**	**15**
	100	0.465	0.554	9
12	20	0.359	0.467	11
	50	0.359	**0.524**	**16**
	100	0.359	0.488	13

In this test, the r-value used is only two, i.e. 5 and 12. While the K-value used are three, namely 20, 50 and 100. Different r-values are intended to see the consistency of conclusions obtained in the previous test. The greater the r-value, the greater the increment of the precision. From this test, it is found that for r = 5, there is a precision increase of 15%. As for r = 12, there is a 16% increase in the precision. It proves that the conclusions obtained in the previous test also apply to this test. This test also obtains the result that KNN method can increase the precision value of MTCD method for all K-values tested.

The next test scenario still uses the corel dataset. To make it different, there are 50 images taken from each category. Thus, it will get 5,000 data train and 5,000 test data. Moreover, the r-value used is 12. The testing scenario like this is also used by the previous research [16]. Table 3 presents the results of the precision testing obtained.

According to this test, the proposed method again outperformed MTCD method with varying increments, ranging from 5% to 9%. It depends on the K-value used. However, of all K-values tested, the proposed method is always superior.

Table 3. Results of the precision testing of corel dataset (5,000 data test, *r* = 12)

K	Precision		Increment (%)
	MTCD	MTCD + KNN	
20	0.288	0.348	6
30	0.288	0.369	8
50	0.288	**0.373**	**9**
100	0.288	0.344	5

IV. CONCLUSION

From the results of the tests performed, it can be seen that the re-ranking method by using KNN can improve the precision of the MTCD method. Moreover, in testing the parameter of r-value, the results show greater values leading to decreases in the precision value from the retrieval results. On the other hand, the increment of r-value will reinforce the KNN's influence in increasing the precision value of the retrieval results. It happens because the greater the r-value, the more likely the irrelevant image in the retrieval result will be higher. This is where the role of the KNN method is more noticeable because it will decrease the possibility by conducting re-ranking model. Meanwhile, the optimal K-value will be difficult to determine because it depends on the r-value used.

REFERENCES

[1] Y. Rui, T. S. Huang, and S. F. Chang, "Image retrieval: Current techniques, promising directions, and open issues," *J. Vis. Commun. Image Represent.*, vol. 10, no. 1, pp. 39–62, 1999.

[2] A. E. Minarno, A. S. Maulani, A. Kurniawardhani, and F. Bimantoro, "Comparison of Methods for Batik Classification Using Multi Texton Histogram," *TELKOMNIKA*, vol. 16, no. 3, 2018.

[3] A. E. Minarno, Y. Minarko, and A. Kurniawardhani, "CBIR of Batik Images Using Micro Structure Descriptor on Android," *Int. J. Electr. Comput. Eng.*, vol. 8, no. 6, 2018.

[4] Y. Munarko and A. E. Minarno, "HII: Histogram Inverted Index For Fast Images Retrieval," *Int. J. Electr. Comput. Eng.*, vol. 8, no. 5, 2018.

[5] J. Huang, S. R. Kumar, M. Mitra, and W.-J. Zhu, "Image Indexing Using Color Correlograms," *U.S. Pat.*, vol. 2, no. 12, pp. 12–15, 2011.

[6] B. Julesz, “Textons, the Elements of Texture Perception, and Their Interactions,” *Nat. 290*, no. 5802, p. 91, 1981.

[7] R. M. Haralick, “Statistical and structural approach to texture,” *Proceeding IEEE vol 67 no 5*, vol. 67, no. 5, pp. 786–804, 1979.

[8] A. K. Jain and A. Vailaya, “Image Retrieval Using Color and Shape,” in *Pattern Recognition 29*, 1996, pp. 1233–1244.

[9] Pradnyana, G. A., Suryantara, I. K. A., & Darmawiguna, I. G. M. (2018). Impression Classification of Endek (Balinese Fabric) Image Using K-Nearest Neighbors Method. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, *3*(3).

[10] Syahrian, N. M., Risma, P., & Dewi, T. (2017). Vision-Based Pipe Monitoring Robot For Crack Detection Using Canny Edge Detection Method as an Image Processing Technique. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, *2*(4), 243-250.

[11] G.-H. Liu and J.-Y. Yang, “Image Retrieval Based on the Texton Co-Occurrence Matrix,” in *Pattern Recognition 41*, 2008, pp. 3521–3527.

[12] G.-H. Liu, L. Zhang, Y.-K. Hou, Z.-Y. Li, and J.-Y. Yang, “Image Retrieval Based on Multi-Texton Histogram,” in *Pattern Recognition 43*, 2010, pp. 2380–2389.

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

[14] G.-H. Liu and J.-Y. Yang, “Content-based image retrieval using color difference histogram,” *Pattern Recognit.*, vol. 46, no. 1, pp. 188–198, 2013.

[15] A. E. Minarno, Y. Munarko, F. Bimantoro, A. Kurniawardhani, and N. Suciati, “Batik Image Retrieval Based on Enhanced Micro-Structure Descriptor,” in *Computer Aided System Engineering (APCASE)*, 2014, vol. 1, no. Feb, pp. 65–70.

[16] A. E. Minarno, & N. Suciati, "Image Retrieval Using Multi Texton Co-Occurrence Descriptor". *Journal Of Theoretical & Applied Information Technology*, vol. *67, no.* 1, 2014.