Color and Texture Features Extraction on Content-based Image Retrieval

Rahmaniansyah Dwi Putri, Harsa Wara Prabawa, Yaya Wihardi
Universitas Pendidikan Indonesia
rahmaniansyah.dwi.p@student.upi.edu, harsawara@upi.edu, yayawihardi@upi.edu

Abstract— The study on Content Based Image Retrieval (CBIR) has been a concern for many researchers. To conduct CBIR study, some essential things should be considered that are determining the image dataset, extraction method, and image measurement method. In this study, the dataset used is the Oxford Flower 17 dataset. The feature extraction employed is the feature extraction of the HSV color, the Gray Level Co-occurrence Matrix (GLCM) texture extraction feature, and the combination of both features. This study is purposely generates precision from CBIR test based on the proposed method. At first, digital image is segmented by applying thresholding. Moreover, the image is converted into vector to be subsequently processed using feature extraction. Further, the similarity level of the image is measured by Euclidean Distance. Tests on the system are based on segmented and unsegmented image. The system test with segmented image yields mean average precision of 83.35% for HSV feature extraction, 83.4% for GLCM feature extraction, and 80.94% for combined feature extraction. Meanwhile, the system test for unsegmented image generates mean average precision of 82.64% for HSV feature extraction, 87.32% for GLCM feature extraction, and 85.73% for extraction of combined features.

Keywords—content-based image retrieval; HSV color extraction; GLCM feature extraction

I. INTRODUCTION

Image Retrieval is one of the methods for tracing information in the picture. Image retrieval is defined as searching for image similarity with the refered image key. The techniques in the image retrieval are divided into two categories, namely text and content-based [1]. Content-Based Image Retrieval (CBIR) possesses two basic approaches to obtain information from images by using low-level features (local) such as lines, textures, colors and spatial information [2] [3].

CBIR study which implements low-level features has been conducted frequently. The study is executed with various purposes such as searching the value of data accuracy with the extraction of certain features, and finding the precision value of the system through data testing. Aside from its purpose, the selection of feature extraction evokes problem in the study. It is because both the implementation of feature extraction along with the procedure employed in CBIR can affect the results in general.

The CBIR study employing flower image as the data has been also executed by many researchers. Oxford Flower (OF) Dataset is the flower image test data provided to examine the CBIR system. Dataset OF consists of two OF102 and OF17 [4]. There are also numerous CBIR studies conducted focusing on the OF datasets with various methods [4-13]. Most of previous studies aim to discover the value of data accuracy. The features extraction used in previous studies include both color and texture features. Based on the previous research, 91.4% [7] is the best accuracy value.

This study focuses on discovering the precision value of CBIR systems using the proposed method, which is tested with the OF17 dataset. The OF17 dataset consists of 17 flower image classes - 80 images for each class. In this study, the image is segmented by using thresholding. Subsequently, the image is converted into vector and extracted using Hue Saturation Value (HSV) color feature, Gray Level Co-occurrence Matrix (GLCM) and feature combination. In the HSV feature, previous to vector conversion, the image is converted into histogram. Furthermore, to measure the image similarity, Euclidean Distance (ED) method is employed.

In this study, CBIR system test is executed based on segmented and unsegmented image. To obtain the Mean Average Precision (MAP) on one flower class, the study examines ten images in one class. One image test will generate one Mean Average Precision (AP).

The study is structured as follows: Section II reviews the related literature. Section III describes the proposed method. Section IV explicates experiments and research results. Section V explain conclusions and suggestions, and Section VI presents the references.

II. CONTENT-BASED IMAGE RETRIEVAL

The term "Content-Based Image Retrieval" is originated in 1992 introduced by T.Kato to describe experiments in automated image retrieval of databases, by color and shape [14]. The CBIR system searches for images or content that have similar characteristics from a set of images. Fig. 1, shows how the CBIR system works. The system requires input in the form of an image that will go through the feature extraction stage. In addition to processing of image query, similar process is performed for existing images in the database. The similarity calculation is performed towards the image and image query on the processed database. As the process of matching is finished, the subsequent process is to select the image in the database that possesses similar match value with the image query.

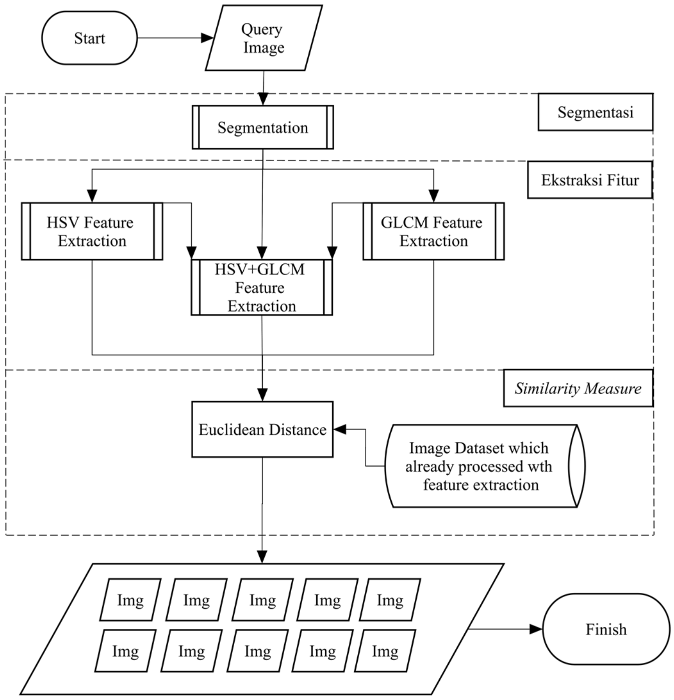

Fig. 1. CBIR system flow diagram.

A. Color Fiture Extraction

The HSV color model is a derivative of RGB color. HSV consists of Hue, Saturation, and Value. Hue represents the dominant wavelength in the light wave mixture. Saturation indicates amount of gray or the intensity level in the color space. Value refers to the brightness level so that HSV is commonly called as Hue Saturation Brightness (HSB) [14]. To get the HSV color model, it is necessary to convert the RGB value into HSV. The converted value is presented in the (1),(2) and (3).

$$H = \begin{array}{ll} 0 & \text{if} \max = min \\ 60^0 \ \frac{g-b}{max-min} + 0^0, & \text{if} \max = r \text{ and } g \ \geq g \\ 60^0 \ \frac{g-b}{max-min} + 360^0, & \text{if} \max = r \text{ and } g \ < \ b \\ 60^0 \ \frac{b-r}{max-min} + 120^0, & \text{if} \max = g \\ 60^0 \ \frac{r-g}{max-min} + 240^0, & \text{if} \max = b \end{array} \tag{1}$$

$$S = \begin{cases} 0, & \text{if} \max = 0 \\ \frac{max-min}{max} = 1 \quad \frac{min}{max} & , ot \ \ erwise \end{cases} \tag{2}$$

$$V = max \tag{3}$$

B. Texture Feature Extraction

In 1973, Haralick et al [15] introduce the co-occurrence matrix and its texture feature as statistical feature of the famous second order. Haralick et al propose two steps in using the texture extract feature: the first step is to compute the GLCMs and the second step is to compute the texture feature using calculated GLCM.

GLCM determines the probability for combining two pixels, i and j, with distance d and orientation angle θ. It is signified by $p_{d,\theta}(i, j)$. The elements *(i, j)* of $p_{d,\theta}$ are the number of occurrences of the gray level pairs i and j which have distance d in the direction θ. GLCM for images with size $N \times M$ with N_g gray level is 2D size of $N_g \times N_g$ size. An example is shown in Fig. 2, portraying GLCM for 4x4 size image with $N_g = 5$. This Array 2D is calculated with d=1 and $\theta = 0$.

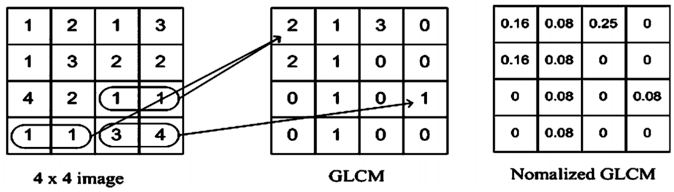

Fig. 2. GLCM for images with size of 4 x 4, with Ng = 5. This Array 2D is calculated with d = 1 and θ = 0, image source [16].

The study employs four orientations, that are 0^0,45^0, 90^0 and 135^0. After performing GLCM computation on orientation, the computation on image texture is executed, which involves contrast, correlation, energy and homogeneity [15].

III. Method

In this section, the study explains the method of developing the CBIR system utilizing the proposed method. There are three stages mentioned in the proposed method: segmentation, feature extraction and similiarity measure. System flow in this study is depicted in Fig.1.

A. Thresholding Segmentation

The first stage is image segmentation using thresholding. Digital imagery is read by system with RGB format. Subsequent, the image is converted into grayscale. Grayscale image is calculated by thresholding value to be converted into binary form. The formula of segmentation with thresholding is presented in (4). In binary image, there is reversal value in which the value of 1 means the object and the value 0 is the background. Furthermore, the image goes through a morphological process. Morphological process aims to eliminate noise in the image. Morphological methods employed are filling holes, areas, and erosion. Hereinafter, the obtained image is created as masking image. This segmentation process returns two values, RGB and segmented grayscale image (Fig. 3).

$$(v) = \begin{cases} 0 & if \ v < t \\ 1 & if \ v \geq t, \end{cases} \tag{4}$$

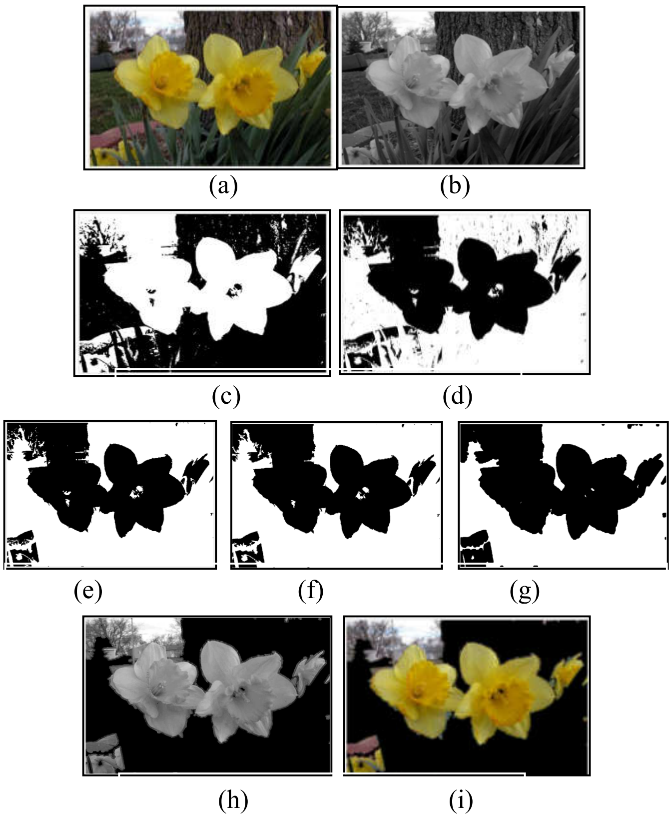

Fig. 3. Thresholding Segmentation: a) RGB image, b) Grayscale image, c) Image of thresholding result in binary form, d) Image of value reversal, e) Image of morphological result with filling holes, f) Image of area morphological result, g) Image of Erosion morphology result, h) Grayscale image with masking process, i) RGB image with masking result.

B. *Feature Extraction*

The second stage, the image is processed using feature extraction. The feature extractions employed in the study are the color feature extraction of HSV, GLCM, and a combination of both features.

- On the HSV color feature extraction stage, the segmented image is converted to HSV value. HSV image is converted to histogram, use (5).

$$h_i = \frac{n_i}{n} \quad , i = 0, 1, ..., L\text{-}1 \tag{5}$$

  Further, the calculation of HSV image's interpolation is performed. The image of the calculation result is then converted into vector.

- On the Extraction of GLCM texture features stage, the segmented image results is computed using the GLCM method. GLCM is calculated with 4 orientations, 00, 450, 900, 1350. Subsequently, the image is calculated using the GLCM properties, including contrast (6), correlation (formula 7), energy (formula 8), homogeneity (formula 9). The extraction process produces a matrix with a 1x4 column. The value is attained based on the average value of the GLCM properties.

  Contrast:

$$\sum_{i,j} |i \quad j|^2 \, p(i,j) \tag{6}$$

  Correlation:

$$\sum_{i,j} \frac{(i-\mu i)(j-\mu j)p(i,j)}{\sigma_i \sigma_j} \tag{7}$$

  Energy:

$$\sum_{i,j} p(i,j)^2 \tag{8}$$

  Homogeneity:

$$\sum_{i,j} \frac{p(i,j)}{1+|i-j|} \tag{9}$$

- Feature extraction of HSV + GLCM, to use the combination of features extraction, query image is duplicated into two. Each image passes through the HSV and GLCM feature extraction stages. The result of the process is two images in vector form. Further, the images are merged into one in vector form.

C. *Euclidean Distance* (ED)

On this stage, the query image and image in the database are in the form of vector. The query image is compared to the image in the database using the ED similiarty measure on (10).

$$\begin{gathered} P = (p_1, p_2, \ldots, p_n) \\ Q = (q_1, q_2, \ldots, q_n) \\ d = \sqrt{(p_1-q_1)^2 + (p_2-q_2)^2 + \ldots + (p_n-q_n)^2} = \sqrt{\textstyle\sum_{i=1}^{n} p_i - q_i} \end{gathered} \tag{10}$$

The result of calculation using ED, is adopted as a parameter of image resemblance. The smaller ED value is the greater its image resemblance. Based on the result of sorting the ED value of image, the study determines the rank owned by the image. The study selects the top ten images that have the smallest ED value.

IV. Experiments and Results

The experiments conducted in this study are system test using OF17 dataset, which are divided based on segmented and unsegmented image. There are total of six experimental scenarios to be performed in this study. The first three scenarios are executed based on segmented imagery on all three extraction features. The three-second scenarios are operated based on unsegmented image on all three extraction features. Experiments are performed on 10 images for each class, thus there are 170 total tests for one class.

To calculate MAP, precision value is required (11). Further, the value is processed using MAP calculation (12)

$$Precision = \frac{Number\ of\ relevant\ document\ found}{Number\ of\ all\ documents\ found} \tag{11}$$

$$\mathrm{MAP}(Q) = \frac{1}{|Q|}\sum_{j=1}^{|Q|}\frac{1}{m_j}\sum_{k=1}^{m_j} Precision(R_{jk}) \tag{12}$$

The MAP is computed based on the experiments performed in five to ten images each class (See Table I).

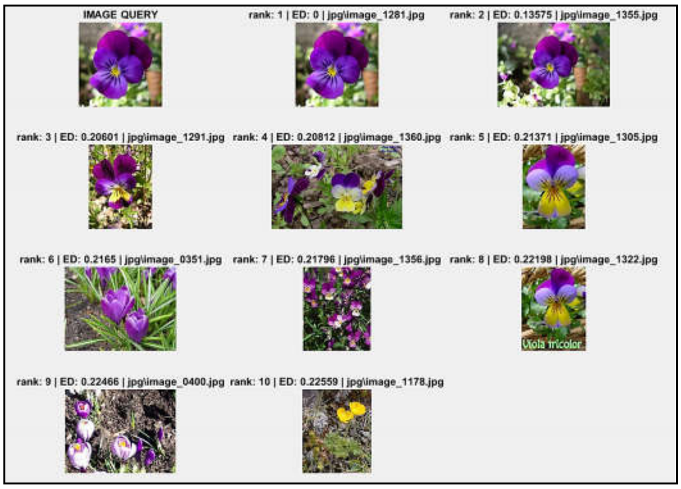

Fig. 4. Illustration of 10 images in pansy flower class, image rank 1 - rank 5, rank 7, and rank 8 is the flower image of the pansy class.

TABLE I. MAP VALUE ON EACH EXPERIMENT SCENARIO PERFORMED IN THE STUDY

Scenario	MAP	
	Top - 5	Top - 10
EF. HSV + Seg	83,3%	83,35%
EF. GLCM + Seg	83,54%	83,5%
EF. HSV+GLCM + seg	84,52%	80,94%
EF. HSV	82,29%	82,64%
EF. GLCM	85,54%	87,32%
EF. HSV+GLCM	84,26%	85,73%

The study assumes that the MAP values on feature extraction using segmented data, provide better value comparing to unsegmented data. However, based on the experimental results presented in Table I, feature extraction with unsegmented data yields greater value, except for feature extraction of HSV. This is presented by at the Average Precision (AP) in the experiment, see Table II.

TABLE II AP VALUE ON SNOWDROP FLOWER CLASS DATA EXPERIMENT, RETRIEVAL ON FIVE IMAGES WITH SEGMENTED IMAGE.

Snowdrop Class					
Feature Extraction	I	II	III	IV	V
HSV	1	0,528	1	1	0,75
GLCM	1	0,6	NaN	1	1
HSV+GLCM	1	1	NaN	1	0,833

Based on Table II, on feature extraction of GLCM and HSV + GLCM, a NaN value is discovered. The value of NaN emerged because there is a 0 operation divided by 0. The returned image segmentation for grayscale yields a nearest-neighbor value of 0. Thus, the result value obtained from the calculation of GLCM is NaN. Further, the value is combined with the image value of HSV extraction result. The result of previous accumulation process is divided by Inf to generate NaN value. However, the NaN value does not appear on the unmarked image feature extraction. In order to avoid errors when calculating the MAP, the NaN value is converted to 0.

NaN value also does not emerge on the segmented image of HSV feature extraction. Based on the assumption, the HSV feature extraction of MAP with segmented image, possesses a higher MAP value, which is 83.35%. While HSV extraction with unsegmented image attains lower result, that is 82,64%. Referred to Table 1, the highest MAP value is generated by GLCM feature extraction with unsegmented image test data of 87.32%. This value is considered as the highest MAP result for all scenarios.

V. CONCLUTIONS AND SUGGESTIONS

The proposed method for CBIR system conducted in this study generates the highest MAP value of 83.35%. The value is obtained from the HSV feature extraction results with segmented image test data. However, experimental results with unsegmented feature extraction yields higher value, which is 87.32%. This value is gained from GLCM feature extraction with unsegmented image test data. It is strongly affected by the segmentation process.

For future research, the study should consider the implementation of k-means segmentation, and addition of feature extraction. The other suggestion is to concern more in overcoming the NaN value generated from GLCM feature extraction.

ACKNOWLEDGMENT

Grateful to the Department of Computer Science Education, Universitas Pendidikan Indonesia who have given support to the writing of this manuscript

REFERENCES

[1] A. Singh, P. Sohoni, and M. Kumar, "A Review of Different Content Based Image Retrieval Techniques," vol. 2, no. 5, pp. 266–275, 2014.

[2] H. Mohamadi, A. Shahbahrami, and J. Akbari, "algorithms," no. February, 2013.

[3] S. Pabboju, "A Novel Approach for Content-Based Image Indexing and Retrieval System using Global and Region Features," vol. 9, no. 2, pp. 119–130, 2009.

[4] M. Nilsback and A. Zisserman, "A Visual Vocabulary for Flower Classification," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2006.

[5] P. Ma, M. Seeland, M. Rzanny, N. Alaqraa, and J. Wa, "Plant species classification using flower images — A comparative study of local feature representations," pp. 1–29, 2017.

[6] L. Y. Yang L, Hu Q, Zhao L, "SALIENCE BASED HIERARCHICAL FUZZY REPRESENTATION FOR OBJECT RECOGNITION," *IEEE Int. Conf. on; 2015. p. 4873–4877.2015.7351733.*, pp. 4873–4877, 2015.

[7] L. Xie, Q. Tian, S. Member, M. Wang, and B. Zhang, "Spatial Pooling of Heterogeneous Features for Image Classification," vol. 23, no. 5, pp. 1994–2008, 2014.

[8] Y. Chai, E. Rahtu, V. Lempitsky, L. Van Gool, and A. Zisserman, "TriCoS : A Tri-level Class-Discriminative Co-segmentation Method for Image Classification," pp. 794–807, 2012.

[9] A. Angelova and Y. Lin, "Image segmentation for large-scale subcategory flower recognition," pp. 39–45, 2013.

[10] J. Liu, C. Zhang, Q. Tian, C. Xu, H. Lu, and S. Ma, "One Step Beyond Bags of Features : Visual Categorization Using Components," pp. 2465–2468, 2011.

[11] Y. Chai, V. Lempitsky, and A. Zisserman, "BiCoS : A Bi-level Co-Segmentation Method for Image Classification," pp. 2579–2586, 2011.

[12] M. Nilsback, "An automatic visual Flora – segmentation and classification of flower images," *Dr. Thesis*, 2009.

[13] M. Nilsback and A. Zisserman, "Automated flower classification over a large number of classes," *Proc. Indian Conf. Comput. Vis.*, 2008.

[14] J. Eakins and M. Grahan, "Content-based Image Retrieval," *Libr. Inf. Briefings*, vol. 85, no. October, pp. 1–15, 1999.

[15] R. Haralick, K. Shanmugan, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 3. pp. 610–621, 1973.

[16] A. Shahbahrami, T. A. Pham, and K. Bertels, "Parallel implementation of Gray Level Co-occurrence Matrices and Haralick texture features on cell architecture," *J. Supercomput.*, vol. 59, no. 3, pp. 1455–1477, 2012.