Color and Texture Feature for Remote Sensing – Image Retrieval System: A Comparative Study

Retno Kusumaningrum[1] and Aniati Murni Arymurthy[2]

**[1] Department of Informatics, Diponegoro University
Semarang, 50275, Indonesia**

**[2] Faculty of Computer Science, University of Indonesia
Depok, 16424, Indonesia**

Abstract

In this study, we proposed score fusion technique to improve the performance of remote sensing image retrieval system (RS-IRS) using combination of several features. The representation of each feature is selected based on their performance when used as single feature in RS-IRS. Those features are color moment using L*a*b* color space, edge direction histogram extracted from Saturation channel, GLCM and Gabor Wavelet represented using standard deviation, and local binary pattern using 8-neighborhood. The score fusion is performed by computing the value of image similarity between an image in the database and query, where the image similarity value is sum of all features similarity, where each of feature similarity has been divided by SVD value of feature similarity between all images in the database and query from related feature. The feature similarity is measured by histogram intersection for local binary pattern, whereas the color moment, edge direction histogram, GLCM, and Gabor are measured by *Euclidean Distance*. The final result shows that the best performance of remote sensing image retrieval in this study is a system which uses the combination of color and texture features (i.e. color moment, edge direction histogram, GLCM, Gabor wavelet, and local binary pattern) and uses score fusion in measuring the image similarity between query and images in the database. This system outperforms the other five individual feature with average precision rates 3%, 20%, 13%, 11%, and 9%, respectively, for color moment, edge direction histogram, GLCM, Gabor wavelet, and LBP. Moreover, this system also increase 17% compared to system without score fusion, simple-sum technique.

Keywords: *Color Moment, Edge Direction Histogram, GLCM, Gabor Wavelet, Local Binary Pattern, Score Fusion.*

1. Introduction

Remote sensing image is a representation of an up-to-date part of the earth surface as seen from the space. The information of remote sensing images is close to the reality of earth surfaces. Hence, remote sensing images are widely used in various fields at the present time, such as agriculture, mineral exploration, military, forestry, fisheries, etc. The high advantage of using remote sensing images as reference not only makes increasing of sensor system technology, but also increases size or volume of remote sensing image. Therefore, it needs to develop a remote sensing – image retrieval system (RS-IRS) that has good performance and easy to use.

Recently, most of proposed RS-IRS automatically extract low-level features (e.g. color, texture, and shape) to measure similarity among images by comparing the feature similarity. Maheswary and Srivastava use combination of color and texture feature [1]. Color feature is represented by color moment of HSV color space, while texture feature is represented by Gray Level Co-occurrence Matrix. On the other hand, Long, et.al. said that the color moment perform better if it is defined by both CIE color spaces, L*a*b* and L*u*v*, as opposed to solely by the HSV color space [2]. Another system proposed by Ruan, et.al. uses combination of color feature and GLCM as texture feature [3]. The color feature vector is defined as combination of ratio between mean value of red and mean value of green and ratio between mean value of blue and mean value of green. Peijun, et.al. develop a prototype of content based remote sensing image retrieval which can be used to retrieve hyperspectral remote sensing images by spectral feature (Normalized Difference of Vegetation Index – NDVI, Normalized Difference of Building Index – NDBI, and Normalized Difference of Water Index – NDWI), color feature (RGB space represented by histogram), texture feature (GLCM, fractal or wavelet), or spatial features (spatial location – coordinate) [4].

However, their approach suffer from a number of weakness. First, they use various combination of features according to the interest of each authors. For example, most of systems use color feature and GLCM as texture representation, whereas there are many other texture representations that probably more suitable for remote sensing images domain. Second, there are some retrieval methods only focus on single low-level feature. Third, some systems define image similarity as simple sum of similarity features. They do not pay attention to the scale of each feature similarity values.

Therefore, this paper will discuss comparison of some representation of color features and texture features for RS-IRS using high-resolution images as domain. In addition, we also proposed a feature fusion technique to combine several features used. The feature representations that will be compared in this study are limited to color features and texture features. Selection of these two features is due to the several reasons. Firstly, the color feature is invariant to the rotation and scale. Secondly, the texture analysis offers the interesting possibilities to characterize the structural heterogeneity of each class. Thirdly, the image retrieval use global low-level model instead of objects-based semantic model. Hence, the shape feature become less appropriate, since this feature usually described after image have been segmented into regions or objects [2].

The rest of this paper follows the following structure. Section 2 describes all related feature analysis methods. Section 3 presents the mechanism of image retrieval system includes feature fusion. Section 4 discusses experimental result and conclusion of the study will be explained in section 5.

2. FEATURES ANALYSIS METHODS

In this section, we will briefly explain about the features used in this study and the techniques for extracting those features.

2.1 Color

Color is the most widely used visual content for image retrieval system. Two points should be considered when using color feature are color space and color description.

There are some color space used to represent an image. First, the most extensively used color space is RGB color space. This color space is called as "additives primaries" since a color in RGB space is produced by adding three color components, i.e. red channel, green channel, and blue channel.

Second, the color space derived from RGB color space is the HSV color space. This color space more intuitive to describe color, invariant to the changes of illumination and the direction of capture, and easy to transform from RGB to HSV and vice versa [2]. The first step of color space conversion from RGB to HSV is find maximum value of RGB triplet, *M*, and minimum value from the RGB triplet, *m*. Saturation, S, is then calculated using this following equation.

$$S = \frac{(M - m)}{M} \quad (1)$$

and Value, V, is equal to *M*. To calculate the Hue, H, we need to normalize the RGB triplet into ranges 0 to 1 as follow.

$$r = \frac{M - R}{M - m} \quad (2)$$

$$g = \frac{M - G}{M - m} \quad (3)$$

$$b = \frac{M - B}{M - m} \quad (4)$$

and Hue, H, is

$$H = \begin{cases} 60(b - g) \ \ if\ R = M \\ 60\,(2 + r - b) \ \ if\ G = M \\ 60(4 + g - r) \ \ if\ B = M \\ H - 360 \ \ if\ H \geq 360 \\ H + 360 \ \ if\ H < 0 \end{cases} \quad (5)$$

Third, the color spaces defined by the Commission Internationale de L'Éclairage (CIE) are L*a*b* (CIELab) and L*u*v* (CIELuv). Both of these color spaces classify color according to the human visual system, so that CIELAB and CIELuv are device independent but suffer from being quite un-intuitive despite the L parameter having a good correlation with perceived lightness [5].

The foundation of CIELab and CIELuv is the XYZ color space, the CIE standard color system, which has a linear relationship with non-gamma corrected RGB. See the equation below :

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 3.240479 & -1.53715 & -0.498535 \\ -0.969256 & 1.875992 & 0.041556 \\ 0.055648 & -0.204043 & 1.057311 \end{bmatrix} \times \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (6)$$

$$x = \frac{X}{(X + Y + Z)} \quad (7)$$

$$y = \frac{Y}{(X + Y + Z)} \tag{8}$$

Subsequently the lightness, denoted by L*, is defined as:

$$L^* = \begin{cases} 116\left(\frac{Y}{Y_n}\right)^{\frac{1}{3}} - 16, & if\ \frac{Y}{Y_n} > 0.008856 \\ \left(\frac{Y}{Y_n} - \frac{4}{29}\right)\frac{108}{841}, & if\ \frac{Y}{Y_n} \leq 0.008856 \end{cases} \tag{9}$$

Where X_n = 0.950456, Y_n = 1, and Z_n = 1.088754. The formula to compute u*, v* and a*, b* can be seen in the following equation.

$$u^* = 13(L^*)(u' - u_n') \tag{10}$$

$$v^* = 13(L^*)(v' - v_n') \tag{11}$$

Where :

$$u' = u = \frac{2x}{(6y - x + 1.5)} \tag{12}$$

$$v' = 1.5v = \frac{4.5y}{(6y - x + 1.5)} \tag{13}$$

One color description that is able to provide efficiency and effectiveness in representing the distribution of colors of an image is color moment [6]. Color moment was introduced by Stricker and Orengo [6]. There are three color moments, i.e. *mean*, *standard deviation*, and *skewness*. These central moments are computed for each channel. Therefore, if an image has three channels, then the dimension of this feature is 9D (9 dimension). Mathematically, those moments for each channels are defined as follow [1].

- *MOMENT 1 – Mean:* The average color value in the image.

$$E_i = \sum_{N}^{j=1} \frac{1}{N} p_{ij} \tag{14}$$

- *MOMENT 2 – Standard Deviation:* The square root of the variance of distribution.

$$\sigma_i = \sqrt{\left(\frac{1}{N}\sum_{N}^{j=1}(p_{ij} - E_i)^2\right)} \tag{15}$$

- *MOMENT 3 – Skewness:* A measure of the degree of asymmetry in the distribution

$$s_i = \sqrt[3]{\left(\frac{1}{N}\sum_{N}^{j=1}(p_{ij} - E_i)^3\right)} \tag{16}$$

Where the p_{ij} is the pixel value of the i^{th} color channel at j^{th} image pixel, N is the number of pixels in the image.

2.2 Textures

Texture is a property to represent the surface and structure of an image which can be defined as a regular repetition of an element or pattern on a surface [7]. Basically, texture representation divide into two major categories, i.e. structural and statistical approach [2]. In this paper, we only discuss some of statistical approaches for texture analysis. In the statistical approach, the texture features are computed from the statistical distibution of observed combination of intensities at specified positions relative to each other position in the image. Based on the number of pixels defining the local feature, statistical approach can be futher classified into first-order (one pixel), second-order (two pixels), and higher-order (three or more pixels) statistics [8].

There are four texture features that will be used in this study, including gray level co-occurrence matrix, edge direction histogram, gabor wavelet, and local binary pattern. The following sub-sections describe the four texture features in detail.

2.2.1 Gray Level Co-occurrence Matrix (GLCM)

GLCM is the two dimensional matrix of joint probabilities between pair of pixels (one with gray level *i* and the other with gray level *j*), separated by a distance *d* in a given direction θ [3]. Hence, GLCM is included in the second-order statistical texture analysis.

The extraction process of GLCM features are divided into two main processes, i.e. the formation of co-occurrence matrix and the extraction of GLCM descriptors against the co-occurrence matrix. The following figure illustrates the formation of co-occurrence matrix.

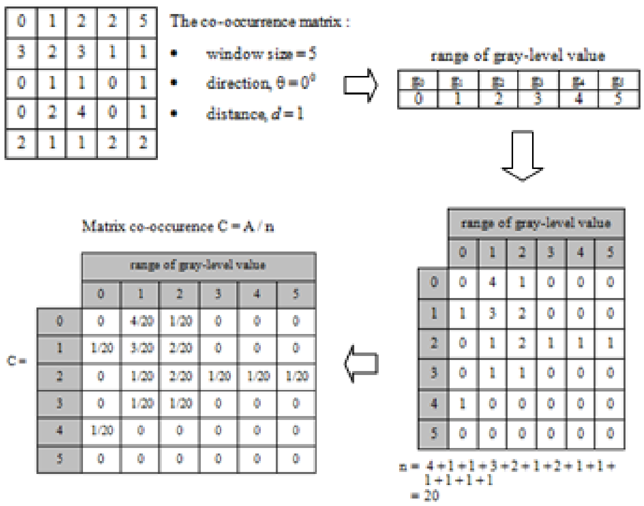


Fig. 3 Illustration of formation of matrix co-occurrence

Based on the co-occurrence matrix, the next step is computing the GLCM descriptors as follows [9]:

- *Angular Second Moment (ASM) / Energy:* shows the texture uniformity or texture homogenity. Energy value will be greater for a homogeneous texture.

$$Energy = \sum_{i=1}^{k}\sum_{j=1}^{k}(C_{ij})^2 \qquad (17)$$

- *Entropy:* shows the degree of randomness. The maximum value of entropy will be reached when all elements C_{ij} has the same value. Inhomogeneous scenes have low entropy, while a homogeneous scene has a high entropy.

$$Entropy = -\sum_{i=1}^{k}\sum_{j=1}^{k} C_{ij}\log(C_{ij}) \qquad (18)$$

- *Contrast / Second Order Element Difference Moment:* shows the contrast texture value. The calculation results in a larger figure when there is great contrast.

$$Contrast = \sum_{i=1}^{k}\sum_{j=1}^{k} C_{ij}(i-j)^2 \qquad (19)$$

- *Cluster Shade:* shows the lack of symmetry in an image.

$$ClusterShade = \sum_{i=1}^{k}\sum_{j=1}^{k}(i-j-2m)^2 C_{ij} \qquad (20)$$

- *Correlation:* A measure of gray level linear dependence between the pixels at the specified positions relative to each other.

$$Correlation = \sum_{i=1}^{k}\sum_{j=1}^{k}(i-m)(j-m)C_{ij} \qquad (21)$$

- *Homogeneity:* shows the first order inverse element difference moment.

$$Homogeneity = \sum_{i=1}^{k}\sum_{j=1}^{k}\frac{C_{ij}}{n} \qquad (22)$$

- *Maximum Probability*: shows the emergence of the gray-level value g_i adjacent to the gray-level value g_j more dominant in the image.

$$MaximumProbability = \max(C_{ij}) \qquad (23)$$

- *Inverse Difference Moment (IDM)*: a low IDM value for inhomogeneous images, and a relatively higher value for homogeneous images

$$Inverse = \sum_{i=1}^{k}\sum_{j=1}^{k}\frac{C_{ij}}{n^2} \qquad (24)$$

where:

C_{ij} is element of matrix co-occurence
$n = i - j$ if $i \neq j$ and $n = 1$ *if* $i = j$
m is mean value of matrix co-occurence

2.2.2 Edge Direction Histogram (EDH)

Initially, we will be performed the Gaussian smoothing against the image channel. After that, perform the edge point's detection using Canny filter. We calculate the gradient of each edge points by utilizing 5-type operators Sobel, i.e. horizontal edge, vertical edge, 45-degree edge, 135-degree edge, and non directional edge. The following figure define those 5 operators Sobel.

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

(a). Horizontal edge

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

(b). Vertical Edge

$$\begin{bmatrix} -2 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

(c). 45-degree edge

$$\begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ -2 & -1 & 0 \end{bmatrix}$$

(d). 135-degree edge

$$\begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix}$$

(e). Non directional edge

Fig. 4 Sobel Operators

Finally, the 5-dimensional edge histogram is calculated by counting the edge pixel in each direction.

2.2.3 Gabor Filter (Wavelet) [10]

For a given image *I(x,y)* with size *P* x *Q* and *s* and *t* are the filter mask size variables, the discrete Gabor wavelet transform is given by a convolution :

$$G_{mn}(x,y) = \sum_s \sum_t I(x-s, y-t)\, \psi^*_{mn}(s,t) \qquad (25)$$

where ψ^*_{mn} is the complex conjugate of ψ_{mn} which is a class of self-similar functions generated from dilation and rotation of the following mother wavelet :

$$\psi(x,y) =$$

$$\frac{1}{2\pi\sigma_x\sigma_y} exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)\right].\exp(j2\pi Wx) \qquad (26)$$

where *W* is called the modulation frequency.

The self-similar Gabor wavelet are obtained through the generating function :

$$\psi_{mn}(x,y) = a^{-m}\psi(\tilde{x},\tilde{y}) \qquad (27)$$

where *m* (*m = 0, 1, ... M-1*) and *n* (*n = 0, 1, ..., N-1)* are the scale and orientation of the wavelet respectively.

$$\tilde{x} = a^{-m}(x\cos\theta + y\sin\theta) \qquad (28)$$

$$\tilde{y} = a^{-m}(-x\cos\theta + y\cos\theta) \qquad (29)$$

where $a > 0$ and $\theta = {n\pi}/{N}$

After applying the Gabor filter on the image with different orientation at different scale, we obtained the energy content at different scale and orientation of the image.

$$E(m,n) = \sum_x \sum_y |G_{mn}(x,y)| \qquad (30)$$

2.2.4 Local Binary Patter (LBP)

The name of "Local Binary Pattern" reflects the functionality of the operator, i.e. a local neighborhood is thresholded at the gray value of centre pixels into a binary pattern [11]. Based on the labels, in the form of binary pattern, we can create histogram of labels as a texture descriptor. See the following figure for an illustration of the basic LBP.

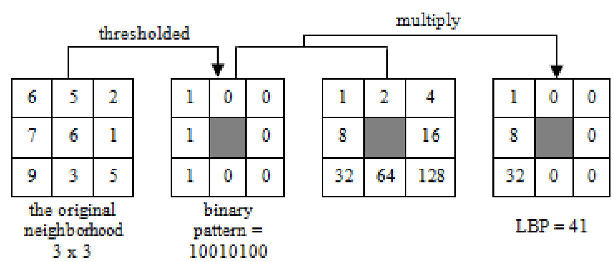


Fig. 5 Illustration of basic LBP8,1.

Matematically, it can be done as follow :

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s\left(g_p - g_c\right).2^p \qquad (31)$$

where *s(x)* is thresholding function

$$s(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases} \qquad (32)$$

The variable in the Eq. (31) are defined as follows.

- P : number of neighborhood
- R : radius
- g_p : gray level value at neighborhood p^{th}
- g_c : gray level value at centre pixel

In practice, Eq. 18 means that the sign of the differences in a neighborhood are interpreted as a *P*-bit binary number, resulting in 2^P distinct value for LBP code and the local gray-scale distribution can thus be approximately described with 2^P-bin discrete distribution of LBP code [12].

3. The Image Retrieval System

We use the global low-level model for the image retrieval system. See the Appendix for the detail diagram of image retrieval system in this study.

Basically, the process of image retrieval system in this study is divided into four main steps, i.e. features extraction, similarity measurement, process of indexing (ranking), and display the results. In the following sub-section, we will explain those steps in detail.

3.1 Features Extraction

This step is divided into two processes, namely pre-processing and feature extraction process itself. The employed features in this study are color moment, EDH, GLCM, Gabor Wavelet, and LBP. For each of features, pre-processing using different approaches. Preprocessing for feature GLCM, Gabor and LBP is the conversion of images into gray-level image. While the color moment features, pre-processing is done by converting the image into 3 different color spaces, i.e. HSV, L*a*b* and L*u*v*. The use of three color spaces for color features is performed to get the most appropriate color space for color moment as a color descriptor, especially in the domain of high-resolution remote sensing images. And pre-processing for EDH is the conversion of images into the HSV color space and gray-level image.

As mentioned before, the result of GLCM extraction is matrix of each descriptor (e.g. energy, entropy, contrast, cluster shade, etc.), while Gabor wavelet is the matrix of Energy. Therefore, we implement two basic statistical analysis, mean and standard deviation, for those descriptor. The parameters used in the process of GLCM and Gabor features extraction are as follows:

- *GLCM:* The GLCM feature extraction uses moving kernel (window) with size 5x5 as recomended in [9]. And the angles used in this process are 0 (horizontal or east-west), 90 (vertical or south-north), 45 (diagonal or southwest-northeast), and 135 (diagonal or southeast-northwest), thus the counting process carried out for all possible of the objects direction.
- *Gabor Wavelet:* The representation of texture feature using Gabor wavelets will provide the best performance by using the scale value is 4 and orientation value is 6. It is recommended in [13].

3.2 Features Extraction

The following tables explain about the formal definition of variables and functions that is used in this study.

Table 1: Variable Formalization in Image Retrieval

Set	Symbol	Element	Size	Description
Image	I	i_j	N_I	Images in the database
Feature	F	f_n	N_F	Features that are extracted from an image
Query	Q	q	-	Feature vector of query

Table 2: Function Formalization in Image Retrieval

Name	Symbol	Mapping
Feature similarity	S_F	$I \; X \; Q \; X \; F \rightarrow [0, \infty]$
Image similarity	S_I	$I \; X \; Q \rightarrow [0, \infty]$

The process of measure similarity between the query image and the images in the database is performed using image similarity function, S_I. The image similarity function is the sum of score fusion for each feature. Mathematically, it can be defined as follow :

$$S_I(i_j, q) = \sum_{f_n \in F} \left[\frac{S_F(i_j, q; f_n)}{svd(S_F(I, q; f_n))} \right] \quad (33)$$

where :

$S_F(i_j, q; f_n)$: feature similarity between the image i_j and query q with the respect to feature f_n.

$S_F(I, q; f_n)$: all of the feature similarity value between the image i_j and query q with the respect to feature f_n. If we have N_I images in the database, then this value is column matrix of feature similarity with the size $N_I \; x \; 1$.

$svd(S_F(I, q; f_n))$: a function to get singular value decomposition of $S_F(I, q; f_n)$

For example, if we have feature vector of an image i_j with the respect to feature f_n is $X_p = (x_1, x_2, \ldots, x_D)$ and the feature vector of query q with the respect to feature f_n is $Y_p = (y_1, y_2, \ldots, y_D)$, then the similarity feature between image i_j and query q with the respect to feature f_n is defined as follow.

$$S_F(i_j, q; f_n) = \begin{cases} 1 - \dfrac{\sum_{p=1}^{D} min(X_p, Y_p)}{min(\sum_{p=1}^{D} X_p, \sum_{p=1}^{D} Y_p)}, if \; f_n \; is \; LBP \\ \left(\sum_{p=1}^{D} (X_p - Y_p)^2 \right)^{1/2}, otherwise \end{cases} \quad (34)$$

In the other word, the feature similarity for local binary pattern is measured by using histogram intersection [14], while color moment, EDH, GLCM, and Gabor are measured by *Euclidean Distance*. We propose the use of SVD (Singular Value Decompostion) function in the score fusion because it is more robust and efficient. The *svd* function in Eq. (33) returns the singular value of all feature similarity value with the respect to feature f_n. This singular value is equal to the operator norm of feature similarity so that we can obtain the same scale for all of the feature similarity values.

3.3 Indexing (Ranking) and Display Result

The indexing process is performed by sorting the images in the database based on image similarity values in ascending. This means the image that has the zero value for image similarity is the most similar image to the query. And the system will display the results, i.e. top-n images with the lowest value of image similarity.

4. EXPERIMENTS AND RESULT

4.1 Dataset

This study uses remote sensing image database. This database contains 200 IKONOS images of resolution 256 x 256 pixel. These images are RGB images.

4.2 Experiment Environment

This study is implemented using Matlab R2010a with the operating system is Windows 7 – 64bit and the hardware specification is as follows:

- Intel Core i5-520M 2.40 GHz
- 4GB of memory (RAM)
- 500 GB of hard disk drive

4.3 Performance Measurement

The performance result of the RS-IRS is presented in the form of Precision-Recall graph (PR-graph) averaged over several queries. We use 30 queries by example in the form of remote sensing images, so that the value of precision in the PR-graph is average precision value.

4.4 Experimental Scenarios

1). Experimental Setup 1: The objective of this scenario is identify the appropriate representation for each features, 1 color descriptor and 4 texture descriptor. The meaning of appropriate representation is a representation which has the best average precision value. Therefore, we have five different comparisons. First, we will compare three types of color space (i.e. HSV, L*a*b, and L*u*v*) for color moment descriptor. Second, we will describe the comparison between gray-level image and each channel of HSV color space for EDH extraction. Third and Forth, we will compare the use of two basic statistical analysis, mean and standar deviation, for texture representation of GLCM and Gabor respectively. The final one is comparison of the appropriate number of neighborhood, 4-neighborhood and 8-neighborhood, for LBP extraction.

2). Experimental Setup 2: The aim of this scenario is compare the simple sum technique and the proposed score fusion technique. This comparison is carried out against two type features combination, i.e. combination of texture features (EDH, GLCM, Gabor, and LBP) and combination of color – texture features (Color Moment, EDH, GLCM, Gabor, and LBP).

4.5 Experiment Results

1). Result of Experimental Setup 1: As mentioned in the previous sub-section, we will compare HSV, L*a*b*, and L*u*v* color space for color moment descriptor. See the following figure.

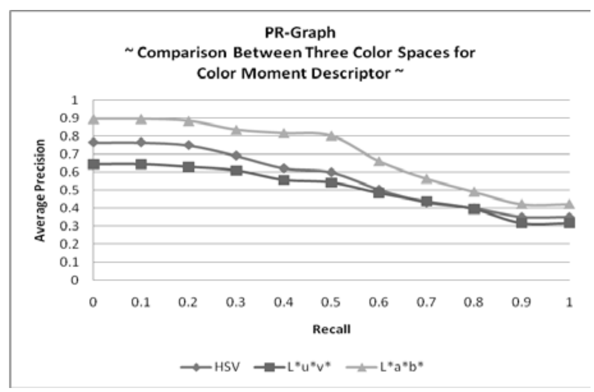


Fig. 6 PR-Graph, Comparison between Three Color Spaces for Color Moment Descriptor.

Based on the figure above, we can conclude that the best color space for color moment in this study is L*a*b*. It can be seen from the average precision of the color moment using the L*a*b* color space (green line) is higher 13% than the HSV color space (blue line) and higher 19% than the L*u*v* color space (red line). L*a*b* color space gives better performance since this space defines colors more closely to the human color perception. In addition, this space also uses three color coordinates includes L* – the lightness coordinate, a* – the red/green coordinate, and b* – the yellow/blue coordinate, thus it can be defined color as combinations of red and yellow, red and blue, green and yellow, and green and blue. Another interesting characteristic of the L*a*b* color space is that the distance can be calculated between two colors and proportional to the difference between the two colors as perceived by the human eye.

Furthermore, see the following figure to get the best pre-processing representation for process of EDH extraction.

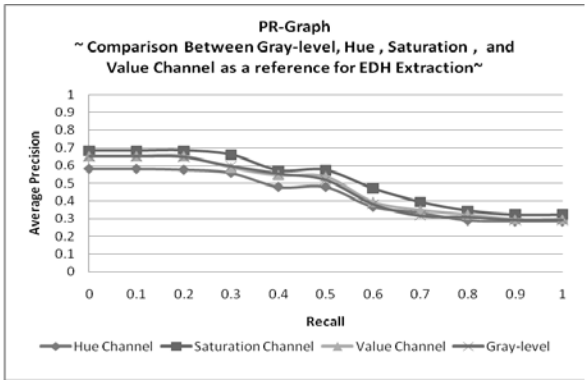
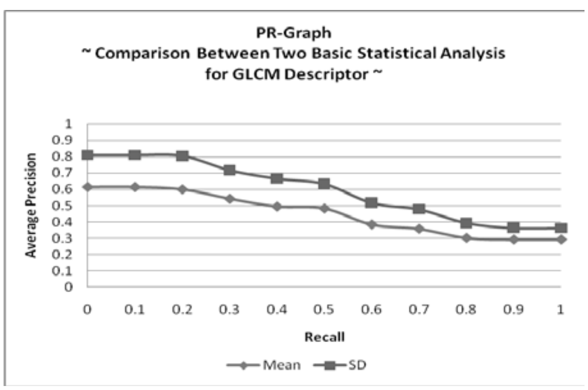
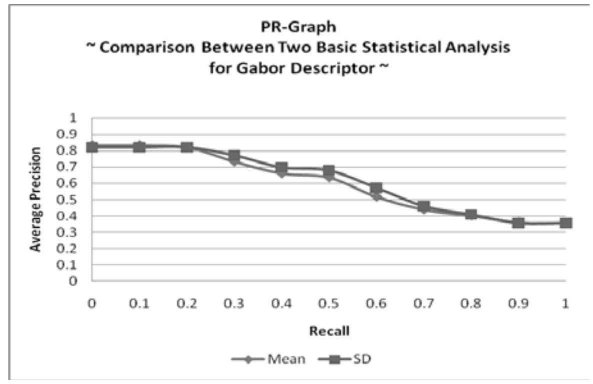


Fig. 7 PR-Graph, Comparison between Four Pre-processing Representation as a reference of EDH extraction.

The figure above shows that for each pre-processing representation, Saturation channel outperforms the other three representations with average precision rates 3%, 10%, and 3%, respectively, for gray-level, hue channel, and value channel representation. The saturation channel in HSV color space indicates the range of grey in the color space, thus it is appropriate to calculate the edge direction histogram. This color space has ranges from 0 to 100% or from 0 to 1. When the value is ’0,’ the color is grey and when the value is ’1,’ the color is a primary color.



(a). GLCM Descriptor



(b). Gabor Descriptor

Fig. 8 PR-Graph, Comparison between Two Basic Statistical Analysis for (a) GLCM Descriptor and (b) Gabor Descriptor.

Both figure above explain that standard deviation is the best representation for GLCM and Gabor. The performance of standard deviation exceeds the mean performance by 14% and 2% for GLCM and Gabor respectively. The standard deviation compared to the mean gives better performance since the standard deviation describe how spread out a set of values are around the mean of that set. A set of values that are closely clustered near the mean will have a low standard deviation, a set of numbers that are widely apart will have a higher standard deviation and a set of numbers that are all the same will have a standard deviation of zero.

The last comparison is compare 4-neighborhood and 8-neighborhood for LBP. The LBP with 8-neighborhood presents the better performace since this approach gives the complete information about the relationship between the centre pixel and all pixels surrounding. In the other word, the 8-neighborhood comparison carried out for all possible of the objects direction. See the following figure to see the comparison between LBP with 4-neigborhood and 8 neighborhood.

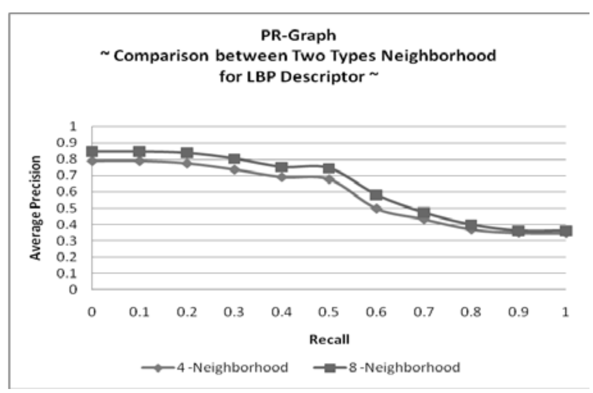


Fig. 9 PR-Graph, Comparison between LBP with 4-neighborhood and LBP with 8-neighborhood.

In summary, the best representation for each color and texture descriptor in this study is color moment descriptor with CIE L*a*b* color space, EDH which is extracted over the Saturation channel, GLCM and Gabor descriptor which is represented using standar deviation, and the LBP with 8-neighborhood.

2). *Result of Experimental Setup 2:* In this scenario, we will compare proposed score fusion technique and simple sum technique (without score fusion). Fig. 10 illustrates the comparison for combination of texture features, whereas Fig. 11 illustrates the comparison for combination of color – texture feature.

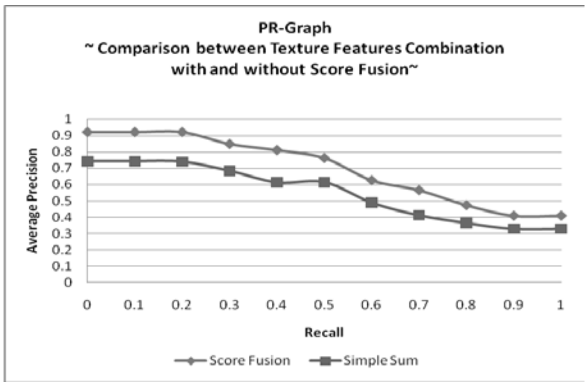


Fig. 10 PR-Graph, Comparison between Texture Feature Combination – with and without score fusion.

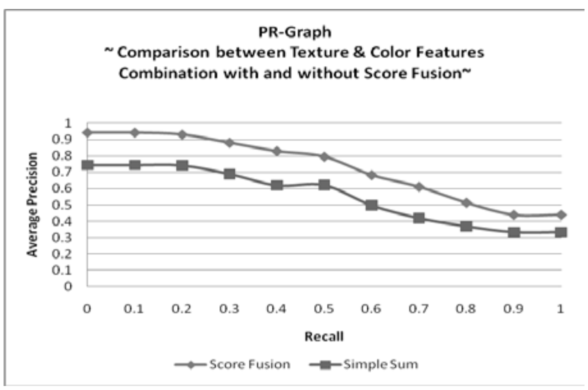


Fig. 11 PR-Graph, Comparison between Combination of Color and Texture Feature – with and without score fusion.

According to Fig. 10 and Fig. 11, score fusion always give better results than simple-sum technique, i.e. an increase of 15% for the combination of textures and 17% for the combination of colors - textures. The condition is due to the use of simple-sum technique is only adding up the value of all features similarity regardless the scale of each feature similarity. Hence, if there are types of features which have bigger feature similarity values than the other features, then these feature become dominant and the addition of other features will have no impact. In contrary, the use of score fusion technique will normalize the values of feature similarity, thus each features will contributes or influences on the image similarity value.

Moreover, Fig. 11 also shows that the addition of color feature gives better performance than simply using the combination of texture features. It is because the texture features can provide good performance on a textured area (heterogeneous), but tend to give unsatisfactory performance in a homogeneous area. On the other hand, color feature is able to distinguish objects in a homogeneous area. Therefore, a combination of both features could complement their respective advantages. In addition, the system which use combination of color and texture feature also exceeds the other five individual feature with average precision rates 3%, 20%, 13%, 11%, and 9%, respectively, for color moment, edge direction histogram, GLCM, Gabor wavelet, and LBP. In conclusion, the best performance of remote sensing image retrieval in this study is a system which uses the combination of color and texture features (i.e. color moment, edge direction histogram, GLCM, Gabor wavelet, and LBP) and uses score fusion in measuring the image similarity between query and images in the database.

4. Conclusions

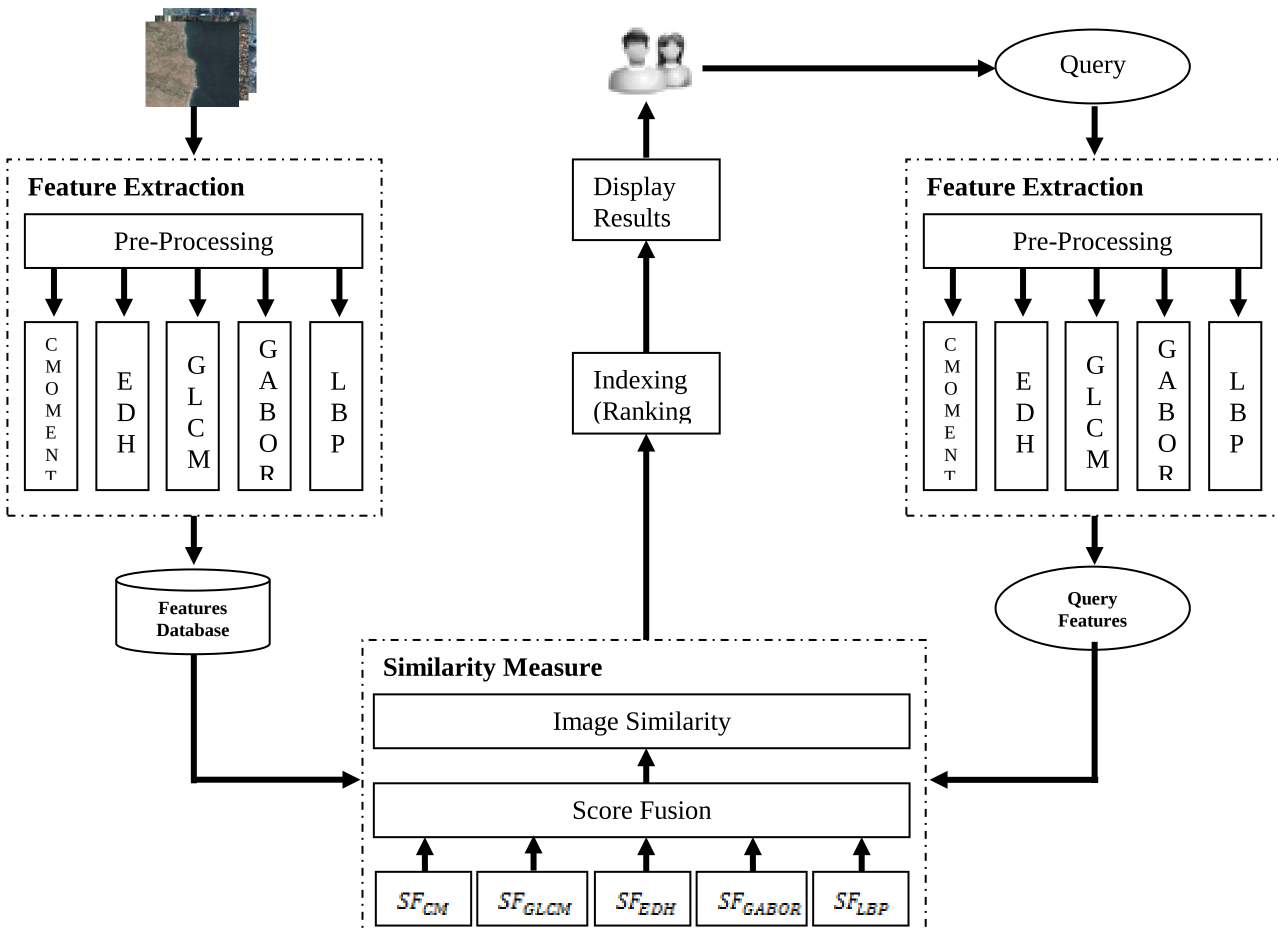
Nowadays, most of proposed RS-IRS use various combination of low-level feature according to the interest of each authors and some of them focus on single low-level feature.

In order to improve the performance of RS-IRS using combination of color and texture features we perform two steps. Firstly, we select the appropriate representation which give the best performance when used as single feature in RS-IRS. Secondly, we proposed score fusion technique to combine several features in the RS-IRS using multiple features.

Those selected features representation are color moment using L*a*b* color space, EDH extracted from Saturation channel, GLCM and Gabor wavelet represented using standard deviation, and LBP using 8-neighborhood. The score fusion is performed by computing the value of image similarity between an image in the database and query, where the image similarity value is sum of all features similarity, where each of feature similarity has been divided by SVD value of feature similarity between all images in the database and query from related feature. The feature similarity is measured by histogram intersection for LBP, whereas the color moment, EDH, GLCM, and Gabor are measured by *Euclidean Distance*.

The final result shows that the best performance of RS-IRS in this study is a system which uses the combination of color and texture features (i.e. color moment, EDH, GLCM, Gabor wavelet, and LBP) and uses score fusion in measuring the image similarity between query and images in the database. This system outperforms the other five individual feature with average precision rates 3%, 20%, 13%, 11%, and 9%, respectively, for color moment, edge direction histogram, GLCM, Gabor wavelet, and LBP. Moreover, this system also increase 17% compared to system without score fusion, simple-sum technique.

Appendix



Acknowledgments

The author would like to thank Indonesian National Institute of Aeronautics and Space - Lembaga Penerbangan dan Antariksa Nasional (LAPAN) which has provided some IKONOS images.

References

[1] P. Maheshwary, and N. Sricastava, "Prototype Sytem for Retrieval of Remote Sensing Images based on Color Moment and Gray Level Co-Occurrence Matrix", IJCSI International Journal of Computer Science Issues, 2009, vol. 3, pp. 20-23

[2] F. Long, H. Zhang, and D.D. Feng, "Fundamentals of Content-Based Image Retrieval", in Multimedia Information Retrieval, 2002

[3] N. Ruan, N. Huang, and. W. Hong, "Semantic-Based Image Retrieval in Remote Sensing Archive : An Ontology Approach", in International Conference on Geoscience and Remote Sensing Symposium, 2006, pp. 2903 – 2906

[4] D. Peijun, C. Yunhao, T. Hong, and F. Tao, "Study on Content-Based Remote Sensing Image Retrieval", in International Conference on Geoscience and Remote Sensing Symposium, 2005, 4 pp.

[5] A. Ford and A. Roberts, "Color Space Conversions", 1998 [available online : http://debut.cis.nctu.edu.tw/~ching/Face/Articles/coloureq.pdf]

[6] M. Stricker and M. Orengo, "Similarity of Color Images," in IS&T and SPIE Storage and Retrieval of Image and Video Databases III, 1995, pp. 381-392

[7] G.N. Srinivasan, and G. Shobha, "Statistical Texture Analysis", in World Academy of Science, Engineering, and Technology, 2008, vol. 36, pp. 1264-1269

[8] T. Ojala and M. Pietikainen, "Texture Classification, Machine Vision, and Media Processing Unit", University of Oulu, Finland

[9] D. Chahyati, "Classification of Radar Images Based on Texture Features of Gray Level Co-occurrence Matrix Semivariogram and Wavelet Stationary", Master Thesis, Department of Computer Science, University of Indonesia , Depok, Indonesia, 2003

[10] D. Zhang, A. Wong, M. Indrawan, and G. Lu, "Content-based Image Retrieval Using Gabor Texture Features", IEEE Transaction on Pattern Analysis and Machine Intelligence – PAMI, 2000, pp. 13-15

[11] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns", IEEE Transaction on Pattern Analysis and Machine Intelligence – PAMI, vol. 24, no. 7, 2002, pp. 971-987

[12] T. Maenpaa and M. Pietikainen, "Texture Analysis with Local Binary Patterns in Handbook of Pattern Recognition and Computer Vision 3rd Edition", 2004, pp. 197-216

[13] B.S. Manjunath, and W.Y. Ma, "Texture Feature for Browsing and Retrieval of Image Data", IEEE Transaction on Pattern Analysis and Machine Intelligence – PAMI, 1996, pp. 837-842

[14] J. Smith and S.F. Chang, "Automated Image Retrieval Using Color and Texture Feature", IEEE Transaction on Pattern Analysis and Machine Intelligence – PAMI, 1995.

Retno Kusumaningrum achieved her undergraduate degree in Department of Mathematics from Diponegoro University, Semarang, Indonesia, where she is currently working toward as lecturer in Department of Informatics, in 2003. She earned her master degree in Faculty of Computer Science from University of Indonesia, Depok, Indonesia, in 2010. Currently, she is studying for her doctoral degree in Faculty of Computer Science, University of Indonesia, Jakarta, Indonesia. Her current research activities are in spatial pattern and image retrieval system, particularly feature extraction, relevance feedback, and objective evaluation, with main application in remote sensing domains.

Aniati Murni Arymurthy is a Professor in Faculty of Computer Science, University of Indonesia. She graduated from Department of Electrical Engineering, University of Indonesia, Jakarta, Indonesia. She earned her Master of Science in Department of Computer and Information Sciences, The Ohio State University (OSU), Columbus, Ohio, USA. She also holds Doktor from Department of Opto-Electronics and Laser Applications, University of Indonesia, Jakarta, Indonesia and a sandwich program at the Laboratory for Pattern Recognition and Image Processing (PRIP Lab), Department of Computer Science, Michigan State University (MSU), East Lansing, Michigan, USA. Her main research activities are image processing and pattern recognition.