

Order Sensitive Scoring for Objective Evaluation of Remote Sensing – Image Retrieval System

R. Kusumaningrum², M.I. Fanany³, and A.M. Arymurthy⁴

Abstract—In order to make the evaluation result become fair and consistent without relying on user judgement, we proposed a new mechanism of objective evaluation for remote sensing image retrieval system. This mechanism is performed by dividing the image database into some subjects and select one subject as a ground truth. Subsequently, we take an image of the ground truth as a query Q . Retrieval accurate ratio (pAR) for query Q will be computed as a comparison between utility value and maximum possible utility. This pAR value can provide the non-degeneracy score. We repeat the retrieval process for all the images in the ground truth as the query and take averaged accurate ratio ($pAAR$) for a subject as the average value of all the accurate ratio for all queries in a subject. These process is performed for all subjects in the database, thus we can obtain the mean averaged accurate ratio ($pmAAR$) as mean value of all averaged accurate ratio for all subjects.

Index Terms—accurate ratio, order sensitive scoring, objective evaluation, remote sensing image retrieval system

I. INTRODUCTION

Remote sensing image provide information about part of the earth surface as seen from space. This information is up-to-date and closed to the reality of earth surfaces. Hence, remote sensing images become more widely used as a reference in many fields, such as agriculture, forestry, military, etc. This wide applicability of remote sensing images make the development of the sensor system technology is increasing rapidly. Furthermore, it increases the volume of remote sensing images. Therefore, a remote sensing – image retrieval system (RS-IRS), that not only has good retrieval performance but also easy to use need to be developed.

Nowadays, most of image retrieval systems are measured by subjective evaluation to evaluate whether the systems successfully retrieved correct images for a given query. For each query, a user is given an opportunity to measure which retrieval results are relevant. Such method has some drawbacks. First, human-based judgment for evaluation is often costly and cumbersome. Second, the evaluation can quickly become unfair and inconsistent. It is prone to human

error. For example, when we use “image A ” as a query by example, we will evaluate that “image I ” is relevant. When we use “image I ” as a query by example, however, we might not evaluate “image A ” as relevant.

Recently, Jing Sun and Ying-Jie Xing use average retrieval accurate ratio (AAR) of several retrieval results to measure performance by using objective evaluation [1]. This evaluation will divide an image database into m subjects, S_1, S_2, \dots, S_m . For one subject, S_m , it includes image $I = I_1, I_2, \dots, I_n$. If we use image I_n as query and we get the retrieval result outputs T in which there are r images that belong to the subject S_m , then the retrieval accurate ratio (AR) is a ratio between r and T . Furthermore, the AAR is an average of AR value for each image in a subject and it is computed for every subject. And the overall performance of retrieval mechanism is mean AAR ($mAAR$) score, i.e. the average value of AAR .

Literatures which discuss objective evaluation are still very few. Recently, Jing Sun and Ying-Jie Xing proposed an objective evaluation score applied to general image retrieval system. The proposed objective evaluation, however, suffers from score degeneracy problem. Score degeneracy is a decline in evaluation scores obtained. For example, two systems provide top-10 images as foremost ranked images. In system A, the relevant images is obtained in the rank 1 to 5, but in system B the relevant images is obtained in the rank 6 to 10. Based on the evaluation method proposed by Jing Sun and Ying-Jie Xing, both systems have the same AR value, i.e. 0.5. Whereas, system A is actually better than system B since it gives more accurate retrieval by providing the relevant images as the foremost ranking.

In this study we proposed a new objective evaluation score applied to the Jing Sun and Ying-Jie Xing method. This objective evaluation is implemented to measure the performance of two types RS-IRS, i.e. basic RS-IRS and RS-IRS which uses feature selection algorithm.

This paper is organized as follows. In section 2, we will review the entire related feature used in this study. Section 3 outlines the framework of our RS-IRS, including the feature extraction, feature selection technique, and image similarity measurement. The proposed objective evaluation technique is described in section 4. Section 5 discusses experimental result and the conclusion of entire study will be explained in section 6.

II. EMPLOYED FEATURES

In this study, we use combination of color and texture features. The use of these features because texture features can provide good performance in heterogeneous area, but

Tresna Dewi, Politeknik Negeri Sriwijaya, email:dee_tresna@yahoo.com

² Department of Informatics, Diponegoro University. Jln. Prof. H. Soedarto, SH Tembalang, Semarang, 50275 (email : retno_ilkom@undip.ac.id)

³ Faculty of Computer Science, University of Indonesia. Kampus UI Depok, 16424 (email : ivan@cs.ui.ac.id)

⁴ Faculty of Computer Science, University of Indonesia. Kampus UI Depok, 16424 (email : aniati@cs.ui.ac.id)

tends to give unsatisfactory performance in homogeneous area. On the other hand, a color feature give good performance in distinguishing objects in homogeneous area and is invariant to the rotation and scale. Therefore, the combination of these features is expected to complement the drawbacks of each feature while combining their strengths.

The following sub-sections explain the features used in this study.

A. Color Feature

Two points should be considered when using color feature, i.e., the selection of color space and color description. We use L*a*b* (CIELab) color space that is represented using color moment in this study. There are three reasons for the selection of this color representation. First, CIELab, color space which is defined by the Commission Internationale de L'Éclairage (CIE) defines colors more closely to the human color perception. It can be seen from the use of three color coordinates, including L* represents the lightness, a* encodes the red-green sensation, and b* encodes the yellow blue-sensation. Second, color moment provides efficiency and effectiveness in representing the distribution of image colors [2]. Third, color moment gives more accurate retrieval result if it is defined by both color spaces, L*a*b* and L*u*v, as opposed to the HSV color space [3].

There are three color moments, i.e. *mean*, *standard deviation*, and *skewness*. Mathematically, the first three moments are defined as follow [4].

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

$$E_i = \sum_{j=1}^N \frac{1}{N} p_{ij} \quad (1)$$

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

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

c) **MOMENT 2 – Skewness:** A measure of the degree of asymmetry in the distribution

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

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, E_i is the mean value for i^{th} color channel, σ_i is the standard deviation for i^{th} color channel, and s_i is the skewness value for i^{th} color channel. These central moments are computed for each channel. Therefore, if we use CIELab which has three channels, then the dimension of this feature is 9-dimension.

B. Texture Feature

Texture is a property to represent the surface and structure of an image and it can be defined as a regular repetition of an element or pattern on a surface [5]. In this study, we only use statistical approach for texture analysis. In the statistical approach, the texture features are computed from the statistical

distribution 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 further classified into first-order (one pixel), second-order (two pixels), and higher-order (three or more pixels) statistics [6].

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

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 and in a given direction θ [7]. 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 steps explain the formation of co-occurrence matrix.

- Defined a co-occurrence matrix at a given offset, i.e. image window with size $m \times m$
- Defined the scale of gray level
- Create $k \times k$ matrix A, where k is the number of gray level and the element of matrix A is a_{ij} . The a_{ij} value describe how often a pixel with gray level value i occurs either horizontally (0°), vertically (90°), or diagonally (45° and 135°) to adjacent pixels with the value j and separated by distance d .
- Normalize the matrix A, by dividing each of values with sum of all element matrix A. The normalized matrix is called the co-occurrence matrix C, with the element c_{ij} .

Based on the co-occurrence matrix, the next step is computing the GLCM descriptors as follows [8]:

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

$$Energy = \sum_{i=1}^k \sum_{j=1}^k (c_{ij})^2 \quad (4)$$

b) **Entropy:** Show 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.

$$Entropy = -\sum_{i=1}^k \sum_{j=1}^k c_{ij} \log(c_{ij}) \quad (5)$$

c) **Contrast / Second Order Element Difference Moment:** Show the contrast texture value and the calculation results in a larger figure when there is great contrast.

$$Contrast = \sum_{i=1}^k \sum_{j=1}^k c_{ij} (i - j)^2 \quad (6)$$

d) *Cluster Shade*: Show the lack of symmetry in an image.

$$\text{ClusterShade} = \sum_{i=1}^k \sum_{j=1}^k (i - j - 2m)^2 c_{ij} \quad (7)$$

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

$$\text{Correlation} = \sum_{i=1}^k \sum_{j=1}^k (i - m)(j - m) c_{ij} \quad (8)$$

f) *Homogeneity*: Show the first order inverse element difference moment.

$$\text{Homogeneity} = \sum_{i=1}^k \sum_{j=1}^k \frac{c_{ij}}{n} \quad (9)$$

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

$$\text{MaximumProbability} = \max(c_{ij}) \quad (10)$$

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

$$\text{Inverse} = \sum_{i=1}^k \sum_{j=1}^k \frac{c_{ij}}{n^2} \quad (11)$$

Where:

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

Edge Direction Histogram (EDH)

For creating EDH, this study uses the saturation channel of HSV color space. Initially, we will be performed the Gaussian smoothing against this channel. After that, perform the edge points 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. 1. Sobel Operators.

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

Gabor Filter (Wavelet) [9]:

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

$$G_{mn}(x, y) = \sum_s \sum_t I(x - s, y - t) \psi_{mn}^*(s, t) \quad (12)$$

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:

$$\psi(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right] \cdot \exp(j2\pi Wx) \quad (13)$$

Where W is the modulation frequency.

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

$$\psi_{mn}(x, y) = a^{-m}\psi(\tilde{x}, \tilde{y}) \quad (14)$$

Where m ($m = 0, 1, \dots, M-1$) and n ($n = 0, 1, \dots, N-1$) are the scale and orientation of the wavelet respectively.

$$\tilde{x} = a^{-m}(x \cos \theta + y \sin \theta) \quad (15)$$

$$\tilde{y} = a^{-m}(-x \cos \theta + y \sin \theta) \quad (16)$$

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.

$$E(m, n) = \sum_x \sum_y |G_{mn}(x, y)| \quad (17)$$

Local Binary Pattern (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 [10]. 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.

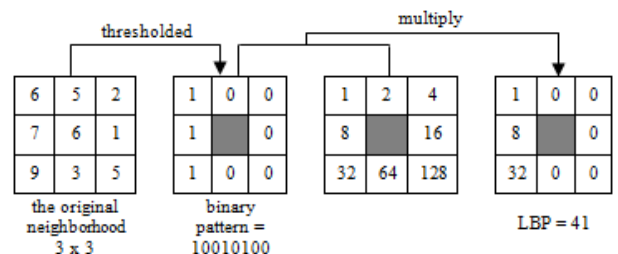


Fig. 2. Illustration of basic $LBP_{8,1}$.

Mathematically, it can be done as follow:

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(g_p - g_c) \cdot 2^p \quad (18)$$

Where $s(x)$ is threshold function

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (19)$$

The variable in the Eq. 18 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 [11].

III. FRAMEWORK OF REMOTE SENSING IMAGE RETRIEVAL SYSTEM

As mentioned before, we use two types of RS-IRS. First, basic RS-IRS uses global low-level model. The process of this RS-IRS is divided into four steps, i.e. features extraction, similarity measurement, indexing process, and displays the final result. Second, improvement of basic RS-IRS implements feature selection. In the following sub-sections, we will explain those steps in detail.

A. Feature Extraction

In this study, we use five features, including color moment, EDH, GLCM, Gabor, and LBP. For each of features, the process uses different pre-processing approaches. In color moment extraction, we implement pre-processing by converting the RGB images into CIELab color space, whereas the GLCM, Gabor, and LBP extraction, we convert the RGB images into gray-level images as pre-processing step. EDH extraction uses Saturation channels, thus the pre-processing step is converting the RGB images into HSV color space.

There are several parameters used in this features extraction process. First, we use 5 x 5 windows for GLCM extraction as recommended in [8]. Second, we use scale and orientation is 4 and 6, respectively, for Gabor extraction. It is recommended in [12].

B. Similarity Measurement

The feature similarity for local binary pattern is measured by using histogram intersection [13], while color moment, EDH, GLCM, and Gabor are measured by Euclidean Distance.

C. Feature Selection

Based on the features used in this study, we have 303-dimension of features vector, i.e. 9-dimension of color moment, 5-dimension of EDH, 8-dimension of GLCM, 24-dimension of Gabor, and 256-dimension of LBP. Therefore, we implement features selection technique to reduce the high dimensional of feature vector.

In this study, we use sequential forward floating selection (SFFS) algorithm, since it gives good performance in the land use classification by choosing optimal features set, as mentioned in [14]. The implementation of feature selection is only performed for color moment and GLCM. In extraction of GLCM, the descriptors are extracted against the same co-occurrence matrix, thus allowing the presence of information redundancy. It is also possible in the extraction of color moments, because the three types of color moments extracted on the same channel.

Basically, SFFS is a combination of three steps, i.e. inclusion, conditional exclusion and continuation of conditional exclusion. For a given m features, SFFS will choose n features out of all candidates. Let $X_k = \{x_1, x_2, \dots, x_k\}$ be the best k sets selected from candidates and Y_{m-k} are the rest features that have not been selected. The procedure of SFFS is as follow [15]:

- First step – Inclusion.
 - In this step, the process will stop if the $k = n$, otherwise select one feature x_{k+1} from Y_{m-k} that guarantee the new subset, $X_{k+1} = \{X_k + x_{k+1}\}$, with the best Criterion Function (CF) value.
- Second step – Conditional Exclusion.
 - Perform Sequential Backward Selection (SBS) to find the feature x_r which is regarded as the worst feature in current subset
 - If $r = k + 1$, let $k = k + 1$ and return to the first step. It means X_k is the best feature subset so far
 - If $r \neq k + 1$ and $C(X_{k+1} - \{x_r\}) \leq C(X_k)$, let $k = k + 1$. It demonstrates the former feature subset is the best so far and X_{k+1} is still used to perform the next step of SFS algorithm
 - If $k = 2$ and $C(X_{k+1} - \{x_r\}) \geq C(X_k)$, let $X_k = \{X_{k+1} - x_r\}$, record $C(X_k) = C(X_{k+1} - \{x_r\})$.
 - Moving to the third step
- Third step – Continuation of Conditional Exclusions
 - $X'_k = \{X_{k+1} - x_r\}$, this subset is the new best feature subset
 - SBS is performed to feature subset X'_k
 - $C(X'_k - \{x_s\}) \leq C(X_{k-1})$, let $X_k = X'_k$ take down $C(X_k) = C(X'_k)$ and return to first step without SBS.
 - Let $X'_k = \{X_{k+1} - x_s\}$, $k = k - 1$
 - If $k = 2$, let $X_k = X'_k$, record $C(X_k) = C(X'_k)$ and return to first step
 - Perform the third step recursively

IV. PROPOSED OBJECTIVE EVALUATION

The aim of this method is to overcome the degeneracy problem, which happens in previous method proposed by Jing Sun and Ying-Jie Xing. The figure in appendix A describes the objective evaluation process.

According to the figure in appendix, the proposed retrieval accurate ratio is measured by using the following equation:

$$pAR_{mn} =$$

$$\frac{\sum_{v \in H} \frac{1}{\text{rank position of } h \text{ in the retrieval system } T_{mn}}}{\sum_{k=1}^t \frac{1}{k}} \quad (20)$$

Where H is a set of retrieval result, T_{mn} , in which belong to the subject S_m and t is the number of retrieval result, T_{mn} . In this case, T_{mn} appropriate to the number of image in the S_m , i.e. 25. The numerator value of pAR uses the utility concept which ensures that if the nearer to the top of ranking is a relevant document, then the value will be greater. Whereas the denominator uses maximum possible utility to ensure that if the number relevant images which is retrieved is greater, then has the value of proposed average ratio will be greater.

V. EXPERIMENT AND RESULTS

A. Dataset

This study use remote sensing image database. This database contains 200 high resolution remote sensing images with size 1024 x 1024 pixel. These images are RGB images.

B. 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

C. Performance Measure

The performance result of the RS-IRS is evaluated using proposed objective evaluation. The image database is divided into 8 subjects, including stadium, farmland, urban, coastal, volcano, forest, junction or cloverleaf junction, and airport.

D. Experimental Result

The experiment is implemented to perform the objective evaluation for the improved RS-IRS which implements feature selection algorithm, i.e. 6-dimension selected color moment, 4-selected GLCM, EDH, Gabor, and LBP. Notice the following retrieval results, to distinguish the previous objective evaluation and proposed objective evaluation. There are two different queries, where each of queries gives the same number of retrieval results but in a different order.

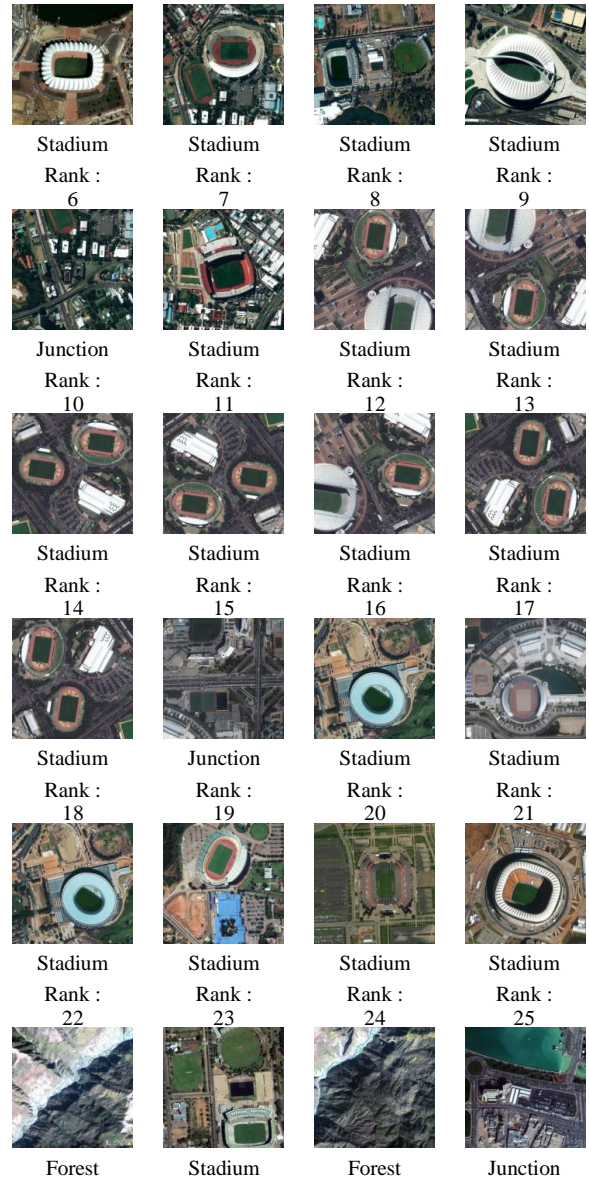
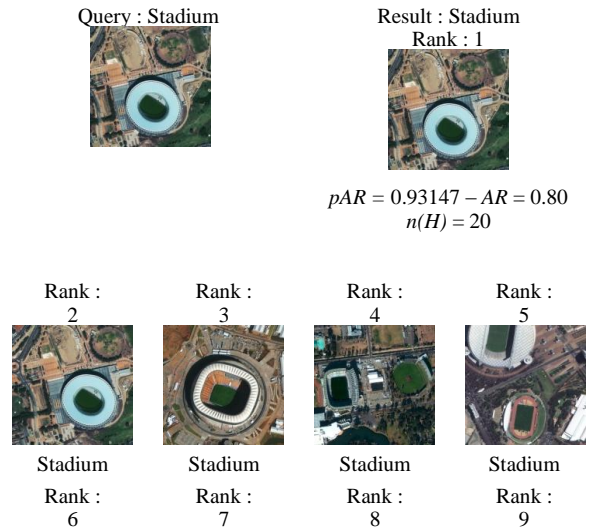
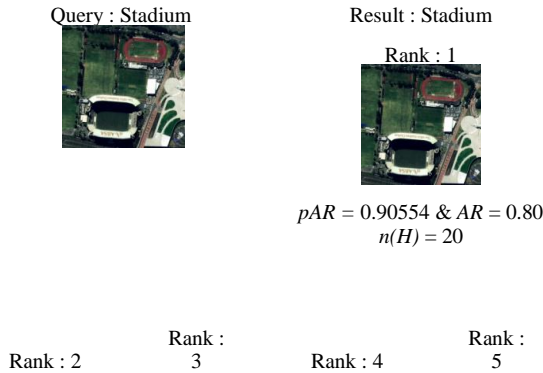


Fig. 3. Retrieval Result for Query Stadium No. 4



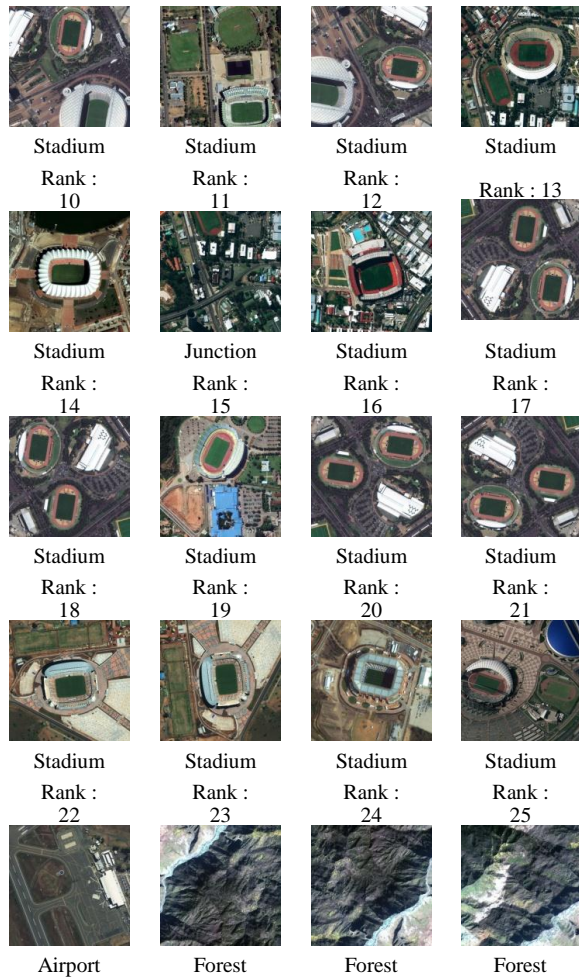
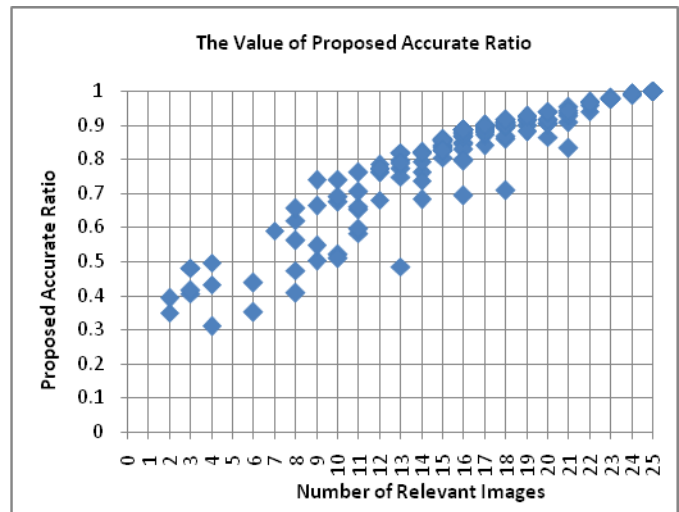


Fig. 4. Retrieval Result for Query Stadium No. 5

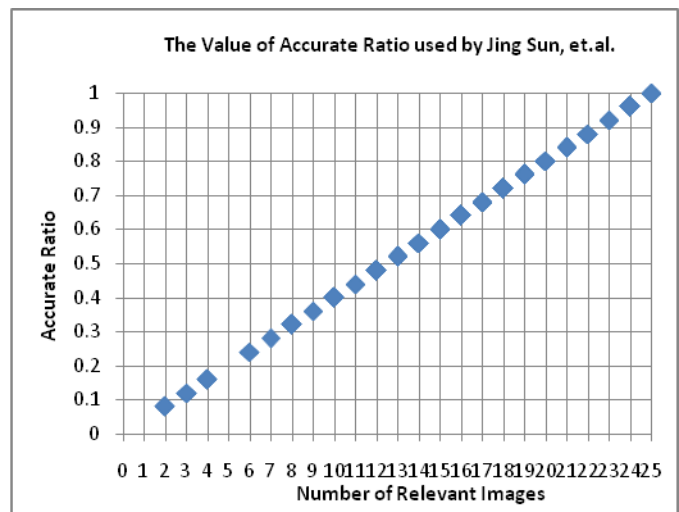
Both retrieval results above show the number of relevant images in the retrieval result is the same, i.e. 20 images. The difference lies in the order of appearance of irrelevant images. In the query - stadium number 4, the irrelevant image firstly appear on ranking of 6 and four other irrelevant images on ranking of 15, 22, 24, and 25. On contrary, in the query - stadium number 5, the irrelevant image first appeared on the ranking of 11 and four other irrelevant images on ranking of 22 to 25, four last retrieved images.

If we use previous objective evaluation, we can conclude that both retrieval results are the same, because of the degeneracy score, i.e. both of retrieval accurate ratio is 0.8. But in fact, the retrieval results for the query stadium - number 5 better than the query stadium - number 4, because the irrelevant image is not shown on the early retrieval results, but on some last images. And this can be distinguished by using the proposed retrieval accurate ratio, i.e. the proposed retrieval accurate ratio for the query stadium – number 5, $pAR = 0.93147$, is higher than query stadium – number 4, $pAR = 0.90554$.

The following figure describes the relationship between the number of relevant images with the value of accurate ratio, i.e. proposed score in figure (a) and score used by Jing Sun, et.al. in figure (b).



(a). Proposed Scoring



(b). Previous Scoring

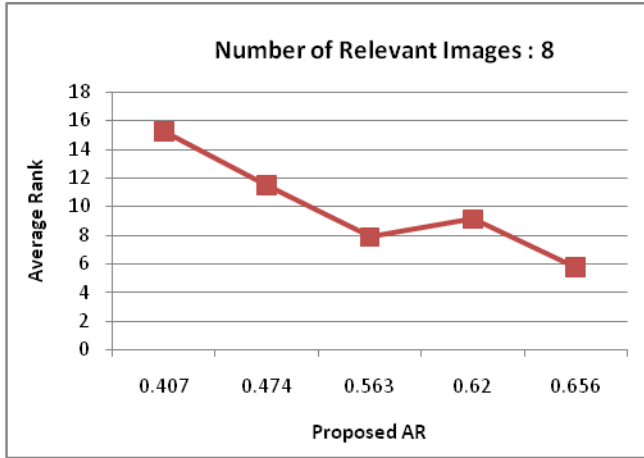
Fig. 5. Graph of the Value of Accurate Ratio for Each Number of Relevant Images

Both figures above show if the number of relevant images is higher, then the value of accurate ratio will greater. The difference lies on the number of accurate ratio value for each of the number of relevant images. In the proposed scoring, n -relevant images can have m -values of accurate ratio, where $m \geq 1$. For example, see the figure on 8-relevant images. This part has 5-value of accurate ratio. It depends on the position of ranking of relevant images. While in the previous scoring, used by Jing Sun, et.al, for all the retrieval result which gives 8-relevant images have one value of accurate ratio, 0.32, without relying on the ranking of the relevant image. See the following table for the detail explanation.

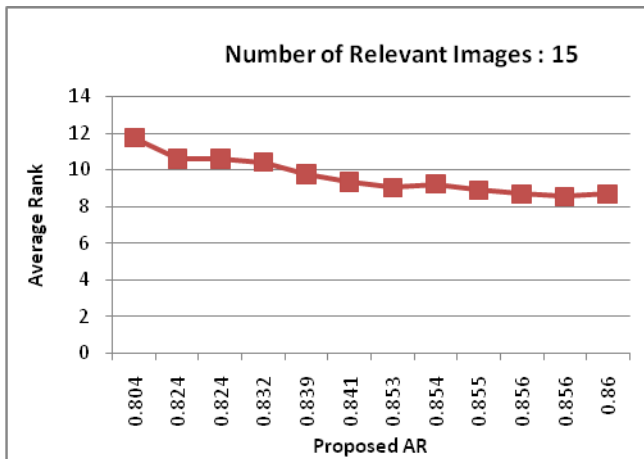
TABLE I
DETAIL EXPLANATION FOR THE RETRIEVAL RESULT WHICH GIVES 8-RELEVANT IMAGES

ID of Image Query	AR	pAR	1st	2nd	3rd	4th	5th	6th	7th	8th
124		0.407	1	5	10	12	22	23	24	25
45		0.474	1	3	7	11	13	15	18	24
123	0.32	0.563	1	3	4	5	7	11	13	19
17		0.620	1	2	3	4	9	14	19	21
43		0.656	1	2	3	5	7	8	9	11

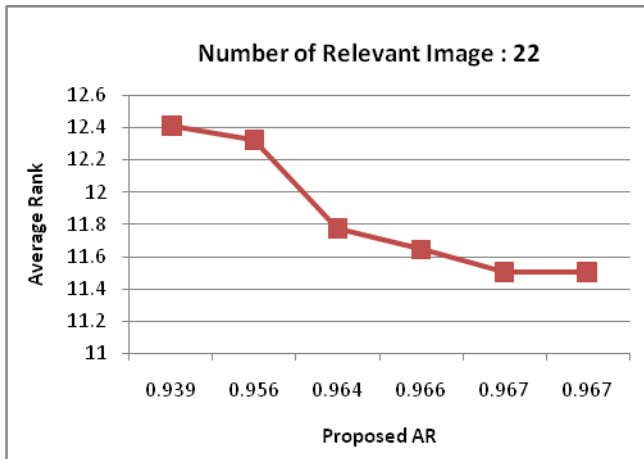
For each number of relevant images, the value of AR is higher if the average of the relevant image ranking gets smaller, or in other words, the more relevant images have the foremost ranking. It can be seen in the following figure.



(a). Number of Relevant Images: 8



(b). Number of Relevant Images: 15



(c). Number of Relevant Images: 22

Fig. 6. Graph of the Average Rank for Each Query

The third figure above show the inclination that the value of the proposed AR is inversely proportional to the average ranking of retrieval result.

The data used in fig. 5 and fig. 6 is an objective evaluation for implementation of improved RS-IRS using 6 selected

feature for color moment (i.e. skewness value of L^* and a^* channel, mean value of a^* and b^* channel, and standard deviation value of a^* and b^* channel), 4 selected feature for GLCM (contrast, cluster shade, homogeneity and maximum descriptor), EDH, Gabor, and LBP. This combination used since it gives the accurate retrieval results as seen on the following figure.

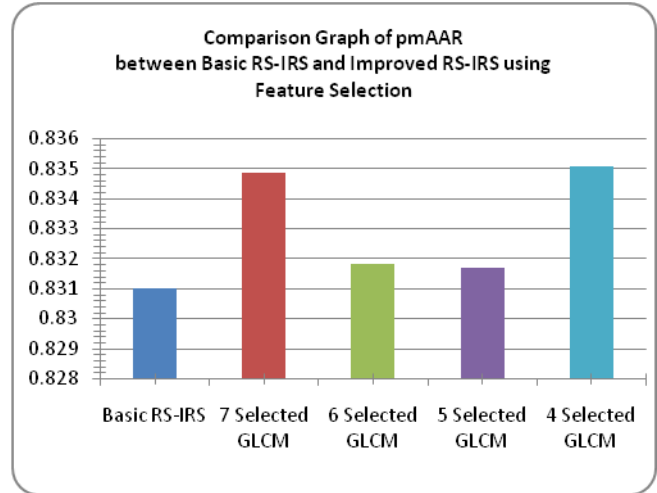


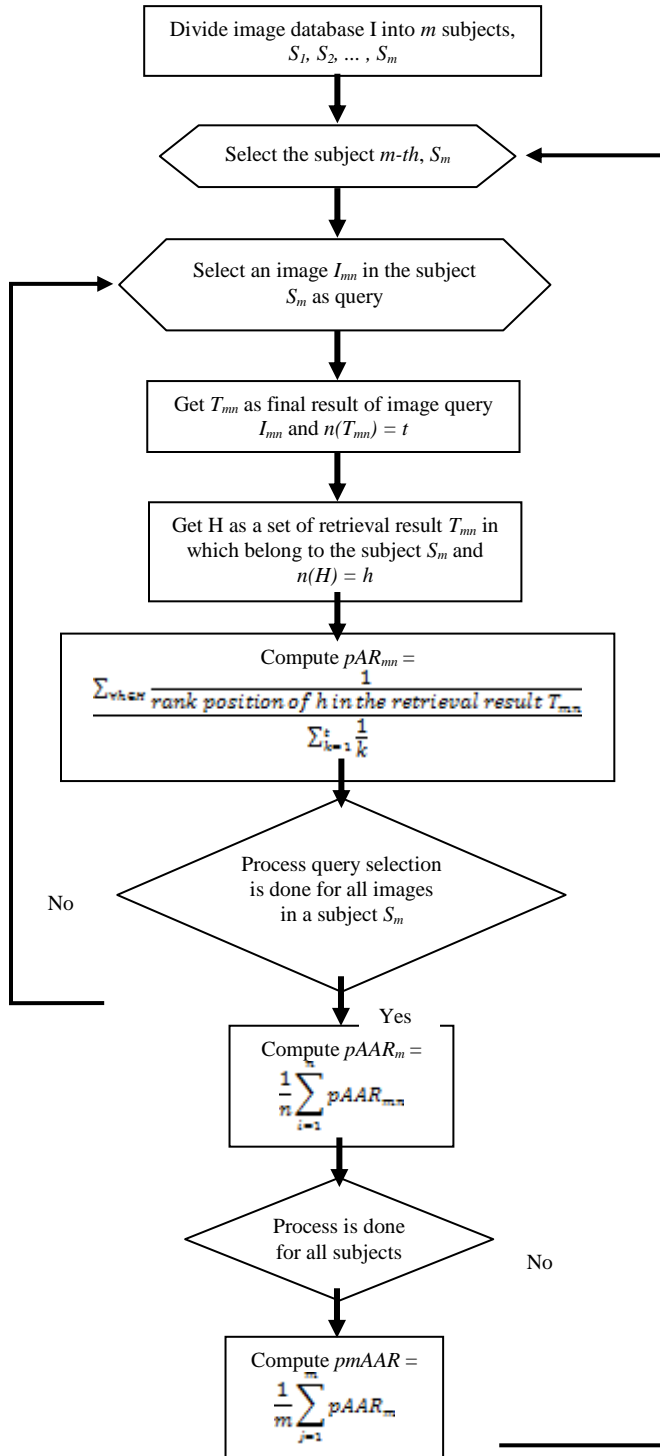
Fig. 7. Comparison Graph of $pmAAR$ between Basic RS-IRS and Improved RS-IRS

Based on the figure above, feature selection can improve the value of $pmAAR$ but in the lower percentage, i.e. lower than 0.5%.

VI. CONCLUSION

In order to make the evaluation result become fair and consistent without relying on the user judgment, we pursued a new mechanism of objective evaluation for remote sensing image retrieval system. This mechanism is performed by dividing the image database into subjects. First, use a subject and make it as ground truth. Subsequently, we take an image of the ground truth as a query Q . Retrieval accurate ratio for query Q will be computed as a comparison between utility value and maximum possible utility. The utility value represents the real utility for ranking in response to the query Q , while the maximum possible utility represent the maximum value of utility can be achieved for the query Q , with respect to ground truth. Repeat the retrieval process for all the images in the ground truth as the query and averaged accurate ratio ($pAAR$) for a subject is the average value of all the accurate ratio of all queries. Furthermore, these process is performed for all subjects in the database, thus we can obtain the mean averaged accurate ratio ($pmAAR$) as mean value of all averaged accurate ratio for all subjects.

APPENDIX



REFERENCES

- [1] J. Sun and Y.-J. Xing, "An Effective Image Retrieval Mechanism Using Family-based Spatial Consistency Filtration with Object Region", *International Journal of Automation and Computing*, vol. 7(1), February 2010, pp. 23-30.

- [2] M. Stricker and M. Orengo, "Similarity of Color Images," *Proceedings of IS&T and SPIE Storage and Retrieval of Image and Video Databases III*, 1995, pp. 381-392
- [3] F. Long, H. Zhang, and D.D. Feng, "Fundamentals of Content-Based Image Retrieval", *Proc. of Multimedia Information Retrieval*, 2002 [available online : [http://twiki.di.uniroma1.it/pub/Estrinfo/Materiale/FUNDAMENTALS_OF_CBIR_\(Long_et_al\).pdf](http://twiki.di.uniroma1.it/pub/Estrinfo/Materiale/FUNDAMENTALS_OF_CBIR_(Long_et_al).pdf)]
- [4] 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*, vol. 3, 2009, pp. 20-23
- [5] G.N. Srinivasan, and G. Shobha, "Statistical Texture Analysis", *Proceedings of World Academy of Science, Engineering, and Technology*, Dec. 2008, vol. 36, pp. 1264-1269
- [6] T. Ojala and M. Pietikainen, "Texture Classification, Machine Vision, and Media Processing Unit", University of Oulu, Finland [available online : http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/OJALA1/texclas.htm]
- [7] N. Ruan, N. Huang, and. W. Hong, "Semantic-Based Image Retrieval in Remote Sensing Archive : An Ontology Approach", *Proceedings of International Conference on Geoscience and Remote Sensing Symposium 2006*, July 31, 2006 – Aug 4, 2006, pp. 2903 – 2906
- [8] D. Chahyati, "Classification of Radar Images Based on Texture Features of Gray Level Co-occurrence Matrix Semivariogram and Wavelet Stationary", Master Thesis, Depok : University of Indonesia
- [9] D. Zhang, A. Wong, M. Indrawan, and G. Lu, "Content-based Image Retrieval Using Gabor Texture Features", *IEEE Transaction on PAMI*, 2000, pp. 13-15
- [10] 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*, vol. 24, no. 7, July 2002, pp. 971-987
- [11] 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
- [12] 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
- [13] J. Smith and S.F. Chang, "Automated Image Retrieval Using Color and Texture Feature", *IEEE Transaction on Pattern Analysis and Machine Intelligence – PAMI*, 1995.
- [14] A. Jain and D. Zongker, "Feature Selection : Evaluation, Application, and Small Sample Performance", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, 1997, pp. 153-158
- [15] PC. Lai, "Operational Knowledge Acquisition of Refuse Incinator Using Data Mining Technique", 2005, National Sun Yat-sen University

R. Kusumaningrum was born on 20 April 1981 in Banyumas. She achieved her undergraduate degree in Department of Mathematics from Diponegoro University, Semarang, Indonesia, in 2003 and 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.

She is currently working toward as lecturer in Department of Informatics, Diponegoro University. 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.

M.I. Fanany Dr. Fanany is an academic staff at Faculty of Computer Science - University of Indonesia. His research interests include 3D perception, reconstruction, recognition, and data-mining for autonomous and assisted driving, multi-modal sensor fusion for real-time decision and planning, combining vision and graphics for the application of advanced machine learning and remote sensing. Before joining the faculty, he worked at Future

Project Div. Toyota Motor Corp, Japan, as a member of middleware development and recognition team; NHK Engineering Services Inc. as a researcher of IT21 Millennium Project on Advanced High Resolution and Highly Sensible Presence 3D Content Creation funded by NICT Japan; and a JSPS Fellow and Research Assistant at Imaging Science and Engineering, Graduate School of Information Science and Engineering, Tokyo Institute of Technology (TIT). He served as the Chairman of TIT IEEE student branch 2002-2003 and member of IAPR, IEEE, and ACM SIGGRAPH.

A.M. Arymurthy 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.

She is a Professor in Faculty of Computer Science, University of Indonesia. Her main research activities are image processing and pattern recognition.