A FLEXIBLE SUB-BLOCK IN REGION BASED IMAGE RETRIEVAL BASED ON TRANSITION REGION

Ahmad Wahyu Rosyadi, Renest Danardono, Siprianus Septian Manek, and Agus Zainal Arifin

Department of Informatics, Faculty of Information Technology, Institut Teknologi Sepuluh Nopember, Jl. Arief Rahman Hakim, Surabaya 60111, East Java, Indonesia

E-mail: wahyu16@mhs.if.its.ac.id, renest16@mhs.if.its.ac.id, septian16@mhs.if.its.ac.id

Abstract

One of the techniques in region based image retrieval (RBIR) is comparing the global feature of an entire image and the local feature of image's sub-block in query and database image. The determined sub-block must be able to detect an object with varying sizes and locations. So the sub-block with flexible size and location is needed. We propose a new method for local feature extraction by determining the flexible size and location of sub-block based on the transition region in region based image retrieval. Global features of both query and database image are extracted using invariant moment. Local features in database and query image are extracted using hue, saturation, and value (HSV) histogram and local binary patterns (LBP). There are several steps to extract the local feature of sub-block in the query image. The first, preprocessing is conducted to get the transition region, then the flexible sub-block is determined based on the transition region. Afterward, the local feature of sub-block is extracted. The result of this application is the retrieved images ordered by the most similar to the query image. The local feature extraction with the proposed method is effective for image retrieval with precision and recall value are 57%.

Keywords: *RBIR, Transition Region, Invariant moment, HSV, LBP*

Abstrak

Salah satu teknik dalam *region based image retrieval* (RBIR) adalah dengan membandingkan fitur global keseluruhan citra dan fitur lokal dari *sub-block* pada citra *database* dan *query*. *Sub-block* yang ditentukan harus dapat mendeteksi objek yang memiliki beragam ukuran dan lokasi. Jadi *sub-block* yang fleksibel pada ukuran dan lokasi sangatlah dibutuhkan. Kami mengusulkan sebuah metode baru untuk ekstraksi fitur lokal dengan menentukan ukuran dan lokasi *sub-block* yang fleksibel berdasarkan *transition region* dalam *region based image retrieval*. Fitur global pada citra *database* dan *query* diekstraksi menggunakan *invariant moment*. Fitur lokal pada citra *database* dan *query* diekstraksi menggunakan histogram *hue, saturation, and value* (HSV) dan *local binary patterns* (LBP). Terdapat beberapa langkah untuk mengekstraksi fitur lokal *sub-block* di citra *query*. Pertama, *preprocessing* dilakukan untuk mendapat *transition region*, kemudian *sub-block* yang fleksibel ditentukan berdasarkan *transition region* tersebut. Setelah itu, fitur lokal *sub-block* diekstraksi. Hasil dari aplikasi ini adalah citra-citra yang ditemu kembali dan diurutkan berdasarkan yang paling mirip dengan citra *query*. Ekstraksi fitur lokal dengan metode ini efektif untuk temu kembali citra dengan *precision* dan *recall* bernilai 57%.

Kata Kunci: *RBIR, Transition Region, Invariant moment, HSV, LBP*

1. Introduction

Content based image retrieval (CBIR) has been a popular research during recent decades. CBIR takes pictures from a large image database based on their visual resemblance [1-2]. Not all images have annotations and not every annotation can resemble images well, looking for images using their content has many advantages than using their annotation text. Furthermore, the weakness of text-based method can be overcome by using CBIR [1] [3].

One of the techniques in the CBIR system that gives a sample image as a query is query by example (QBE). This technique uses the extracted query feature and compared with the database image feature. This feature includes global and local features. Global features use the entire image to extract their features without considering the user interest[1][4][5], while local features only use some relevant regions of the images with respect to user interest known as region based image retrieval (RBIR). The region of interest (ROI) in the query image must be defined by the user to select the relevant region. Chosen ROI by the user is more relevant to be used in RBIR. However, it becomes troublesome if there are a lot of query images that
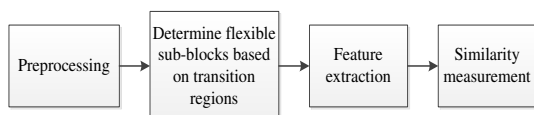
Figure 1. Design System

(a)   (b)

Figure 2. (a) Original image (b) Transition region image

have to be dealt by the user. Another method is to divide an image into several sub-blocks in a certain size to obtain the ROI [6]. The ROI is determined by selecting the sub-block which is overlapped with the object. Another approach is using Region Important Index (RII) and Saliency Region Overlapping Block (SROB) to select the salient region as a sub-block automatically [7]. Another method is using the ratio of proportional overlapping object (RPOO) which uses a threshold value to determine the degree of overlapping between object and sub-block [1]. In addition, there is a method that make the region of image flexible. This method uses semantic meaningful region (SMR) that is RII with size equal or greater than threshold [8].

Fix location and size sub-block is commonly used in Region Based Image Retrieval. It divides an image into several regions or sub-blocks $b \times b$ in a certain size. Fix location and size sub-block selects the sub-block based on the region of image. If there are several detected objects in all regions of image, sub-blocks will be made in all regions. These sub-blocks may be irrelevant because all regions of image will be used. Fix location and size sub-block also has a weakness in finding the relevant sub-block in image containing small object. It occurs because the size of an object does not meet the minimum threshold to become the relevant sub-block. Another reason is because the use of fix location sub-block. Fix location sub-block cannot adapt the location of an object which means that it can slice and divide one object into several objects if the location or size of the object exceeds the region. If the size of a divided object does not meet the minimum threshold, its region will not be selected as a sub-block even if the actual size of an object meets the minimum threshold. In addition, although method [8] uses the flexible region, that region has to be equal or greather than the threshold to be the ROI. It makes this method not efficient if there is no ROI in the image because the method will use the entire image to campare. So, the flexible sub-block which is able to adapt the size and the location based on the detected object is needed.

In this paper, we propose a new method for local feature extraction by determining the flexible size and location of sub-block based on the transition region in region based image retrieval. Extracted local feature from the flexible sub-block can satisfy the user because it can handle any size of the detected object in the query image and im-

prove the precision and recall value of the retrieving results.

The rest of this paper is divided into section 2-5. In section 2, proposed method is described. Section 3 reports the experimental results. Evaluations are described in Section 4, and conclusions are described in Section 5.

2. Methods

Query image is an example image that is compared with the database images to get the similar images to it. The database images are images in the database that will be retrieved based on their similarity to the query image. In this study, the query image is compared with the database image by measuring the distance of local and global features. The distance of global and local feature of query and database image is calculated by using Euclidean distance. The global feature is extracted from the entire image. However, the local feature is extracted from sub-block of image. The sub-block has to be the relevant sub-block that has object inside. The proposed method has to select the region or sub-block in flexible size and location based on the transition region.

There are four main steps as we can see in Figure 1 that are conducted in retrieving the images. The first, preprocessing is conducted to get the transition region. After that, the transition region is used to determine the flexible sub-block. Then, local feature is extracted from the determined sub-block. After the features of query image are obtained, the similarity measurement is conducted to determine the similar images.

Preprocessing

Preprocessing is conducted to obtain the transition region image that will be used in the following process. The noise in the query image has to be reduced by using the Gaussian filter. Query image is also required to have a one-color channel in order to speed up the computing process in the next process. We use the grayscale function to change the image with three-color channel into the one-color channel.

To detect the object, some research use image segmentation [1][6]. Image segmentation is a tech-
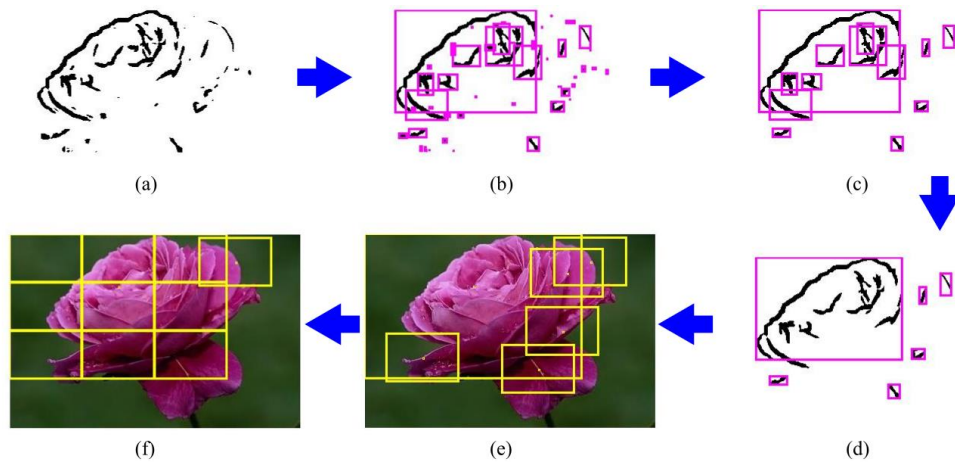
Figure 1. (a) Transition region image (b) Transition region image with its area or block (c) Normalized transition region with its area or block (d) Image after merging the transition region (e) Original image with adjusted block size

nique for separating an object from the background [8–13]. In this study, transition region is used as the salient region which is able to find the transition zone near the contour of the true object [11]. The transition region extraction process is performed after filtering and grayscale converting process to obtain the transition region as we can see in Figure 2(b).

A transition region is a structure of image that is similar with the edge. The transition region has three characteristics [15]. The first is the transition region usually has a width of several pixels near an edge. The second is that it is around the object and should be located between the object and the background. The last characteristic is that there is a dramatically and frequently change in gray levels within pixels in the transition region. Thus, the information for extracting transition regions can be obtained by transitional pixels which have larger magnitude and frequency of gray level changes than non-transitional pixels.

There are many descriptors [15–17] that have been developed for extracting transition regions. In this paper we use local variance to extract the transition region. Local variance can distinguish the area contains edge or not [12]. Edge generally exists in the area that has high value of variance. Hence, local variance is used because it is more important for finding transition region than local complexity [18].

For a center pixel $p(i,j)$ of $m \times m$ local neighborhood, the local variance can be calculated using equation(1).

$$LV(i,j) = \frac{1}{m^2-1}\sum_{x=1}^{m}\sum_{y=1}^{m}(f(x,y)-\bar{f})^2 \quad (1)$$

where $f(x,y)$ defines gray level value of a local coordinate in a neighborhood and the mean of gray level of that neighborhood is denoted as $\bar{f}$.

By sliding the window from left to right and top to bottom, the local variance measurement is conducted throughout the image to achieve the matrix of variance as shown in equation(2).

After local variance matrix is obtained, it is converted into a vector and sorted all values in descending order. Transitional pixels are generated by choosing the first αn pixels from sorted vector, where n is the total pixel number in the image and the value of α is 0.05 based on [11]. A label matrix TP is defined by using equation(3) to denote transitional pixels.

$$LV = \begin{bmatrix} LV(1,1) & LV(1,2) & ... & LV(1,N) \\ LV(2,1) & LV(2,2) & ... & LV(2,N) \\ ... & ... & ... & ... \\ LV(M,1) & LV(M,1) & ... & LV(\ddot{M},N) \end{bmatrix} \quad (2)$$

where, M denotes the height of the image and N denotes the width of the image.

$$TP(i,j) = \begin{cases} 0, \text{ if } p(i,j) \text{ is a transitional pixel} \\ 1, \qquad\qquad \text{otherwise} \end{cases} \quad (3)$$

The transition region is composed by grouping connected transitional pixels and making every different group into the different transition region TR. Then all transition regions are labeled with different number k.

Determine flexible sub-blocks based on transition regions

Flexible sub-block is used for the query image. However, fix sub-block is used for the database image. Database image is divided into several sub-blocks in a certain size $b \times b$. Certain size $b \times b$ is also used for the query image to obtain the standard size of sub-block. The standard size of sub-block
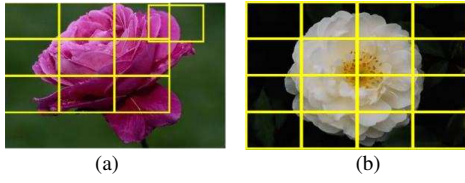
Figure 2. (a) Sub-block in query image (b) Sub-block in database image

can be obtained by dividing the query image into several sub-blocks in a certain size $b \times b$, then the width and the height of a sub-block are used as a standard size of sub-block.

The goal of determining flexible sub-blocks is to make a sub-block adjust to the size and the place of the transition region. Flexible sub-block can focus on the transition region and minimize the selected background in the sub-block. Determining flexible sub-block can be conducted via following steps:

1. Measure the standard size of sub-block's area LSB by multiplying the standard size of sub-block's width WSB and the standard size of sub-block's height HSB using equation(4) and (5) respectively.

$$WSB = N/b \qquad (4)$$

$$HSB = M/b \qquad (5)$$

2. Make an area for every transition region as shown in Figure 3(b). The area is based on the top-left and the bottom-right of the transition region. The top-left consist of two values: minimum index of row from transition region k bt_k and minimum index of column from transition region bl_k. And the bottom-right consist of maximum index of row from transition region k bb_k and maximum index of column from transition region br_k.

3. Measure the wide of every transition region LTR_k by multiplying its width WTR_k and its height HTR_k. It can be calculated as:

$$WTR_k = br_k - bl_k \qquad (6)$$

$$HTR_k = bb_k - bt_k \qquad (7)$$

$$LTR_k = HTR_k * WTR_k \qquad (8)$$

4. Delete a transition region that has an area less than 3% of a standard sub-block's area LSB, the result is shown in Figure 3(c).
5. Merge a transition region with another 7transition region if its center is in area of another transition region as illustrated in Figure 3(d).
6. Make a block BTR_k for every transition region that its width WTR_k and its height HTR_k

adjust their size according to the standard size of sub-block's width WSB and height HSB as shown in Figure 3(e). It can be defined as:

$$WTR_k = \begin{Bmatrix} WSB * z(v)| \\ \left(WSB * (z(v) - 1)\right) \\ < WTR_k < \left(WSB * z(v)\right)| \\ \forall z(v) \in z \end{Bmatrix} \quad (9)$$

$$HTR_k = \begin{Bmatrix} HSB * z(v)| \\ \left(HSB * (z(v) - 1)\right) \\ < HTR_k < \left(HSB * z(v)\right)| \\ \forall z(v) \in z \end{Bmatrix} \quad (10)$$

where z is an array having values ranging from 1 to b and defined as follows

$$z = 1, 2, 3, \dots, b \qquad (11)$$

The width WTR_k and the height HTR_k of block are updated but the center is not. It can adjust the top-left, top-right, bottom-left, and bottom-right automatically. Record the probability number of sub-block that can be made vertically $SHTR_k$ and horizontally $SWTR_k$ from HTR_k and WTR_k. It defined as follows.

$$SWTR_k = \begin{Bmatrix} z(v)| \\ \left(WSB * (z(v) - 1)\right) \\ < WTR_k < \left(WSB * z(v)\right)| \\ \forall z(v) \in z \end{Bmatrix} \quad (12)$$

$$SHTR_k = \begin{Bmatrix} z(v)| \\ \left(HSB * (z(v) - 1)\right) \\ < HTR_k < \left(HSB * z(v)\right)| \\ \forall z(v) \in z \end{Bmatrix} \quad (13)$$

7. Delete the block that its center is in area of another not smaller block.
8. Record the index of row iy_k and column ix_k where the block BTR_k exists. The index is based on matrix $b \times b$ of image. It can be defined as

$$iy_k = \begin{Bmatrix} z(v)| \\ \left(HSB * (z(v) - 1)\right) \\ < bt_k < \left(HSB * z(v)\right)| \\ \forall z(v) \in z \end{Bmatrix} \quad (14)$$

$$ix_k = \begin{Bmatrix} z(v)| \\ \left(WSB * (z(v) - 1)\right) \\ < bl_K < \left(WSB * z(v)\right)| \\ \forall z(v) \in z \end{Bmatrix} \quad (15)$$

TABLE 1
PRECISION VALUE OF THRESHOLD 1%, 2%, 3%, 4%, AND 5%

Threshold	Precision
1%	0.55
2%	0.56
3%	**0.57**
4%	0.56
5%	0.56

TABLE 2
PRECISION OF RETRIEVING 100 IMAGES WITH DIFFERENT B

Label query	Precision	
	b=4	b=5
Africa	0.55	**0.57**
Beaches	**0.42**	0.34
Buildings	**0.51**	0.34
Bus	**0.84**	**0.84**
Dinosaur	**0.90**	0.74
Elephant	**0.25**	0.23
Flowers	0.43	**0.46**
Horse	**0.74**	0.70
Mountains	0.26	**0.34**
Food	**0.75**	0.58
Average	**0.57**	0.51

9. Divide block BTR_k into sub-block $STR_{k,it}$. BTR_k is divided into a tb_k number of sub-block. The value of tb_k is obtained by multiplying $SWTR_k$ and $SHTR_k$ that is shown in equation(16). Sub-block $STR_{k,it}$ has width equal to WSB and height equal to HSB as we can see in Figure 3(f).

$$tb_k = SWTR_k * SHTR_k \quad (16)$$

$$it = 1,2, \dots, tb_k \quad (17)$$

$$STR_{k,it} \subseteq BTR_k \quad (18)$$

10. Record the index of row $iys_{k,it}$ and column $ixs_{k,it}$ where the sub-block exists. The index is based on the BTR_k's index, $SWTR_k$, and $SHTR_k$. It can be defined as

$$ixs_{k,it} = ix_k + ic(v) - 1 \quad (19)$$

$$iys_{k,it} = iy_k + ir(v) - 1 \quad (20)$$

where

$$ic = 1,2, \dots, SWTR_k \quad (21)$$

$$ir = 1,2, \dots, SHTR_k \quad (22)$$

11. Use the block as the sub-block SQ_k in image query, and the sub-block of a block as sub of sub-block $SSQ_{k,it}$

Feature Extraction

Local and global feature are used in the similarity seasurement process. Local feature of database image is extracted from all sub-blocks in it. On the other hand, local feature of query image is extracted from the selected sub-block only. Color and texture are used as local feature, while shape of object is used as descriptor for the global feature.

HSV is used as the color feature in this research. HSV color space is developed to adapt to the visual characteristic of human, considering hue, saturation, and value [19]. In this research, color is generated from every determined sub-block in query and database image.

Local binary pattern (LBP) is used as texture feature because it is efficient and has a good classification ability [20]. LBP has the ability to define the surface of an object and its relationship with the surrounding area [7]. In this research, LBP is calculated for every pixel in the determined sub-block. To represent the sub-block's texture, then the histogram of LBP is created.

Shape is used as the global feature in this research. Shape descriptor can be divided into two types that is region-based shape descriptor and contour-based shape descriptor. One of the important shape descriptors is moment invariants. Hu invariant moments is used as the global feature because it is invariant under rotation, translation, and also changes in scale [1].

Similarity Measurement

Similarity between query image and database image is computed by using Euclidean distance for every feature. The similarity distance of local feature is obtained by calculating the distance for each selected sub-blocks of query and database image as seen in Figure 4. Sub-blocks of query image are convoluted in database image in its own area to measure the minimum distance of all iterations. The distance of sub-block in query image that has some subs of sub-block is done by convoluting the sub-block of query image in database image, but the distance value of every iteration is determined by taking the average of all distance values of sub-block's subs. The distance between qu-ery image sub-block and database image sub-block can be calculated using equation(24), where $FSQ_{k,it}$ defined the local feature of sub-block's sub $SSQ_{k,it}$, $FDB_{q+isy,z+isx}$ defined the local feature of sub-block in database image, $q = \{-1, 0, 1\}$, $z = \{-1, 0, 1\}$, isy defined $iys_{k,it}$ and isx defined $ixs_{k,it}$.

The final distance of a local feature is described as average distance for every sub-block shown as
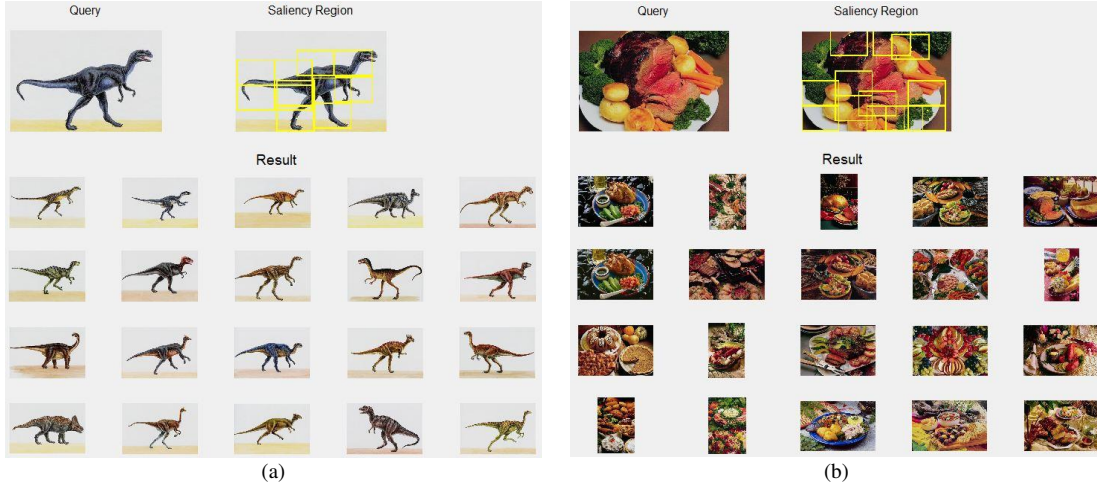
Figure 3. Example query and results for retrieval 20 images from category (a) Dinosaur (b) Food

$$d(SQ, SDB) = \frac{\sum_{k=1}^{n} d(SQ_k - SDB_{q,z})}{n} \quad (23)$$

$$d(SQ_k, SDB_{q,z})$$
$$= \min_{q,z} \left(\frac{\sum_{it=1}^{tb_k} \sqrt{(FSQ_{k,it} - FDB_{q+isy,z+isx})^2}}{tb_k} \right) \quad (24)$$

Euclidean distance is applied directly to measure the similarity of the global feature of query and database image. A weight is assigned to every feature before calculating the total distance. Based on [1], the optimal weight is 0.1, 0.4, 0.5 for the weight of color feature w_c, the weight of texture feature w_t, and the weight of shape feature w_s, respectively. The distances of two local features and one global feature are combined to obtain the total distance as shown in equation(25).

$$d(q, db) = w_c. d_c + w_t. d_t + w_s. d_s \quad (25)$$

The sorting by ascending order is performed after obtaining the distances to all images in the database. The smaller the distance, indicate that two images have higher similarity.

3. Results and Analysis

This research used Wang's image datasets that is usually used in image retrieval research. This dataset has 10 categories. Each category consists of 100 images. So there are 1000 images in this dataset.

Parameter to Delete Transition Region

In this study, the threshold used for determining the transition region is $0.05 \times n$ pixels as explained in [11]. The determined transition region is not al-

ways big, there will be some small transition regions. The small transition region may represent part of object or even part of background. We assume that the very small transition region represents the part of background. So it is important to delete the very small transition region, by deleting the very small transition region, we can reduce the number of transition region and reduce the computing time for retrieving the images. In this study we use a threshold value to delete the transition region. The threshold value is 3% of a sub-block's area based on our experiment and shown in Table 1.

Performance Measure Using Precision and Recall

Precision is used to measure the number of selected images are relevant, while recall is used to measure the number of the relevant images are selected. Precision and recall can be calculated by equation(26) and (27) respectively.

$$\text{precision} = \frac{TP}{TP+FP} \quad (26)$$

$$\text{recall} = \frac{TP}{TP+FN} \quad (27)$$

The number of relevant images that can be retrieved by the system is defined as *TP* (True Positive), the number of relevant images that cannot be retrieved by the system is defined as *FN* (False Negative), the number of irrelevant images but retrieved by the system is defined as *FP* (False Positive),and the number of irrelevant images that is not retrieved by the system is defined as *TN* (True Negative). The results of a query image are shown based on their ranking. The distance between query and image database is calculated to obtain the ranking of each image in database.
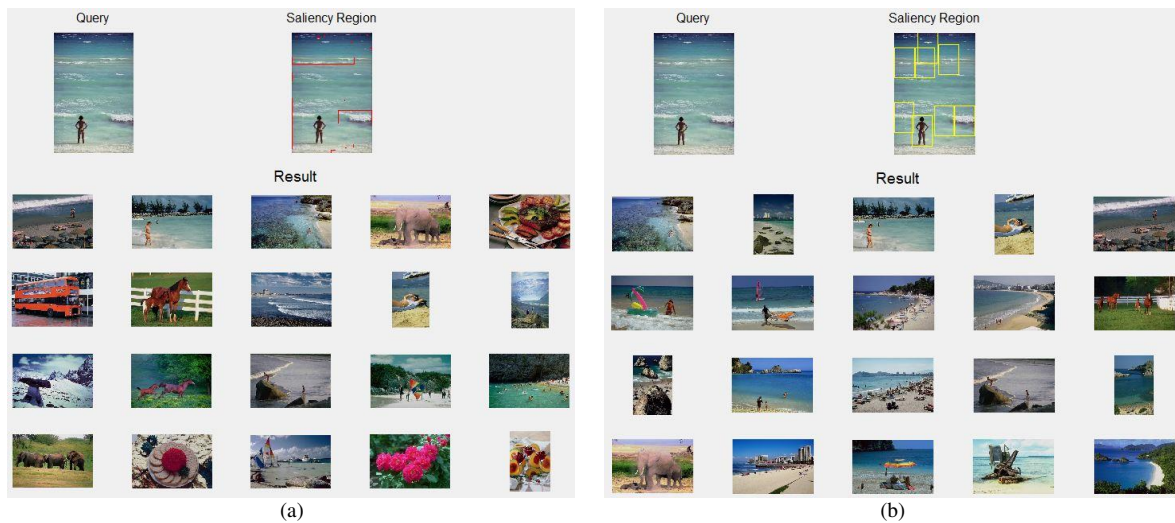
Figure 4. Example query and results for retrieval 20 image on query image with small object, (a) Query and results of RPOO, (b) Query and results of proposed method

The number of sub-block is very important for retrieving images. Parameter b with value of 4 has better precision value than b with value of 5 as shown in Table 2. So in this study, b with value of 4 is used.

It can be seen from Table 3 that the proposed method has better average precision value than Vimina & Jacob [6] and RPOO [1] in retrieving 20 images. The proposed method also has better precision value than RPOO [1] in most of cases except Elephant and Horse in retrieving 100 images as shown in Table 4. In addition, we also conducted image retrieval in image containing small object as shown in Table 5 and Figure 6. From Table 5 we can see that the proposed method has better average precision and recall value than RPOO in image retrieval of small object.

Evaluation

There are several points that could be discussed from the experimental results. As we can see in Table 2 that b with value of 4 has better precision value than 5. It is because the value of b affects the wide of sub-block. Bigger value of b means smaller wide of sub-block. When the wide of sub-block is small, the similarity measurement process is conducted in the small area because the convolution process is done in surrounding area of sub-block instead of in the entire image. Bigger value of b can focus the sub-block on small object well. However, it may not be good for similarity measurement process if the similar object is outside the area because the similarity value will be low. Moreover, bigger value of b may result bigger number of sub-block that can make the computing time is longer because there will be many sub-

blocks to compare by convoluting each sub-block in its surrounding area of sub-block. Based on our experiment, b equals 4 has better precision value than 5 because the size of sub-block with b equals 4 is fit for most of objects of the dataset image even for object that has small size.

Table 3 showed that Africa has precision value of 100% in retrieving 20 images and 55% in retrieving 100 images because the proposed method can determine sub-block from the parts of object. The determined sub-block has unique color and texture which is similar in its own category in database images.

Bus has precision value of 100% in retrieving 20 images and 84% in retrieving 100 images because Bus has unique color in the body and the glass, so it can determine a good transition region. Moreover, the shape of bus is similar in bus category, so the global feature shape can identify it well.

Dinosaur images also has precision value of 100% in retrieving 20 images and 90% in retrieving 100 images, it is because the selected sub-blocks in the query image are determined by the transition regions surrounding the body of dinosaur that is shown in Figure 5(a). In dinosaur, all the determined transition regions are from the dinosaur body because the background image is relatively homogeneous.

Although the horse category has heterogeneous background, the proposed method can result a good precision value of 100% in retrieving 20 images and 74% in retrieving 100 images because the sub-block is made surrounding the object. It is happened because the transitional pixels which have big value of local variance are determined from pixels surrounding the horse body. Another cate-

TABLE 1
PRECISION OF RETRIEVING 20 IMAGES USING SAME IMAGES IN DIFFERENT METHODS

Label query	Precision of retrieving 20 images		
	Vimina and Jacob	RPOO	Proposed method
Africa	0.71	0.80	**1.00**
Beaches	0.43	0.45	**0.85**
Buildings	0.53	0.70	**0.90**
Bus	0.85	0.95	**1.00**
Dinosaur	0.99	**1.00**	**1.00**
Elephant	**0.59**	0.55	0.20
Flowers	**0.90**	0.75	0.80
Horse	0.92	**1.00**	**1.00**
Mountains	0.38	0.55	**0.65**
Food	0.72	0.65	**1.00**
Average	0.70	0.74	**0.84**

TABLE 3
PRECISION AND RECALL OF RETRIEVING 100 IMAGES USING SAME IMAGES IN DIFFERENT METHODS

Label query	RPOO		Proposed method	
	Precision	Recall	Precision	Recall
Africa	0.42	0.42	**0.55**	**0.55**
Beaches	0.23	0.23	**0.42**	**0.42**
Buildings	0.39	0.39	**0.51**	**0.51**
Bus	0.67	0.67	**0.84**	**0.84**
Dinosaur	0.80	0.80	**0.90**	**0.90**
Elephant	**0.29**	**0.29**	0.25	0.25
Flowers	0.25	0.25	**0.43**	**0.43**
Horse	**0.79**	**0.79**	0.74	0.74
Mountains	0.21	0.21	**0.26**	**0.26**
Food	0.38	0.38	**0.75**	**0.75**
Average	0.44	0.44	**0.57**	**0.57**

TABLE 2
PRECISION AND RECALL OF RETRIEVING 100 IMAGES OF IMAGE CONTAINING SMALL OBJECT USING DIFFERENT METHODS

Label query	RPOO		Proposed method	
	Precision	Recall	Precision	Recall
Beach-49	0.26	0.26	**0.30**	**0.30**
Beach-62	0.21	0.21	**0.26**	**0.26**
Beach-64	0.23	0.23	**0.42**	**0.42**
Beach-74	0.24	0.24	**0.40**	**0.40**
Average	0.24	0.24	**0.35**	**0.35**

gory that has precision value of 100% in retrieving 20 images is food, it is because the determined sub-blocks from query image is not overall of image. Although the object is big, the determined sub-blocks from query image are part of the object, so it just compare the saliency regions.

In elephant category the proposed method has worst precision value, it is caused by the convolution process. Even the determined transition region is good, the convolution process makes it like a separate object and many images in database have similar color of elephant category.

The proposed method successfully retrieved the images with varying sizes of object as shown in the Table 3–5. Table 3 and 4 showed the result of retrieval in all categories with varying sizes of object, while the Table 5 showed the results of image retrieval containing small object.

For general image with varying sizes of object, the proposed method has better average precision and recall value than the result of RPOO as shown in Table 3 and 4. Furthermore, in cases of image retrieval of image contains small object, the proposed method also has the better result than the result of RPOO in all cases. It means that the proposed method has the ability in retrieving the images from image containing varying sizes of object.

The proposed method is capable of retrieving the images of image containing varying sizes of object because it uses the flexible sub-block. It has the ability to create sub-block in any size and any location based on the determined transition region. It can adapt the size of sub-block in the determined transition region by fitting the transition region's size with the standard size of sub-block. It is different from the fix sub-block that cannot adapt the size of sub-block and cannot make a sub-block in the region of a small object. The proposed method also places the sub-block based on the transition region instead of placing sub-block based on the region of an image. Placing sub-block based on the region of an image that is used in the fix sub-block will extract the irrelevant local feature if the object is not in the right place because the sub-block will contain too much background. Consequently, the extracted local feature from the flexible sub-block will be more relevant because the sub-block contains all the transition regions of object and minimizes the selected background.

Extracted local feature from the sub-block that has a lot of object inside is a good local feature. It is explained in [1] and [6] that the better segmentation result from the segmentation process will select sub-block based on the detected object more effectively and obtain a good image retrieval result.

Moreover, convolution process can detect the similar sub-block of database image in surrounding area of query image's sub-block. In this proposed method, sub-blocks of query image maybe a lot, but the selected sub-blocks is not overall image, so it can affect the performance of local feature. However, the convolution process sometimes determines the bad result although the determined sub-block is good, it occurs because many images in database have similar color or texture.

In addition, the generated sub-block may be irrelevant if the object is too small because the sub-block contains too much background. It happens

because the proposed method still needs a standard size to generate the flexible sub-block. The size of its flexibility has to be set by fitting its size with the standard size of sub-block.

Sometimes the deleted sub-block is a useful sub-block. It happens because the proposed method uses the size of sub-block as the only condition to delete sub-block. So the improvement in selecting the sub-block that will be deleted is needed. And there will be improvement in the convolution process, because in case of elephant category, the selected sub-block was good but the convolution process produced the bad precision value.

4. Conclusion

The proposed method successfully retrieved the images with varying sizes of object as shown in the experimental results. In most of categories in dataset, the proposed method has better average precision and recall value than the result of RPOO. Moreover, the proposed method showed superior results than RPOO in retrieving images from an image contains small object by surpassing the result of RPOO in all testing. However, the proposed method is not good enough to retrieve images from an image contains too small object. In addition, the proposed method does not have a good condition to determine which sub-block to be deleted.

References

[1] A. Z. Arifin, R. W. Sholikah, D. F. H. P., and D. A. Navastara, "Region Based Image Retrieval Using Ratio of Proportional Overlapping Object," *TELKOMNIKA (Telecommunication Comput. Electron. Control.*, vol. 14, no. 4, p. 1608, 2016.

[2] M. Athoillah, I. Irawan, M., and M. Imah, Elly, "Study Comparison of SVM-, K-NN- and Backpropagation-Based Classifier for Image Retrieval," *J. Ilmu Komput. dan Inf. (Journal Comput. Sci. Information)*, vol. 8, no. 1, pp. 11–18, 2015.

[3] H. Maghfirah, F. Arnia, and K. Munadi, "Temu Kembali Citra Busana Muslimah Berdasarkan Bentuk Menggunakan Curvature Scale Space (CSS)," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 6, no. 1, pp. 74–83, 2017.

[4] X.-Y. Wang, Y.-J. Yu, and H.-Y. Yang, "An effective image retrieval scheme using color, texture and shape features," *Comput. Stand. Interfaces*, vol. 33, no. 1, pp. 59–68, 2011.

[5] X. Wang, Z. Chen, and J. Yun, "An effective method for color image retrieval based on texture," *Comput. Stand. Interfaces*, vol. 34,

no. 1, pp. 31–35, 2012.

[6] E. R. Vimina and K. P. Jacob, "A Sub-block Based Image Retrieval Using Modified Integrated Region Matching," vol. 10, no. 1, pp. 686–692, 2013.

[7] N. Shrivastava and V. Tyagi, "Content based image retrieval based on relative locations of multiple regions of interest using selective regions matching," *Inf. Sci. (Ny).*, vol. 259, pp. 212–224, 2014.

[8] X. Yang and L. Cai, "Adaptive region matching for region-based image retrieval by constructing region importance index," *IET Comput. Vis.*, vol. 8, no. 2, pp. 141–151, 2014.

[9] G. P. Cahyono, A. Y. Aprilio, and H. Ramadhan, "Multithresholding in Grayscale Image Using Peak Finding Approach and Hierarchical Cluster Analysis," *J. Ilmu Komput. dan Inf. (Journal Comput. Sci. Information)*, vol. 7, no. 2, pp. 83–89, 2014.

[10] M. Hani'ah, C. S. kusuma Aditya, A. Harto, and A. Z. Arifin, "Cortical Bone Segmentation Using Watershed and Region Merging Based on Statistical Features," *J. Ilmu Komput. dan Inf. (Journal Comput. Sci. Information)*, vol. 8, no. 2, pp. 76–82, 2015.

[11] Z. Li, G. Liu, D. Zhang, and Y. Xu, "Robust single-object image segmentation based on salient transition region," *Pattern Recognit.*, vol. 52, pp. 317–331, 2016.

[12] P. Parida and N. Bhoi, "Transition region based single and multiple object segmentation of gray scale images," *Eng. Sci. Technol. an Int. J.*, vol. 19, no. 3, pp. 1206–1215, 2016.

[13] N. P. Husain and C. Fatichah, "Segmentasi Citra Sel Tunggal Smear Serviks Menggunakan Radiating Component Normalized Generalized GVFS," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 6, no. 1, pp. 107–114, 2017.

[14] D. Tuwohingide and C. Fatichah, "Spatial Fuzzy C-means dan Rapid Region Merging untuk Pemisahan Sel Kanker Payudara," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 6, no. 1, pp. 51–57, 2017.

[15] Y. J. Zhang and J. J. Gerbrands, "Transition region determination based thresholding," Pattern *Recognit. Lett.*, vol. 12, no. 1, pp. 13–23, 1991.

[16] C. Yan, N. Sang, and T. Zhang, "Local entropy-based transition region extraction and thresholding," *Pattern Recognit. Lett.*, vol. 24, no. 16, pp. 2935–2941, 2003.

[17] Z. Li and C. Liu, "Gray level difference-based transition region extraction and thresholding," *Comput. Electr. Eng.*, vol. 35,

no. 5, pp. 696–704, 2009.

[18] Z. Li, D. Zhang, Y. Xu, and C. Liu, "Modified local entropy-based transition region extraction and thresholding," *Appl. Soft Comput. J.*, vol. 11, no. 8, pp. 5630–5638, 2011.

[19] H. Zhang and X. Jiang, "An Improved Algorithm Based on Texture Feature Extraction for Image Retrieval," *Int. Conf. Intell. Human-Machine Syst. Cybern.*, vol. 8, pp. 281–285, 2016.

[20] G. Cheng and J. Chen, "A local feature descriptor based on Local Binary Patterns," pp. 251–258, 2016.