

A Object-Based Image Retrieval Using Feature Analysis and Fractal Dimension

Jeong-Bong Lee[†], Chang-Chun Park^{**}

ABSTRACT

This paper proposed the content-based retrieval system as a method for performing image retrieval through the effective feature extraction of the object of significant meaning based on the characteristics of man's visual system. To allow the object region of interest to be primarily detected, the region, being comparatively large size, greatly different from the background color and located in the middle of the image, was judged as the major object with a meaning. To get the original features of the image, the cumulative sum of the declination difference vector the segment of the object contour had and the signature of the bipartite object were extracted and used in the form of being applied to the rotation of the object and the change of the size after partition of the total length of the object contour of the image into the normalized segment. Starting with this form feature, it was possible to make a retrieval robust to any change in translation, rotation and scaling by combining information on the texture sample, color and eccentricity and measuring the degree of similarity. It responded less sensitively to the phenomenon of distortion of the object feature due to the partial change or damage of the region. Also, the method of imposing a different weight of similarity on the image feature based on the relationship of complexity between measured objects using the fractal dimension by the Boxing-Counting Dimension minimized the wrong retrieval and showed more efficient retrieval rate.

특징 분석과 프랙탈 차원을 이용한 객체 기반 영상검색

이정봉[†], 박장춘^{**}

요약

영상 검색의 수행 방법으로 사람의 시각 시스템의 특성을 기반으로 주요 의미를 갖는 객체의 효과적인 특징 추출을 통한 내용기반 영상 검색 시스템을 제안한다. 관심 객체 영역이 우선적으로 검출되도록 하기 위해 영상 내에서 비교적 면적이 크고 배경색상과의 차이가 크면서 영상의 가운데 위치하는 영역을 의미물 갖는 주요 객체로 판단하였다. 영상 고유의 특징을 얻기 위해서는 영상의 객체 윤곽선의 전체 길이를 정규화된 일정한 세그먼트로 분할한 후에 객체 윤곽선의 세그먼트가 갖는 편각차분 벡터들의 누적 합과 양분된 객체의 시그니처를 추출하여 물체의 회전과 크기 변화에 적용적인 형태 특징으로 사용한다. 이와 같은 형태 특징을 필두로 해서 질감 샘플과 칼라, 그리고 이심률 정보를 결합하여 유사도를 측정함으로써 이동, 회전 크기 변화에 강건한 검색이 가능했으며 영역의 부분적인 변화나 손상으로 인한 객체 특성의 왜곡 현상에 덜 민감하게 반응하였다. 또한 Box-Counting Dimension에 의한 프랙탈 차원을 이용하여 측정된 객체간 복잡도 관계를 기반으로 하여 영상 특징에 서로 다른 유사도 가중치를 부여하는 방법이 잘못된 검색을 최소화 하여 더욱 효율적인 검색율을 보였다.

Key words: image retrieval(영상 검색), object segmentation(객체 분할), feature(특징), signature(시그니처), fractal dimension(프랙탈 차원), complexity(복잡도)

※ 교신저자(Corresponding Author) : 이정봉, 주소 : 서울 성북구 길음1동 길음1차 삼성 래미안 아파트 102동 1302호 (136-111), 전화 : 02)450-3534, FAX : 02)450-3269, E-mail : jblee@cse.konkuk.ac.kr

접수일 : 2003년 4월 22일, 완료일 : 2003년 7월 28일
† 정회원, 건국대학교 대학원 컴퓨터공학과 박사과정

** 정회원, 건국대학교 컴퓨터공학과 교수
(E-mail : ccpark@cse.konkuk.ac.kr)

1. Introduction

Recently, As the multimedia processing application increases rapidly by going on increasing multimedia data, the efficient retrieval method of image information is required in many fields of

application and becoming the matter of major concern. Studies on content-based approach using information on the image for this purpose are actively done.

The initial retrieval method needs to extract the efficient feature representing image information because it is to retrieve the image by getting the similarity between images using the feature values of the overall image. It has the disadvantage that it is difficult to perform the retrieval of the image similar to the particular part of the query image.

At present, the image retrieval system is focused on the description of segmenting the object or region within the image as the previous work of image retrieval.

The object of segmentation in which this study is interested is the one of important meaning. But it can be said that it is a very difficult problem to judge the meaning extent of the image, because it is a very subjective problem to calculate the degree of importance that the object of the image has. To extract this object, this study further extended the initial hypothesis based on the hypothesis[1] that the object is generally located around the middle of the image. This study began from the perspective that the object of significant meaning is comparatively broad in area, greatly different from the background color and located in the middle of the image, and used the improved watershed algorithm to get a candidate region likely to have a certain meaning.

One of the most important matters in the content-based retrieval method can be said to be the method of choosing the feature set used to describe an image or an object. It should be robust to any change in the rotation, translation and scaling of an object and determine its feature, taking into consideration the efficiency of the processing time and the storage space.

For this purpose, contour information drawn from the divided object based on the characteristic of man's visual cognition that comes to identify an image is used as the main feature element to

describe the feature of an object. In addition, when a person sees an image, the person comes to see the shape of overall edge construction of the image at first, then see the more detailed part and come to identify the image.

Describing shape of an object is one of the most important work in image analysis. It should be applicable to any change in translation or scaling and easy to use and saved as best suited feature. And it should afford the local feature as well as global feature.

Such diverse methods as chain code string, curvature, polygon approximation, Fourier descriptor, and moment have been proposed as the contour-based method to describe this shape and applied to several fields. Of them, only chain code fails to express a form to the rotation and scaling of the object, and modified code such as derivative chain code[2] have been proposed.

In addition, Jian[3] put the lines along the boundary on the 2-dimensional plane, made those lines the 2-dimensional signal and expressed as the vector the coefficient obtained through Fourier transformation. But the method of use this contour has the disadvantage that it is likely to respond sensitively to a little transformation of the contour and Fourier descriptor or moment has the disadvantage of failing to provide the local information of a region.

On the one hand, Jain[4] used the normalized edge histogram intersection that applied invariant moment and smoothing as the method of combining shape information with color information. This method may be robust to any change in the scaling of the image but has the disadvantage that it fails to recognize a different image as a similar one due to more normalization or smoothing than needed.

This study attempted to define the shape feature that does not respond sensitively to a little local change of the contour while meeting the condition that it should be robust to any change in translation, scaling and rotation in order to complement

this problem and used it as the feature reflective of the overall shape of an object. And this shape information is composed of the regional descriptor and the boundary descriptor.

To overcome this problem, it is necessary to have the shape description that does not respond sensitively to a little local change of the contour while meeting the condition that it should be robust to any change in translation, scaling and rotation. Accordingly, an attempt was made to complement the problem the existing shape description had by defining the boundary descriptor and the regional descriptor and mutually combining the texture feature and eccentricity with each other. And this study determined and compared similarity by adding detailed information such as color histogram and defining the scale of weight-based similarity between the query image and database image, thereby enabling the image retrieval to be performed.

Especially, this study attempted to designate the weighted value of similarity between features according to the complexity of object shape quantitatively expressed by the fractal dimension and the segmented contour variance and reduce the possibility of retrieving the unwanted image in case of the quality of the image with the less complex form. Fig. 1 shows our retrieval system.

This paper is organized as follows. Section 2 explained object segmentation proposed by this

study. Section 3 explained the method for extracting the feature from the image. Section 4 explained the approach proposed about similarity weight. Section 5 investigated the actual experiment results and analysis. Section 6 referred to the conclusion and the future direction.

2. Proposed Object Segmentation Method

2.1 Preprocessing

The work of averaging is carried out after applying the median filter to the image. In this process, remarkable dots such as the noise existing within the image is not only eliminated but changed into the blurred image and forms the partial regionalization, and the work of regional split can be easily carried out. On the one hand, an attempt is made to get the color at the left uppermost corner region closely related to the background color. The color space is divided into 16 hues, 3 saturations and 3 intensities into 3 by transforming RGB colors into the HSI color space similar to the human visual traits using the equ.(1), and is divided into 146 bins by adding two value of intensity for the achromatic color of undefined hue values. Like this, the special value of the color could be simplified by quantizing the color space. Fig. 2 shows the result of averaging the image.

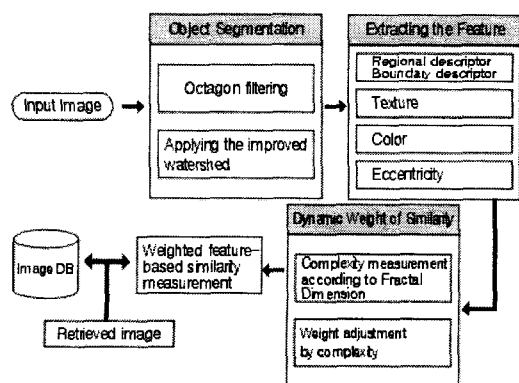


Fig. 1. Block diagram of the retrieval system



(a) original image (b) averaging image

Fig. 2. Result of averaging transformation

2.2 Object Segmentation of the 1-level

This study splitted the region by using the edge ingredient taking into consideration the gradient-

directional component of the image and the value of the modified color. First of all, the luminance of the pixel was calculated by applying the equ.(1) independently of the I value and RGB color image was transformed into the grey image in order to detect the direction of the gradient.

$$grey = 0.3R(x, y) + 0.59G(x, y) + 0.11B(x, y) \quad (1)$$

The direction of the gradient vector $a(x,y)$ is calculated by using 8 adjacent pixels and can be defined ∇f , the degree of gradient, and $a(x,y)$ can be defined as in the equ.(2).

$$a(x, y) = \tan^{-1}\left(\frac{G_y}{G_x}\right), \nabla f = mag(\nabla F) = [G_x^2 + G_y^2]^{1/2}$$

$$G_x = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)$$

$$G_y = (z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7) \quad (2)$$

Then, pixels within the small region including line patterns or dots shown in the object like tigers are either merged into the large region by applying the 5×5 octagon filter in (a) of Fig. 3 to averaged image and the remarkably small regions included within the large region is eliminated.

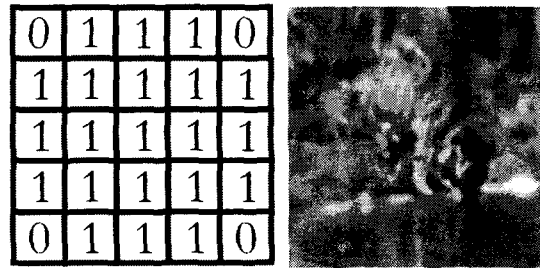
In the filtering process, this study obtained the histogram of all the pixels except the central pixel in the pixels existing within the filter. Then, it obtained the component of the gradient direction of the color region of the mode pixel with the highest color frequency and the corresponding color region in the histogram[9]. When the difference between these values and the color of the central pixel and the component of direction was compared with the threshold, the central pixel was changed into the color of adjoining pixels with the most colors. The equ.(3) is the condition for changing the central pixels within the filter into the color of mode pixels.

$$|Cc(x, y) - Cmax(x, y)| \leq T_c \text{ and } |Gc(x, y) - Gmean(x, y)| \leq T_g$$

$Cc(x, y)$: central pixel, $Cmax(x, y)$: mode color pixel

$Gc(x, y)$: gradient direction of central pixel (3)

$Gmean(x, y)$: mean value of the gradient direction of mode pixels.



(a) octagon filter (b) filtering image

Fig. 3. Octagon filtering result

2.3 Object Segmentation of the 2-level

2.3.1 Improved watershed algorithm

Because the number of catchment basins is directly related to the number of regions generated from the marker in the watershed method, this study introduced the method of [6] to choose the optimal marker as well as to reduce oversegmentation.

The marker image (x, y) set as 1 in the gradient image (x, y) with the relevant gradient comes to shape the 'n' regions in which markers with the pixel size 1 are mutually separated from each other. And because all the markers decided in this way are used as the beginning point of inundation, it segments the only label.

2.3.2 Region merging

Many meaningless small-scale regions not still merged into the large region exist in the image even though regional segmentation occurred through the implementation of the watershed process. On the premise that the small-scale region has a weak meaning, these regions are merged into the large region with similar attributes existing around and proceed to expand their regions.

First of all, the seed region is decided for merging the largest region of the regions still not segmented into a region of new objects and to extend into the surrounding regions.

In the process of merging regions, the region R attempts to merge with the adjoining region using the mean m and variance σ^2 . Each region decides

whether to merge with other regions based on the conditional equ.(5) by comparing the measurement value of its region and the measurement value of its surrounding region.

$$m_i = \frac{1}{n} \sum_{x,y} R_i(x,y), \sigma_i = \sqrt{\frac{1}{n} \sum_{x,y} (R_i(x,y) - m_i)^2}, xy \in R_i \quad (4)$$

$$|m_i - m_{i+1}| < k\sigma_i \quad (5)$$

Here, k is the constant that can decide the range of merging, and if k is small, the merge is subdivided and if it is large, the merging is extensively done. The number of regions that can be segmented to the maximum for the efficiency of calculation is limited to 30. The threshold value increases the number of regions and performs the process of merging each stage till the number of regions becomes less than 30. Fig. 4 shows the segmented regions.



Fig. 4. Region segmentation result

2.4 Extracting the Object of Main Interest

A different regional label is assigned to each region, and then since the object of meaningful interest of the image occupies the comparatively great area, a region in which the number of pixels accounts for less than 30 is judged as the region of the noise and not marked as the region. The algorithm extracting the major object comprises the following three steps:

(1) Three regions of broad area are selected from the image segmented into the regional unit and designated as the candidate object region.

(2) The color histogram of a region is drawn from the candidate region, and then the most frequently appearing bin is decided as C_f , the prevailing color representing the region. And the prevailing color, C_b , is obtained as to the corner region of the background in the same way.

(3) The window having one-third as large as the center of the image is set. Lastly, W_s , the size of the window region, and R_s , the candidate region of each object, and the overlapping rate S_{op} of common regions are calculated using the equ.(6) and by calculating $d[i]$, i , the candidate region with the greatest value is selected as the final object of main interest.

$$d[i] = \max(k_1 * S_{op}[i] + k_2 * C_R[i]) \quad (6)$$

$$i = 1, 2, 3 \quad k_1 = 0.6 \quad k_2 = 0.4$$

$$S_{op}[i] = \frac{R_s[i] \cap W_s}{W_s}, C_R[i] = \frac{|C_f[i] - C_b|}{C_f[i] + C_b}$$

(a) in Fig. 5 shows the extracted candidate region such as A, B, C, and (b),(c) show the finally extracted 'tiger' object region, contour.



Fig. 5. Final object of main interest

3. Extracting the Feature Vector

The feature vector is defined as the vector composed of two shape features s_b, s_r and texture t , color c , eccentricity e and complexity $comp$. An attempt was made to extract the feature vector F from the object and key-index these feature values in the database.

$$F_{vector} = (s_b, s_r, t, c, e, comp)$$

3.1 Shape Descriptor

In the modeling of the shape of the object, this

study defined the proper regional descriptor, proposed the standardized shape function unrelated to the change and translation of rotation and scale. And to heighten the discriminate power of the shape, it newly defined and combined the function of the applicable declination difference vector and the contour vector based on the method[7,10]. By doing so, this study approximated the structural shape of the region to the straight line more usefully and designated the boundary descriptor of the object.

3.1.1 Regional descriptor

This study proposed the radius-based bipartite signature into which the radius-based signature was improved. Signature expresses the contour as the primary function. The radius-based signature constitutes the distance from the centroid of the region to the contour. For this purpose, the oblique angle of the shape should be calculated to hold an object unchanged to its rotation. the oblique angle is defined as the angle of the axis in which the inertia moment becomes minimal. This angle(θ) is defined from the equ.(7).

$$\theta = \frac{1}{2} \arctan\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\right) \tag{7}$$

$$\mu_{p,q} = \sum_x \sum_y f(x,y)(x-\bar{x})^p (y-\bar{y})^q, \bar{x} = \frac{m_{10}}{m_{00}} \quad \bar{y} = \frac{m_{01}}{m_{00}}$$

$m_{p,q} = \sum_x \sum_y f(x,y)(x)^p (y)^q$, $m_{p,q}$ is the moment of the (p+q) order to the 2-dimensional function $f(x,y)$ and $\mu_{p,q}$ is the central moment of the (p+q) order.

The main axis of rotation to the object region is calculated from this angle and eccentricity is calculated. The width and height of the square including the object region are calculated using the following equ.(8) and (9) defined by calculating the eigen value of the matrix composed of moments, and eccentricity is calculated as the ratio of the major axis to the minor axis by designating the longer of two lengths as the major axis and the short as the minor axis[8].

$$height = \sqrt{2(a+c) - 2\sqrt{b^2 + (a-c)^2}} \tag{8}$$

$$width = \sqrt{2(a+c) + 2\sqrt{b^2 + (a-c)^2}} \tag{9}$$

$$a = \mu_{20}/\mu_{00} - (\bar{x})^2$$

$$b = 2(\mu_{11}/\mu_{00} - \mu_{10}(\bar{y})^2) \quad c = \mu_{02}/\mu_{00} - (\bar{y})^2$$

Then, the major axis is given as the dividing line of the object, and then the surrounding distribution to each region to the right and left of the line is defined as the statistical shape measurer.

If the segmented interval of the distance from the central point is long, it may be insensitive to any local change of the contour and if it becomes short, it may become sensitive to any local change of the contour conversely. Therefore, this study attempted to enable both the global feature and the local feature to be accommodated and the extent to which the amount of changes in distance affects changes in this feature to be reduced.

For this purpose, the centroid is designated as the center of the object, which in turn is divided into the right and left regions. Then the entire length of the contour in each of the left and right regions is moved up to 15° and 30° and the point of segmentation on the contour pertinent to each angle is located and the distance between the center and the point of segmentation is calculated. Fig. 6 exhibits the radius-based bipartite signature that has this structure. ($\theta = 45^\circ$ in figure)

The standard deviation of distances, Std(A) and Std(B), are calculated in the each of the region 'A'

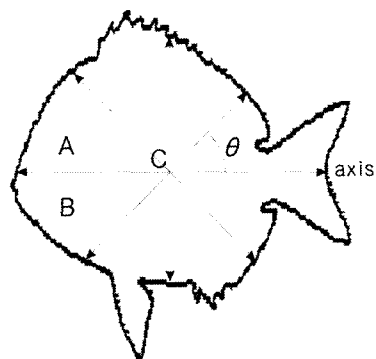


Fig. 6. radius-based bipartite signature

and 'B' to the right and left of this object and the regional descriptor is designated as the mean of two standard deviations.

At this time, the starting point of segmentation of each region is the two points of segmentation on the contour meeting the main axis, from which segmentation is carried out in a clockwise direction.

Even though the region is rotated to the original region by 180° and the direction of segmentation at the starting point is changed, any change in the direction of segmentation does not have an effect on the value of the shape descriptor because the distance is calculated by dividing it into two regions in this study.

The distance within the object is allowed not to be influenced by any change in its size by normalizing it as the ratio of the major axis length.

3.1.2 Boundary descriptor

Since only the bipartite signature is not enough to classify images, the boundary descriptor by overlapping contour segment is combined and defined as the shape property of another image. In the same way as the region descriptor, the contour is segmented at the angle of a given length(15°, 30°). Each applicable declination difference $DS(l, S_i)$ was calculated using the equ.(10) at each angle at the division point after segmentation.

The function of declination difference is the method that it is not influence by rotation and the angel θ the straight line connecting the point between the front and back of the 'N' point with a given point, P, as the center is defined as the Declination difference at the 'P' point. If the sign of the declination difference is negative, 360 is added to it.

This study attempted to enable the location of the division point to move within the segment section so that the declination difference between straight lines more approximate to the direction of the object contour.

The method was to get the values similar to the

change rate of the boundary direction by seeking the new declination difference by moving the division point to the point (P+1)' on the curvilinear line intersecting the normal line at the central point of the straight line like in Fig. 7 if the ratio of length between the curvilinear and straight lines of the divided segment line exceed a particular threshold and the curve of the contour become large.

$$\arccos(i) = \frac{\overrightarrow{P_{i-j}P_i} \cdot \overrightarrow{P_{i+j}P_i}}{|\overrightarrow{P_{i-j}P_i}| |\overrightarrow{P_{i+j}P_i}|} \quad (10)$$

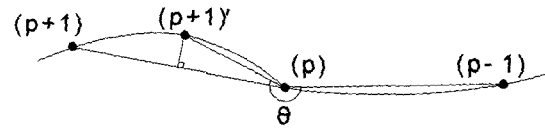


Fig. 7. Declination difference in segment point movement

Then, it was to transform $DS(l, S_i)$ into the declination vector $V_i(S_i)$ of the 10-dimension segment composed of 10 bins in according to the equ.(11) and get the feature vector V_j of the entire contour each division angle as the cumulative sum of the declination the individual contour segment had. The boundary descriptor is obtained by averaging two vectors.

$$V_i(S_i) = [v_k] \quad k=0, 1, \dots, 8, 9$$

$$v_j = \begin{cases} 1, & \text{if } \frac{DS(S_i)}{360} \times 10 = j \\ 0, & \text{otherwise} \end{cases}, \quad V_j = \sum_{i=0}^L V_i(S_i) \quad (11)$$

Because the cumulative frequency or the declination vector each segment is adopted as the feature, the cumulative frequency is not changed though the object is rotated. Therefore, one more robust feature value can be extracted in spite of a little local change of rotation or contour.

3.2 Extracting the Texture Feature

This study used the method that The bounding box of the standardized object to rotation was divided into 5×5, namely 25 regions and the block crossing the center of weight of the object was

selected to obtain the square texture sample of the same size having the texture quality of the same region in the random shape of object region.

To obtain the texture feature, the block was quantized into 32 grey-level spaces and then GLCM was calculated in 4 directions(0°, 45°, 90°, 135°). Then, the texture feature was obtained by extracting the *Contrast* from these matrixes and averaging it. The equ.(12) presents the equation to calculate the *Contrast*.

$$\frac{1}{N} \sum_a \sum_b |a - b|^2 P_{\phi, d}(a, b) \quad (12)$$

$P_{\phi, d}(a, b)$: function of angle ϕ and distance d
 N : total frequency of GLCM.

3.3 Extracting the Color Feature

The color histogram of the object was used as the color characteristic. This method has the advantage that it is not comparatively sensitive to the rotation, small translation and some clouding of the image. This characteristics may have the mutually complementary relationship to the shape descriptor because it can be used to identify any change in the contour of the object to some extent.

The histogram $h[i]$ normalized as O_s , the area of the object is used. The following histogram intersection was used as the yardstick for judging whether two objects have the similar color histogram.

$$S(Q, D) = \frac{\sum_{i=0}^{n-1} \min[Q(h[i]), D(h[i])]}{\sum_{i=0}^{n-1} Q(h[i])}, h[i] = \frac{H[i]}{O_s} \quad (13)$$

Here, $H[i]$ is the number of pixels of the n -th bin of the color histogram. Q and D exhibit the query image and the database image, respectively. n represents the entire number of bins.

4. Similarity Measurement

This study used as the function of similarity the linear combination of Euclidean distances between

characteristic vectors as in equ.(14) to calculate the degree of similarity between the query image f_q and the intra-database f_d .

$$D_s(f_q, f_d) = [w_1 (s_s(f_q) - s_s(f_d))^2 + w_2 (s_r(f_q) - s_r(f_d))^2 + w_3 (t(f_q) - t(f_d))^2 + w_4 (c(f_q) - c(f_d))^2 + w_5 (e(f_q) - e(f_d))^2]^{\frac{1}{2}} \quad (14)$$

Such weights as w_1, w_2, w_3, w_4, w_5 are defined as the relative weight between features, and their sum is equal to one. The degree of similarity between feature vectors of two objects was expressed as the value between zero and one, and they represent the high degree of similarity as the sum of Euclidean distances between feature vectors approaches one.

Here, w_1 and w_2 are the weight of shape features, and w_3, w_4, w_5 are the weights of texture, color and eccentricity, respectively. The efficiency of retrieval can be improved by applying the weight in accordance with the object to be retrieved and its purpose.

This study adjusted the weight on the basis of the relationship of shape complexity between two comparable images in order to raise a little more the weight of non-shape information such as texture and color characteristics as the degree of complexity becomes low. Fractal dimension and segmented contour variance were used to make a quantitative expression of the complexity.

4.1 Fractal Dimension

4.1.1 Box-Counting dimension

Because the objects existing in this world contain irregular changes in a detailed shape, there is some case that the dimension of statistical self-similarity can not be used as it is. In this case, the complexity of the object can be measured using the lattice of diverse sizes[11].

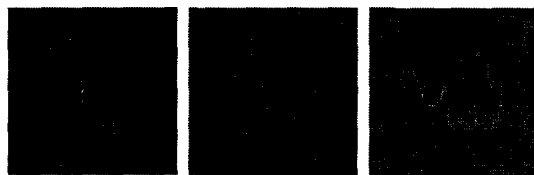
To calculate the Box-Counting dimension, D the entire edge image is segmented into partial images in the predetermined size of 's', and then this region

is defined as one, and the number of segmented regions including the object can be counted. If it is assumed that the number of segmented regions including the object like this is $N(s)$, D can be calculated as follows:

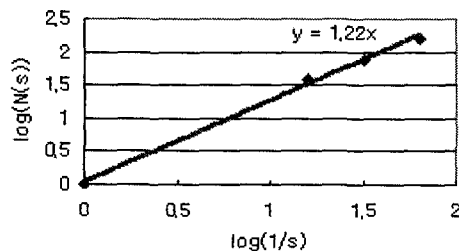
$$D = \lim_{s \rightarrow \infty} \frac{\log N(s)}{\log(1/s)} \quad (15)$$

This study made the size of 's' 1/16, 1/32, 1/64 to absorb both the properties of the global change and the local change. Fractal dimension was calculated by the liner gradient passing the mean point of $N(s)$ and $(1/s)$. If the fractal dimension has a considerably intricate structure and fills more space due to its complicated and rugged structure, it would have the value approaching two. Fig. 8 is the one representing the relationship between the segmented region of the image based on s and $N(s)$ and $(1/s)$ with the use of the double log graph $(\log(1/s), \log N(s))$, in which fractal dimension can be seen to be 1.22.

Here, dimension normalization is not considered as to any change in the size of the object. The reason is as follows. whether the image has the same object or not, in case comparison was made of the degree of similarity between the images



(a) $s=1/16$ (b) $s=1/32$ (c) $s=1/64$



(d) double log graph

Fig. 8. Fractal dimension

whose length of the object is different, the image whose length is short becomes simple because its complexity become smaller due to the dissipation of the pixel of the contour arising from its contraction. The power of discriminating the shape may come short. But though the scale is not standardized, the state of change in complexity arising from the scale element and the regularity of the shape can be compensated as the weight of non-shape information is raised a little high and the segmented contour variance is combined.

4.2 Segmented Contour Variance

The contour of the object is first divided by a given angle, and then the value of standard deviation $Std(C)$ of the contour length of the divided segment is calculated. This value, combined with the fractal dimension, helps to judge whether the shape of the object is regular or irregular. It is attributed to the fact that the length of lines of the divided segments would be not regular because the irregular contour has the possibility of having the intricate and rugged shape compared to the simple contour with regular pattern.

On the one hand, the length of the segment is calculated by overlapping the angle of segmentation by 30° and 15° because regularity may vary according to the angle of segmentation even in the case of the same contour. That is, the above-mentioned signature segmentation process is applied here.

4.3 Dynamic Weight

The weight of similarity can be adjusted for when C , the state of complexity of the object shape is obtained. For this purpose, the state of complexity can be obtained by using the equ.(16), and the image is divided into three states such as "SIMPLEX", "ORDINARY", "COMPLEX".

$$C = \begin{cases} \text{SIMPLEX} & \text{if } (D + Std(C)) > T_a \\ \text{ORDINARY} & \text{if } T_b \leq (D + Std(C)) \leq T_a \\ \text{COMPLEX} & \text{otherwise} \end{cases} \quad (16)$$

Finally, the weight of similarity is determined in accordance with the state of complexity of the image comparable to the query image. This method designates the image having the low state of complexity, of the two images, as zero and the image of high complexity as one, and one of the weights in Table 1 is chosen on the basis of the state of complexity of the image relevant to the value obtained from the OR operation. The weight and threshold of similarity was determined to be most effective through the test on the object of several experimental images.

$$C_i \wedge C_j \rightarrow C_{i,j} = \{SIMPLEX, ORDINARY, COMPLEX\} \quad (17)$$

Table 1. Dynamic weighted value

weight \ complexity	w ₁	w ₂	w ₃	w ₄	w ₅
SIMPLEX	0.2	0.2	0.3	0.3	0.0
ORDINARY	0.25	0.25	0.2	0.25	0.05
COMPLEX	0.3	0.3	0.15	0.2	0.05

5. Experiment and Analysis

5.1 Experimental Environment

To analyze the method proposed in this study, it was implemented by using Pentium III-700 PC in the Windows 98 environment and using the visual C++ 6.0, MS Access 2000 DBMS as the program.

The experimental image generally used in this study was the color image in which the object existed in the center of the image, and 256×256 256-color raw format image. This study selected 300 diverse kinds of images, added the images with noise and translation, scaling and rotation changed to those images and composed a total of 400 experimental images.

5.2 Experiment Result & Performance Analysis

The performance of image retrieval was eval-

uated by using the Recall rate and the Precision rate was generally used. Both recall rate and precision rate were measured using the equ.(18).

$$Recall = \frac{R_r}{T} \quad Precision = \frac{R_r}{T_r} \quad (18)$$

Here, T refers to the total number of items relevant to query of the database to be retrieved, and R_r refers to the number of items relevant to query of the retrieved items. And T_r refers to the total number of the retrieved items. This study randomly sampled the query image with the object of interest in the image to make the performance evaluation of the retrieval method and made a comparative experiment of the histogram method of the existing entire image and the method proposing the Jain method using the query by example method.

Table 2 exhibits the efficiency of the existing proposed method and the method proposed in this study. Table 3 and 4 shows the result of comparison the ranking order within five ranks on top and the time required to extract features in Jain's method and the proposed method.

The results presented in Tables 2, 3 and 4 shows that the proposed method had the better retrieval rate than the existing method. In addition, the proposed method shows the considerably improved performance in the top ranking order. On the other hand, our method is a little lower in the extraction time than Jain's method.

Table 2. Comparison of the retrieval rate

method	recall	precision
color histogram	0.54	0.39
Jain's method	0.89	0.82
proposed method	0.92	0.84

Table 3. Comparison of the recall rate ranking from the 1st to the 5th places on top

method	recall
Jain's method	0.80
proposed method	0.89

Table 4. Comparison of the time of extraction features

method	average time(sec)/image
Jain's method	2.13
proposed method	2.25

Fig. 9 and 10 presents the example of the images retrieved by the existing histogram method and

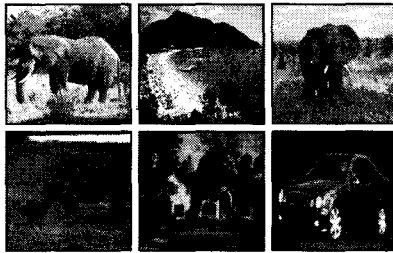


Fig. 9. Histogram result

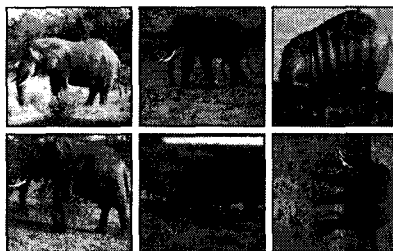


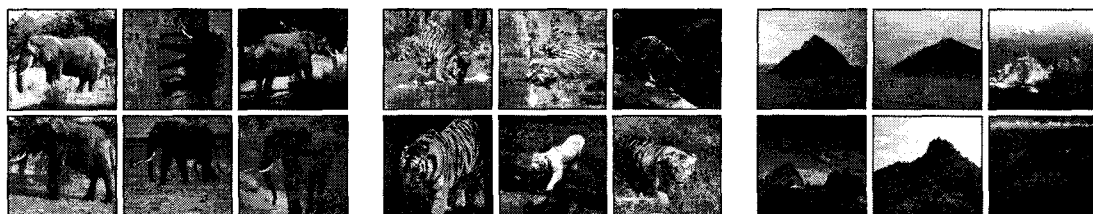
Fig. 10. Jain's method result

Jain-proposed method. They shows the images retrieved to the 1st to 5th rank from left to right according to the degree of similarity when the image of the 'elephant' at the leftmost side was given the query image.

It can be seen that the efficiency, when retrieved in the color histogram method, is very low. The reason is that the image is retrieved to be similar though the objects may be different from each other in case that the entire color distribution is similar.

In case of Jain's method, the applied invariant moment revealed a more or less vulnerable problem in retrieving the image with rotation change and noise added as can be seen in the result of recognition rate in Fig. 11.

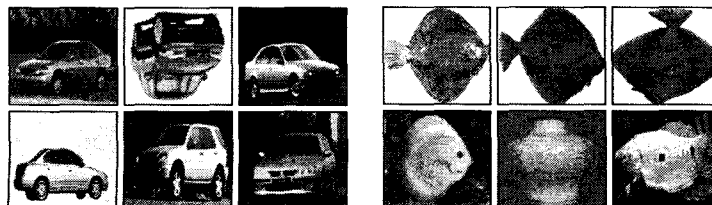
To complement this point, Jain proposed the combination of the invariant moment with the normalized edge histogram, and this method could raise the efficiency of retrieval performance. But it was found that the method has the problem of responding sensitively to the element of changing the histogram like a noise. Fig. 11 presents the examples of the results by the retrieval method proposed in this study to diverse query images.



(a)

(b)

(c)



(d)

(e)

Fig. 11. Retrieval results for proposed method

As can be seen in the images retrieved, the method proposed in this study could be found to be applicable to the images to which a noise is added or the image in which changes are made to the brightness of light as in the images such as 'elephant' or 'island'.

In addition, the method proposed in this study showed the good retrieval performance in case of the similar image with diverse changed added such as changes in scaling or location as well as in case that the shape of the object within the image is uniform compared to the existing method.

Fig. 12 presents the graphical representation of the recognition rate of images by feature in case translation, change in scaling, rotation and noise were added to the original image. Here, it made a comparative experiment to 10 diverse images, added the same image with any change. The signature used in the experiment is equally segmented in 30° directions on the basis of the central point.

It can be found that all the features are robust to any change in translation or scaling except for signature. But while the invariant moment and the signature method are lowered in recognition performance in case translation or change in scaling occurs, the shape feature proposed in this study can be found to have a much better recognition outcome.

The radius-based bipartite signature can interact with texture, color and eccentricity features and absorb both the global properties and the local

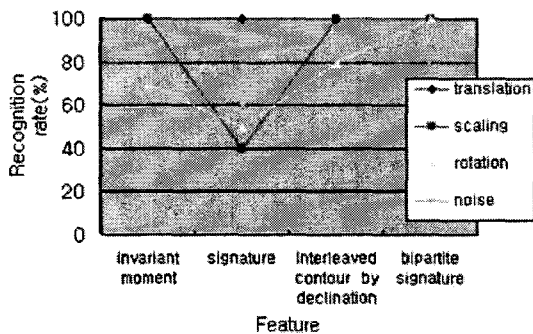


Fig. 12. Recognition rate of features

properties of the object. For this reason, unless any change in the viewpoint occurs, it showed the very good retrieval rate due to the tendency for inanimate things to maintain a given shape in general without any change in pose and had the comparatively robust retrieval though the shape of the image is partially distorted by any change of the contour or pose.

Especially, in case the boundary descriptor, combined with this method, was used as shape information, the similar image resulted in ranking top. And as seen in Table 5, the dynamic application of the weighted value by the degree of complexity showed the better retrieval result than the imposition of the weighted value on the feature in the complex state.

As shown in the fig. 13, in case the query image takes a shape of the simple triangle such as 'island' even though the shape may have the object that is generally kept constant in the image, it comes to include the image such as the 'thatched roof' due to the similarity of the roof shape. In consequence, it had the bad effect on the precision rate performance of the retrieval process.

Table 5. Comparison of the retrieval rate according to the method of imposing the weighted value

method	recall	precision
non-application of the dynamic weight	0.91	0.81
application of the dynamic weight	0.92	0.84

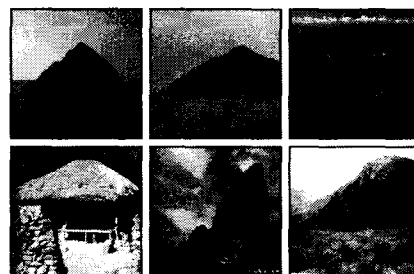


Fig. 13. Retrieval result by non-application of the weighted value

Therefore, as can be seen in (c) of Fig. 11, adjusting for and comparing the relative importance of features in accordance with the state of complexity of the image prior to comes to be able to minimize the possibility of wrong retrieval.

On the one hand, in case information on the color of the object of interest appears to very similar to the background and in case of the complex image composed of many small objects, it caused the segmentation process to be improperly carried out and acted as the factor of causing performance to be lowered.

6. Conclusion

Taking into additional consideration relationship to the background color as well as the size of the area in the hypothesis that the object is generally located around the middle of the image, this study attempted to allow the region of the object of main interest to be primarily retrieved by applying the improved watershed.

Especially, this study merged the overlapping contour by differential declination and the radius-based bipartite signature and proposed it as the shape properties of the object.

As a result of the experiment, the proposed method responded less sensitively to the distortion phenomenon of the object properties such as partial change or damage of the region and it was possible to make a retrieval robust to any change in translation, rotation and scaling. Also, it was found that the method of imposing the different weight of similarity on the image based on the relationship of complexity between objects led to the excellent recall and precision rate by minimizing the wrong retrieval on the object similar in shape.

Further study is required to improve the system performance more so that retrieval may be made possible on a real-time basis over the network and it would be necessary to conduct the study on more accurate image segmentation cues and needed to perform the more robust retrieval of the image in

case the partial region of the object is clouded or in case the amount of translation of the part of the object region is large.

References

- [1] W. Hsu, T. S. Chua, H. K. Pung. "An Integrated Color Spatial Approach to Content Based Image Retrieval," *Multimedia 95*, pp. 305-313. 1995.
- [2] E. Bribiesca and A. Guzman "Shape Description and Shape Similarity for Two Dimensional Region," *International Conference on Pattern Recognition*, 1978.
- [3] A. Jian, *Fundamentals of Digital Image Processing*. Prentice Hall, 1989.
- [4] A. K. Jain, A. Vailaya. "Image Retrieval using Color and Shape," *Pattern Recognition*, Vol. 29, No.8, pp. 1233-1244, 1996.
- [5] A. K. Jain, A. Vailaya, "Shape-Based Retrieval: A Case Study with Trademark Image Databases", *Pattern Recognition*, Vol. 31, No. 9, pp. 1369-1390, 1998.
- [6] M. C. Lim, W. S. Kim, "Image Segmentation By Watersheds Using Analysis of Co-occurrence Matrix," *Journal of KIPS*, Vol. 8-B, No. 1, pp. 59-65, 2001.
- [7] M. K. Kim, "Two-dimensional Shape Description of Objects using The Contour Fluctuation Ratio," *Journal of KMMS*, Vol. 5, No. 2, pp. 158-166, 2002.
- [8] C. H. Tch, R. T. Chin, "On Digital Approximation of Moment Invariants," *Computer Vision, Graphics, and Image Processing*, Vol. 33. pp. 318-326, 1986.
- [9] B. G. Ko, H. S. Lee, H. R. Byun, "FRIP System for Region-based Image Retrieval," *Journal of KISS: Computing Practices*, Vol 7, No 3 2001.
- [10] E. Milios and Euripides G. M. Petrakis "Shape Retrieval Based on Dynamic Programming", *IEEE Trans. on Image Processing*, Vol. 9, No. 1 pp. 142-147, 2000.

[11] Jonathan M. Blackledge, Martin J. Turner, Fractal Geometry in Digital Imaging, AP Professional, 1998.



이 정 봉

1993년 울산대학교 수학과 (이학사)
1996년 건국대학교 대학원 전자계산학과(공학석사)
1998년~현재 건국대학교 대학원 컴퓨터공학과 박사과정

관심분야: 영상검색, 패턴인식, 신경망



박 장 춘

1965년 연세대학교 전기공학과 (공학사)
1970년 연세대학교 대학원 전기공학과(공학석사)
1984년 캘리포니아 주립대학 대학원 전기공학과(공학석사)
1990년 연세대학교 대학원 전기공학과(공학박사)

1975년~1976년 광운 공과대학 전자계산학과 조교수
1976년~1981년 홍익대학교 전자계산학과 부교수
1984년~1988년 건국대학교 전자계산학과 부교수
1988년~현재 건국대학교 컴퓨터공학과 교수
관심분야: 영상처리, 컴퓨터 구조, 영상검색