On the Use of Edge Orientation and Distance for Content-Based Image Retrieval

Hwei-Jen Lin, Yang-Ta Kao, Feng-Ming Liang, Ta-Wei Liu, and Yi-Chun Pai
Department of Computer Science and Information Engineering
Tamkang University, Taipei, Taiwan, China
112992@mail.tku.edu.tw

***Abstract*—Recently, various features for content-based image retrieval (CBIR) have been proposed, such as texture, color, shape, and spatial features. In this paper we propose a new feature, called orientation-distance histogram for CBIR. Firstly, we transform the RGB color model of a given image to the HSI color model and detect edge points by using the H-vector information. Secondly, we evaluate the orientation-distance histogram from the edge points to form a feature vector. After normalization of feature, our proposed method can cope with most problems of variations in image. Finally, we show some results of query for real life images with the precision and recall rates to measure the performance. The experimental results show that the proposed retrieval method is efficient and effective.**

I. INTRODUCTION

Because of recent advances in computer technology and the revolution in the way information is processed, increasing interest has developed in automatic information retrieval from huge databases. In particular, content-based retrieval has become a hot research topic and, consequently, improving the technology for content-based querying systems is more challenging [1].

The content of image can be classified into two categories [1]: one is the characteristic of pixel value, such as color distribution, texture information of image, and the shape property of object in an image, another is the spatial relationship between objects. Current technologies of CBIR have combined different features to achieve the matching process [2], such as the combination of color and shape [3, 4], the combination of color, texture, and shape [5, 6] etc. Different image features are integrated in order to design a suitable query system to agree with the human's vision. Such systems usually include image segmentation, feature selections, and similarity matching. Each system has its advantages and defects and is only suitable for retrieving certain types of images. Despite years of extensive research, however, assisting users to find their desired images accurately and efficiently is still an open problem. In this paper, we propose a new feature, called the orientation-distance histogram, to achieve the requirement of some affine transformations. Because of the complexity of segmentation works, it is difficult to obtain a good segmentation results. Therefore, trending the segmentation based towards edge point based maybe a different thinking. We analyzed the contour of different images on their orientation-distance histograms, and found that the relation between edge points and their neighboring edge points possess lots of information. Similar images tend to be alike in this feature. Consequently, this feature can be used to effectively retrieve the relevant images.

This paper is organized as follows: In Section 2, we briefly survey the related works of CBIR. In Section 3, we briefly review our previously proposed edge detection method [7], and describe orientation-distance feature extraction, feature normalization, and a similarity matching scheme used in this research. In Section 4 the experimental results are illustrated with query results and precision-recall curve compared with F. Mahmoudi et al. [6]. The final section makes a conclusion for this paper.

II. RELATED WORKS

The QBIC (Query By Image Content) is a representative image retrieval system, which was supported and constructed by IBM corp. The system uses color, texture, shape, and sketch content to describe an image. The user can choose the characteristics of an image to query the system. The QBIC searches the database according to the features given by the user. The QBIC provides three ways for querying. The first one uses the dominated color of an image to match the user's query. The second one uses the percentage of color existing in the image to retrieve relevant images. The last one allows user to sketch the shape, size, and position of the image his or her queries. QBIC system uses the low-level features instead of high-level ones. For this reason, it cannot distinguish the meaning/semantic content of image.

While closing the semantic gap is still a far cry given today's technology. Traditional pattern recognition has proposed a lot of methods in this area. But it is still hard to find a robust and efficient method today. Grey-level distribution [8, 9] is used to describe an image which is effective only when the illumination is a constant factor. Y. Chen et al. [10] used the one-class SVM method for

0-7803-9422-4/05/$20.00 ©2005 IEEE

learning in image retrieval, but it is not practicable due to its time complexity. F. Mahmoudi et al. [6] uses an edge orientation autocorrelogram to represent the shape of an object. This approach is effective even when the query image is somewhat distorted, however, it is very time-consuming. In the work of T. Bernier et al. [11], distances and angles of the contour points relative to the centroid of an object form a string representation, which is invariant to translation, rotation, and scaling. However, this representation has a low tolerance for occlusion. J. Zhang et al. [12] uses a shape space method to deal with both noise and occlusion problems, but again, the process of feature point selection is very time-consuming. H. Nishida [13] proposed a representation method that tolerates some contour deformation but the structure features are very sensitive to noise. Medial axis [14, 15] or skeletal representation [16] methods have traditionally been computed from discrete boundary-based descriptions by using skeletonization algorithms. Such a boundary description, however, yields medial loci with a complex branching structure. Also, the boundary-to-medial transformation is inherently unstable as the resulting branching topology is sensitive to even slight boundary perturbations, especially in the regions known as ligatures.

III. THE PROPOSED METHOD

This section illustrates the proposed method. A simple and effective method used to detect edge points [7] is briefly reviewed in Subsection 3.1. The feature point extraction is described in Subsection 3.2. Subsection 3.3 indicates that the normalization of feature vector can cope with the variance of affine transformations. Finally, similarity measurement is given in Subsection 3.4.

A. Prompt Edge Detection

One of the common edge detection methods is the Sobel operation [17], which computes the partial derivatives of a local pixel to form the gradient. For the gray-level function f(x,y), the gradient of f at pixel (x,y) is a vector as defined in Equation (1). The magnitude of ∇f can be approximated as shown in Equation (2).

$$\nabla f = \left[\partial f / \partial x , \partial f / \partial x \right] \tag{1}$$

$$\begin{aligned} \|\nabla f\| \approx & \left|(z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)\right| \\ & + \left|(z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)\right| \\ & + \left|(z_6 + 2z_8 + z_9) - (z_1 + 2z_2 + z_4)\right| \\ & + \left|(z_4 + 2z_7 + z_8) - (z_2 + 2z_3 + z_6)\right| \end{aligned} \tag{2}$$

Although Sobel's edge detector generally works well, it must repeatedly go over the whole image many times. As a result, this method requires W*H*32 time to operate on an image of size W*H. In this subsection, we use our proviously proposed method [8] to perform edge detection, called the prompt edge detection, which is not only simple but also efficient.

In the real life image, the color space is in RGB domain, which mixed the color and intensity information. Therefore it is difficult to analysis in a single vector; such as color or intensity vector. In view of this, we transform RGB to HSV color model, and apply our edge detection in H vector. The prompt edge detector detects edge points by checking the differences in the gray value for each point from those of its neighbors. For a pixel (x,y) with gray value g(x,y), let g0(x,y), g1(x,y), ..., and g7(x,y) denote the gray values of its neighbors in 8 directions as shown in Figure 1.

g_3	g_2	g_1
g_4	g	g_0
g_5	g_6	g_7

Figure 1. The 8 neighbors of the center point.

Let $h_d(x,y) = |g(x,y) - g_d(x,y)|$ be the difference in the gray value of pixel (x, y) from that of its neighbor in direction d. Let Bd(x,y) be the number of differences that exceed a threshold T, where T = a + c, c is a constant, and a is the average difference between all pairs of adjacent pixels' gray values in the image. In this work, we set the value of c to 2 and, instead of taking a single value for T, we take the local average differences for T by dividing the whole image into $M \times M$ blocks and computing the average difference for each block.

The pixel (x,y) is indicated as an edge point if inequalities (3) hold.

$$3 \le B_d(x,y) \le 6 \tag{3}$$

These inequalities can also avoid indicating noisy points as edge points. Further details can be found in [7].

B. Features Extraction

A good representation should cope with the affine transformations such as translation, scaling, and rotation, and should have high tolerance to noise and occlusion. The translation-invariance is fixed in our proposed method, which is achieved by the use of the orientation-distance sequence. In this subsection, we describe how the proposed method extracts the feature of an image. Our feature extraction includes two steps, as given in the following:

(1)Edge orientation quantization: This stage quantizes all edge orientations into n segments (We set n = 36 in this work).

(2)Computing orientation-distance histogram: For each orientation *i* and distance *d*, for every edge point having orientation *i*, we count the points of *d*-distance surrounding it and in the same orientation. The counted values are accumulated to form a so-called orientation-distance histogram. The points of *d*-distance are the outermost points in the $(2d+1)\times(2d+1)$ window sorrounding the current point. As illustated in Figures 2(a) and 2(b), the points p_1, p_2, …, and p_8 are points of *1*-distance and q_1, q_2, …, and q_{24} are points of *3*-distance for the center point. Figure 3 shows the orientation-distance histogram of a real image. The orientation-distance histrogram is denoted by $h_{i,j}$, which indecates the amount of j-distance accumulation in orientation i, where $0 \le i \le n-1$ and $j \in \{1,3\}$. In this research, we combine this histogram sequence ($h_{1,0}$, $h_{1,1}$, …,$h_{1,n-1}$, $h_{3,0}$, $h_{3,1}$, …, $h_{3,n-1}$) and the values, c_1 and c_2, of the two most dominated colors to form the (2n+2)-dimensional feature vector ($h_{1,0}$, $h_{1,1}$, …,$h_{1,n-1}$, $h_{3,0}$, $h_{3,1}$, …, $h_{3,n-1}$,c_1,c_2).

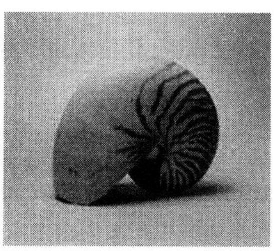
p_1	p_2	p_3
p_8	*	p_4
p_7	p_6	p_5

(a)

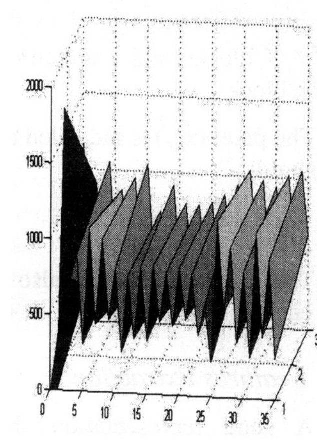
q_1	q_2	q_3	q_4	q_5	q_6	q_7
q_{24}						q_8
q_{23}						q_9
q_{22}			*			q_{10}
q_{21}						q_{11}
q_{20}						q_{12}
q_{19}	q_{18}	q_{17}	q_{16}	q_{15}	q_{14}	q_{13}

(b)

Figure 2. (a) points of 1-distance, (b) points of 3-distance.



(a)



(b)

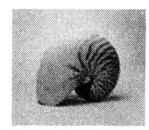
Figure 3. (a) A source image, (b) The orientation-distance histogram of 3(a).

C. Feature Normalization

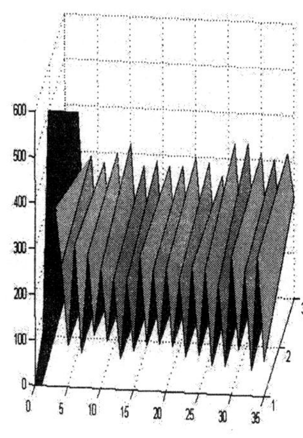
This subsection describes the feature normalization to cope with the scaling and rotation invariance. Subsection 3.3.1 illustrates how the proposed method achieves scaling invariance and Subsection 3.3.2 deals with the problem of rotation invariance.

C.1 Scaling Invariance

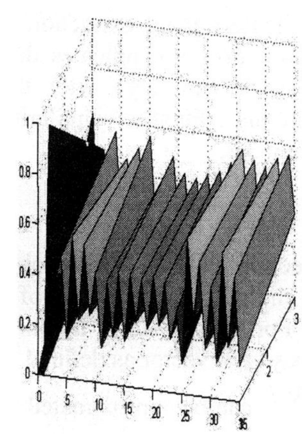
Since the amount of edge points depends on the size of an image. Generally, the larger an image is, the more its edge points can be detected. In dealing with the scaling invariance, we normalized each bin by dividing the total number of edge points accumulated in the orientation-distance histogram, so that all population values fall between 0 and 1. Figure 4 shows the normalized results. Figure 4(a) shows the 50%-scaled-down version of the image given in Figure 3(a), and Figure 4(b) shows its histogram. Figures 4(c) and 4(d) show the normalized versions of the histograms given in Figures 3(b) and 4(b), respectively. These normalized versions are quite similar.



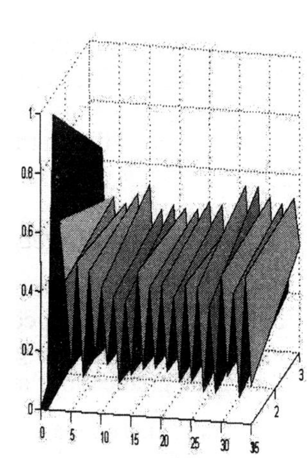
(a)



(b)



(c)



(d)

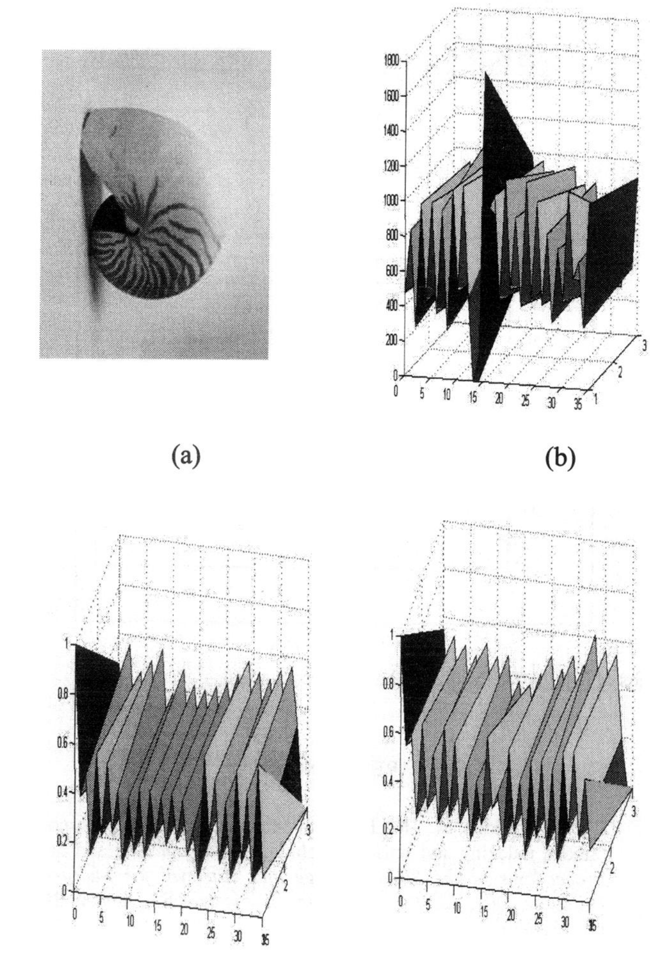
Fiugre 4. (a) The 50%-scaled-down version of 3(a), (b) The orientation-distance histogram of 4(a), (c) The normalized version of 3(b), (d) The normalized version of 4(b).

C.2 Rotation Invariance

In consideration of rotation invariance, we compute the MCS [18] sequences to shift the orientation sequence to achieve the best affinement by using the ordering-consistency function $c_{i,j}$ of distance j in orientation i, defined in Equation (4), to obtain a shifting factor s, defined in Equation (5), where $S_D = (p_{0,j},p_{1,j},\ldots p_{n-1,j})$ is the orientation sequence of j-distance [18]. With the shifting factor s, the MCS shifted version of the feature vector ($h_{1,0}$, $h_{1,1}$, ..., $h_{1,n-1}$, $h_{3,0}$, $h_{3,1}$, ..., $h_{3,n-1}$, c_1, c_2) comes out to be ($h_{1,s}$, $h_{1,(s+1) \bmod n}$, ..., $h_{1,(s+n-1) \bmod n}$, $h_{3,s}$, $h_{3,(s+1) \bmod n}$, ..., $h_{3,(s+n-1) \bmod n}$, c_1, c_2). A rotated version of the image in Figure 3(a) and its orientation-distance are given in Figure 5. The shifted and normalized versions of the histograms given in 3(b) and 5(b) are quite similar, as shown in 5(c) and 5(d), respectively. The shifting factors for results shown in Figures 5(c) and 5(d) are s = 3 and s =12, respectively.

$$c_{i,j}(S_D) = \sum_{k=i}^{i+(n-1)} (n-(k-i)) * p_{k,j} \tag{4}$$

$$s = \arg \min_{0 \le i < n, j=\{1,3\}} \{c_{i,j}(S_D)\} \tag{5}$$



Fiugre 5. (a) A rotated version by 90 degrees of the image in 3(a), (b) The oreintation-distance histogram of 5(a), (c) The shifted and normalized result of 3(b) with s = 3, (d) The shifted and normalized result of 5(b) with s = 12.

D. Similarity Measure

The final part of the proposed method is similarity measure. As discussed in the previous subsection, the orientation-distance histogram and the amounts of points in two dominated color are combined to represent the characteristic of an image. We use the distance metric expressed in Equation (6) to measure the distance between two images based on this characteristic.

$$d(U,V) = \alpha \sqrt{\sum_{i=1}^{2n} (u_i - v_i)^2} + \beta \sqrt{\sum_{i=2n+1}^{2n+2} (u_i - v_i)^2} \tag{6}$$

In Equation (6), U and V represent feature vectors of the query image a database image, v_i, u_i are the respect i-th features, α and β are weights (in this work, we set α = 0.5 and β = 0.5).

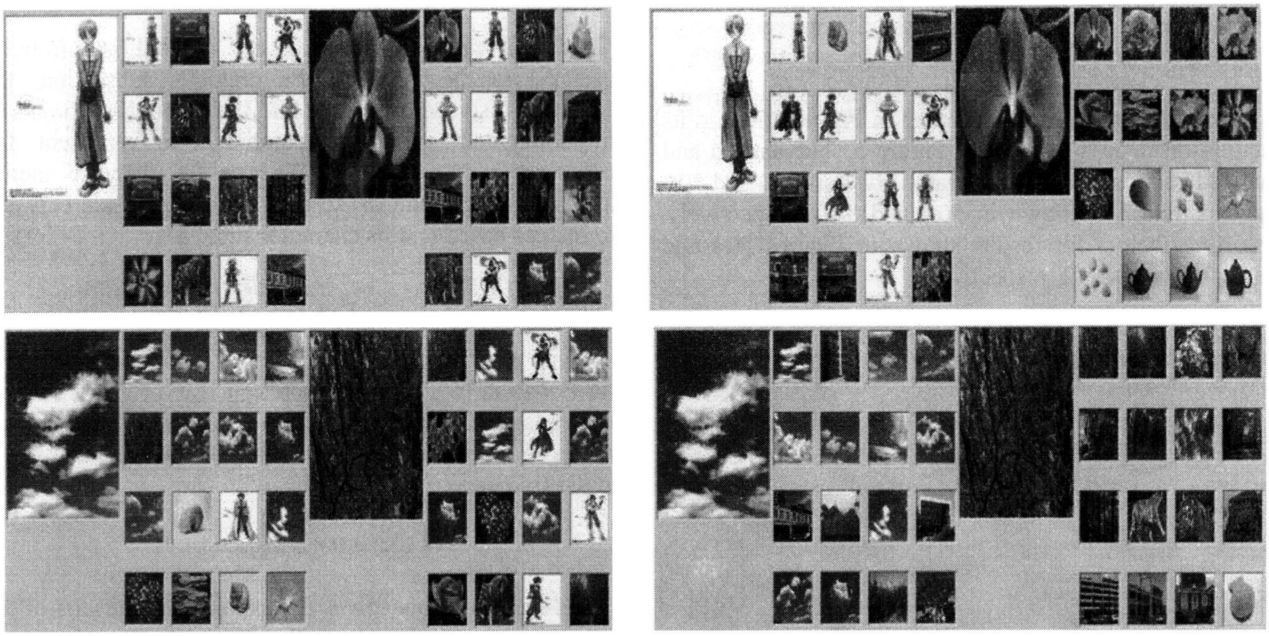
IV. EXPERIMENTAL RESULTS

This section demonstrates the experimental results of the proposed mehtod. We show the query results of various types of images, including cartoons, flowers, sky, and forest. The experimental images are collected from internet, Ulead Pick-a-Photr 01-05, and book scan, with totally 10 categories including tigers, cars, flowers, airplanes, buildings, forest, sky, seashell, cartoons, and teapot in our database. The developed environment of our system is on a desktop PC with celeron 700 MHz CPU, 256 MB memory, and the OS is Windows 2000.

Figure 6 shows some query results of the two compared methods. The left-top corner in each query example is the query image, and the part to the right shows the query results. We can find that our proposed method retrieve more similar images than the method proposed in [6]. Another most widely used performance measures for information retrieval are precision (Pr) and recall (Re) rates. Pr is defined as the number of retrieved relevant objects over the number of total retrieved objects. Re is defined as the number of retrieved relevant objects over the total number of relevant objects in the image collection. In general, Pr decreases when Re increases. The performance of an ideal system should have a higher precision rate at any certain recall value. As a result, Pr-Re curves are often used to measure the performance of a retrieval system. Figure 7 shows the Pr-Re curves of the methods proposed by us and by F. Mahmoudi et al. [6]. The result shows that our proposed method is more effective than that proposed by F. Mahmoudi et al..

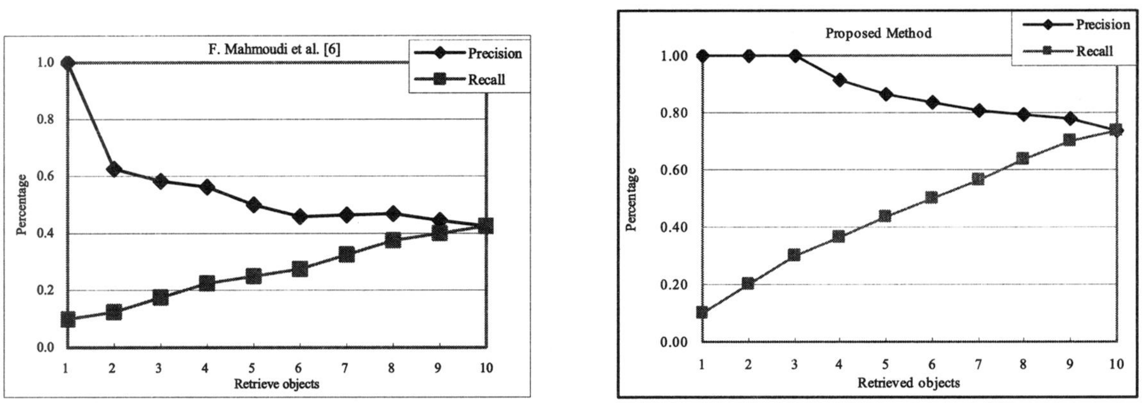
V. CONCLUSIONS AND FUTURE DIRECTION

This paper proposes a new method of image representation using the orientation-distance histogram. It manifests the contour trend of images. The histograms are different from images to images. The experimental results



(a) (b)

Figure 6. Examples of query results. (a) The results of [6]. (b) The results of the proposed method.



(a) (b)

Figure 7. Retrieval performance of Pr-Re curve. (a) The results of [6]. (b) The results of the proposed method.

show that our method achieves good performance by using the following schemes: color domain transform, prompt edge detection, orientation-distance feature extraction, feature normalization, and similarity matching. We have also shown the precision and recall curves to prove its robustness and effectiveness.

As a general future direction, we believe that considering more on color feature and integrating the shape and spatial features will be a significant step towards narrowing down the semantic gap. And furthermore, combining the relevant feedback with these features is expected to provide a more robust and adaptive retrieval experience.

REFERENCES

[1]. J. Shanbehzadeh, A. –M. Eftekhari-Moghadam, and F. Mahmoudi, "Image indexing and retrieval techniques: past, present and next", *Proceedings of the SPIE: Storage and Retrieval for Multimedia Database 2000*, Vol. 3972, San Jose, California, USA, January 2000, pp. 461-470.

[2]. M. S. Chor and W. Y. Kim, "A novel two stage template matching method for rotation and illumination invariance", *Pattern Recognition 35*, 2002, pp. 119-129.

[3]. H. Wei and D. Y. Y. Yun, "Illumination-Invariant Image indexing using Directional gradient angular histogram". *Proceedings of the IASTED International Conference Computer Graphics and Imaging, Honolulu, Hawaii, USA, August 2001*, pp. 13-16.

[4]. Y. Chahir and L. Chin, "Efficient Content-Based Image Retrieval based on color homogeneous objects segmentation and their spatial relationship characterization", *IEEE International Conference on Multimedia Computing and Systems*, Vol. 2, 1999, pp. 705-709.

[5]. H. Nishida, "Structural feature indexing for retrieval of partially visible shapes". *Pattern Recognition 35*, 2002, pp. 55-67.

[6]. F. Mahmoudi, J. Shanbehzadeh, and A. M. Eftekhari-Moghadam, "Image retrieval based on shape similarity by edge orientation autocorrelogram". *Pattern Recognition 36*, 2003, pp. 1725-1736.

[7]. H. J. Lin and Y. T. Kao, "A prompt contour detection method", *International Conference on the Distributed Multimedia Systems*, 2001.

[8]. J. H. Chang, K. C. Fan, and Y. L. Chang, "Multi-modal gray-level histogram modeling and decomposition", *Image and Vision Computing 20*, 2002, pp. 203-216.

[9]. R. Brenulli and O. Mich, "Histograms analysis for image retrieval", *Pattern Recognition 34*, 2001, pp. 1625-1637.

[10]. Y. Chen, X. Zhou, and T. Huang, "One-class SVM for learning in image retrieval", Proceedings of the IEEE *International Conference on Image Processing, Thessaloniki, Greece*, 2001, pp. 815-818.

[11]. T. Bernier, J. A. Landry, "A new method for representing and matching shapes of natural objects". *Pattern Recognition 36*, 2003, pp. 1711-1723.

[12]. J. Zhang, X. Zhang, H. Krim, and G. G. Walter, "Object representation and recognition in shape spaces", *Pattern Recognition 36*, 2003, pp. 1143-1154.

[13]. H. Nishida, "Structural feature indexing for retrieval of partially visible shapes", *Pattern Recognition 35*, 2002, pp. 55-67.

[14]. A. Bonnassie, F. Peyrin, and D. Attali, "A new method for analyzing local shape in three-dimensional images based on medial axis transformation", *IEEE Transactions on Systems, Man, and Cybernetics, Part B, Cybernetics 33*, 2003, pp. 700-705.

[15]. P. Yushkevich, P. Fletcher, S. Joshi, A. Thall, and S. M. Pizer, "Continuous medial representations for geometric object modeling in 2D and 3D", *Image and Vision Computing 21*, 2003, pp. 17-27.

[16]. C. D. Ruberto, "Recognition of shapes by attributed skeletal graphs", *Pattern Recognition 37*, 2004, pp. 21-23.

[17]. G. Rafel and E. G. Richard Woods, *Digital Image Processing*.

[18]. H. J. Lin, Y. T. Kao, S. H. Yen, and C. J. Wang, "A study of shape-based image retrieval", *Proceedings of the 6th International Workshop on Multimedia Network Systems and Applications*, 2004, pp. 118-123.