Semantic Evaluation and Efficiency Comparison of the Edge Pixel Neighboring Histogram in Image retrieval

Abdolah Chalechale and Alfred Mertins

School of Electrical, Computer and Telecommunications Engineering
University of Wollongong, Wollongong, NSW 2522, Australia
Email: {ac82, mertins}@uow.edu.au

Abstract

A novel approach in image retrieval based on middle-level features, instead of low-level features, is applied in trademark matching problem. The approach is developed utilizing edge pixel neighboring histograms for capturing the overall image structure. The closeness of the method to human perceptual judgment is evaluated by a subjective test along with the efficiency in comparison with four other methods known from the literature. The query set and the image database used in the tests are taken from the MPEG-7 dataset and the experimental results show a significant supremacy in semantic compatibility over the histogram of edge angles, MPEG-7's edge histogram descriptor (EHD), and the invariant moments method. The proposed method is more efficient than the correlation method which is closer to human perception.

1 Introduction

Incorporating semantics in visual information retrieval is one of the most active research areas in growing multimedia technology, and the main goal is to close the results of search engines, for example in World Wide Web, to human expectations as much as possible.

Although most current image retrieval systems [1, 2, 3] attempt to find similar images based on the content rather than textual annotations (filename, comments, keywords) they still suffer from the semantic gap. In other words, despite the encouraging results of applying low level features in certain applications there are still situations where some/many of the output results are far from the user's intuition [4]. Therefore, for capturing high-level concepts, it is necessary to move toward more instinctive methods by employing new and robust image representations and similarity measures that are more compatible with human reasoning.

The most important image content clues, as suggested by the MPEG-7 standard, are color, texture and shape [5, 6], while it is possible to take advantage of spatial relationship of objects in some applications as well. Color, texture and shape descriptors have been shown to be powerful in many application domains, however, their low level characteristic makes the output of image retrieval methods based on these attributes unsatisfactory for users in some instances. In addition, when the query or the database image has no such attributes the above descriptors would lose their original ability in finding good answers. This is for example the case when the query image is a fast drawn, rough sketch with only some black and white [7], or when the aim is to search in thousands of black and white trademarks without a well defied object contour, to find logos similar to the given one in a trademark registration process [8].

As there is an immerse demand for visual information retrieval systems nowadays and in the near future, existing methods need to be evaluated for their ability in finding similar images which are more acceptable by users and/or to exploit new methods which produce results not only compatible with human understanding (effective) but also low in computation cost (efficient).

Bridging the gap between human perception and low level features in visual information retrieval has been studied, for example, by applying neural networks [9, 10], fuzzy methods [11, 12], designing hierarchical structure of data layer to semantic layer [13] and utilizing experts knowledge in the form of a set of linguistic rules [14], and also by relevance feedback [15].

When the query is a drafted sketch or when the database images have no discriminating color and texture features and there is also no well-defined object contour, the following methods are able to produce reasonable results in the retrieving process:

1. Correlation approach [3, 16],
2. Histogram of Edge directions (HED) [8, 17, 18],
3. Edge Histogram Descriptor (EHD), proposed in MPEG-7 standard [5, 19] and
4. Invariant Moments [8, 20].

In this study we apply a novel approach in image retrieval that is based on middle-level features rather than low-level ones. The features are given by an Edge Pixel Neighboring Histogram (EPNH). Its closeness to human perceptual judgment is assessed by a subjective test along with its efficiency in comparison with the above four methods. The evaluation, without losing generality, is made in solving the problem of trademark matching, where the query set and the image database are taken from the known and accessible MPEG-7 dataset.

The method utilizing EPNH for capturing the overall image structure and its comprehensive details is given in the next section. In Section 3 the description of the other four methods, used in our tests, is briefly provided. Section 4 shows the experimental results. An analysis and discussion of the results is given in Section 5. Section 6 concludes the paper and suggests some further research directions in this area.

2 The Edge Pixel Neighboring Histogram (EPNH)

The main objective of the proposed method is to transform the image data into a new structure that supports measurement of the similarity between images in a correct, easy and fast way with emphasis on capturing more conceptual meaning. Using histograms is recognized as a powerful tool in image retrieval. Several kinds of histograms and related similarity measuring techniques are proposed for comparing images based on color, texture and shape [5, 6]. Moreover, using an edge direction histogram has been used as a discriminating cue especially in the absence of color information or in the presence of images with similar colors [17, 18].

EPNH is a novel approach that focuses on higher-level edge distribution rather than edge-based or pixel-distribution based methods. In this approach, the information about neighbors of edge pixels is obtained first and then used to produce the neighboring histogram.

Initially, the images may have color attributes, and they may be composed of single or multiple objects. Color images are converted to a gray intensity by eliminating the hue and saturation while retaining the luminance in a pre-processing stage. Gray scale, black & white and sketched images are not needed go through this process. Then, applying the Canny edge operator [21] on the image will result in an edge map, I, that is the basic platform for obtaining the edge pixel neighboring histogram. The histogram originally has 240 bins and the frequencies of edge pixels with similar neighboring code are stored in appropriate bins. The neighboring code is defined using the following edge pixel-neighboring diagram (see Fig. 1). In this diagram the center is an edge pixel in I. Considering 8-connectivity, each pixel has up to 4 neighbors in most cases, as the pixel and the neighbors are all edge points.

By numbering the directions as indicated in Fig. 1 each pixel neighborhood is coded with a number n with $0 \leq n \leq 240$; that is the sum of all direction numbers of its neighbors. For instance, $n = 0$ means a singular point (without any neighbor), a pixel point with two horizontal neighbors has $n = 17$ (1+16) neighboring code

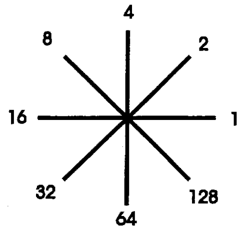

Figure 1: Edge Pixel-Neighboring Diagram.

and $n = 240$ means a point with 4 neighbors in the directions represented by 128, 64, 32 and 16. This code indicates the structure of the neighbor pixels for each pixel in the edge map. Figure 2 shows an example. An edge pixel with one edge neighbor in the northern direction (4), one neighbor in the western direction (16) and another neighbor in the southeastern direction (128) has the unique number 148 (4+16+128) for its neighboring code. The existence and the frequency of this structure in an edge map is an important feature and can be used for measuring the similarity between the corresponding images.

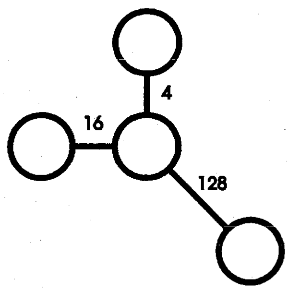

Figure 2: An edge point with three neighbors.

The frequencies of the neighborhood codes ($n \neq 0$) form the EPNH bins. The singular points ($n = 0$) are scarce and not considered, and the sum of other edge pixels with the same neighboring code is stored in the appropriate bins.

The proposed numbering scheme simplifies the code finding process since we only set a bit in the appropriate position in the code byte indicated by the direction of the neighbor. Also, we can easily recognize the number of neighbors by counting the number of ones in the corresponding codes. The emphasis in this histogram is on the number of occurrences (frequencies) of each code and not on the code itself. Therefore, the overall structure of the edge map image is captured in a histogram representation which can be used for measuring the similarity between images by any histogram intersection techniques or simply by ℓ_1 or ℓ_2 distances.

The histogram entries depend on the size of the image and a normalization is required. Normalization of EPNH is made by considering the size of the image rather than the number of edge points. Brandt et al. used this type of normalization in their work and they argued that it is better than normalizing by the number of edge points [22]. The normalized histogram is produced as follows:

$$EPNH_i = \frac{\sum_i P_i}{\text{size } I}$$

$1 \leq i \leq 240$ and P_i is any edge pixel in I with $n = i$.

The EPNH is translation and scale invariant, and because of its middle-level feature, it captures more semantics in the image than the low-level methods. Because the edge neighborhood feature is local, it is robust to partial occlusion and local disturbance in the image. The rotation invariance property is achievable in this method by grouping rotation-similar bins of the histogram.

The histogram was recently used together with a vicinity table and a specific similarity measure in sketch based image retrieval [7]. It demonstrated a significant improvement in retrieving images in a prototype image dataset. In Section 4, it is shown that the EPNH can also produce results that are much closer to human perception than the results obtained from the edge direction histogram, invariant moments and the MPEG-7's edge histogram descriptor.

3 Alternative Methods

In this section, the description of three edge-based methods and the invariant moments method are addressed. In Section 4, they will be compared with the EPNH method.

3.1 Correlation Approach

The Query by Visual Example (QVE) system developed by Hirata and Kato [16] performs retrieval by computing the correlation between the query sketch and database edge images. In this approach, the query and target images are resized to 64×64 pixels and then the proposed gradient operator extracts their edges. The resulting edge maps are called pictorial indexes and used for image-to-image matching. After dividing each pictorial index image into 64 blocks of equal size, the correlation between corresponding blocks in the query index q and the database index p is calculated by the following bit-wise summation with shifting blocks of q by δ and ε over blocks of p:

$$C_{\delta\varepsilon} = \sum_s \sum_r \quad (\alpha p_{rs} \cdot q_{r+\delta s+\varepsilon} + \beta \overline{p_{rs}} \cdot \overline{q_{r+\delta s+\varepsilon}} \\ + \gamma p_{rs} \oplus q_{r+\delta s+\varepsilon}) \tag{1}$$

The coefficients α, β and γ are control parameters used to estimate matching and mismatching patterns and take up values 10, 1 and -3, respectively. The maximum value of $C_{\delta\varepsilon}$ for all δ and ε for each block is called the local correlation factor C:

$$C = \max(C_{\delta\varepsilon}) \text{ for } -4 \leq \delta \leq 4 \text{ and } -4 \leq \varepsilon \leq 4.$$

Finally, a global correlation factor C_t that is the sum of all 64 C's, is calculated and used as the similarity measurement. A modified version of this method is used in QBIC system by IBM [3].

Although the method has a good ability to find similar images in small datasets [16] it does not allow indexing, and because of the expensive computational cost, using that method in a large image databases is time consuming. While the method can tolerate a small local rotation, it is not rotation invariant and does not allow for large global rotation.

3.2 Histogram of Edge Directions

Histogram of edge directions (HED) for representing image information is one of the well-known methods in the image retrieval literature. M. Abdel-Mottaleb [18] used this method by applying the Canny edge operator to find strong edges in an image and then quantized them into only 4 directions (horizontal, vertical, and the two diagonals) to build histograms of edge directions for different image regions. The histograms are then used as hash values in a hash table indexing mechanism. Jain and Vailaya [17] used edge directions as an image attribute for shape description. They show that in the absence of color information or in images with similar colors this histogram is a significant tool in searching for similar images. They also exploited the histogram together with invariant moments in a case study using a trademark image database [8]. The edge information contained in the database images is extracted off-line using the Canny edge operator and then the corresponding edge directions are quantized into 72 bins of 5° each. To reduce the effect of rotation, they smooth the histogram as follows:

$$I_s[i] = \frac{\sum_{j=i-k}^{i+k} I[j]}{2k+1}$$

where I_s is the smoothed histogram, I is the original normalized histogram, and the parameter k determines the degree of smoothing. In their experiments they used $k = 1$.

3.3 MPEG-7 Edge Histogram Descriptor (EHD)

The MPEG-7 standard defines the edge histogram descriptor (EHD) in its texture part [5]. The distribution of edges is not only a good texture signature it is also useful for image-to-image matching in the absence of any homogeneous texture. A given image is first divided into 16 sub-images (4×4), and local edge histograms are computed for each sub-image. Edges are grouped into five classes: vertical, horizontal, 45° diagonal, 135° diagonal, and isotropic (non-directional). The directional edge strengths are obtained for each edge class using 5 corresponding 2×2 filter masks. If the maximum is grater than a threshold value (Th_{edge}) then the underlying block is designated to belong to the corresponding edge class. To compute the edge histogram, each of the 16 sub-images is further subdivided into image blocks. The size of each image block is proportional to the size of the original image and is assumed to be a multiple of two. The number of image blocks, independent of the original image size, is constant (*desired_num_of_blocks*) and the block size is figured as follows:

$$x = \sqrt{\frac{image\_width * image\_height}{desired\_num\_of\_blocks}}$$

$$block\_size = \left\lfloor \frac{x}{2} \right\rfloor * 2$$

where *image_width* and *image_height* represent horizontal and vertical size of the image, respectively. Each image block is then partitioned into four (2×2) blocks of pixels and the pixel intensity for these four divisions are computed by averaging the luminance values of the existing pixels.

The histogram for each sub-image represents the frequency of occurrence of the five classes of edges in the corresponding sub-image. As there are 16 sub-images and each has a 5-bin histogram, a total of $16 \times 5 = 80$ bins in the histogram is achieved. For normalization, the number of edge occurrences for each bin is divided by the total number of image blocks in the sub-image. For minimizing the overall number of bits, the normalized bins are nonlinearly quantized and fixed-length coded with 3 bits per bin, resulting in a descriptor of size 240 bits.

S. Won et al. proposed the efficient use of this descriptor by extending the histogram to 150 bins [19]. The extended histogram is obtained by grouping the image blocks in 13 clusters (4 vertical, 4 horizontal and 5 square). Each cluster contains 4 sub-images. In addition to this semi-global histogram with $13 \times 5 = 65$ bins, another 5-bin global histogram is computed by combining all the 16 local bins. This results in a 150 (80+65+5) bin histogram that is used for measuring the similarity between images. The global and the semi-global histograms could be produced directly from the local histogram at the matching time.

3.4 Invariant Moments

The shape of an image could be represented in terms of seven invariant moments ($\phi_1 - \phi_7$). They have been widely used in a number of applications [8, 15, 20]. The first six functions ($\phi_1 - \phi_6$) are invariant under rotation and the last one ϕ_7 is both skew and rotation invariant. They are based on the central i, j-th moments (μ_{ij}) of a 2-D image $f(x, y)$, which are defined as follows:

$$\mu_{ij} = \sum_x \sum_y (x - \bar{x})^i (y - \bar{y})^j f(x, y).$$

We have

$$\begin{aligned}
\phi_1 &= \eta_{20} + \eta_{02} \\
\phi_2 &= (\eta_{20} + \eta_{02})^2 + 4\eta_{11}^2 \\
\phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
\phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
\phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) \\
&\quad \cdot \left[3(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2\right] \\
&\quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \\
&\quad \cdot \left[3(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2\right] \\
\phi_6 &= (\eta_{20} - \eta_{02}) \left[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2\right] \\
&\quad + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
\phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) \\
&\quad \cdot \left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2\right] \\
&\quad - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03}) \\
&\quad \cdot \left[3(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2\right]
\end{aligned}$$

where $\eta_{ij} = \frac{\mu_{ij}}{\mu_{00}^{\gamma}}$ and $\gamma = \frac{i+j}{2} + 1$. The ϕ values make a seven-entries feature vector that is used for measuring the similarity between images.

4 Experimental Results

For the experiments, we used the S8 part of the MPEG-7 still images content set, which contains 2810 trademark images captured by a scanner (B&W images). 30 queries were chosen from the database. Starting from the image number 240 (the database images are numbered from 232) every hundredth image was chosen as a query (29 queries) and then the last image in the database (number 3084) was added to make 30 fairly chosen queries. In our test, for each of the 30 query images, 50 images (5 sets, each containing 10 images) that were retrieved by the 5 methods ($30 \times 50 = 1500$) were given to 17 observers (students and staffs in the University of Wollongong). They were asked to mark (1 to 10) each set in terms of the similarity of each set to the given query. In other words, 510 (30×17) questions are uniformly distributed to the subjects. Each question contained one query image and 5 sets of similar images to the query. Each set contained 10 images that were found by one of the methods. Statistical results of marking by subjects and the efficiency comparison are summarized in Table I.

Table I:
Statistical results of subjective and efficiency test.

Method	Highest Score%	Mean Value	Searching Time
correlation	44.90	5.15	$37000\, t_0$
EPNH	38.24	4.73	$9.5\, t_0$
EHD	23.73	4.36	$6\, t_0$
HED	22.55	4.17	$3.3\, t_0$
Inv. Mom.	13.33	3.33	t_0

The first column shows the name of the method. In the second column the percentage ratio gaining the highest score is shown for each method. If two or more methods shared the highest score then this score was considered for all participating methods. The next column exhibits the average value of the marks (out of 10) that each method has gained. The last column is the efficiency and shows the symbolic time consuming factor for each method. This time is computed for the matching process phase (on-line phase) and is the time needed to find 10 similar images to one arbitrary query.

We used $k = 1$ in HED method, and in EHD method desired_num_of_blocks is set to 1100 (the default value) and Th_{edge} is set to 0 (because the images are all black and white). We ignored the 3-bit quantization in EHD to put all methods in the same situation. The Euclidean distance was used for measuring the similarity between histograms in EPNH, HED and in invariant moments, while for EHD method the ℓ_1 distance with a weighting factor of 5 for global bins, as recommended in [19], was applied. We computed the correlation for the 64 (8×8) blocks in steps of two to decrease the computation cost. Then the C_t factor was considered as the similarity measure in the correlation approach. Computing the correlation in steps of two did not affect the accuracy of the method.

5 Analysis and Discussion

As indicated in the last column of Table I, the searching times for all methods, except for the correlation method, are close to each other. The time t_0 turned out to be 0.5 seconds for searching through 2810 database images, using a Pentium-III, 1000 MHz machine. The difference among the matching times originates from the number of comparisons in each method and the type of the distance measure used in that method. In the invariant moments method, only 7 values have to be compared while in HED 70 values, in EHD 150 values and in the current version of EPNH 240 values need to be considered. The ℓ_1 distance, which is used in EHD, has less computation cost than the ℓ_2 distance (Euclidean) which was employed for the other methods. On the other hand, the correlation approach convolves all points in pictorial indexes (64×64 images) by the correlation formula given in (1). This computation is more complex and more time consuming than a simple distance measure.

Although the interval that the correlation method needs to find the similar images is unacceptable in many applications (more than 5 hours in this case), its advantages in gathering the best mean value (column 3) and the top percentage of highest score (column 2) are clear. Therefore, we can consider the correlation technique as a benchmark for evaluating the other methods and then, the EPNH obtains the best rank in the highest score percentage (85.17% of the benchmark) and the mean value (91.84%) while the search time is sensible (4.75 sec.).

The EHD, HED and invariant moments acquire 52.85%, 50.22% and 29.69% of the benchmark in the highest score and 84.66%, 80.97% and 64.66% of the benchmark in the mean value, respectively. It is also remarkable that with an increasing search time the results become much closer to human perceptual judgment.

As a critical note, the third column in Table I shows that the underlying methods produce results that are earning at most 5.15 (out of 10) average mark in the subjective test. This indicates that filling the semantic gap in image retrieval domain needs more research and encourages further investigations in semantic-based techniques.

6 Conclusion

The new EPNH approach, correlation, HED, EHD and invariant moments methods were evaluated in this study for a) their closeness to human's perception and b) their search time interval in finding similar logos in the selected database. While the search time of EPNH, HED, EHD and invariant moments methods were in an acceptable range, the correlation approach produced its results in an unreasonable interval. On the other hand, the subjective test ranked the compatibility with human's assessment as follows: 1. correlation; 2. EPNH; 3. EHD; 4. HED, and 5. invariant moments. By choosing the correlation results as a benchmark, EPNH obtained 85.17% in the highest score criteria and 91.84% in the mean value criteria of the benchmark. The mean value criteria showed that even the correlation method could gain only the average mark 5.15 (out of 10), which means that more researches are needed to fill the semantic gap between human's expectation and the output of the current methods. For the future, we intent to improve the EPNH efficiency by grouping similar bins to decrease the number of bins while retaining its ability in finding similar images. In addition, for sub-image search purposes, dividing the whole image into parts and applying multiple local histograms is essential. Using relevance feedback for the grater satisfaction of the users in EPNH is the other research path.

Acknowledgments. The authors would like to acknowledge Dr. Ian Burnett, University of Wollongong, Australia, and Dr. Fernando Pereira, Instituto Superior Técnico, Portugal, for their help in providing the MPEG-7 still image dataset used in this research. The first author is financially supported by the Ministry of Science, Research and Technology of I.R. Iran.

References

[1] J. Laaksonen, M. Koskela, S. Laakso, and E. Oja, "Picsom-content-based image retrieval with self-organizing maps," *Pattern Recognition Letters*, vol. 21, no. 13-14, pp. 1199–1207, Dec. 2000.

[2] M. Beigi, A. B Benitez, and S. Chang, "Metaseek: a content-based meta-search engine for images," in *Proceedings of Spie*, USA, 1997, vol. 3312, pp. 118–128.

[3] W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Petkovic, P. Yanker, C. Faloutsos, and G. Taubin, "The QBIC project: querying images by content using color, texture, and shape," in *Proceedings of Spie*, USA, 1993, vol. 1908, pp. 173–187.

[4] H. W. Yoo, D. S. Jang, S. H. Jung, J. H. Park, and K. S. Song, "Visual information retrieval system via content-based approach," *Pattern Recognition*, vol. 35, pp. 749–769, 2002.

[5] B. S. Manjunath, J.-R. Ohm, and V. V. Vasudevan, "Color and texture descriptors," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 6, pp. 703–715, June 2001.

[6] M. Bober, "Mpeg-7 visual shape descriptors," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 6, pp. 716–719, June 2001.

[7] A. Chalechale and A. Mertins, "An abstract image representation based on edge pixel neighborhood information (EPNI)," in *first EurAsian conference on advances in information and communication technology*, 2002, In press.

[8] A. J. Jain and A. Vailaya, "Shape-based retrieval: a case study with trademark image databases," *Pattern Recognition*, vol. 31, pp. 1369–1390, 1998.

[9] H. K. Lee and S. I. Yoo, "A neural network-based image retrieval using nonlinear combination of heterogeneous features," in *IEEE proceedings of the 2000 Congress on Evolutionary Computation.*, 2000, vol. 1, pp. 667–674.

[10] Y. Kageyama and H. Saito, "Image retrieval system capable of learning the user's sensibility using neural networks," in *IEEE International Conference on Neural Networks*, 1997, vol. 3, pp. 1563–1567.

[11] Z. Wang, Z. Chi, and D. Feng, "Fuzzy integral for leaf image retrieval," in *IEEE International Conference on Fuzzy Systems*, Hong Kong, 2002, pp. 372–377.

[12] C. Vertan and N. Boujemaa, "Embedding fuzzy logic in content based image retrieval," in *IEEE 19th International Conference of the North American Fuzzy Information Processing Society*, 2000, pp. 85–89.

[13] T. S. Chua, W.C Low, and T.M. Sim, "visual information retrieval," *The Journal of the Institiute of Image and Electronic Engineers of Japan*, vol. 27, pp. 10–19, 1998.

[14] C. Colombo, A. D. Bimbo, and P. Pala, "Semantics in visual information retrieval," *IEEE Multimedia*, vol. 6, no. 3, pp. 38–53, 1999.

[15] S. J. Yoon, D. K. Park, S. Park, and C. S. Won, "Image retrieval using a novel relevance feedback for edge histogram descriptor of Mpeg-7," in *proceedings IEEE International Conference on Consumer Electronics*, Piscataway, NJ, USA, 2001, pp. 354–355.

[16] K. Hirata and T. Kato, "Query by visual example-content based image retrieval," in *Advances in Database Technology - EDBT '92*, Berlin, Germany, 1992, pp. 56–71.

[17] A. K. Jain and A. Vailaya, "Image retrieval using color and shape," *Pattern Recognition*, vol. 29, no. 8, pp. 1233–1244, Aug. 1996.

[18] M. Abdel-Mottaleb, "Image retrieval based on edge representation," in *Proceedings 2000 International Conference on Image Processing*, Piscatway, NJ, USA, 2000, vol. 3, pp. 734–737.

[19] C. S. Won, D. K. Park, and S Park, "Efficient use of Mpeg-7 edge histogram descriptor," *Etri Journal*, vol. 24, no. 1, pp. 23–30, Feb. 2002.

[20] D. Mohamad, G. Sulong, and S. S. Ipson, "Trademark matching using invariant moments," in *Proceedings second Asian Conference on Computer Vision, [ACVV'95].*, Singapore, 1995, vol. 1, pp. 439–444.

[21] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986.

[22] S. Brandt, J. Laaksonen, and E. Oja, "Statistical shape features in content-based image retrieval," in *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*, Los Almaitos, CA, USA, 2000, vol. 2, pp. 1062–1065.