An Efficient Region-Based Approach for Object Recognition and Retrieval Based on Mathematical Morphology and Correlation Coefficient

Jehad Alnihoud
Department of Computer Science, Al al Bayt University, Jordan

Abstract: *In this paper, a new technique based on slicing the images to equally sub-areas, then applying the density slicing to the color histogram of these areas combined with the color pair technique is presented as a filtration phase based on the color features of images. This allows dramatic reduction in the candidate sets of images, which needs to be proceed through the time-consuming process of shape feature extraction and comparison algorithms. After that, shape recognition method is proposed as a higher level phase in the proposed system. The proposed approach for shape-based retrieval relies on detecting the edges of the objects, using laplacian of gaussian, in the image query as well as the set of filtered images from the proceeding layer of color approach. Next, the morphological techniques, using dilate and erosion are applied, in order to solidify the objects within the image and to increase the chance of correlation coefficient to be used as means of similarity between images. Furthermore, transformation of objects is applied to overcome the rotation based problem. The experimental results show that the proposed approach is fast, efficient, and feasible.*

Keywords: *Morphological operations, density slicing, color histogram, correlation coefficient.*

Received September 27, 2006; accepted February 15, 200

1. Introduction

Shape-based image retrieval is a challenging research topic. Among the difficulties facing shape-based retrieval are the issues of comparing the shape similarity, isolating objects from the background, and obtaining shape description despite occlusion, clutter, and deformation. Smeulders *et al*. in [1] stated that object segmentation for broad domain of general image is not likely to succeed, with a possible exception for sophisticated techniques in a narrow domain.

The most common categories of shape-based retrieval approaches are global shape feature-based methods, region-based querying, and modal-based querying. Some researchers use shape-based features to create a global histogram for image matching [2]. In [3], an approach, which combines color and shape in the retrieval process, is proposed. Integrating color and shape is a promising area in image retrieval. The Canny operator is used for edge detection. Histogram normalization is used to make the algorithm scale invariant. In addition, to handle the rotation invariance problem for edge direction, they use a histogram match across all possible shifts.

The region-based querying methods use segmentation techniques to segment the image to a homogeneous region and use these regions to extract the shape features [4]. In [5], a method for representing and matching trademarks, using positive and negative shape features is proposed. The region-based approaches do not take into consideration the spatial layout information between regions. The layout information is important in describing complex objects and their relations in the image. To overcome this limitation of the region-based approach, some researchers propose region graph-based which takes into account the spatial relationships between regions of interest. In [6], a hybrid graph representation for color based image retrieval is presented. They used modified color adjacency graph and spatial variance graph to handle the spatial similarity. In the matching stage, the histogram intersection is extended into graph intersection. The algorithm works well even if the image is corrupted. The model based querying is proposed to overcome the problems of shape deformation and automatic matching. All methods belonging to this approach have taken advantage of the model information of the object of interest. In [7] and [8], an image similarity based on elastic template matching is proposed. They used 1-dimensional template model of the object contour and a 2-dimensional template of the gray level pattern within the region.

In this paper, the proposed approach for shape-based retrieval falls under region-based methods. The approach is invariant under translation, rotation, and starting point. The rest of the paper is organized as follows. Section 2 is an overview of the similarity and evaluation techniques. System overview is presented in

section 3. In this section, the proposed color and shape based approaches are illustrated in details. Evaluation of the proposed approaches is presented in section 4. In section 5, the conclusion is presented.

2. Similarity Measurements and Evaluation Techniques

2.1. Similarity Measurements

There are different methods to measure the similarity or dissimilarity between feature vectors. The difference can be measured by *distance measure* in the *n*-dimensional space where the bigger the distance between two vectors, the greater the difference [9]. Given two vectors *A* and *B*, where $A = [a_1 a_2 a_3 ... a_n]$ and $B = [b_1 b_2 b_3 ... b_n]$. The distance between *A* and *B* can be calculated, based on city block as follows:

$$D_{ist} = \sum_{i=1}^{n} |a_i - b_i| \quad (1)$$

or based on euclidean distance function as follows:

$$D_{ist} = \sqrt{\sum_{i=1}^{n} (a_i - b_i)^2} \quad (2)$$

Or based on maximum value metric as follows:

$$D_{ist} = max\{|a_1 - b_1|, |a_2 - b_2|, |a_3 - b_3|, ..., |a_n - b_n|\} \quad (3)$$

The second type of metric used for comparing two feature vectors is the similarity measure. The most common is the vector inner product, which may be defined as:

$$Sim = \sum_{i=1}^{n} a_i b_i \quad (4)$$

2.2. Evaluation Techniques

A. Recall and Precision

To measure the performance of any retrieval system, precision and recall are still the most prominent techniques. Suppose a data set *DB* and a query *q* is given. The data set may be divided into two sets: the relevant set to the query $q, R(q)$, and, the set of irrelevant images $\overline{R}(q)$. Suppose that the query *q* is given to a data set and that it returns a set of images *A(q)* as the answer. Then, the precision of the answer is the fraction of the returned images that is indeed relevant to the query [1].

$$Precision = \frac{|A(q) \cap R(q)|}{|A(q)|} \quad (5)$$

The recall is the fraction of relevant images that is returned by the query with respect to the whole set of relevant images in the database.

$$Recall = \frac{|A(q) \cap R(q)|}{|R(q)|} \quad (6)$$

The above techniques are valid for a single query. In general, retrieval algorithms are evaluated by running them for several queries. Therefore, to evaluate the performance of an algorithm over all test queries, the average precision at each recall level should be considered [10]. The average precision may be defined as:

$$\overline{P}(r) = \frac{\sum_{i=1}^{N_q} P_i(r)}{N_q} \quad (7)$$

where $\overline{P}(r)$ is the average precision at recall level r, N_q is the number of queries, and $P_i(r)$ is the precision at recall level *r* for the *i*th query.

B. Threshold vs. Cut-off Techniques

In [11], an image retrieval system for coats of arms is proposed and the deployment of the threshold technique is preferred compared to the *cut off* technique. They stated that the threshold technique is hardly affected the precision of query model and when average precision is used then the recall level may improved even if the threshold value is reduced. They spell that the threshold technique suit images with slightly greater distance if they exceed the number of images to be return more than the *cut off* technique. Based on that, threshold technique may be considered as a major component of similarity measurement. In this research, the threshold technique is adopted rather than the *cut off* technique.

3. System Overview

3.1 System Flow Diagram

The methodology of implementation relies on incremental development process combined with iterative enhancement method. This method allows the ease of implementation and utilizes reusability of components in a flexible manner. The system flow diagram is presented in Figure 1.

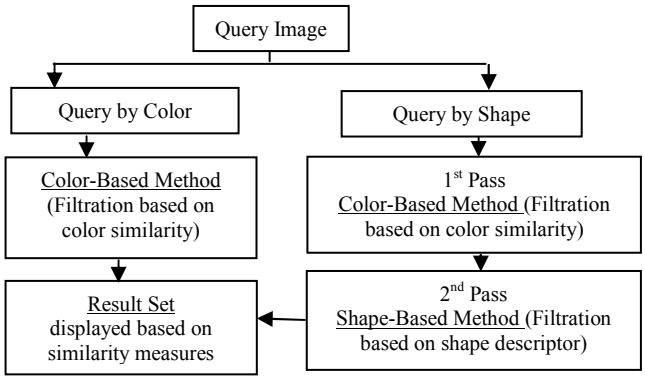


Figure 1. System flow diagram.

3.2. Method of Color Features Extraction and Comparison

A. Density Slicing (Multiple Thresholding)

The density slicing technique [12] can be summarized as follows:

1. *Let [0, L-1] represent the gray scale.*
2. *Suppose that M planes are defined at levels* $l_1, l_2, l_3, l_4, ..., l_M$ *and let* l_0 *represent black [f(x,y)=0] and* l_{L-1} *represent white [f(x,y)= L – 1].*
3. *Assuming that 0<M<L-1, the M planes partition the gray level scale to M+1 intervals (regions) [*R_1, R_2, ..., R_{M+1}*] gray levels to color assignments are made according to the relation:*

$$f(x,y) = c_k \quad if \quad f(x,y) \in R_k \qquad (8)$$

where c_k *is the color associated with the k-*th *intensity interval* R_k *defined by the partitioning planes at l=k-1 and l=k.*

B. Density Slicing of Color Histogram and Color-Pair Technique

Color histogram extracts information about color distribution not colors location [13]. This will lead to undesirable result when we need to examine the spatial distribution of colors in the image. As a solution to this problem, we may consider dividing the image to sub-areas. Locality information is captured for each sub area, more sub areas more accuracy in capturing the information related to the image. Density slicing of color histogram combined with color-pair technique is proposed as follows:

1. *Divide the image to 4x4 sub areas.*
2. *Calculate the color histogram for each sub area. This can be numbered from 1 to 16 in left-right, top-down sequence. Calculation of the color histogram will be based on the density slicing. We choose to use 8 bit images with 255 shades, based on density slicing the pixels reside before the particular bin is added to that bin, this will make the computational less expensive than using all color bins.*
3. *Choose the most distinct two colors -based on the number of pixels holding these colors- from each sub area.*
4. *Store the resulting color histogram value for each image in a database. At the end of this stage, for each image we should obtain 32 color pairs, two for each area in that image.*
5. *Repeat 1 to 2 for image query.*
6. *Compare the two color pairs for each sub area in the image query with the relevant sub area in the database using city block. At the end of this stage, the number of similar regions is identified.*
7. *Retrieve an image from the database if the similar regions are more than or equal to 8.*

As a summary for the proposed similarity and ranking approach, the system measured the distance between two vectors to decide if the region of interest is similar or not. If so, it will be labeled, otherwise it will be ignored. Then the decision to retrieve an image from the database will be based on how many regions from that image labeled as similar. If more than 8 regions labeled then retrieve that image, otherwise ignore. Given that we have 16 regions in an image, to rank the retrieved image we extract the following technique:

1. *If the number of similar regions between the image query* I_q *and images in the database* I_d *are different then the similarity ranking will be based on the following formula:*

$$Sim(I_q, I_d) = \frac{No.\ of\ similar\ regions}{16}\ X\ 100\ \% \qquad (9)$$

2. *If there are two or more images having the same number of similar regions then the ranking technique need to consider the distance between the query image and the database rather than just considering the number of regions.*

3.3. Shape-Based Retrieval Approach

The following steps show the major procedures of the proposed shape-based retrieval Approach.

1. *Create proper structuring element.*
2. *For each candidate image retrieved through the first layer*
   Do {
   Resize to 200 x 200 pixels
   Convert to gray scale
   Apply LOG (Laplacian of Gaussian)
   Transform images (create aliases)
   Apply mathematical morphology operation
   }
3. *Apply step 2 to query image - {Transformation}*
4. *Find absolute correlation coefficient (|r|) between query and set of images as well as their aliases.*
5. *Sort and display result set, based on a selected threshold value.*

3.3.1. Laplacian of Gaussian

Laplacian edge enhancement filter is one of the most common edge enhancement filters. Laplacian masks are rotationally symmetric, which means edges at all orientations contribute to the result. The only non-directional (rotationally independent) linear 2nd derivative operator (in 2D) is the Laplacian. The Laplacian of two-variable function *f*(*x*, *y*) is a second-order derivative defined as:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \tag{10}$$

$\nabla^2 f$ is calculated using convolution filter. Since the input image is represented as a set of discrete pixels, a discrete convolution kernel may be used to approximate the second derivative in the laplacian definition. Laplacian filters are derivative filters that are used to find areas of rapid change (edges) in images. Since derivative filters are very sensitive to noise, it is common to smooth the image (e.g., using a gaussian filter) before applying the laplacian. This two-step process is called the Laplacian of Gaussian (LoG) operation. As stated in [14], to include a smoothing gaussian filter, the laplacian and gaussian functions can be combined into one filter, the LoG operator, which is defined by

$$LoG\ (x, y) = -\frac{1}{\pi\sigma^4}\left[1 - \frac{x^2 + y^2}{2\sigma^2}\right] e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{11}$$

where *x* and *y* are the images co-ordinates and σ is the gaussian standard deviation.

LoG edge detection involves filtering, enhancement, and detection. Gaussian filter is used for smoothening and the second derivative of it is used for the enhancement step. The detection criterion is the presence of a zero crossing in the second derivative, with the corresponding large peak in the first derivative. The discrete kernel that approximates the LoG function depends on the gaussian standard deviation σ value. The default σ is equal to 0.5 and the size of the filter is *N* x *N*, where *N*= *ceil* (σ x 3) x 2 + 1. LoG edge detector is the only detector, which satisfies the three criteria of Marr and Hilbert at the same time. This is because laplacian, by itself, is considered as a solution to criteria number three and the gaussian filter is considered as optimal for the trade-off between localization in space and frequency [15].

LoG is essential in the proposed technique for shape retrieval. This is because it is used to detect and enhance or smooth the edges. This characteristic of the LoG operator made it the optimal solution for one of the major problems, which may be arisen when the low level image processing operation is used in image retrieval that is "noise" or unwanted edges.

3.3.2. Definitions of Mathematical Morphology

Morphology in general, is related to the structure of objects. Morphological filters are created to simplify a segmented image and to facilitate searching for object of interest. This can be carried out through various operations, such as smoothing the outlines of object, filling holes and eliminating small projections. The main concern within this research is to apply morphological *dilate and erosion* on binary images. Theses operations are considered after implementing the edge detection technique. Erosion after dilation solidify the object and fill the holes, that yield to isolate the object from the background and enhance the correlation coefficient based similarity measurement to be realistic and accurate.

In general, a dilation of set *A* by set *B* (structuring element) can be represented by the following formula:

$$D(A, B) = A \oplus B = \bigcup_{\beta \in B} (A + \beta) \tag{12}$$

while, the erosion may define as:

$$E(A, B) = \bigcap_{\beta \in B} (A - \beta) \tag{13}$$

where $-B = \{-\beta \mid \beta \in B\}$

Properties of dilation and erosion:

- Commutative $D(A, B) = A \oplus B = B \oplus A$
- Non-commutative $E(A, B) \neq E(B, A)$
- Associative $A \oplus (B \oplus C) = (A \oplus B) \oplus C$
- Translation invariant $A \oplus (B + h) = (A \oplus B) + h$

We can combine dilation and erosion to build two important higher order operations:

- Opening: O (A, B) = A ○ B = D (E (A, B), B).
- Closing: C (A, B) = A ● B = E (D (A, -B)- B)

A structuring element (also known as kernel) is a set of point coordinate that is represented as a binary image. The differences between the structuring element and the input image is that the structuring element is smaller in size, and its origin is not in the corner, and because of that, some coordinate elements may have negative values. In implementing morphological operation, the structuring element represents shapes (i.e., line, square and diamond). Dilation, in general, causes objects to dilate or grow in size; erosion causes objects to shrink. The initial dilation fills the holes with 1's, but also widens the boundary of objects. So, to reduce the object size, the erosion is used. The amount and the way that they grow or shrink depend upon the choice of the structuring element. Dilating or eroding without specifying the structural element makes no more sense than trying to low-pass filter an image without specifying the filter. The two most common structuring elements (given a cartesian grid) are the 4-connected and 8-connected sets, N_4 and N_8. They are convex, bounded, symmetric sets. Because of that the Vincent's theorem may applied in case of using one of these sets. When the theorem is applied to an image consisting of discrete pixels, states that for a bounded, symmetric structuring element *B* that contains no holes and contains its own center $[0, 0] \in B$, then $D(A,B) = A \oplus B = A \cup (\Im A \oplus B)$, where $\Im A$ the contour of the object is. That is, $\Im A$ is the set of pixels that have a background pixel as a neighbor. The implication of this theorem is that it is

not necessary to process all the pixels in an object in order to compute dilation or erosion. We only have to process the boundary pixels. This also holds for all operations that can be derived from dilations and erosions. The processing of boundary pixels instead of object pixels means that, except for pathological images, computational complexity can be reduced from $O(n^2)$ to O(n) for an *n* x *n* image. Choosing or defining the structuring element has a great effect on the performance of the morphological operations – dilate and erosion- because the output of these operations depends on the type of the structuring element. In this research, the structuring element is chosen based on image understanding. This implies that different structuring elements are chosen for the different query in hands.

3.3.3. Transformation of Objects

The object can be transformed through applying a transformation matrix to each point representing the object. One of the major problems facing shape-based similarity is the location of the object within the image, when the objects in the query image as well as the image database are translated to a fixed point within the image then the location-based problem will be eliminated. Using transformation matrices the process of translation and rotation may be expressed as shown in Figure 2.

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \qquad \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

(a) Translation matrix. (b) Rotation matrix.

Figure 2. The process of translation and rotation.

Through changing the angle of rotation, different aliases of the object may be created. In the proposed approach the objects in the images database are translated to a fixed point (pivot point) then rotated using $\theta = \{90^o, 180^o, 240^o, and\ 360^o\}$. The original image as well as the aliases is linked together. Any of these aliases may represent the original image if it is matching the query more than the others. This technique reduced the effect of one of the major problems of shape-based similarity that is object rotation-based problem.

3.3.4. Matching by Correlation Coefficient

One of the principle applications of correlation in image processing is in the area of template matching or prototype matching, where the problem is to find the closest match between an unknown image and a set of known images. Computing the correlation between the unknown and each of the known images is a way to find this matching. The higher the correlation value, the closer the match is. If we consider finding the matches between an image $f(x, y)$ of size *M* x *N* and a sub image $w(x, y)$ of size *J* x *K,* as shown in Figure 5, then, the correlation has a disadvantage of being sensitive to the changes of amplitude of the image $f(x, y)$ and the sub image $w(x, y)$. An approach to overcome this difficulty is to perform matching via correlation coefficient, which may be defined as:

$$\gamma(st) = \frac{\sum_x \sum_y (f(x,y) - \bar{f}(x,y))(w(x-s,y-t) - \bar{w})}{\sqrt{\left(\sum_x \sum_y (f(x,y) - \bar{f}(x,y))^2\right)\left(\sum_x \sum_y (w(x-s,y-t) - \bar{w})^2\right)}} \quad (12)$$

where s=0, 1, 2, …, M-1, t=0, 1, …, N-1, $\bar{w}$ is the average value of the pixels in $w(x, y)$, and $\bar{f}(x,y)$ is the average value of $f(x, y)$ in the region coincident with the current location of *w*. The correlation coefficient is scaled in the range of -1 to 1, independent of scale changes in the amplitude of $f(x, y)$ and $w(x, y)$ [12].

Correlation can be normalized for amplitude changes through the correlation coefficient but for changes in size and rotation, it is difficult to use. Normalizing for size involves spatial scaling a process, which needs a lot of computation. Normalizing for rotation is more difficult. If a clue for rotation can be extracted from rotation in $f(x, y)$, then we rotate $w(x, y)$ so that it aligns itself with the degree of rotation in $f(x, y)$. If the nature of rotation is unknown, then looking for the best match requires exhaustive rotation of $w(x, y)$, which is an impractical procedure. The rotation and scaling problems associated with correlation coefficient are addressed through resizing the images and object transformation.

In the proposed shape retrieval technique, the concern is given to the correlation coefficient between two matrices of the same size, which may be defined as follows. Given that A is a matrix of size $m \times n$ and B is another matrix of the same size, then the correlation coefficient may be interpreted as follows.

$$r = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{\left(\sum_m \sum_n (A_{mn} - \bar{A})^2\right)\left(\sum_m \sum_m (B_{mn} - \bar{B})^2\right)}} \quad (13)$$

where $\bar{A} = \frac{\sum_m \sum_n A_{mn}}{m \times n}$ and $\bar{B} = \frac{\sum_m \sum_n B_{mn}}{m \times n}$

The correlation coefficient is applied as the last step in the proposed approach, as a similarity mean to rank the retrieved images (result set). After applying the edge detection operator and dilate function for both image query and images database, it is beneficial to compare the images based on the correlation coefficient *r* with respect to the query. In general, there are three scenarios based on the value of *r*.

1. *r* is close to one, which indicates that the two images are very much similar.
2. *r* is close to zero, which suggests that the two images are dissimilar.
3. *r* is close to minus one, which suggests that the two images are similar, but the objects located in opposite locations.

To retain the similarity based on shape, considering the previous scenarios, the best option to use is the absolute value of *r* and then sorts the values in descending order. This is because applying the absolute value of *r* will ensure that the images, which cope with the first and the third scenarios, will be ranked in the top of the retrieved set. This is due to the concern in retaining images with similar objects, even if the location of the objects is not the same due to rotation or shifting.

4. System Evaluation

4.1. Experimental Setup

To evaluate the performance of the proposed shape-based retrieval approach, the following image queries shown in Figure 3 have been used for testing. The database to test the shape approach is composed of 1000 images. The images database contains different images with different objects and colors. The query images contain different shapes in different positions.

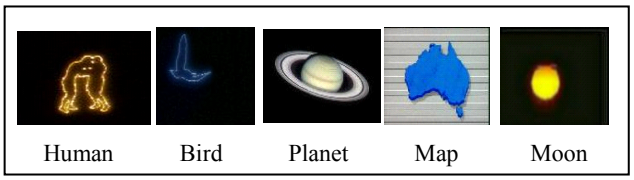


Figure 3. Sample objects.

Each of these images is filtered in a preprocessing stage, using the Paint Shop Pro 5.0 software and saved in the database. This is carried out to examine the reliability of the shape technique in dealing with the different shape and orientation of objects within the query images. Filtration includes rotation, flipping, mirror, noise, and edge enhancement. The totaled number of filtered image is 139 images, that is around 14% of the database, while the rest of the images in the database are chosen randomly without applying any preprocessing operation. The only constraint considered in examining the shape technique through these query images is that 'the object should be in contrast with the background'.

The interface allows the user to interact with the system in a flexible manner, through three options of retrieval. Based on these options, the threshold value will be changed automatically. Basically, the available options are as follows.

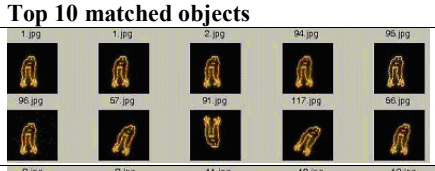
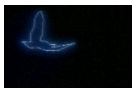
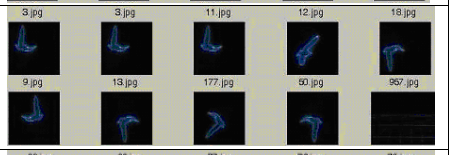
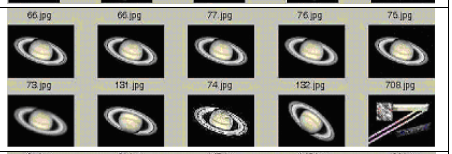
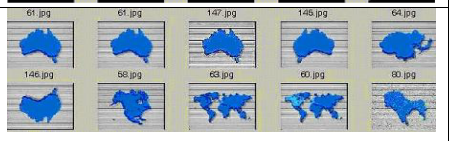
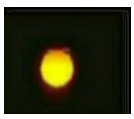
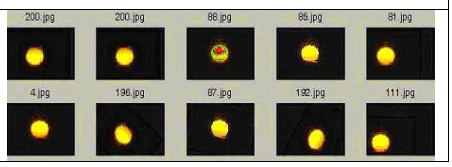
1. *1st Option*: in which the correlation coefficient value (*r*) is chosen to be ≥ 0.50. This option may be considered as the standard option to be used in the system.
2. *2nd Option*: in which the correlation coefficient value (*r*) is chosen to be ≥ 0.8. This option may be used if the user is looking for precise matching.
3. *3rd Option*: in which the correlation coefficient value (|*r*|) is chosen to be ≥ 0.1. This option may be considered when the user is looking to increase the number of retrieved set.

These options meet the different needs of the user community, and at the same time, the user interference is minimal, since the threshold values are changed automatically.

4.2. Experiments and Results

Table 1 shows the result retained by the system when the 1st option is used. The query image is seen in the upper left corner of the retrieved set and the retrieved images are ranked in the top-down, left-right sequence. Table 2 shows the recall versus the average precision for the query images shown in Figure 3.

Table 1. Sample queries and retrieved sets.

Query image	Top 10 matched objects
	
	
	
	
	

The average precision has been calculated, taking into consideration that some precision values have been interpolated. Table 2 proves that the shape approach is able to achieve high accuracy if the 1st option is used. As it can be clearly seen, the highest average precision value achieved is at recall level 10% and the lowest average precision value for the seven query images is 0.32 at 100% recall level. In between

these two values, the average precision is decreasing slightly. Applying the 1st option proves that with $r \geq 0.50$ as the threshold value, the shape approach

Table 2. Recall versus average precision ($r \geq 0.50$).

	Human	Bird	planet	Map	Moon	
Recall	Prec.	Prec.	Prec.	Prec.	Prec.	Average precision
10%	0.97	0.96	0.95	0.98	0.99	0.97
20%	0.96	0.94	0.95	0.98	0.98	0.96
30%	0.95	0.93	0.94	0.97	0.96	0.95
40%	0.94	0.93	0.93	0.97	0.96	0.95
50%	0.94	0.93	0.77	0.96	0.95	0.91
60%	0.91	0.86	0.70	0.96	0.93	0.87
70%	0.90	0.86	0.67	0.96	0.90	0.86
80%	0.90	0.85	0.67	0.95	0.94	0.86
90%	0.89	0.85	0.00	0.89	0.89	0.70
100%	0.00	0.81	0.00	0.78	0.00	0.32

achieves high average precision values up to 90 %recall level. After that, the same set of images is tested, using $r \geq 0.80$. Using a high correlation value implies that fewer images are retained. This may explain the results achieved in this experiment. The average precision was very high up to 60% recall level. However, it begins to decrease rapidly with higher recall levels. That because, some of the query images are not able to reach these high recall levels. This has a major impact on the average precision at high recall levels (more than 60%). Applying the 2nd option, the maximum recall level achieved is 90% with a relatively low average precision of 0.18.

The final test of recognition rate is done based on $r \geq 0.1$ as threshold value. This yields to increase the number of images retained, which implies that most of the similar images in the database may retain. At the same time, it gives the chance to some false hits to be retained. The experiment shows that even with these false hits, the approach still achieves very high average precision. The highest average precision 0.93 is achieved at recall level 10 % and subsequently the average precision starts to decrease gradually to reach 0.31 at recall level 100 %.

We may conclude that the 1st option and 3rd option are behaving almost the same way, where the precision decreases gradually as the recall level increase. It is clear that the 1st option performs slightly better than the 3rd option, with regards to all recall levels. The intersection point between the 2nd option and the 1st option is at recall level close to 50%. At this recall level, the precision of the 2nd option starts to be less than the 1st and 3rd options. In brief, the three options are valid and it is left to the user to identify, which one is to be used.

Through an extensive testing of the proposed approach, the overall recognition rate is found to be 96%. The proposed approach shows an improvement in the recognition rate compared to [16], in which the recognition rate was 91.5%. Table 3 shows a comparison between the proposed approach and other approaches as mentioned in the previous reference.

Table 3. Recognition rate comparison.

	Recognition Rate
Proposed approach	96 %
K-Means	86.42%
C-Means	84.64%

The comparison proves that the proposed approach shows significant improvements compared to a well known approaches for shape-based retrieval. Finally, the speed of performance is crucial to the success of any shape retrieval technique. To enhance the speed of performance of the proposed system, two techniques are employed. These are the most time consuming process is calculated off-line, in which the color features from the image database is extracted and filtration through the first layer. To test the speed of performance of the shape approach, the execution time for the query images is measured, taking into account the different options available. Table 4 shows the result of this testing. The execution time of the 1st option is less than 3rd option, but it is more than the time needed, if the 2nd option is used. This result is expected and realistic, since the execution time increases or decreases, based on the number of images retained through the color layer and executed based on the correlation coefficient value in the shape-based layer. The run-time testing is done, using Pentium 4 with 2.2GHZ speed and 256 MB of RAM.

Table 4. Run-time testing.

	Execution time in seconds		
	$r \geq 0.50$	$r \geq 0.80$	$R \geq 0.10$
Human	4.70	2.73	6.03
Bird	6.86	2.13	8.05
Planet	6.48	4.23	9.18
Map	3.42	1.89	4.08
Moon	5.05	3.57	7.41
Average	5.302	2.91	6.95

5. Conclusion

A new color and shape approaches is proposed in this paper. The new approach may be categorized under the region-based methods for shape-based retrieval. The approach is capable of solving the most prominent drawback of using region-based methods, that is, the problems related to the unrelated intensity edges to the boundary of the objects within the image. The use of *Laplacian of Gaussian* enables to detect and smooth the edges of the objects in the images. This removes the unwanted intensity edges formulated through noise or other artifacts, which are not related to the boundary of the object of interest. Furthermore, applying

morphological operations –Dilate and Erosion- fills the holes within the boundaries that solidify the objects. This enables the use of correlation coefficient as bases of shape similarity. Moreover, the proposed double-layered system based on the color feature extracted through the first layer to discriminate the irrelevant images from being tested through the shape layer is proposed. This provides a significant solution to the speed problem, since it enables the use of an off-line process in extracting and saving the color features from the images database. Finally, the proposed system is fully automated, interactive, and user friendly.

References

[1] Anil K. and Aditya V., "Image Retrieval Using Color and Shape," *Pattern Recognition*, vol. 29, no. 8, pp. 1233-1244, 1996.

[2] Breitender C. and Eidenberger H., "A Retrieval System for Coats of Arms," *in Proceedings of the ISIMADE'99*, pp. 91-99, 1999.

[3] Carson C., Belongie S., Greenspan H., and Jitendra M., "Region Based Image Querying," *CVPR'97 Workshop on Content-Based Access of Image Library*, San Juan, pp. 42-51, 1997.

[4] DelBimbo A. and Pala P., "Visual Image Retrieval by Elastic Matching of User Sketches," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 2, pp. 121-132, 1997.

[5] Fangwei Z. and Christopher J., "Use of the Laplacian of Gaussian Operator in Prostate Ultrasound Image Processing," *in Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 20, no. 2, pp. 812-815, 1998.

[6] Gong Y., Zhang H., Chuan H., and Sakauchi M., "An Image Database System with Content Capturing and Fast Image Indexing Abilities," IEEE 0-8186-5530-5/9, 1994.

[7] Marr D. and Hildreth E., "Theory of Edge Detection," *in Proceedings of Royal Society London*, vol. B207, pp. 187-217, 1980.

[8] Pala P. and Santini S., "Image Retrieval byShape and Texture," *Pattern Recognition*, vol. 32, no. 3, pp. 517-527, 1999.

[9] Park I., Yun I., and Lee S., "Color Image Retrieval Using Hybrid Graph Representation," *Image and Vision Computing*, vol. 17, no. 7, pp. 465, 1999.

[10] Rafael C. and Richard E., *Digital Image Processing*, Addison Wesley Publishing Company, 1992.

[11] Scott U., *Computer Vision and Image Processing*, Prentice-Hall, 1998.

[12] Smeulders A., Worring M., Sanati S., Gupta A., and Jain R., "Content- Based Image Retrieval at the End of the Early Years," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 12, pp.395-406, 2000.

[13] Soffer A. and Samet H., "Using Negative Shape Features for Logo Similarity Matching," *in Proceedings of International Conference on Pattern Recognition*, pp. PR21, 1998.

[14] Swain M. and Ballard D., "Color Indexing," *International Journal of Computer Vision*, vol. 7, no. 1, pp. 11-32, 1991.

[15] Yang S., Gao L., Wang R., and Bian L., "The Automatic Recognizable Method of TOE Shape Based on Fuzzy Comprehensive Evaluation," Computer Vision and Image Understanding, vol. 103, pp. 139-154, 2006.

[16] Yates B., *Modern Information Retrieval*, Addison-Wesley, 1999.

Jehad Al-Nihoud is an associate professor at Prince Hussein bin Abdullah Information Technology College, Al-al Bayt University, Jordan. He received the BSc in mathema tical statistics from Yarmouk University in 1990, MSc in computer science from University Putra Malaysia in 1999, and completed his PhD in computer science in 2004 at University Putra Malaysia.