Multimedia Tools and Applications, 15, 247–268, 2001

Spatial Color Indexing Using Rotation, Translation, and Scale Invariant Anglograms*

YI TAO yitao@compuware.com
Testing and Implementation Lab, Compuware Corporation, 31440 Northwestern Highway, Farmington Hills, MI 48334

W.I. GROSKY
Computer and Information Science Department, University of Michigan-Dearborn, Dearborn, MI 48128

Abstract. As color plays an essential role in image composition, many color indexing techniques have been studied for content-based image retrieval. This paper examines the use of a computational geometry-based spatial color indexing methodology for effective and efficient image retrieval. In this scheme, an image is evenly divided into a number of $M * N$ non-overlapping blocks, and each individual block is abstracted as a unique feature point labeled with its spatial location and dominant colors. For each set of feature points labeled with the identical color, we construct a Delaunay triangulation and then compute the feature point histogram by discretizing and counting the angles produced by this triangulation. The concatenation of all these feature point histograms serves as the image index, the so-called *color anglogram*. An important contribution of this work is to encode the spatial color information using geometric triangulation, which is rotation, translation, and scale invariant. We have compared the proposed approach with two of the best performing of recent spatial color indexing schemes, Color-WISE and the color correlogram approaches, respectively, at image block and pixel levels of different granularity. Various experimental results demonstrate the efficacy of our techniques.

Keywords: spatial color indexing, content based image retrieval, color anglogram, color autoanglogram, Delaunay triangulation, feature point, feature point histogram, point feature map, image database

1. Introduction

Multimedia information systems may be viewed as storage and retrieval systems where large volumes of multimedia data such as audios, images, and videos are created, indexed, modified, searched, and retrieved [7]. With the explosive advancement in imaging technologies, image retrieval has attracted the increasing interests of researchers in the fields of digital libraries, image processing, and database systems [21]. Traditional indexing for image retrieval is text-based [13]. For example, an image is manually annotated by identifying the photographer, time, place, and participating objects. Although text annotation is a practical technique, however, this task is labor intensive, language dependent, vocabulary controlled, and subjective in nature. In some cases, it is rather difficult to characterize certain important real-world concepts, entities, and attributes by means of text only. The shape of a single object and the various spatial constraints among multiple objects in an image are examples of such concepts.

*We would like to acknowledge the support of NSF grant 97-29818.

Motivated by the ultimate goal of automatically computing efficient and effective descriptors which symbolize various properties of images, recent research on image retrieval systems has been directed towards the development of content-based retrieval techniques for the management of visual information such as color, texture, shape and spatial constraints [28]. As color plays an important role in image composition, many color indexing techniques have been studied. Although global color histograms and moments have been proven to be very useful for image indexing [24, 26], they do not take color-based spatial information into account. Thus, when the image collection becomes very large, many false hits frequently occur. In order to incorporate both color and its spatial layout for image retrieval, the latest work attempts to characterize finer details of the color distribution, so-called *spatial color indexing*. For properties not concerned with color, the consideration of spatial similarity in content-based image retrieval has been widely studied with respect to the absolute location of an image object, the shape of an image object [15, 17, 30] and the relative locations of multiple image objects [1, 3, 9].

In this paper, we examine the use of computational geometry-based spatial color indexing for effective and efficient image retrieval. In this scheme, an image is evenly divided into a number of $M * N$ non-overlapping blocks, and each individual block is abstracted as a unique feature point labeled with its spatial location and dominant colors. For each set of feature points labeled with an identical color, we construct a Delaunay triangulation and then compute the feature point histogram by discretizing and counting the angles produced by this triangulation. The concatenation of all these feature point histograms serves as the image index, the so-called *color anglogram*. An important contribution of this work is to encode the spatial color information using geometric triangulation, which is translation, rotation, and scale invariant.

The remainder of this paper is organized as follows. In the next section, we briefly review various spatial color indexing techniques. In Section 3, some Delaunay triangulation-related concepts in computational geometry are introduced. Section 4 presents our methodology of computing the spatial color index for image retrieval as well as the similarity function for image comparison. In Section 5, we compare the proposed approach with two of the best performing of recent spatial color indexing schemes, Color-WISE [22] and the color correlogram [11] approaches, over two databases that comprise a total of 3388 images, to demonstrate the effectiveness and efficiency of our proposed approach. Finally, we give some concluding remarks.

2. Related work

In this section, we present a few descriptions that snapshot various current approaches of spatial color indexing for content-based image retrieval.

One of the earliest image retrieval projects was QBIC [17]. Provided with a visual query interface, the user can manually outline an image object to facilitate image analysis in order to acquire an object boundary, and then request images containing objects whose color is similar to the color of the object in the query image. In the QBIC system, each image object is indexed by a union of area, circularity, eccentricity, major axis orientation and some algebraic moment invariants as its shape descriptors, along with color moments such as the average (R, G, B), (Y, i, q), (L, a, b) and MTM (Mathematical Transform to Munsell) coordinates, as well as a k element color histogram. Other research groups have also tried to combine color and shape features for improving the performance of image retrieval. In [12], the color in an image is represented by three 1-D color histograms in (R, G, B) space, while a histogram of the directions of the edge points is used to represent the general shape information. A composite feature descriptor is proposed in [16] based on a clustering technique, and it combines the information of both the shape and color clusters, which are characterized by seven invariant moments and color cluster means, respectively. In [2], a system which uses a so-called *blobworld* representation to retrieve images is described, and it attempts to recognize the nature of images as combinations of objects so as to make both query and learning in the blobworld more meaningful to the user. In this scheme, each blob (region) in the image is described by the two dominant colors, the centroid for its location and a scatter matrix for its basic shape representation.

Due to the uncontrolled nature of the images available, automatically and precisely extracting image objects is still beyond the reach of the state-of-the-art in computer vision. Moreover, each image object may appear differently, depending on viewpoint, occlusion, and deformation. Therefore, researchers have introduced the notion of a rectangular cover of an image object or region for approximation. For example, spatial knowledge of a set of selected colors is obtained using a maximum entropy discretization with an event covering method in [10, 18]. It is noted in [10] that image comparison could be accomplished by matching objects of the same color (i.e., direct matching), or matching similar objects with different colors (i.e., indirect matching). The analysis in [18] showed that the retrieval effectiveness in terms of precision and recall [32] is good enough with only four regions considered within an image. In [23], a color-set back-projection technique is developed to extract color regions. By indexing the attributes of regions, such as sizes, locations and color sets, a variety of complex joint spatial and color queries can be efficiently computed. In order to circumvent the exhaustive search for spatial arrangements, the 2-D string is used for the representation of image spatial relationships and their comparisons.

Though it is more meaningful to represent the spatial distribution of color information based on image objects or regions, various fixed image partitioning techniques have also been proposed because of their simplicity and acceptable performance. In [25], an image is divided into five partially overlapped, fuzzy regions, with each region indexed by its three moments of the color distribution. In [4], the inter-hierarchical distance (IHD) is defined as the color variance between two different hierarchical levels (i.e., an image region and its subregions). Based on a fixed partition of the image, an image is indexed by the color of the whole image and a set of IHD's which encode the spatial color information. The system Color-WISE is described in [22], and it partitions an image into $8 * 8$ blocks with each block indexed by its dominant hue and saturation values.

Instead of partitioning an image into regions, there are other approaches for the representation of spatial color distribution. A histogram refinement technique is described in [20] by partitioning histogram bins based on the spatial coherence of pixels. A pixel is coherent if it is a part of some *sizable* similar-colored region, and incoherent otherwise. In [11], a statistical method is proposed to index an image by color correlograms which is actually a table containing color pairs, where the k-th entry for $<i, j>$ specifies the probability of locating a pixel of color j at a distance k from a pixel of color i in the image.

We note that both the histogram refinement and correlogram approaches do not recognize the nature of images as combinations of objects. As for meaningful region-based image representations, two image objects are usually considered similar only if the corresponding regions they occupy overlap. Along with the position dependence of similar image objects, the fixed image partition strategy does not allow image objects to be rotated within an image. In addition, in order to check whether these image objects are in the requisite spatial relationships, even 2D-strings and its variants suffer from exponential time complexity in terms of the number of concerned image objects.

3. Delaunay triangulation in computational geometry

Let $P = \{ p_1, p_2, \ldots, p_n\}$ be a set of points in the two-dimensional Euclidean plane, called *sites*. Partition the plane by labeling each point in the plane to its nearest site. All those points labeled as p_i form the *Voronoi region* $V(p_i)$. $V(p_i)$ consists of all points x at least as close to p_i as to any other site:

$$V(p_i) = \{x\colon |p_i - x| \leq |p_j - x|, \quad \forall \mathrm{j} \neq i\}.$$

Some points do not have a unique nearest site. The set of all points that have more than one nearest site form the *Voronoi diagram* $V(P)$ for the set of sites.

Construct the *dual* graph G for a Voronoi Diagram $V(P)$ as follows: the nodes of G are the sites of $V(P)$, and two nodes are connected by an arc if their corresponding Voronoi polygons share a Voronoi edge. In 1934, Delaunay proved that when the dual graph is drawn with straight lines, it produces a planar triangulation of the Voronoi sites P, the so-called *Delaunay triangulation* $D(P)$. Each face of $D(P)$ is a triangle, a so-called *Delaunay triangle*.

Among various algorithms for constructing the Delaunay triangulation of a set of N points, we note that there are O($NlogN$) algorithms [5, 6] for solving this problem. Amazingly, the Delaunay triangulation can be computed with less than 30 lines of C code as shown in [19], but with a costly algorithm efficiency of O(n^4). Theoretically, from the definition of the Delaunay triangulation, it is easily shown that those angles of the resulting Delaunay triangles of a set of sites (points) remain the same under uniform translations, scalings, and rotations of the point set. The proof of Delaunay's theorems and properties is beyond the scope of this paper, but can be found in [19]. An example is illustrated in figure 1. Figure 1(a) shows an MRI brain image with the boundary of one segmented lesion highlighted. Figure 1(b) shows the extracted lesion with its shape characterized by a set of 42 corner points. Figure 1(c) shows the resulting Delaunay triangulation for the set of 42 points shown in figure 1(b). Figure 1(d) shows the resulting Delaunay triangulation of a transformed (translation, rotation, and scaled-up) set of the 42 points in figure 1(b). Figure 1(e) shows the resulting Delaunay triangulation of another transformed (translation, rotation, and scaled-down) set of the 42 points in figure 1(b).

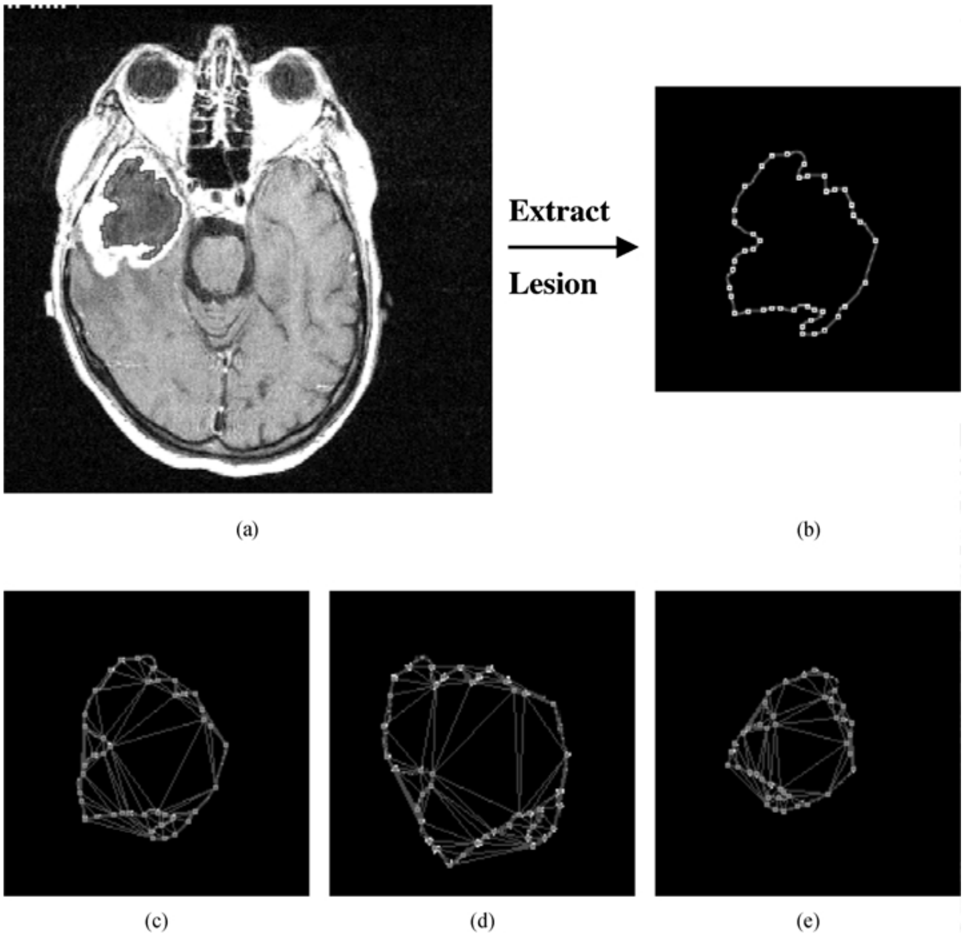


Figure 1. (a) An MRI brain image. (b) A set of 42 points. (c), (d) and (e) Delaunay triangulations.

4. Delaunay triangulation for spatial color indexing

Humans are much better than computers at extracting semantic information from images. We believe that complete image understanding should start from interpreting image objects and their relationships. Therefore, it is necessary to move from image-level to object-level interpretation in order to deal with the rich semantics of images and image sequences. An *image object* is either an entire image or some other meaningful portion of an image that could be a union of one or more disjoint regions. Typically, an image object would be a semcon (iconic data with semantics) [8]. For example, consider an image of a seashore scene shown in figure 2, consisting of some seagulls on the coast, with the sky overhead and a sea area in the center. Examples of image objects for this image would include the entire scene (with textual descriptor *Life on the Seashore*), the seagull region(s), the sand

Figure 2. An image of seashore scene.

regions(s), the water region(s), the sky region(s), and the bird regions (the union of all the seagull regions). Now, each image object in an image database contains a set of unique and characterizing features $F = \{f_1, \ldots, f_k\}$. We believe that the nature as well as the spatial relationships of these various features can be used to characterize the corresponding image objects [1, 2, 10, 23].

In 2-D space, many features can be represented as a set of points. These points can be tagged with labels to capture any necessary semantics. Each of the individual points representing some feature of an image object we call a *feature point*. The entire image object is represented by a set of labeled feature points $\{p_1, \ldots, p_k\}$. For example, a corner point of an image region has a precise location and can be labeled with the descriptor *corner point,* some numerical information concerning the nature of the corner in question, as well as the region's identifier. A color histogram of an image region can be represented by a point placed at the center-of-mass of the given region and labeled with the descriptor *color histogram,* the histogram itself, as well as the region's identifier. We note that the various spatial relationships among these points are an important aspect of our work, which sets our work apart from the other approaches described in Section 2.

Effective semantic representation and retrieval requires labeling the feature points of each database image object. The introduction of such feature points and associated labels effectively converts an image object into an equivalent symbolic representation, called its *point feature map*. We have devised an indexing mechanism to retrieve all those images from a given image database which contain image objects whose point feature map is similar to the point feature map of a particular query image object [1]. An important aspect of our approach is that it is rotation, translation, and scale invariant when matching images containing multiple semcons.

4.1. Point feature map representation

Digital images can be represented in different color spaces such as RGB, HSI, YIQ, or Munsell. Since a very large resolution of millions of colors is unwanted for image retrieval, the color space is usually quantized to a much coarser resolution. For example, HSI (Hue-Saturation-Intensity) color space is designed to resemble the human perception of color in which hue reflects the dominant spectral wavelength of a color, saturation reflects the purity of a color, and intensity reflects the brightness of a color. It is noted in [31] that humans are less sensitive to differences in either saturation or intensity than to differences in the hue component, so that, in general, hue is quantized more finely than the saturation or intensity component for image retrieval when HSI is used for image representation. As the process of grouping low-level image features into meaningful image objects and then automatically attaching semantic descriptions to these image objects is still an unsolved problem in image understanding, our work intends to combine both the simplicity of fixed image partition and the nature of images as combinations of objects into spatial color indexing so as to facilitate image retrieval.

Based on the assumption that salient image constituents generally tend to occupy relative homogeneous regions within an image, we expect that one or more meaningful image constituents may be composed of some image blocks with a particular color. Regardless of whether these image blocks are connected or not, they approximate the composition of the nature of images as combinations of objects. In our spatial color-indexing scheme, an image is first divided evenly into a number of $M * N$ non-overlapping blocks. Then each individual block is abstracted as a unique feature point labeled with its spatial location and dominant colors. After we adjust all two neighboring feature points to a fixed distance, all the normalized feature points form a *point feature map* of the original image for further analysis.

By representing an image as a point feature map, we capture not only the color information of the image, but also the spatial information about color. We can flexibly manipulate sets of feature points instead of dealing with image blocks. In order to compute our spatial color index of an image, we construct a Delaunay triangulation for each set of feature points in the point feature map labeled with the identical color, and then compute the feature point histogram by discretizing and counting the angles produced by this triangulation. An O(*max* (N, *#bins*)) algorithm is necessary to compute the feature point histogram corresponding to the Delaunay triangulation of a set of N points. The final image index is obtained by concatenating all the feature point histograms together. We note that in our spatial color indexing scheme, feature point histograms are not normalized, as a drawback of normalized histograms is its inability to match parts of image objects. For example, if region A is a part of region B, then, in general, the normalized histogram H_A is no longer a subset of the normalized histogram H_B.

An example is shown in figure 3. Figure 3(a) shows a pyramid image of size $192 * 128$; by dividing the image evenly into $16 * 16$ blocks, figure 3(b) and (c) show the image approximation using dominant hue and saturation values to represent each block, respectively. Figure 3(d) shows the corresponding point feature map perceptually, and we note that the distance between any two neighboring feature points is fixed, as images of different sizes undergo normalization. Figure 3(e) highlights the set of feature points labeled with hue 2, and figure 3(f) shows the resulting Delaunay triangulation. Figure 3(g) shows the resulting Delaunay triangulation of a set of feature points labeled with saturation 5, and figure 3(h) shows the corresponding feature point histogram obtained by counting only the two largest angles out of each individual Delaunay triangle with bin size of 10°. Our work in [30] has concluded that such a feature point histogram provides a sufficient and effective way for image object discrimination.

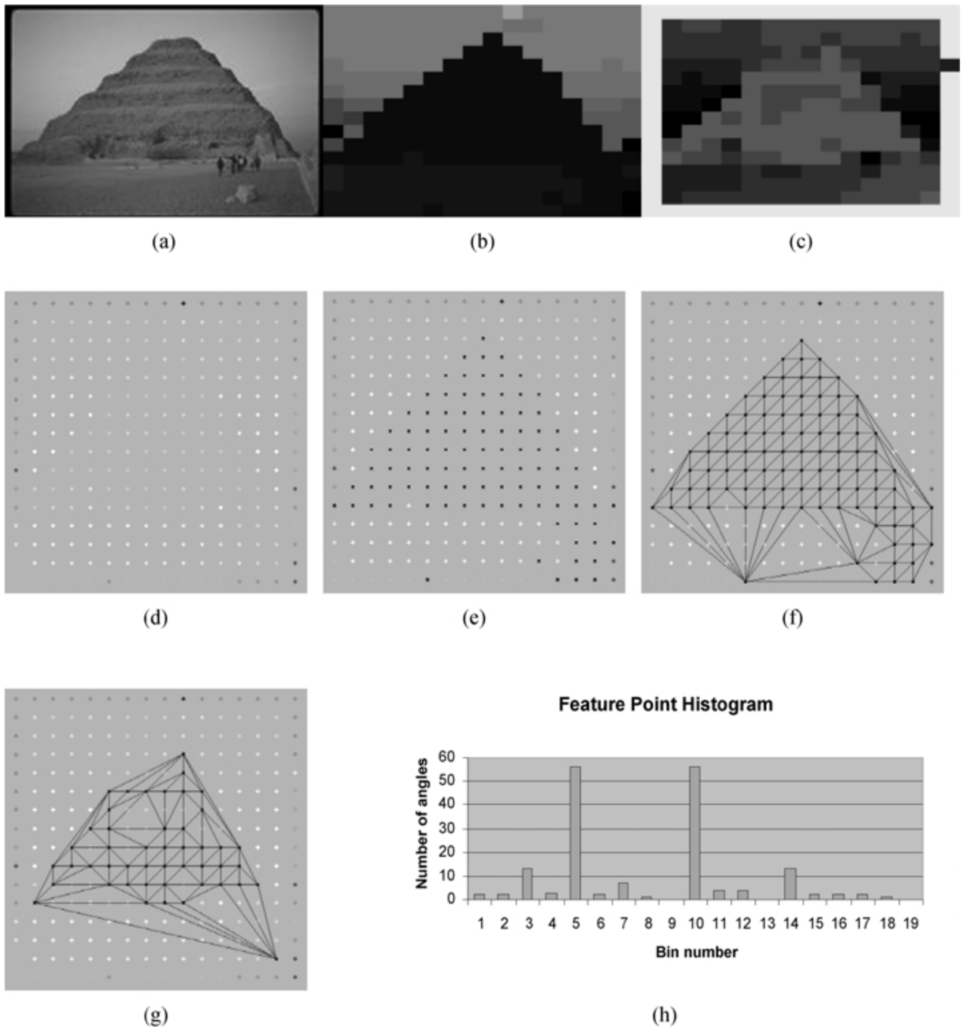


Figure 3. (a) A pyramid image. (b) Hue component. (c) Saturation component. (d) Point feature map. (e) Feature points of hue 2 (f) Delaunay triangulation of hue 2. (g) Delaunay triangulation of saturation 5. (h) Resulting feature point histogram of saturation 5.

4.2. Similarity function for image comparison

Histogram intersection was originally proposed in [27] for comparing color histograms of query and database images. It was shown that histogram intersection is especially suited to comparing histograms for recognition. Additionally, histogram intersection is an efficient way of matching histograms, and its complexity is linear in the number of elements in the histograms. The intersection of the histograms I_{query} and $M_{database}$, each of n bins, is defined as follows.

$$D(I_{query}, M_{database}) = \frac{\sum_{j=1}^{n} \min(I_j, M_j)}{\sum_{j=1}^{n} I_j}$$

Suppose that Q is the query image index consisting of $\sum Q_i$ for m color-related feature point histograms, D is the database image index with corresponding m color-related feature point histograms as $\sum D_i$, and w_i is the ith of m of variables which define the relative importance of color-related feature point histograms in our similarity calculation. For example, if HSI is used for image representation, hue-related feature point histograms are often assigned a larger weight value than saturation-related ones, as humans are more sensitive to hue variation. The similarity measure function used in this study is histogram intersection-based; it is given below.

$$Distance(Q, D) = \frac{\sum_{i=1}^{m} w_i D(Q_i, D_i)}{\sum_{i=1}^{m} w_i}$$

Each $D(Q_i, D_i)$ uses histogram intersection to obtain a fractional value between 0 and 1. Before being normalized by the number of angles in the query image, the result of histogram intersection is the number of angles from the database image that have the same corresponding angles in the query image. Therefore, we can think about the spatial color index of an image meaningfully. Any non-zero feature point histogram represents some image objects of a particular color, while any all-zero feature point histogram, called an *empty* histogram, means that there are no image objects of that color. Based on the histogram intersection-based similarity function, the comparison of query and database images using spatial color indices can be taken as a query-by-objects-appearing [29].

5. Experiments

We have implemented the proposed approach and have tested it over two image collections comprising 3388 images. The first database consists of 2008 JPEG images of size 192×128. Most of these images are outdoor scenes containing trees, flowers, mountains, clouds, waters, boats, sculptures, pyramids, bricks, buildings, peoples, and animals. The second database is a collection of 1380 GIF images in a variety of sizes, consisting of various scenes related to earth and space science.

5.1. Efficiency of retrieval

In this study, we use the *efficiency of retrieval* for the formal evaluation of the retrieval effectiveness [14, 16]. The efficiency of retrieval for a given output list of size T is defined as:

$$E_T = \begin{cases} \dfrac{n}{N}, & \text{if } N \leq T \\ \dfrac{n}{T}, & \text{otherwise} \end{cases}$$

where n is the number of similar images retrieved and N is the total number of similar images in the database. We note that if $N \leq T$, then E_T becomes the traditional *recall* value of information retrieval, while if $N > T$, then E_T becomes the traditional *precision* value.

5.2. Comparison with Color-WISE

As mentioned in Section 2, the image retrieval system Color-WISE [22] captures image color distribution information by employing a fixed partitioning technique, with each block indexed by its hue and saturation area-peak values. It is noted that Color-WISE achieves better performance than the use of global and local histograms [22]. Due to the simplicity of implementing the Color-WISE scheme and the acceptability of its overall retrieval performance, we initially compare our approach with it. In the Color-WISE approach using an $8 * 8$ image partitioning scheme, we found that indexing each block by its average hue and saturation values sometimes outperformed the approach of indexing each block by its dominant hue and saturation values. However, using a $16 * 16$ partitioning scheme in combination with dominant hue and saturation values for image indexing always yielded better image retrieval performance than using the average hue and saturation values. With a $16 * 16$ partition, a larger image index is generated with 256 hue and 256 saturation values per image. In addition, for the Color-WISE approach, we had the best results when using 2.5 as the weight for the hue component and 0.5 as the weight for the saturation component.

For our proposed approach, each image is evenly divided into 16 by 16 non-overlapping blocks, the hue component of each pixel is quantized to 24 values, the saturation component of each pixel is quantized to 12 values, and the feature point histogram is computed with a bin size of 10°. Therefore, an image index consists of 24 feature point histograms for the sampled hue constituents, and 12 feature point histograms for the sampled saturation constituents, each feature point histogram consisting of a sequence of 18 integers. We note that the size of our image indices depends only on the quantization of color components, and is independent of the image partitioning scheme. To search for similar images, we always use 1.5 and 1 as the weights for hue-related and saturation-related feature point histograms, respectively.

If any feature point histogram contains all 0's, it can be represented by a single special integer to reduce the image index size. Compared with a generic image index, which contains 648 integers, the actual image indices for the two image databases we use have an average of 325 and 176 integers, respectively, as shown in Table 1. These statistical data confirm that images in the first database have more variation in color than those in the second database. In general, outdoor scene images are quite complex, while images in earth and space science

Table 1. The comparison of original and improved image indices in our two test databases.

Image databases	Hue-related histograms		Saturation-related histograms		Image index size	
	Generic	Actual	Generic	Actual	Generic	Actual
IDB 1	$24 * 18$	182.4	$12 * 18$	142.6	648	325
IDB 2	$24 * 18$	96.8	$12 * 18$	78.7	648	176

usually have clean background plus simple foreground objects such as comets, rockets, and satellites.

Through many sample queries over the two image databases, our comparisons, in terms of the acceptability of the retrieved images, concluded that geometric triangulation-based spatial color indexing brings out more relevant images. This is because this technique allows similar image objects to be translated and rotated within the image. As an example, consider figure 4, where the database query is the pyramid image appearing in the upper-left. Figure 4(a) contains 5 pyramid images, and figure 4(b) contains 6 pyramid images. Moreover, geometric triangulation-based spatial color indexing allows similar image objects to be relatively scale-independent in proportion to their corresponding image sizes. Another example, where the database query is the rocket launch image appearing in the upper-left, is shown in figure 5. It is seen that figure 5(b) contains more relevant images than figure 5(a).

In the first database, there are 17 images containing various pyramids of different scales, positions, directions, shapes and colors. In the second database, there are 84 images of rocket launches of different scales, positions, directions, shapes and colors, even including several gray-level images. For output lists of various sizes T, we show the corresponding comparisons in Tables 2 and 3 for the two image queries previously mentioned.

In this study, a Delaunay triangulation is constructed on the set of feature points labeled with the same hue or saturation values. The corresponding image blocks approximate the spatial layout of meaningful image constituents, as salient image objects are usually large in size and relatively homogeneous in color within an image. Due to the simplicity of the color-based separation of the various image constituents, there are sometimes false positive and missed feature points. However, the nature of our spatial encoding and our geometric triangulation-based feature point histogram representation tends to suppress these from occurring.

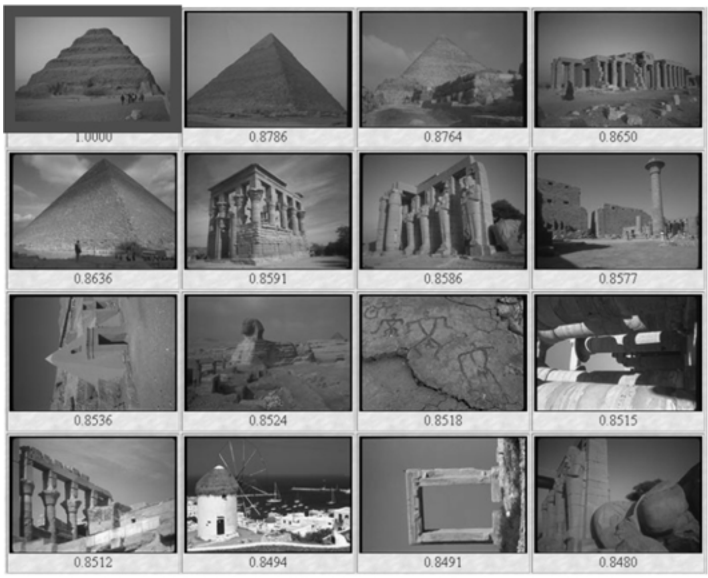
For the earth science and space image database, there are 541 subdirectories in all. All the images stored in any subdirectory are prefixed with their corresponding folder name. There is one subdirectory containing 323 images, while the other subdirectories contain a maximum of 19 images. Images within the same subdirectory are semantically similar, based on their textual annotations, but may not be visually similar. However, based on the assumption that images are relevant to each other if they are stored within the same

Table 2. The comparison of the pyramid image query in terms of the efficiency of retrieval.

	$T = 16$	$T = 32$	$T = 48$	$T = 64$	$T = 80$	$T = 96$	$T = 112$
Color-WISE	31%	29%	35%	35%	35%	47%	53%
Our work	38%	35%	35%	35%	41%	47%	53%

Table 3. The comparison of the rocket launch image query in terms of the efficiency of retrieval.

	$T = 16$	$T = 32$	$T = 48$	$T = 64$	$T = 80$	$T = 96$	$T = 112$
Color-WISE	50%	44%	42%	38%	33%	35%	38%
Our work	88%	78%	73%	64%	58%	57%	60%



(a)

(b)

Figure 4. (a) Color-WISE retrieved images. (b) Color anglogram retrieved images.

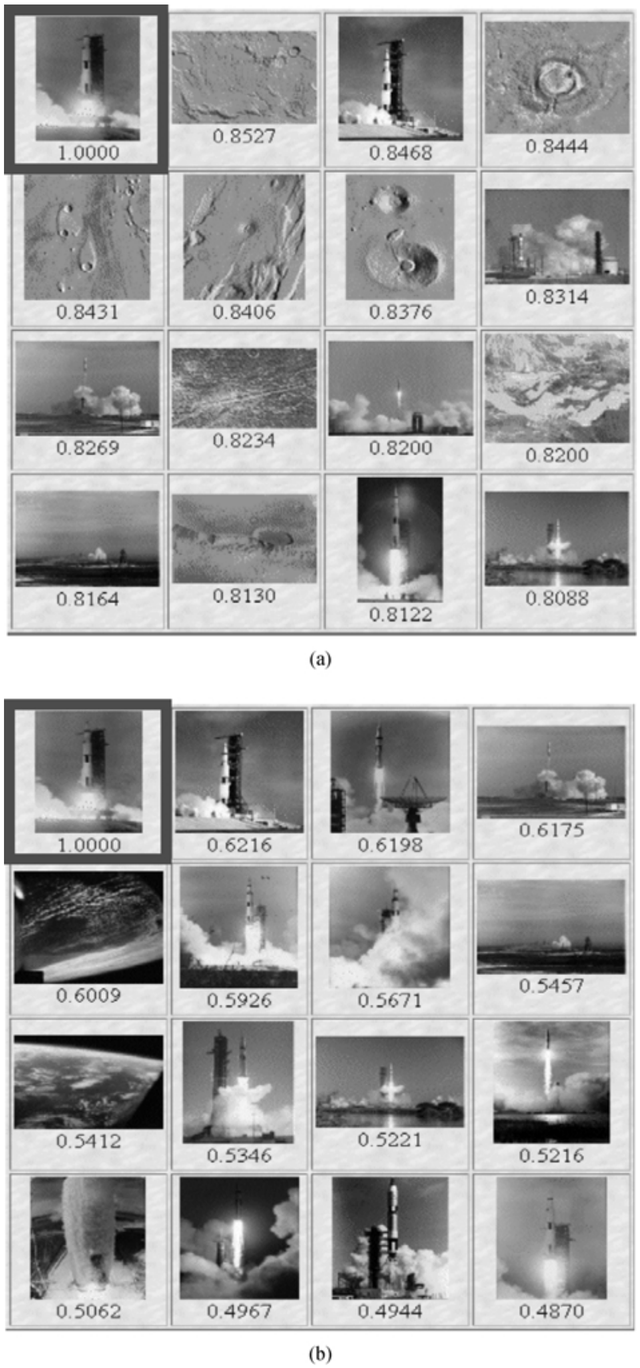


Figure 5. (a) Color-WISE retrieved images. (b) Color anglogram retrieved images.

Table 4. The comparison of four random queries in terms of the efficiency of retrieval.

	$T=4$		$T=8$		$T=12$		$T=16$	
Name of subdirectory	Color WISE	Our work	Color WISE	Our work	Color WISE	Our work	Color WISE	Our work
74s0	50%	50%	50%	100%	50%	100%	50%	100%
dion	33%	67%	33%	67%	33%	67%	33%	67%
huyg	67%	67%	67%	100%	100%	100%	100%	100%
s74e	20%	40%	40%	60%	60%	60%	60%	60%

subdirectory, experimental results show that our spatial color indexing has the ability to rank images in a more flexible way to achieve relatively high retrieval effectiveness. Because of the limited space available, a list of 4 randomly chosen queries with at least two relevant images is shown in Table 4.

5.3. Comparison with the correlogram approach

As mentioned in Section 2, image indexing using color a correlogram characterizes how the spatial correlation of color changes with distance. For simplicity, the color correlogram of an image is a table indexed by color pairs and distance, where the k-th entry for $<i, j>$ specifies the probability of having a pixel of color j at a distance k from a pixel of color i in the image. Given that an image is quantized into m colors with d being the cardinality of a set D of different distances, the size of the correlogram is m^2d. [11] define a color autocorrelogram as a color correlogram in which only the spatial correlation between identical colors is considered. Obviously, a color autocorrelogram is a subset of a color correlogram with a smaller size of md. With image color quantized into 64 colors in RGB space and the distance set $D=\{1, 3, 5, 7\}$, the authors computed the autocorrelogram, histogram, and two refined histograms [20], and then conducted experiments on a database of 14554 color JPEG images. Their experiments showed that the color autocorrelogram outperforms both the traditional histogram method and the histogram refinement methods for content-based image retrieval.

Up to this point, we have constructed a Delaunay triangulation for each set of feature points labeled with the identical color in a point feature map, and then compute the final image index. Similar to the color correlogram approach, we can construct the Delaunay triangulation for a set of feature points labeled with different colors in a point feature map, such that the spatial correlation between more than one color can be encoded, too. As the nature of our computational geometry-based spatial color indexing consists of angle histograms, we call the image index a *color anglogram* (vs. *color correlogram*). Similarly, a color autoanglogram (vs. color autocorrelogram) is a subset of a color anglogram if only the spatial correlation of the identical color is encoded.

With the two image databases described before, we quantized each image into 64 colors ($R=4$, $G=4$, $B=4$) as in [11]. For the color autocorrelogram approach, an image is indexed by 576 integers resulting from the distance set $D = \{1, 3, 5, 7, 9, 11, 13, 15, 17\}$ (i.e., $m = 64$, $d = 9$, and $m * d = 576$). In order to keep the same size of spatial color index, our color autoanglogram is obtained by concatenating all the feature point histograms computed by counting the two smallest angles of each Delaunay triangle with a histogram bin size of 10°. As discussed in [30], the resulting two-smallest-angle histogram has all zero values on those histogram bins equivalent to or greater than 90°. Thus, for the color autoanglogram approach, an image is also indexed by 576 integers resulting from the number of bins $b = 90^\circ/10^\circ$ of each feature point histogram (i.e., $m = 64$, $b = 9$, $m * b = 576$). We note that, as the color autocorrelogram is pixel-based, in this case our color autoanglogram is also computed at the pixel level to characterize finer details of the color distribution. To be specific, each pixel of an image serves as a feature point normalized to be part of the point feature map for further autoanglogram computation.

To search for similar images, we always use equivalent weights for all color-related feature point histograms. Through many sample queries over the two image databases, our comparisons in terms of the acceptability of the retrieved images concluded that image retrieval using color autoanglograms brings out more relevant images in general. Therefore, with same size of image index, the autoanglogram is more effective in encoding the spatial correlation of image color than the autocorrelogram. Although we have not compared these two approaches for encoding the spatial correlation of two or more colors, it is believed that our geometric triangulation-based spatial color indexing scheme will outperform the color correlogram approach under the same conditions.

We show the pyramid image query over the first database in figure 6, and the rocket launch image query over the second database in figure 7. For output lists of various sizes T, we show the corresponding comparisons in Tables 5 and 6, respectively, for these two image queries.

From figures 8–11 is a list of 8 sample query images used in [11]. These 4 pairs of images belong to our first database with the same format and size as the rest of the first database images. We use each of these images as a query over the first database of 2008 images for an additional experiment. Based on the assumption that each query image is relevant only to itself and to its counterpart in the pair, we compare the actual positions of relevant images in the ranking lists resulting from the two indexing schemes (color autocorrelogram vs.

Table 5. The comparison of the pyramid image query in terms of the efficiency of retrieval.

	$T = 16$	$T = 32$	$T = 48$	$T = 64$	$T = 80$	$T = 96$	$T = 112$
Autocorrelogram	13%	12%	18%	24%	29%	35%	35%
Our work	19%	18%	18%	24%	29%	35%	35%

Table 6. The comparison of the rocket launch image query in terms of the efficiency of retrieval.

	$T = 16$	$T = 32$	$T = 48$	$T = 64$	$T = 80$	$T = 96$	$T = 112$
Autocorrelogram	69%	56%	44%	39%	34%	35%	39%
Our work	88%	72%	56%	47%	46%	48%	50%

(a)

(b)

Figure 6. (a) Color correlogram retrieved images. (b) Color anglogram retrieved images.

(a)

(b)

Figure 7. (a) Color correlogram retrieved images. (b) Color anglogram retrieved images.

Table 7. The comparison of actual positions of the counterpart images for each of the 8 queries using autocorrelogram and autoanglogram respectively.

Query Image	Actual position of its counterpart	
	Autocorrelogram	*Autoanglogram*
bat1 (*bat2 as counterpart*)	8	9
bat2 (*bat1 as counterpart*)	58	6
pad1 (*pad2 as counterpart*)	79	16
pad2 (*pad1 as counterpart*)	92	16
rock1 (*rock2 as counterpart*)	13	42
rock2 (*rock1 as counterpart*)	12	6
woman1 (*woman2 as counterpart*)	39	28
woman2 (*woman1 as counterpart*)	12	4

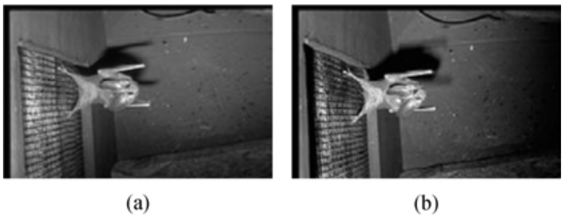


Figure 8. (a) Bat1. (b) Bat2.

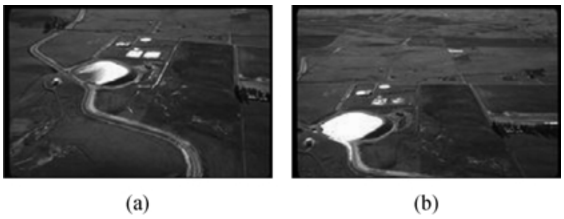
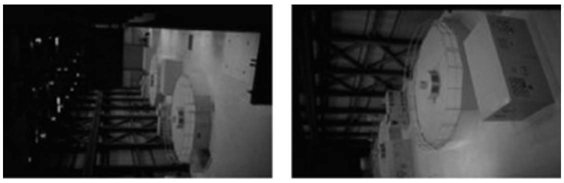


Figure 9. (a) Pad1. (b) Pad2.

color autoanglogram). Because a query image is always top ranked in the two methods, we only show the actual positions of their counterparts in the answer images for comparison in Table 7. For these 8 image queries, the average position of relevant images improves by approximately 23 ranks by using our color anglogram approach. These experimental results indicate that the color autoanglogram usually outperforms the color autocorrelogram.

6. Conclusion

In this work, we examine the use of computational geometry-based spatial color indexing for efficient and effective image retrieval. Our preliminary experiments show that the efficacy



(a) (b)

Figure 10. (a) Rock1. (b) Rock2.

(a) (b)

Figure 11. (a) Woman1. (b) Woman2.

of the proposed approach is quite promising. Based on this current scheme, we are initiating further investigations on several related issues.

We are refining our spatial color indexing approach so as to integrate more features, such as texture and the spatial encoding of image constituents of different hue and saturation values, into our image representation. This integration should only improve our results.

Also, we intend to store separately the indices of salient image constituents so as to support more complicated image queries such as query-by-spatial-objects-appearing and query-by-subimage [29].

Additionally, we also plan to incorporate relevance feedback into our system so as to reduce the gap between low-level visual features and high-level image semantics, and take into account the subjectivity of human perception in image retrieval.

Acknowledgments

We would like to acknowledge Dr. Ishwar Sethi and the members of his group for some stimulating discussions concerning the Color-WISE system. We would also like to acknowledge Dr. Sang Lee for his suggestion to compare our approach with the color correlogram approach. In the first image database, 2000 JPEG images were downloaded from http://darney.usc.edu/~xiawan/demo/cbir.html, and the other 8 JPEG images shown in figures 8–11 were downloaded from http://simon.cs.cornell.edu/Info/People/huang/images.html. The second image database of 1380 GIF images were downloaded from ftp://ftp.eecs.umi ch.edu/groups/ai/dberwick/essbthm.zip. We would like to give special thanks for allowing all these images to be used in our research work.

References

1. I. Ahmad and W.I. Grosky, "Spatial similarity-based retrievals and image indexing by hierarchical decomposition," in Proceedings of the International Database Engineering and Application Symposium (IDEAS'97), Montreal, Canada, August 1997, pp. 269–278.
2. S. Belongie, C. Carson, H. Greenspan, and J. Malik, "Color- and texture-based image segmentation using EM and its application to content-based image retrieval," in Proceedings of the International Conference on Computer Vision (ICCV'98), 1998.
3. S.-K. Chang, Q.-Y. Shi, and C.-W. Yan, "Iconic indexing by 2-D strings," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 9, No. 3, pp. 413–428, 1987.
4. A. Dimai, "Spatial encoding using differences of global features," in Proceedings of SPIE Storage and Retrieval for Image and Video Databases, Vol. 3022, February 1997.
5. R.A. Dwyer, "A Faster divide-and-conquer algorithm for constructing delaunay triangulations," Algorithmic, Vol. 2, No. 2, pp. 127–151, 1987.
6. S. Fortune, "A sweepline algorithm for voronoi diagrams," Algorithmic, Vol. 2, No. 2, pp. 153-174, 1987.
7. W.I. Grosky, "Managing multimedia information in database systems," Communications of the ACM, Vol. 40, No. 12, pp. 73–80, 1997.
8. W.I. Grosky, F. Fotouhi, and Z. Jiang, "Using metadata for the intelligent browsing of structured media objects," in Managing Multimedia Data: Using Metadata to Integrate and Apply Digital Data, A. Sheth and W. Klas (Eds.), McGraw Hill Publishing Company: New York, 1998, pp. 67–92.
9. V.N. Gudivada and V.V. Raghavan, "Design and evaluation of algorithms for image retrievals by spatial similarity," ACM Transactions on Information Systems, Vol. 13, No. 1, pp. 115–144, 1995.
10. W. Hsu, T.S. Chua, and H.K. Pung, "An integrated color-spatial approach to content-based image retrieval," in Proceedings of ACM Multimedia, San Francisco, California, November 1995, pp. 305–313.
11. J. Huang, S.R. Kumar, M. Mitra, W.-J. Zhu, and R. Zabih, "Image indexing using color correlograms," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, San Juan, Puerto Rico, June 1997, pp. 762–768.
12. A.K. Jain and A. Vailaya, "Image retrieval using color and shape," Pattern Recognition, Vol. 29, No. 8, pp. 1233–1244, 1996.
13. C. Jorgensen, "Attributes of images in describing tasks," Information Processing & Management, Vol. 34, No. 2/3, pp. 161–174, 1998.
14. M.S. Kankanhalli, B.M. Mehtre, and J.K. Wu, "Cluster-based color matching for image retrieval," Pattern Recognition, Vol. 29, No. 4, pp. 701–708, 1996.
15. B.M. Mehtre, M.S. Kankanhalli, and W.-F. Lee, "Shape measures for content based image retrieval: A comparison," Information Processing & Management, Vol. 33, No. 3, pp. 319–337, 1997.
16. B.M. Mehtre, M.S. Kankanhalli, and W.-F. Lee, "Content-based image retrieval using a composite color-shape approach," Information Processing & Management, Vol. 34, No. 1, pp. 109–120, 1998.
17. W. Niblack, R. Barder, W. Equitz, M. Flickner, E. Glasman, D. Petkovic, P. Yanker, C. Faloutsos, and G. Yaubin, "The QBIC Project: Querying images by content using color, texture, and shape," in Proceedings of SPIE Storage and Retrieval for Image and Video Databases, Vol. 1908, January 1993, pp. 173–181.
18. B.C. Ooi, K.-L. Tan, T.S. Chua, and W. Hsu, "Fast image retrieval using color-spatial information," The VLDB Journal, No. 7, pp. 115–128, 1997.
19. J. O'Rourke, Computational Geometry in C, Cambridge University Press: Cambridge, England, 1994.
20. G. Pass and R. Zabih, "Histogram refinement for content-based image retrieval," in IEEE Workshop on Applications of Computer Vision, 1996, pp. 96–102.
21. Y. Rui, T.S. Huang, and S.-F. Chang, "Image retrieval: Current techniques, promising directions, and open issues," Journal of Visual Communication and Image Representation, Vol. 10, No. 1, pp. 39–62, 1990.
22. I.K. Sethi, I. Coman, B. Day, F. Jiang, D. Li, J. Segovia-Juarez, G. Wei, and B. You, "Color-WISE: A system for image similarity retrieval using color," in Proceedings of SPIE Storage and Retrieval for Image and Video Databases, Vol. 3312, February 1998, pp. 140–149.
23. J.R. Smith and S.-F. Chang, "Integrated spatial and point feature map query," ACM Multimedia Systems Journal, Vol. 7, No. 2, pp. 129–140, 1999.

24. M. Stricker, "Bounds for the discrimination power of color indexing techniques," in Proceedings of SPIE Storage and Retrieval for Image and Video Databases, Vol. 2185, February 1994, pp. 15–24.
25. M. Stricker and A. Dimai, "Color indexing with weak spatial constraints," in Proceedings of SPIE Storage and Retrieval for Image and Video Databases, Vol. 2670, February 1996, pp. 29–39.
26. M. Stricker and M. Orengo, "Similarity of color images," in Proceedings of SPIE Storage and Retrieval for Image and Video Databases, Vol. 2420, February 1995, pp. 381–392.
27. M.J. Swain and D.H. Ballard, "Color indexing," International Journal of Computer Vision, Vol. 7, No. 1, pp. 11–32, 1991.
28. Y. Tao and W.I. Grosky, "Image matching using the OBIR system with feature point histograms," in Proceedings of 4th IFIP 2.6 Working Conference on Visual Database Systems (VDB4), L'Aquila, Italy, May 1998, pp. 192–197.
29. Y. Tao and W.I. Grosky, "Object-based image retrieval using point feature maps," in Proceedings of the 8th IFIP 2.6 Working Conference on Database Semantics (DS8), Rotorua, New Zealand, January 5–8, 1999, pp. 59–73.
30. Y. Tao and W.I. Grosky, "Delaunay triangulation for image object indexing: A novel method for shape representation," in Proceedings of IS&T/SPIE's Symposium on Storage and Retrieval for Image and Video Databases VII, San Jose, California, January 23–29, 1999, pp. 631–642.
31. X. Wan and C.-J. Kuo, "Color distribution analysis and quantization for image retrieval," in Proceedings of SPIE Storage and Retrieval for Image and Video Databases, Vol. 2670, February 1996, pp. 8–16.
32. I.H. Witten, A. Moffat, and T.C. Bell, Managing Gigabytes, Van Nostrand Reinhold: New York, 1994.

Yi Tao is a software developer at QACenter Products Division, Testing and Implementation Lab, Compuware Corporation, Farmington Hills, MI. Prior to that, he worked as a technical lead at an Internet startup, Logic Solutions, Inc, Ann Arbor, MI. His research interests are Web technologies, multimedia information systems, and automated software testing.

He received his B.S. in 1993 and M.S. in 1995 from Beijing Polytechnic University, and Ph.D. in 1999 from Wayne State University (all in Computer Science). He is a member of IEEE Computer Society.

William I. Grosky is currently Professor and Chair of the Computer and Information Science Department at University of Michigan-Dearborn. Previously to that, he was professor and Acting Chair of the Computer Science Department at Wayne State University in Detroit, Michigan. Before joining Wayne State in 1976, he was an Assistant Professor of Information and Computer Science at Georgia Institute of Technology in Atlanta. His current research interests are in visual information management, web-enabled multimedia information systems, hypermedia, databases, and web technology.

Grosky received his B.S. in mathematics from MIT in 1965, his M.S. in Applied Mathematics from Brown University in 1968, and his Ph.D. from Yale University in 1971. He has given many short courses in the area of database management for local industries and has been invited to lecture on multimedia information systems world-wide. Serving also on many database and multimedia conference program committees, he is currently the Editor-in-Chief of IEEE Multimedia, and on the editorial boards of the Journal of Database Management and Pattern Recognition.