Advances in Visual Information Systems, R. Laurini (edt.), Lecture Notes in Computer Science 1929, 2000 (14-25)

An Image Data Model

William I. Grosky and Peter L. Stanchev

Department of Computer Science, Wayne State University
Detroit, MI 48202
{grosky, stanchev}@cs.wayne.edu

Abstract. In this paper, we analyze the existing approaches to image data modeling and we propose an image data model and a particular image representation in the proposed model. This model establishes a taxonomy based on the systematization on the existing approaches. The image layouts in the model are described in semantic hierarchies. The representation is applicable to a wide variety of image collections. An example for applying the model to a plant picture is given.

1 Introduction

Images are becoming an essential part of the information systems and multimedia applications. The image data model is one of the main issues in the design and development of any image database management system. The data model should be extensible and have the expressive power to present the structure and contents of the image, their objects and the relationships among them. The design of an appropriate image data model will ensure smooth navigation among the images in an image database system. The complexity of the model arises because images are richer in information than text, and because images can be interpreted differently, according to the human perception of the application domain.

There are different models for representing the semantic richness of an image, but most of them are more or less dependent of the image application domain. In this paper we analyze some existing tools and approaches to image data modeling and we propose a new image data model. It can be applied to a wide variety of image collections. The model employs multiple logical representations of an image. The logical image representation can be viewed as a multiple level abstraction of the physical image view. The model is based on the analysis of different image application domains such as: medical images, house furnishing design plans [27], electronic schema catalogues, and geographical information systems [28]. The proposed model could be used as a frame for designing and building a wide range of image database systems and could be proposed as a standard to the MPEG committee. It can be treated as an extension of the general image database model [25, 26]. A particular representation based on this model is also discussed.

2 Image Data

Before we analyze the various existing approaches to image data modeling and the proposed tools, we introduce some of the basic methods using for description of the image and the image contents. The image data can be treated as a physical image representation and their meaning as a logical image representation. The logical representation includes the description of the image, image-objects characteristics, and the relationships among the image objects. In the following sections some of the main techniques for image representation are shown.

2.1 Physical Image Representation

The most common form of the physical image representation is the ***raster form***. It includes the image header and image matrix. The ***image header*** describes the main image parameters such as image format, image resolution, number of bits per pixel, and compression information. The ***image matrix*** contains the image data.

2.2 Logical Image Representation

An ***image object*** is a meaningful portion (consisting of a union of one or more disjoint regions) of an image, which we have called a ***semcon*** [6]. The image description includes meta, semantic, color, texture, shape, and spatial attributes. In the proposed image representation we use the concept of ***object-based point feature maps*** [31]. In 2-D space, many of the image features can be represented as sets 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.***

Meta attributes are attributes related to the process of the image creation. These attributes can be image acquisition date, image identification number and name, image modality device, image magnification, etc.

Semantic attributes contain subjective information about the analyzed image. A specialist in the field of the specific image collection gives the values of such attributes.

Color attributes could be represented as a histogram of intensity of the pixel colors. Based on a fixed partition of the image, an image could be indexed by the color of the whole image and a set of inter-hierarchical distances, which encode the spatial color information. The system Color-WISE is described in [24], and it partitions an image into 8*8 blocks with each block indexed by its dominant hue and saturation values. A histogram refinement technique is described in [23] 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 [12] 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.

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 can represent a color histogram of that region. For the proposed representation, each image object is evenly divided into equal non-overlapping blocks, and the representation captures the spatial distribution of the dominant hue and saturation values for each block. This spatial distribution is captured by our previously devised ***anglogram*** data structure [32]. In more detail, 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 anglogram 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.

Texture attributes. According to Amadasun and King [2]: "Literally, texture refers to the arrangement of the basic constituents of a material. In the digital image, texture is depicted by the spatial interrelationships between, and/or spatial arrangement of the image pixels. Visually, these spatial interrelationships, or arrangement of image pixels, are seen as changes in the intensity patterns, or gray tones". The most used set of texture features is Haralick's gray level co-occurrence features [11]. This is based on the calculation of the co-occurrence matrix $P_{d,\alpha}(g, g')$, which counts the number of pixel pairs in an image that have values g and g' and are separated by a pixel distance *d* in a relative direction α. Using $\alpha = 0^o, 45^o, 90^o$, and 135^o, we get the so-called 4 neighbors of a pixel. Using the co-occurrence matrix, the following coefficients are calculated: angular second momentum, contrast, correlation, sum of squares, inverse difference moment, sum average, sum variance, sum entropy, entropy, difference variance, difference entropy, information measures of correlation, maximum correlation coefficient. Other often-used texture measurements are (1) Tamura features [29]. He suggested six basic textural features; namely, coarseness, contrast, directionality, line-likeness, regularity, and roughness; (2) Unser's sum and difference histogram [33]. He proposed 32 features based on calculations over different sums and histograms of the pixel gray levels; (3) Galloway's run-length based features [5]. He calculated 20 coefficients on the basic on run-length matrixes; (4) Chen's geometric features form binary image sequences [4]. He proposed 16 coefficients, based on threshold images; (5) Laine's texture energy from Daubechies wavelet [15]. He suggested 21 features, based on Daubechies wavelet transformed image; (6) Gabor's filter. Wagner [34] summarized 18 different methods including 318 different features and gave the exact formulas for every single such feature.

For the proposed representation, for each image object, the Tamura features [28] of coarseness, contrast, directionality, line-likeness, and regularity, are calculated as texture object characteristics. These values are normalized.

Shape attributes techniques can be represented in two distinct categories: ***measurement-based methods*** ranging from simple, primitive measures such as ***area*** and

circularity [21] to the more sophisticated measures of various ***moment invariants*** [19]; and ***transformation-based methods*** ranging from functional transformations such as ***Fourier descriptors*** [18] to structural transformations such as ***chain codes*** [17] and ***curvature scale space feature vectors*** [20]. An attempt to compare the various shape representation schemes is made in [18]. Those features, which characterize the shape of any image object, can be classified into the following two categories.

- ***Global shape features*** are general in nature and depend on the characteristics of the entire image object. ***Area, perimeter,*** and ***major axis*** direction of the corresponding image region are examples of such features.
- ***Local shape features*** are based on the low-level characteristics of image objects. The determination of local features usually requires more involved computation. ***Curvatures, boundary segments,*** and ***corner points*** around the boundary of the corresponding image region are examples of such features.

For the proposed representation, an object shape is represented by the spatial distribution of its ***corner points*** [1], again using our ***anglogram*** data structure [30].

Spatial attributes could be presented in different ways: (1) as a ***topological set of relations*** between two image-objects, containing the relations *in, disjoint, touch,* and *cross;* (2) as a ***vector set of relations*** which considers the relevant positions of the image-objects. These include E, S, W, N, SE, SW, NW, NE in terms of the four world directions East, South, West, North; (3) as a ***metric set of relations*** based on the distance between the image-objects, containing the relations *close, far, very close, very far;* (4) 2D-strings [3]. Each image is considered as a matrix of symbols, where each symbol corresponds to an image object. The corresponding 2D-string is obtained by symbolic projection of these symbols along the horizontal and vertical axes, preserving the relative positions of the image objects. In order to improve the performance of this technique, some 2D-string variants have been proposed, such as the extended 2D-string [14], 2D C-string [16], and 2D C^+-string [13]; (5) geometry-based ***θR-string*** approach [8]; (6) the ***spatial orientation graph*** [7], (7) the ***quadtree-based spatial arrangements of feature points*** approach [1].

For the proposed anglogram representation, the global description of the feature maps is obtained by constructing a Delaunay triangulation [22] of all relevant feature points. This feature point histogram is obtained by discretizing the angles produced by this triangulation and counting the number of times each discrete angle occurs in the image object of interest, given the selection criteria of which angles will contribute to the final feature point histogram. For example, the feature point histogram can be built by counting the two largest angles, the two smallest angles, or all three angles of each individual Delaunay triangle. 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 shape and color description of the object-based point feature maps with the help of these techniques are described in [30, 31].

3 Image Data Models

An ***Image Data Model*** is a type of image data abstraction that is used to provide a conceptual image representation. It is a set of concepts that can be used to describe the structure of an image. The process of image description consists of extracting the global image characteristics, recognizing the image-objects and assigning a semantics to these objects. Approaches to image data modeling can be categorized based on the views of image data that the specific model supports.

Some valuable proposals for image data models are: VIMSYS image data model, model, where images are presented as four plane layers [10]; $EMIR^2$- an extended model for image representation and retrieval [18]; and AIR - an adaptive image retrieval model [9].

The ***AIR (Adaptive Image Retrieval) model*** claims that it is the first comprehensive and generic data model for a class of image application areas that coherently integrates logical image representations. It is a semantic data model that facilitates the modeling of an image and the image-objects in the image. It can be divided into three layers: physical level, logical level and semantic or external level representation. There are two kinds of transformations that occur in the model. The first is the transformation from the physical to the logical representation, such as a spatial oriented graph. The second transformation involves the derivation of the semantic attributes from the physical representation.

The ***VIMSYS (Visual Information Management System) model*** views the image information entities in four planes. This model is based on the image characteristics and the inter-relations between those characteristics in an object-oriented design. These planes are the *domain objects and relations* (DO), the *domain events and relations* (DE), the *image objects and relations* (IO) and the *image representations and relations* (IR). An object in this model has a set of attributes and methods associated with them. They are connected in a class attribute hierarchy. The attribute relationships are spatial, functional and semantic. The ***IO plane*** has three basic classes of objects: images, image features and feature organizations. These objects are related to one another through *set-of, generalization (is-a)*, and relations. Image feature is further classified into texture, color, intensity and *feature-of* geometric feature. The ***DO plane*** consists of a semantic levels specification of domain entities, built upon the two previous levels. The ***DE plane*** has been included in the model to accommodate the event definition over image sequences. The ***IR plane*** is clearly functional.

The ***$EMIR^2$ (Extended Model for Image Representation and Retrieval) model*** combines different interpretations of an image in building its description. Each interpretation is presented by a particular view. An image is treated as a multiple-view object and is described by one physical view and four logical views: structural, spatial, perceptive and symbolic. For the description of the view, a context-free grammar formalism is used. The ***structural*** view defines the set of image objects. The ***spatial*** view of an image object is concerned about the shape of the image objects (contour) and their spatial relations (far, near, overlap, etc.), which indicates their relative positions inside the image. The ***perceptive*** view includes all the visual attributes of the image and/or image objects. In this model, these attributes describe color, brightness and texture. The ***symbolic*** view associates a semantic description to an image and/or image object. In this model, two subsets of attributes are used: one associated with the image, e.g. size, date, author, etc., and the other associated with the image objects, e.g. identifier, name, etc.

4 The Image Model

The proposed model establishes a taxonomy based on a systematization of the existing approaches. The main requirement to the proposed model could be summarized as: (1) powerfulness - to be applicable to a wide variation of image collections; (2) efficiency - the obtained image description to be easy used for image retrieval in an image database.

The image is presented at two levels - logical and physical. The logical level contains the global description and content-based sublevels. The global description level consists of the meta and the semantic attributes of the image. The content-based layout contains the object features connected with color, texture, shape, and spatial object characteristics. The physical level contains the image header and the image matrix. A semantic schema of the proposed model is shown in Figure 1.

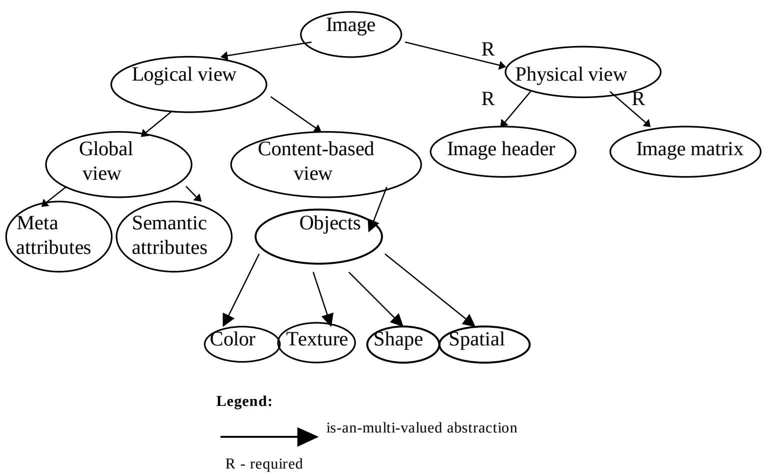


Figure 1. Semantic schema of the model

In the proposed representation the contents based view is given with the help of the Delaunay triangulation. Our experiments confirm that this technique is very useful for color, shape and spatial distribution. Theoretically, from the definition of the Delaunay triangulation, it is easily shown that the angles of the resulting Delaunay triangles of a set of points remain the same under uniform translations, scalings, and rotations of the point set. An example is illustrated in Figure 2 as follows: Figure 2b shows the resulting Delaunay triangulation for a set of 26 points shown in Figure 2a for a airplane image; Figure 2(c) shows the resulting Delaunay triangulation of the transformed (translation, rotation, and scaled-up) set of 26 points in Figure 2a; Figure 2d shows the resulting Delaunay triangulation of the transformed (translation, rotation, and scaled-down) set of 26 points in Figure 2a.

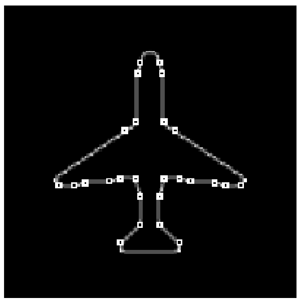


Figure 2a – A Set of 26 points

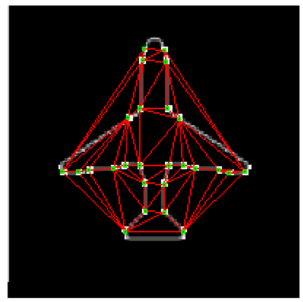
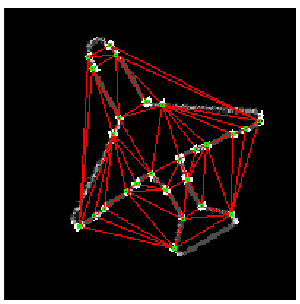


Figure 2a – Delauney Triangulation



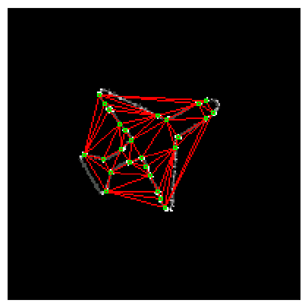
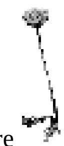
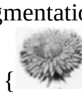


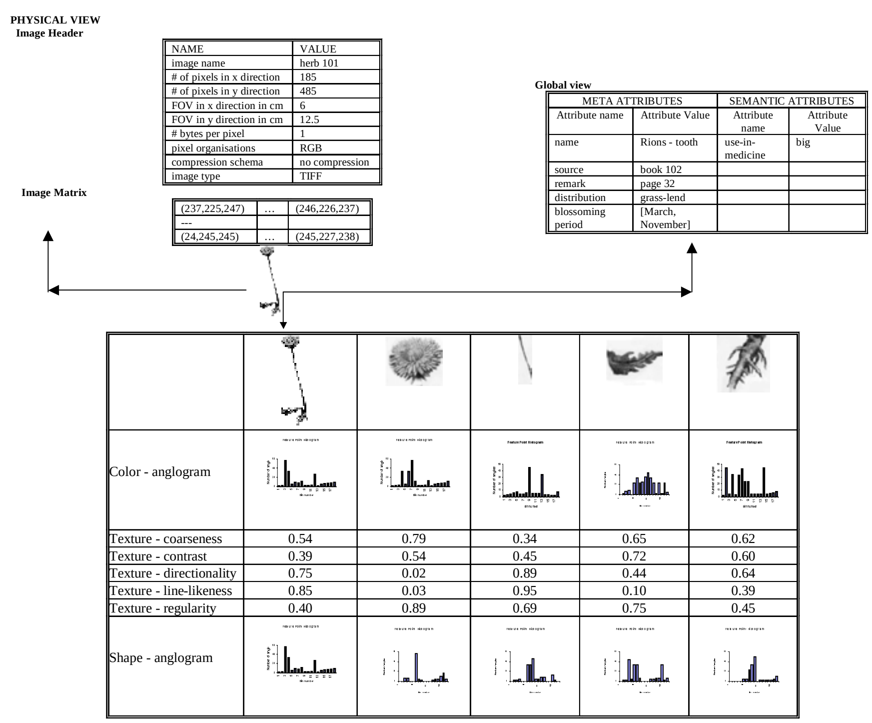
Figure 2c and Figure 2d – Other Delauney Triangulations

4.1 Image Retrieval

The images are searched by their image description representation, and it is based on similarity retrieval. Let a query be converted through the image data model in an image description $Q(q_1, q_2, \ldots, q_n)$ and an image in the image database has the description $I(x_1, x_2, \ldots, x_n)$. Then the retrieval value (RV) between Q and I is defined as: $RV_Q(I) = \Sigma_{i=1,\ldots,n}(w_i * sim(q_i, x_i))$, where w_i ($i = 1,2, \ldots, n$) is the weight specifying the importance of the i^{th} parameter in the image description and $sim(q_i, x_i)$ is the similarity between the i^{th} parameter of the query image and database image and is calculated in different way according to the q_i, x_i values. They can be: *symbol, numerical* or *linguistic* values, *histograms, attribute relational graphs, pictures or spatial representations characters.*

5.Applying the Model Via an Example

Let's consider the used in the following plant picture . After the segmentation procedure the image is partitioned in the following image-objects: blossom {}, stalk {}, leaf {}, and root {}. The model manipulation capabilities are realized in the Detroit Image Database Management System. A possible view as a result of applying the suggested representation to the example image is given in Figure 3. A possible definition of the proposed representation in the IDBMS Detroit is given in Figure 4.



PHYSICAL VIEW
Image Header

NAME	VALUE
image name	herb 101
# of pixels in x direction	185
# of pixels in y direction	485
FOV in x direction in cm	6
FOV in y direction in cm	12.5
# bytes per pixel	1
pixel organisations	RGB
compression schema	no compression
image type	TIFF

Image Matrix

(237,225,247)	…	(246,226,237)
...		
(24,245,245)	…	(245,227,238)

Global view

META ATTRIBUTES		SEMANTIC ATTRIBUTES	
Attribute name	Attribute Value	Attribute name	Attribute Value
name	Rions - tooth	use-in-medicine	big
source	book 102		
remark	page 32		
distribution	grass-lend		
blossoming period	[March, November]		

Texture - coarseness	0.54	0.79	0.34	0.65	0.62
Texture - contrast	0.39	0.54	0.45	0.72	0.60
Texture - directionality	0.75	0.02	0.89	0.44	0.64
Texture - line-likeness	0.85	0.03	0.95	0.10	0.39
Texture - regularity	0.40	0.89	0.69	0.75	0.45

Figure 3. Image representation through the proposed model

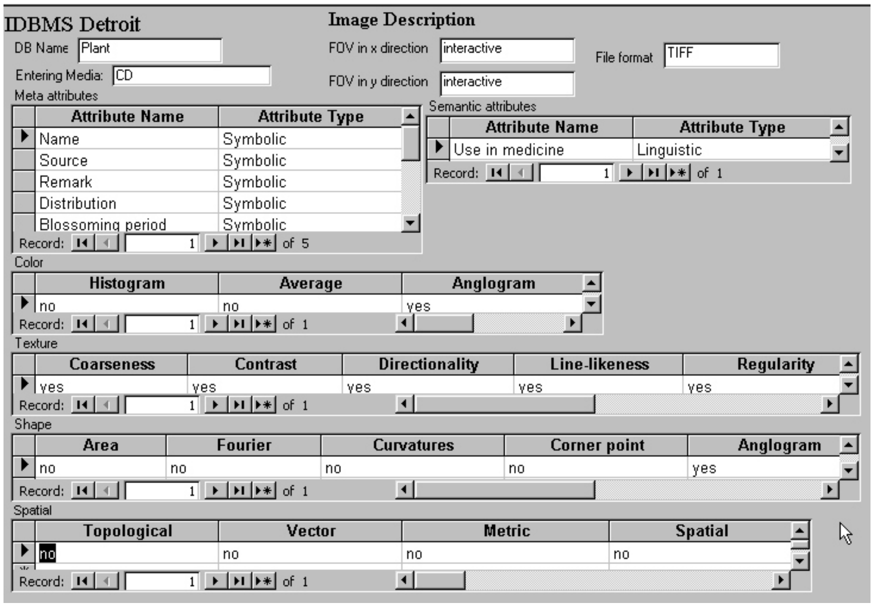


Figure 4. The application domain definition in the IDBMS Detroit

6 Conclusions

The main advantages of the proposed model and particular image representation could be summarized as follows:

- Generality. The model uses the main techniques from the existing image data models and it is applicable to a wide variety of image collections;
- Practical applicability. The model can be used as a part of image retrieval and image database system;
- Flexibility. The model could be customized when used with a specific application;
- Robustness of the proposed representation. The chosen methods for image description allow similar descriptions again minor variation on the image.

The proposed image representation is very easy to be implemented and is very convenient for image retrieval in an image database system.

References

1. Ahmad, I., and Grosky, W., ‘Spatial Similarity-based Retrievals and Image Indexing By Hierarchical Decomposition,’ *Proceedings of the International Database Engineering and Application Symposium (IDEAS’97)*, Montreal, Canada, (August 1997), pp. 269-278.
2. Amadasun, M., and King, R., ‘Textural Features Corresponding to Textural Properties,’ *IEEE Transactions on Systems, Man, and Cybernetics*, 19 (1989), pp. 1264-1274.
3. Chang, S., Shi Q., and Yan, C., ‘Iconic Indexing by 2-D Strings,’ *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume 9, Number 3 (May 1987), pp. 413-428.
4. Chen, Y., M. Nixon, and Thomas, D., ‘Statistical Geometrical Features for Texture Classification,’ *Pattern Recognition*, 28 (1995), pp. 537-552.
5. Galloway, M., ‘Texture Analysis Using Gray Level Run Lengths,’ *Computer Graphics and Image Processing*, 4 (1975), pp. 172-179.
6. Grosky, W., Fotouhi, F., and Jiang, Z., Using Metadata for the Intelligent Browsing of Structured Media Objects, *Managing Multimedia Data: Using Metadata to Integrate and Apply Digital Data,* Sheth A., and Klas W., (Eds.), McGraw Hill Publishing Company, New York, (1997), pp. 67-92.
7. Gudivada V. and Raghavan, V., ‘Design and Evaluation of Algorithms for Image Retrievals By Spatial Similarity,’ *ACM Transactions on Information Systems*, Volume 13, Number 1 (January 1995), pp. 115-144.
8. Gudivada, V., ‘θR-String: A Geometry-Based Representation for Efficient and Effective Retrieval of Images By Spatial Similarity,’ *IEEE Transactions on Knowledge and Data Engineering*, Vol. 10, No. 3, (May/June 1998), pp. 504-512.
9. Gudivada, V., Raghavan, V., and Vanapipat, K., ‘A Unified Approach to Data Modeling and Retrieval for a Class of Image Database Applications,’ *IEEE Transactions on Data and Knowledge Engineering,* (1994).
10. Gupta, A., Weymouth, T., and Jain, R., ‘Semantic Queries with Pictures: The VIMSYS Model,’ Proceedings of the 17^{th} Conference on Very Large Databases, Palo Alto, California (1991), pp. 69-79.
11. Haralick, R., Shanmugam K., and I. Dinstein, H., ‘Texture Features for Image Classification,’ *IEEE Transactions on Systems, Man, and Cybernetics,* SMC-3 (1973).
12. Huang J., Kumar, S., Mitra, M., Zhu, W., and Zabih, R., ‘Image Indexing Using Color Correlograms,’ *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, San Juan, Puerto Rico, (June 1997), pp. 762-768.
13. Huang, P. and Jean, Y., ‘Using 2D C^+-Strings as Spatial Knowledge Representation for Image Database Systems,’ *Pattern Recognition*, Volume 27, Number 9 (September 1994), pp. 1249-1257.
14. Jungert E. and Chang, S., ‘An Algebra for Symbolic Image Manipulation and Transformation,’ *Proceedings of the IFIP TC 2/WG 2.6 Working Conference on*

Visual Database Systems, Elsevier Science Publishing Company, Amsterdam, The Netherlands (1989), pp. 301-317.
15. Laine A. and Fan, J., 'Texture Classification by Wavelet Packet Signatures,' *IEEE Transactions on Pattern Recognition and Machine Intelligence*, 15 (1993).
16. Lee, S. and Hsu, F., '2D C-String: A New Spatial Knowledge Representation for Image Database System,' *Pattern Recognition,* Volume 23, Number 10 (October 1990), pp. 1077-1087.
17. Lu, G., 'An Approach to Image Retrieval Based on Shape,' *Journal of Information Science*, Volume 23, Number 2 (1997), pp. 119-127.
18. Mechkour, M., 'EMIR^2. An Extended Model for Image Representation and Retrieval,' in Revell, N. and Tjoa, A. (Eds.), *Database and Expert Systems Applications,* Berlin, (Springer -Verlag, 1995), pp. 395-414.
19. Mehtre, B., Kankanhalli, M., and Lee, W., 'Shape Measures for Content Based Image Retrieval: A Comparison,' *Information Processing & Management,* Volume 33, Number 3 (June 1997), pp. 319-337.
20. Mokhtarian, F., Abbasi, S., and Kitter, J., 'Efficient and Robust Retrieval by Shape Content through Curvature Scale Space,' *Proceedings of the International Workshop on Image Database and Multimedia Search,* Amsterdam, The Netherlands, (August 1996), pp. 35-42.
21. Niblack, W., Barder, R., Equitz, W., Flickner, M., Glasman, E., Petkovic, D., Yanker, P., Faloutsos, C., and Yaubin, G., 'The QBIC Project: Querying Images by Content Using Color, Texture, and Shape,' *Proceedings of SPIE Storage and Retrieval for Image and Video Databases*, Volume 1908, (January 1993), pp. 173-181.
22. O'Rourke, J., *Computational Geometry in C,* Cambridge University Press, Cambridge, England, 1994.
23. Pass, G. and Zabih, R., 'Histogram Refinement for Content-Based Image Retrieval,' *IEEE Workshop on Applications of Computer Vision*, (1996), pp. 96-102.
24. Sethi, I., Coman, I., Day, B., Jiang, F., Li, D., Segovia-Juarez, J., Wei, G., and You, B., 'Color-WISE: A System for Image Similarity Retrieval Using Color,' *Proceedings of SPIE Storage and Retrieval for Image and Video Databases*, Volume 3312, (February 1998), pp. 140-149.
25. Stanchev, P., 'General Image Database Model,' *Visual Information and Information Systems, Proceedings of the Third Conference on Visual Information Systems,* Huijsmans, D. Smeulders A., (Eds.) Lecture Notes in Computer Science, Volume 1614 (1999), pp. 29-36.
26. Stanchev, P., 'General Image Retrieval Model,' *Proceedings of the 27th Spring Conference of the Union of the Bulgarian Mathematicians,* Pleven, Bulgaria, 1998, pp. 63-71.
27. Stanchev, P., and Rabitti, F., GRIM_DBMS: a GRaphical IMage DataBase Management System,' in T. Kunii (Ed.), Visual Database Systems, (North-Holland, 1989), pp. 415-430.

28. Stanchev, P., Smeulders, A., and Groen, F., 'An Approach to Image Indexing of Document,' in E. Knuth and L. Wegner (Eds.), Visual Database Systems II, (North Holland, 1992), pp. 63-77.
29. Tamura, H., Mori, S., and Yamawaki,, T., 'Textural Features Corresponding to Visual Perception,' *IEEE Transaction on Systems, Man, and Cybernetics*, SMC-8 (1978), pp. 460-472.
30. Tao Y. and Grosky, W., 'Image Indexing and Retrieval Using Object-Based Point Feature Maps,' Journal of Visual Languages and Computing, To Appear
31. Tao Y. and Grosky, W., 'Shape Anglograms for Image Object Representation: A Computational Geometry. Submit.
32. Tao Y. and Grosky, W., 'Spatial Color Indexing using Rotation, Translation, and Scale Invariant Anglograms,' *Multimedia Tools and Applications*, To Appear.
33. Unser, M., 'Sum and Difference Histograms for Texture Classification,' *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8 (1986), pp. 118-125.
34. Wagner T., 'Texture Analysis,' in Jahne, B., Haussecker, H., and Geisser P., (Eds.), *Handbook of Computer Vision and Application*, Academic Press, San Diego, (1999), pp. 275-308.