Journal of Visual Communication and Image Representation **10**, 39–62 (1999)
Article ID jvci.1999.0413, available online at http://www.idealibrary.com on IDEAL®

Image Retrieval: Current Techniques, Promising Directions, and Open Issues

Yong Rui and Thomas S. Huang

Department of ECE & Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801
E-mail: yrui@ifp.uiuc.edu, huang@ifp.uiuc.edu

and

Shih-Fu Chang

Department of EE & New Media Technology Center, Columbia University, New York, New York 10027
E-mail: sfchang@ee.columbia.edu

Received January 29, 1998; accepted January 7, 1999

This paper provides a comprehensive survey of the technical achievements in the research area of image retrieval, especially content-based image retrieval, an area that has been so active and prosperous in the past few years. The survey includes 100+ papers covering the research aspects of image feature representation and extraction, multidimensional indexing, and system design, three of the fundamental bases of content-based image retrieval. Furthermore, based on the state-of-the-art technology available now and the demand from real-world applications, open research issues are identified and future promising research directions are suggested. © 1999 Academic Press

1. INTRODUCTION

Recent years have seen a rapid increase in the size of digital image collections. Everyday, both military and civilian equipment generates giga-bytes of images. A huge amount of information is out there. However, we cannot access or make use of the information unless it is organized so as to allow efficient browsing, searching, and retrieval. Image retrieval has been a very active research area since the 1970s, with the thrust from two major research communities, database management and computer vision. These two research communities study image retrieval from different angles, one being text-based and the other visual-based.

The text-based image retrieval can be traced back to the late 1970s. A very popular framework of image retrieval then was to first annotate the images by text and then use text-based

Work of the first author was supported in part by a CSE Fellowship, College of Engineering, UIUC; work of the second author was supported in part by ARL Cooperative Agreement DAAL01-96-2-0003; and work of the third author was supported in part by the National Science Foundation under a CAREER award (IRI-9501266), a STIMULATE award (IRI-96-19124), and industrial sponsors of Columbia's ADVENT project.

1047-3203/99 $30.00

database management systems (DBMS) to perform image retrieval. Representatives of this approach are [20, 21, 24, 26]. Two comprehensive surveys on this topic are [157, 25]. Many advances, such as data modeling, multidimensional indexing, and query evaluation, have been made along this research direction. However, there exist two major difficulties, especially when the size of image collections is large (tens or hundreds of thousands). One is the vast amount of labor required in manual image annotation. The other difficulty, which is more essential, results from the rich content in the images and the subjectivity of human perception. That is, for the same image content different people may perceive it differently. The perception subjectivity and annotation impreciseness may cause unrecoverable mismatches in later retrieval processes.

In the early 1990s, because of the emergence of large-scale image collections, the two difficulties faced by the manual annotation approach became more and more acute. To overcome these difficulties, content-based image retrieval was proposed. That is, instead of being manually annotated by text-based key words, images would be indexed by their own visual content, such as color and texture. Since then, many techniques in this research direction have been developed and many image retrieval systems, both research and commercial, have been built. The advances in this research direction are mainly contributed by the computer vision community. Many special issues of leading journals have been dedicated to this topic [55, 105, 97, 4, 130].

This approach has established a general framework of image retrieval from a new perspective. However, there are still many open research issues to be solved before such retrieval systems can be put into practice. Regarding content-based image retrieval, we feel there is a need to survey what has been achieved in the past few years and what are the potential research directions which can lead to compelling applications.

Since excellent surveys for text-based image retrieval paradigms already exist [157, 25], in this paper we will devote our effort primarily to the content-based image retrieval paradigm. There are three fundamental bases for content-based image retrieval, i.e. visual feature extraction, multidimensional indexing, and retrieval system design. The remainder of this paper is organized as follows. In Section 2, we review various visual features and their corresponding representation and matching techniques. To facilitate fast search in large-scale image collections, effective indexing techniques need to be explored. Section 3 evaluates various such techniques, including dimension reduction and multidimensional indexing. State-of-the-art commercial and research systems and their distinct characteristics are described in Section 4. Based on the current situation and what is demanded from real-world applications, promising future research directions and suggested approaches are presented in Section 5. Section 6 gives some concluding remarks.

2. FEATURE EXTRACTION

Feature (content) extraction is the basis of content-based image retrieval. In a broad sense, features may include both text-based features (key words, annotations) and visual features (color, texture, shape, faces). However, since there already exists rich literature on text-based feature extraction in the DBMS and information retrieval research communities, we will confine ourselves to the techniques of visual feature extraction. Within the visual feature scope, the features can be further classified as general features and domain-specific features. The former include color, texture, and shape features while the latter is application-dependent and may include, for example, human faces and finger prints. The domain-specific features are better covered in pattern recognition literature and may involve

much domain knowledge which we will not have enough space to cover in this paper. Therefore, the remainder of the section will concentrate on those general features which can be used in most applications.

Because of perception subjectivity, there does not exist a single best presentation for a given feature. As we will see soon, for any given feature there exist multiple representations which characterize the feature from different perspectives.

2.1. Color

The color feature is one of the most widely used visual features in image retrieval. It is relatively robust to background complication and independent of image size and orientation. Some representative studies of color perception and color spaces can be found in [90, 95, 163].

In image retrieval, the color histogram is the most commonly used color feature representation. Statistically, it denotes the joint probability of the intensities of the three color channels. Swain and Ballard proposed histogram intersection, an L_1 metric, as the similarity measure for the color histogram [152]. To take into account the similarities between similar but not identical colors, Ioka [68] and Niblack *et al.* [99] introduced an L_2-related metric in comparing the histograms. Furthermore, considering that most color histograms are very sparse and thus sensitive to noise, Stricker and Orengo proposed using the cumulated color histogram. Their research results demonstrated the advantages of the proposed approach over the conventional color histogram approach [151].

Besides the color histogram, several other color feature representations have been applied in image retrieval, including color moments and color sets. To overcome the quantization effects, as in the color histogram, Stricker and Orengo proposed using the color moments approach [151]. The mathematical foundation of this approach is that any color distribution can be characterized by its moments. Furthermore, since most of the information is concentrated on the low-order moments, only the first moment (mean), and the second and third central moments (variance and skewness) were extracted as the color feature representation. Weighted Euclidean distance was used to calculate the color similarity.

To facilitate fast search over large-scale image collections, Smith and Chang proposed color sets as an approximation to the color histogram [139, 140]. They first transformed the (R, G, B) color space into a perceptually uniform space, such as HSV, and then quantized the transformed color space into M bins. A color set is defined as a selection of colors from the quantized color space. Because color set feature vectors were binary, a binary search tree was constructed to allow a fast search. The relationship between the proposed color sets and the conventional color histogram was further discussed [139, 140].

2.2. Texture

Texture refers to the visual patterns that have properties of homogeneity that do not result from the presence of only a single color or intensity [141]. It is an innate property of virtually all surfaces, including clouds, trees, bricks, hair, and fabric. It contains important information about the structural arrangement of surfaces and their relationship to the surrounding environment [59]. Because of its importance and usefulness in pattern recognition and computer vision, there are rich research results from the past three decades. Now, it further finds its way into image retrieval. More and more research achievements are being added to it.

In the early 1970s, Haralick *et al.* proposed the co-occurrence matrix representation of texture features [59]. This approach explored the gray level spatial dependence of texture. It first constructed a co-occurrence matrix based on the orientation and distance between image pixels and then extracted meaningful statistics from the matrix as the texture representation. Many other researchers followed the same line and further proposed enhanced versions. For example, Gotlieb and Kreyszig studied the statistics originally proposed in [59] and experimentally found out that *contrast*, *inverse deference moment*, and *entropy* had the biggest discriminatory power [52].

Motivated by the psychological studies in human visual perception of texture, Tamura *et al.* explored the texture representation from a different angle [156]. They developed computational approximations to the visual texture properties found to be important in psychology studies. The six visual texture properties were *coarseness*, *contrast*, *directionality*, *linelikeness*, *regularity*, and *roughness*. One major distinction between the Tamura texture representation and the co-occurrence matrix representation is that all the texture properties in Tamura representation are visually meaningful, whereas some of the texture properties used in co-occurrence matrix representation may not be (for example, entropy). This characteristic makes the Tamura texture representation very attractive in image retrieval, as it can provide a more user-friendly interface. The QBIC system [44] and the MARS system [64, 102] further improved this texture representation.

In the early 1990s, after the wavelet transform was introduced and its theoretical framework was established, many researchers began to study the use of the wavelet transform in texture representation [138, 31, 73, 54, 72, 159]. In [138, 141], Smith and Chang used the statistics (mean and variance) extracted from the wavelet subbands as the texture representation. This approach achieved over 90% accuracy on the 112 Brodatz texture images. To explore the middle-band characteristics, a tree-structured wavelet transform was used by Chang and Kuo in [31] to further improve the classification accuracy. The wavelet transform was also combined with other techniques to achieve better performance. Gross *et al.* used the wavelet transform, together with KL expansion and Kohonen maps, to perform texture analysis in [54]. Thyagarajan *et al.* [159] and Kundu *et al.* [72] combined the wavelet transform with a co-occurrence matrix to take advantage of both statistics-based and transform-based texture analyses.

There also were quite a few review papers in this area. An early review paper, by Weszka *et al.*, compared the texture classification performance of Fourier power spectrum, second-order gray level statistics (co-occurrence matrix), and first-order statistics of gray level differences [165]. They tested the three methods on two sets of terrain samples and concluded that the Fourier method performed poorly while the other two were comparable. In [100], Ohanian and Dubes compared and evaluated four types of texture representations, namely Markov random field representation [40], multichannel filtering representation, fractal-based representation [107], and co-occurrence representation. They tested the four texture representations on four test sets, with two being synthetic (fractal and Gaussian Markov random field) and two being natural (leather and painted surfaces). They found that co-occurrence matrix representation performed best in their test sets. In a more recent paper [82], Ma and Manjunath evaluated the texture image annotation by various wavelet transform representations, including orthogonal and bi-orthogonal wavelet transforms, the tree-structured wavelet transform, and the Gabor wavelet transform. They found that the Gabor transform was the best among the tested candidates which matched human vision study results [141].

2.3. Shape

In image retrieval, depending on the applications, some require the shape representation to be invariant to translation, rotation, and scaling, while others do not. Obviously, if a representation satisfies the former requirement, it will satisfy the latter as well. Therefore, in the following we will focus on shape representations that are transformation invariant.

In general, the shape representations can be divided into two categories, boundary-based and region-based. The former uses only the outer boundary of the shape while the latter uses the entire shape region [125]. The most successful representatives for these two categories are Fourier descriptor and moment invariants.

The main idea of a Fourier descriptor is to use the Fourier transformed boundary as the shape feature. Some early work can be found in [170, 108]. To take into account the digitization noise in the image domain, Rui *et al.* proposed a modified Fourier descriptor which is both robust to noise and invariant to geometric transformations [125].

The main idea of moment invariants is to use region-based moments which are invariant to transformations, as the shape feature. In [62], Hu identified seven such moments. Based on his work, many improved versions emerged. In [169], based on the discrete version of Green's theorem, Yang and Albregtsen proposed a fast method of computing moments in binary images. Motivated by the fact that most useful invariants were found by extensive experience and trial-and-error, Kapur *et al.* developed algorithms to systematically generate and search for a given geometry's invariants [71]. Realizing that most researchers did not consider what happened to the invariants after image digitization, Gross and Latecki developed an approach which preserved the qualitative differential geometry of the object boundary, even after an image was digitized [71]. In [36, 75], a framework of algebraic curves and invariants is proposed to represent complex objects in a cluttered scene by parts or patches. Polynomial fitting is done to represent local geometric information, from which geometric invariants are used in object matching and recognition.

Some recent work in shape representation and matching includes the finite element method (FEM) [106], the turning function [9], and the wavelet descriptor [35]. The FEM defines a stiffness matrix which describes how each point on the object is connected to the other points. The eigenvectors of the stiffness matrix are called modes and span a feature space. All the shapes are first mapped into this space and similarity is then computed based on the eigenvalues. Along a similar line of the Fourier descriptor, Arkin *et al.* developed a *turning function*-based method for comparing both convex and concave polygons [9]. In [35], Chuang and Kuo used the wavelet transform to describe object shape. It embraced the desirable properties such as multiresolution representation, invariance, uniqueness, stability, and spatial localization. For shape matching, chamfer matching attracted much research attention. Barrow *et al.* first proposed the chamfer matching technique which compared two collections of shape fragments at a cost proportional to the linear dimension, rather than area [12]. In [15], to further speed up the chamfer matching process, Borgerfos proposed a hierarchical chamfer matching algorithm. The matching was done at different resolutions, from coarse to fine.

Some recent review papers in shape representations are [77, 93]. In [77], Li and Ma showed that the geometric moments method (region-based) and the Fourier descriptor (boundary-based) were related by a simple linear transformation. In [93], Babu *et al.* compared the performance of boundary-based representations (chain code, Fourier descriptor, UNL Fourier descriptor), region-based representations (moment invariants, Zernike

moments, pseudo-Zernike moments), and combined representations (moment invariants and Fourier descriptor, moment invariants and UNL Fourier descriptor). Their experiments showed that the combined representations outperformed the simple representations.

In addition to 2D shape representations, there were many methods developed for 3D shape representations. In [161], Wallace and Wintz presented a technique for normalizing Fourier descriptors which retained all shape information and was computationally efficient. They also took advantage of an interpolation property of Fourier descriptor which resulted in efficient representation of 3D shapes. In [160], Wallace and Mitchell proposed using a hybrid structural/statistical local shape analysis algorithm for 3D shape representation. Further, Taubin proposed using a set of algebraic moment invariants to represent both 2D and 3D shapes [158], which greatly reduced the computation required for shape matching.

2.4. Color Layout

Although the global color feature is simple to calculate and can provide reasonable discriminating power in image retrieval, it tends to give too many false positives when the image collection is large. Many research results suggested that using color layout (both color feature and spatial relations) is a better solution to image retrieval. To extend the global color feature to a local one, a natural approach is to divide the whole image into subblocks and extract color features from each of the subblocks [46, 34]. A variation of this approach is the quadtree-based color layout approach [80], where the entire image was split into a quadtree structure and each tree branch had its own histogram to describe its color content. Although conceptually simple, this regular subblock-based approach cannot provide accurate local color information and is computation- and storage-expensive. A more sophisticated approach is to segment the image into regions with salient color features by color set back-projection and then to store the position and color set feature of each region to support later queries [139]. The advantage of this approach is its accuracy while the disadvantage is the general difficult problem of reliable image segmentation.

To achieve a good trade-off between the above two approaches, several other color layout representations were proposed. In [116], Rickman and Stonham proposed a color tuple histogram approach. They first constructed a code book which described every possible combination of coarsely quantized color hues that might be encountered within local regions in an image. Then a histogram based on quantized hues was constructed as the local color feature. In [150], Stricker and Dimai extracted the first three color moments from five predefined partially overlapping fuzzy regions. The usage of the overlapping region made their approach relatively insensitive to small region transformations. In [104], Pass *et al.* classified each pixel of a particular color as either coherent or incoherent, based on whether or not it is part of a large similarly colored region. By using this approach, widely scattered pixels were distinguished from clustered pixels, thus improving the representation of local color features. In [63], Huang *et al.* proposed a color correlogram-based color layout representation. They first constructed a color co-occurrence matrix and then used the auto-correlogram and correlogram as the similarity measures. Their experimental results showed that this approach was more robust than the conventional color histogram approach in terms of retrieval accuracy [63].

Along the same line of the color layout feature, the layout of texture and other visual features can also be constructed to facilitate more advanced image retrieval.

2.5. Segmentation

Segmentation is very important to image retrieval. Both the shape feature and the layout feature depend on good segmentation. In this subsection we will describe some existing segmentation techniques used in both computer vision and image retrieval.

In [81], Lybanon *et al.* researched a morphological operation (opening and closing) approach in image segmentation. They tested their approach in various types of images, including optical astronomical images, infrared ocean images, and magnetograms. While this approach was effective in dealing with the above scientific image types, its performance needs to be further evaluated for more complex natural scene images. In [58], Hansen and Higgins exploited the individual strengths of watershed analysis and relaxation labeling. Since fast algorithm exists for the watershed method, they first used the watershed to subdivide an image into catchmen basins. They then used relaxation labeling to refine and update the classification of catchment basins initially obtained from the watershed to take advantage of the relaxation labeling's robustness to noise. In [78], Li *et al.* proposed a fuzzy entropy-based segmentation approach. This approach is based on the fact that local entropy maxima correspond to the uncertainties among various regions in the image. This approach was very effective for images whose histograms do not have clear peaks and valleys. Other segmentation techniques based on Delaunay triangulation, fractals, and edge flow can be found in [50, 134, 85].

All the above-mentioned algorithms are automatic. A major advantage of this type of segmentation algorithms is that it can extract boundaries from a large number of images without occupying human time and effort. However, in an unconstrained domain, for nonpreconditioned images, the automatic segmentation is not always reliable. What an algorithm can segment in this case is only regions, but not objects. To obtain high-level objects, which is desirable in image retrieval, human assistance is needed.

In [128], Samadani and Han proposed a computer-assisted boundary extraction approach, which combined manual inputs from the user with the image edges generated by the computer. In [41], Daneels *et al.* developed an improved method of active contours. Based on the user's input, the algorithm first used a greedy procedure to provide fast initial convergence. Second, the outline was refined by using dynamic programming. In [124], Rui *et al.* proposed a segmentation algorithm based on clustering and grouping in spatial–color–texture space. The user defines where the attractor (object of interest) is, and the algorithm groups regions into meaningful objects.

One last comment worth mentioning in segmentation is that the requirements of segmentation accuracy are quite different for shape features and layout features. For the former, accurate segmentation is highly desirable while for the latter, a coarse segmentation may suffice.

2.6. Summary

As we can see from the above descriptions, many visual features have been explored, both previously in computer vision applications and currently in image retrieval applications. For each visual feature, there exist multiple representations which model the human perception of that feature from different perspectives.

What features and representations should be used in image retrieval is application dependent. There is a need of developing an image content description (model) to organize the features. The features should not only be just associated with the images, but also they

should be invoked at the right place and the right time, whenever they are needed to assist retrieval. Such research effort has taken place. MPEG has started a new work item called MPEG-7, whose formal name is multimedia content description interface [2, 1, 5, 120]. It will specify a standard set of descriptors (feature representations) that can be used to describe various types of multimedia information. The descriptions shall be associated with the content itself, to allow fast and efficient searching for information of a user's need [2].

3. HIGH DIMENSIONAL INDEXING

To make the content-based image retrieval truly scalable to large size image collections, efficient multidimensional indexing techniques need to be explored. There are two main challenges in such an exploration for image retrieval:

• *High dimensionality.* The dimensionality of the feature vectors is normally of the order of 10^2.

• *Non-Euclidean similarity measure.* Since Euclidean measure may not effectively simulate human perception of a certain visual content, various other similarity measures, such as histogram intersection, cosine, correlation, need to be supported.

Towards solving these problems, one promising approach is to first perform dimension reduction and then to use appropriate multidimensional indexing techniques, which are capable of supporting non-Euclidean similarity measures.

3.1. Dimension Reduction

Even though the dimension of the feature vectors in image retrieval is normally very high, the *embedded dimension* is much lower [166, 167]. Before we utilize any indexing technique, it is beneficial to first perform dimension reduction. At least two approaches have appeared in the literature, i.e. Karhunen–Loeve transform (KLT) and column-wise clustering.

KLT and its variation in face recognition, eigenimage, and its variation in information analysis, principal component analysis (PCA), have been studied by researchers in performing dimension reduction. In [98], Ng and Sedighian followed the eigenimage approach to carry out the dimension reduction, and in [47] Faloutsos and Lin proposed a fast approximation to KLT to perform the dimension reduction. Experimental results from their research showed that most real data sets (visual feature vectors) can be considerably reduced in dimension without significant degradation in retrieval quality [98, 47, 167]. Recently, Chandrasekaran *et al.* developed a low-rank singular value decomposition (SVD) update algorithm which was efficient and numerically stable in performing KLT [19]. Considering that the image retrieval system is a dynamic system and new images are continuously added to the image collection, a dynamic update of indexing structure is indispensably needed. This algorithm provides such a tool.

In addition to KLT, clustering is another powerful tool in performing dimension reduction. The clustering technique is used in various disciplines such as pattern recognition [43], speech analysis [115], and information retrieval [127]. Normally it is used to cluster similar objects (patterns, signals, and documents) together to perform recognition or grouping. This type of clustering is called row-wise clustering. However, clustering can also be used column-wise to reduce the dimensionality of the feature space [127]. Experiments show that this is a simple and effective approach.

One thing worth pointing out is that blind dimension reduction can be dangerous, since information can be lost if the reduction is below the embedded dimension. To avoid blind dimension reduction, a postverification stage is needed. Among different approaches, Fisher's discriminant analysis can provide useful guidance [138].

3.2. Multidimensional Indexing Techniques

After we identify the *embedded dimension* of the feature vectors, we need to select appropriate multidimensional indexing algorithms to index the reduced but still high dimensional feature vectors. There are three major research communities contributing in this area, i.e. computational geometry, database management, and pattern recognition. The existing popular multidimensional indexing techniques include the bucketing algorithm, k-d tree, priority k-d tree [167], quad-tree, K-D-B tree, hB-tree, R-tree and its variants R^+-tree and R^*-tree [57, 132, 53, 13, 118]. In addition to the above approaches, clustering and neural nets, widely used in pattern recognition, are also promising indexing techniques [43, 171].

The history of multidimensional indexing techniques can be traced back to the middle 1970s, when cell methods, quad-tree, and k-d tree were first introduced. However, their performances were far from satisfactory. Pushed by then urgent demand of spatial indexing from GIS and CAD systems, Guttman proposed the R-tree indexing structure in 1984 [57]. Based on his work, many other variants of R-tree were developed. Sellis *et al.* proposed R^+ tree in [132]. Greene proposed her variant of R-tree in [53]. In 1990, Beckman and Kriegel proposed the best dynamic R-tree variant, R^*-tree [13]. However, even for R^*-tree, it was not scalable to dimensions higher than 20 [46].

Very good reviews and comparisons of various indexing techniques in image retrieval can be found in [167, 98]. The research goal of White and Jain in [167] was to provide general purpose and domain-independent indexing algorithms. Motivated by k-d tree and R-tree, they proposed VAM k-d tree and VAMSplit R-tree. Experimentally they found that the VAMSplit R-tree provided the best performance, but the trade-off is the loss of the dynamic nature of R-tree. In [98], Ng and Sedighian proposed a three-step strategy towards image retrieval indexing, i.e. dimension reduction, evaluation of existing indexing approaches, and customization of the selected indexing approach. After dimension reduction using the eigenimage approach, the following three characteristics of the dimension-reduced data can be used to select good existing indexing algorithms:

- the new dimension components are ranked by decreasing variance,
- the dynamic ranges of the dimensions are known,
- the dimensionality is still fairly high.

On their test data sets, they found that the BA-KD-tree gave the best performance.

Considering that most of the tree indexing techniques were designed for traditional database queries (point queries and range queries) but not for the similarity queries used in image retrieval, there was a need to explore the new characteristics and requirements for indexing structures in image retrieval. Such a technique was explored in [155], where Tagare developed a tree adaptation approach which refined the tree structure by eliminating inefficient tree nodes for similarity queries.

So far, the above approaches only concentrated on how to identify and improve indexing techniques which are scalable to high dimensional feature vectors in image retrieval. The other nature of feature vectors in image retrieval, i.e. non-Euclidean similarity measures,

has not been deeply explored. The similarity measures used in image retrieval may be non-Euclidean and may even be nonmetric. There are two promising techniques towards solving this problem, i.e. clustering and neural nets. In [32], Charikar *et al.* proposed an incremental clustering technique for dynamic information retrieval. This technique had three advantages, i.e. dynamic structure, capable of handling high dimensional data, and the potential to deal with non-Euclidean similarity measures. In [118], Rui *et al.* further extended this technique in the directions of supporting non-Euclidean similarity measure and faster and more accurate search strategies.

In [171], Zhang and Zhong proposed using self-organization map (SOM) neural nets as the tool for constructing the tree indexing structure in image retrieval. The advantages of using SOM were its unsupervised learning ability, dynamic clustering nature, and the potential of supporting arbitrary similarity measures. Their experimental results over the Brodatz texture collection demonstrated that SOM was a promising indexing technique.

4. IMAGE RETRIEVAL SYSTEMS

Since the early 1990s, content-based image retrieval has become a very active research area. Many image retrieval systems, both commercial and research, have been built. Most image retrieval systems support one or more of the following options [22]:

- random browsing
- search by example
- search by sketch
- search by text (including key word or speech)
- navigation with customized image categories.

We have seen the provision of a rich set of search options today, but systematic studies involving actual users in practical applications still need to be done to explore the trade-offs among the different options mentioned above. Here, we will select a few representative systems and highlight their distinct characteristics.

4.1. QBIC

QBIC [99, 48, 46, 44, 129, 74, 41], standing for query by image content, is the first commercial content-based image retrieval system. Its system framework and techniques have profound effects on later image retrieval systems.

QBIC supports queries based on example images, user-constructed sketches and drawings, and selected color and texture patterns, etc. The color feature used in QBIC are the average (R,G,B), (Y,i,q), (L,a,b), and MTM (mathematical transform to Munsell) coordinates, and a k-element color histogram [46]. Its texture feature is an improved version of the Tamura texture representation [156]; i.e. combinations of coarseness, contrast, and directionality [44]. Its shape feature consists of shape area, circularity, eccentricity, major axis orientation, and a set of algebraic moment invariants [129, 46]. QBIC is one of the few systems which takes into account the high dimensional feature indexing. In its indexing subsystem, KLT is first used to perform dimension reduction and then R^*-tree is used as the multidimensional indexing structure [74, 46]. In its new system, text-based key word search can be combined with content-based similarity search. The on-line QBIC demo is at http://wwwqbic.almaden.ibm.com/.

4.2. Virage

Virage is a content-based image search engine developed at Virage Inc. Similar to QBIC, Virage [11, 56] supports visual queries based on color, composition (color layout), texture, and structure (object boundary information). But Virage goes one step further than QBIC. It also supports arbitrary combinations of the above four atomic queries. The users can adjust the weights associated with the atomic features according to their own emphasis. In [11], Jeffrey *et al.* further proposed an open framework for image management. They classified the visual features ("primitive") as general (such as color, shape, or texture) and domain specific (face recognition, cancer cell detection, etc.). Various useful "primitives" can be added to the open structure, depending on the domain requirements. To go beyond the query-by-example mode, Gupta and Jain proposed a nine-component *query language* framework in [56]. The corresponding demos of Virage are at http://www.virage.com/cgi-bin/query-e.

4.3. RetrievalWare

RetrievalWare is a content-based image retrieval engine developed by Excalibur Technologies Corp. [42, 3]. From one of its early publications, we can see that its emphasis was in neural nets to image retrieval [42]. Its more recent search engine uses color, shape, texture, brightness, color layout, and aspect ratio of the image, as the query features [3]. It also supports the combinations of these features and allows the users to adjust the weights associated with each feature. Its demo page is at http://vrw.excalib.com/cgi-bin/sdk/cst/cst2.bat.

4.4. Photobook

Photobook [106] is a set of interactive tools for browsing and searching images developed at the MIT Media Lab. Photobook consists of three subbooks from which shape, texture, and face features are extracted, respectively. Users can then query, based on the corresponding features in each of the three subbooks.

In its more recent version of Photobook, FourEyes, Picard *et al.* proposed including *human* in the image annotation and retrieval loop [112, 94, 110, 113, 109, 111, 114, 79]. The motivation of this was based on the observation that there was no single feature which can best model images from each and every domain. Furthermore, a human's perception is subjective. They proposed a "society of model" approach to incorporate the human factor. Experimental results show that this approach is effective in interactive image annotation [94, 114].

4.5. VisualSEEk and WebSEEk

VisualSEEk [135, 144] is a visual feature search engine and WebSEEk [137] is a World Wide Web oriented text/image search engine, both of which are developed at Columbia University. Main research features are spatial relationship query of image regions and visual feature extraction from compressed domain [162, 27, 28, 29].

The visual features used in their systems are color set and wavelet transform based texture feature [138–141]. To speed up the retrieval process, they also developed binary tree based indexing algorithms [136, 30, 142, 143].

VisualSEEk supports queries based on both visual features and their spatial relationships. This enables a user to submit a "sunset" query as red-orange color region on top and blue or green region at the bottom as its "sketch." WebSEEk is a web oriented search

engine. It consists of three main modules, i.e. image/video collecting module, subject classification and indexing module, and search, browse, and retrieval module. It supports queries based on both keywords and visual content. The on-line demos are at http://www.ee.columbia.edu/~sfchang/demos.html.

4.6. Netra

Netra is a prototype image retrieval system developed in the UCSB Alexandria Digital Library (ADL) project [86]. Netra uses color, texture, shape, and spatial location information in the segmented image regions to search and retrieve similar regions from the database. Main research features of the Netra system are its Gabor filter based texture analysis [7, 82, 89, 88], neural net-based image thesaurus construction [87, 84, 83] and edge flow-based region segmentation [85]. The on-line demo is at http://vivaldi.ece.ucsb.edu/Netra/.

4.7. MARS

MARS (multimedia analysis and retrieval system) was developed at University of Illinois at Urbana-Champaign [64, 91, 92, 102, 124, 125, 121, 119, 122, 123]. MARS differs from other systems in both the research scope and the techniques used. It is an interdisciplinary research effort involving multiple research communities: computer vision, database management system (DBMS), and information retrieval (IR). The research features of MARS are the integration of DBMS and IR (exact match with ranked retrieval) [64, 102], integration of indexing and retrieval (how the retrieval algorithm can take advantage of the underline indexing structure) [118], and integration of computer and human. The main focus of MARS is not on finding a single "best" feature representation, but rather on how to organize various visual features into a meaningful retrieval architecture which can dynamically adapt to different applications and different users. MARS formally proposes a relevance feedback architecture in image retrieval [123] and integrates such a technique at various levels during retrieval, including query vector refinement [119], automatic matching tool selection [121], and automatic feature adaption [122, 123]. The on-line demo is at http://jadzia.ifp.uiuc.edu:8000.

4.8. Other Systems

ART MUSEUM [61], developed in 1992, is one of the earliest content-based image retrieval systems. It uses the edge feature as the visual feature for retrieval. Blob-world [18], developed at UC-Berkeley, provides a transformation from the raw pixel data to a small set of localized coherent regions in color and texture space. This system allow the user to view the internal representation of the submitted image and the query results and therein enables the user to know why some "nonsimilar" images are returned and can therefore modify his or her query accordingly. The distinct feature of CAETIIML (http://www/videolib.princeton.edu/test/retrieve), built at Princeton University, is its combination of the on-line similarity searching and off-line subject searching [60]. More image retrieval systems can be found in [49, 10, 131, 154, 51, 168, 33, 153, 101].

5. FUTURE RESEARCH DIRECTIONS

From the above review, we can see that many advances have been made in various research aspects, including visual feature extraction, multidimensional indexing, and system design

[148, 6, 66, 65, 67, 69, 70]. However, there are still many open research issues that need to be solved before the current image retrieval can be of practical use.

5.1. Human in the Loop

A fundamental difference between a computer vision pattern recognition system and an image retrieval system is that a human is an indispensable part of the latter system. We need to explore the synergy of a human and a computer [69, 70]. This research trend has already been reflected in the evolution of content-based image retrieval. Early literature emphasizes "fully automated systems" and tries to find a "single best feature." But such an approach does not lead to success, as the computer vision technique is not there yet. More recent research emphasis is given to "interactive systems" and "human in the loop." For example, the QBIC team uses interactive region segmentation [41]. Based on the observation that each of the different texture representations, MRSAR, EV, and Wold-decomposition [94], has its own advantages in different domains, the MIT team moves from the "automated" Photobook to the "interactive" FourEyes [112, 94]. The WebSEEk system allows for dynamic feature vector recomputation based on the user's feedback [146]. The UCSB team incorporates supervised learning in texture analysis [87, 84]. The MARS team formally proposes a *relevance feedback* architecture in image retrieval, where human and computer can interact with each other to improve the retrieval performance [119, 121, 123]. Other relevance feedback-based approaches include PicHunter [38, 63].

5.2. High-level Concepts and Low-level Visual Features

Humans tend to use high-level concepts in everyday life. However, what current computer vision techniques can automatically extract from image are mostly low-level features. In constrained applications, such as the human face and finger print, it is possible to link the low-level features to high-level concepts (faces or finger prints). In a general setting, however, the low-level features do not have a direct link to the high-level concepts. To narrow down this semantic gap, some off-line and on-line processing is needed. The off-line processing can be achieved by using either supervised learning, unsupervised learning, or the combination of the two. Neural nets, genetic algorithms, and clustering are such learning tools [87, 84, 112, 94]. For on-line processing, a powerful and user-friendly intelligent query interface is needed to perform this task. It should allow the user to easily provide his or her evaluation of a current retrieval result to the computer. The relevance feedback technique proposed in MARS is one possible tool [119, 123].

5.3. Web Oriented

The expansion of the World Wide Web is astonishing. Each day thousands of documents, among which many are images, are added to the web. To better organize and retrieve the almost unlimited information, web-based search engines are highly desired. Such a solution exists for text-based information. The fact that Alta Vista, Inforseek, etc. are among the most frequently visited web sites indicates the need for a web-based search engine [2]. For images on the web, even though some good work has taken place [137, 49, 10, 131, 76], technical breakthroughs are needed to make the image search engines comparable to their text-based counterpart.

One major technical barrier lies in linking the low-level visual feature indexes used in most systems today to more desired semantic-level meanings. Based on preliminary on-line

experiments, we have observed that subject browsing and text-based matching are still more popular operations than feature-based search options [145]. That is partly the reason that commercial image retrieval systems on the web typically use customized subject categories to organize their image collection. Usually, different image retrieval systems focus on different sections of users and content. As a result, the indexing features and the subject taxonomies are also different, causing the concern of interoperability. Several recent efforts in standards have started to address this issue [2, 164]. Several research systems on image metaservers [14, 23, 96] have also investigated frameworks for integrated access to distributed image libraries.

5.4. High Dimensional Indexing

A by-product of the web expansion is the huge collection of images. Most currently existing research prototype systems only handle hundreds or at most a few thousand images; therefore a sequential scan of all images will not degrade the system's performance seriously. Because of this very reason, only a few existing systems explored the multi-dimensional indexing aspect of image retrieval [99, 48, 64, 137]. However, as the image collections are getting larger and larger, the retrieval speed is becoming a bottle neck. Although some progress has been made in this area, as described in Section 3, effective high dimensional indexing techniques are still in urgent need of being explored.

5.5. Performance Evaluation Criterion and Standard Testbed

Any technique is pushed forward by its domain's evaluation criterion. SNR is used in data compression, and precision and recall are used in text-based information retrieval. Good metrics will lead the technique in the correct direction while bad ones may mislead the research effort. Currently, some image retrieval systems measure performance based on the "cost/time" to find the right images [144]. Others evaluate performance using precision and recall, terms borrowed form text-based retrieval.

Although these criteria measure the system's performance to some extent, they are far from satisfactory. One major reason causing the difficulty of defining a good evaluation criterion is the perception subjectivity of image content. That is, the subjectivity of image perception prevents us from defining objective evaluation criteria. But still, we need to find a way of evaluating the system performance to guide the research effort in the correct direction [69, 70].

An equally important task is to establish a well-balanced large-scale testbed. For image compression, we have the Lena image, which has a good balance in various textures. For video compression, the MPEG community developed well-balanced test video sequences. For text-based information retrieval, a standard large-scale testbed also exists. For the image retrieval testbed, the MPEG-7 community has recently started to collect test data. For a testbed to be successful, it has to be large in scale to test the scalability (for multidimensional indexing), to be balanced in image content to test image feature effectiveness and overall system performance.

5.6. Human Perception of Image Content

The ultimate end user of an image retrieval system is human; therefore the study of human perception of image content from a psychophysical level is crucial. This topic is

closely related to the topics in Sections 5.1 and 5.2. Because of its importance, we will re-emphasize it here.

This topic is gaining increasing attention in recent years, aiming at exploring how humans perceive image content and how we can integrate such a "human model" into the image retrieval systems. The early research was conducted independently by the MIT team [112, 94], the NEC team [39, 38, 37, 103], and the UIUC team [119, 123]. Interesting enough, these teams are also the teams who initiated the study of relevance feedback in image retrieval. This is because, after realizing the difficulty in interpreting human perception subjectivity of image content, they naturally resorted to relevance feedback to "decode" the human perception.

More recent study of human perception focuses on the psychophysical aspects of human perception [103, 117]. In [103], Papathomas *et al.* conducted experiments studying the importance of using (a) semantic information, (b) memory of previous input, and (c) relative versus absolute judgement of image similarities, using PicHunter [38] as the underlying image retrieval system. Results show that the best performance is achieved when it uses only semantic cues, with memory and relative similarity judgement. The combination of semantic and visual cues only achieves a second-best result. We feel that one of the reasons that visual cues did not help in this case may due to the limited test dataset size—if the dataset is not big enough, the system may not be able to utilize the additional information from the visual cues. Our conjecture matches the experimental results from another group. In [117], Rogowitz *et al.* conducted a series of experiments analyzing human psychophysical perception of image content. According to their results, even though visual features do not capture the whole semantic meaning of the images, they do correlate a lot with the semantics. This result encourages us to develop perceptually based image features and metrics to achieve semantically meaningful retrievals.

5.7. Integration of Disciplines and Media

Both the database community literature and the computer vision community literature have used "image database" as the title of many articles [33]. However, in reality, most database community systems are non-image (text-based key words or graphics-based icons) databases, while most computer vision systems are image nondatabases (just a large file containing thousands of images is not a database, since most fundamental database units such as a data model and indexing, are not addressed at all). To the authors' knowledge, even though there are ongoing research efforts to build true image databases [102, 45], the systems are not at the complete stage yet.

A successful image database system requires an interdisciplinary research effort. Besides the integration of database management and computer vision, research from the traditional information retrieval area [17, 127, 133, 126, 16, 8] is also an indispensable part. Although the traditional information retrieval area's research focus was in text-based document retrieval, many useful retrieval models and techniques can be adapted to image retrieval. Some successful examples of such research effort include the adaption of Boolean retrieval models in image retrieval [92, 102], and the utilization of relevance feedback in image retrieval [119, 123].

Another observation is that integration of multimedia, multi-modalities provides great potential for improved indexing and classification of images in general domains. Research in [149, 147, 145] has shown promising results in using both textual and visual features

in automatic indexing of images. More sophisticated techniques for cross-mapping image classification between the high level using textual cues and the low level using the visual cues will bear fruit.

6. CONCLUDING REMARKS

In this paper, past and current technical achievements in visual feature extraction, multidimensional indexing, and system design are reviewed. Open research issues are identified and future research directions suggested. From the previous section, we can see that a successful image retrieval system requires the seamless integration of multiple research communities' efforts. Progress in each individual research community and an overall system architecture are equally important. We propose one possible integrated system architecture, as shown in Fig. 1.

There are three databases in this system architecture. The image collection database contains the raw images for visual display purpose. During different stages of image retrieval, different image resolutions may be needed. In that case, a wavelet-compressed image is a good choice [64]. Image processing and compression research communities contribute to this database.

The visual feature database stores the visual features extracted from the images using techniques described in Section 2. This is the information needed to support content-based image retrieval. Computer vision and image understanding are the research communities contributing to this database.

The text annotation database contains the key words and free-text descriptions of the images. It is becoming clear in the image retrieval community that content-based image retrieval is not a replacement of, but rather a complementary component to, the text-based

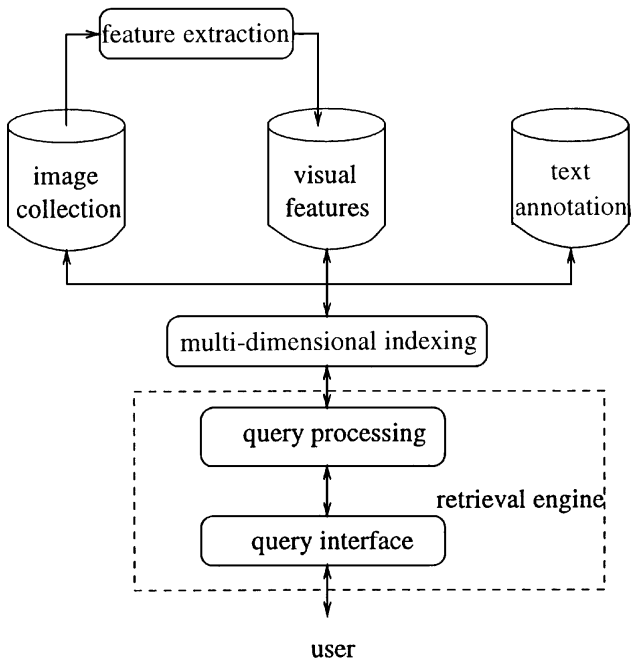

FIG. 1. An image retrieval system architecture.

image retrieval. Only the integration of the two can result in satisfactory retrieval performance. The research progress in IR and DBMS is the main thrust to this database.

As discussed earlier, to achieve fast retrieval speed and make the retrieval system truly scalable to large-size image collections, an effective multidimensional indexing module is an indispensable part of the whole system. This module will be pushed forward by computational geometry, database management, and pattern recognition research communities.

The retrieval engine module includes a query interface submodule and a query processing submodule. To communicate with the user in a friendly manner, the query interface is graphics-based. The interface collects the information need from the users and displays back the retrieval results to the users in a meaningful way. Research progress in user psychology and user interface helps to improve the interface design. Furthermore, the same query from a user can be processed in different ways. The query processing submodule manipulates the user query into the best processing procedures. This techniques is advanced by the database management community.

There are two major characteristics of this system architecture. One is its multidiscipline and interdisciplie nature, as demonstrated in the above discussion. The other is its interactive nature between human and computer. (Note that from the user to the three databases, the arrows are bi-directional). In all, integration of various disciplines, of multiple information sources, and of the human and computer will lead to a successful image retrieval system.

REFERENCES

1. MPEG-7 applications document, *ISO/IEC JTC1/SC29/WG11 N1922, MPEG97*, Oct. 1997.
2. MPEG-7 context and objectives (v.5), *ISO/IEC JTC1/SC29/WG11 N1920, MPEG97*, Oct. 1997.
3. Retrievalware, demo page. *http://vrw/excalib.com/cgi-bin/sdk/cst/cst2.bat*, 1997.
4. Special issue on visual information management (R. Jain, Guest Ed.), *Comm. ACM*, Dec. 1997.
5. Third draft of MPEG-7 requirements, *ISO/IEC JTC1/SC29/WG11 N1921, MPEG97*, Oct. 1997.
6. *Proc. Int. Conf. on Multimedia, 1993–1997*, ACM, New York, 1997.
7. A. D. Alexandrov, W. Y. Ma, A. El Abbadi, and B. S. Manjunath, Adaptive filtering and indexing for iamge databases, in *Proc. SPIE Storage and Retrieval for Image and Video Databases, 1995*.
8. J. Allan, Relevance feedback with too much data, in *Proc. of SIGIR'95, 1995*.
9. E. M. Arkin, L. Chew, D. Huttenlocher, K. Kedem, and J. Mitchell, An efficiently computable metric for comparing polygonal shapes, *IEEE Trans. Patt. Recog. Mach. Intell*. **13**(3), 1991.
10. V. Athitsos, M. J. Swain, and C. Frankel, Distinguishing, photographs and graphics on the world wide web, in *Proc. IEEE Workshop on Content-Based Access of Image and Video Libraries, 1997.*
11. J. R. Bach, C. Fuller, A. Gupta, A. Hampapur, B. Horowitz, R. Humphrey, R. Jain, and C. F. Shu, The Virage image search engine: An open framework for image management, in *Proc. SPIE Storage and Retrieval for Image and Video Databases*.
12. H. G. Barrow, Parametric correspondence and chamfer matching: Two new teachniques for image matching, in *Proc. 5th Int. Joint Conf. Artificial Intelligence, 1977*.
13. N. Beckmann, H.-P. Kriegel, R. Schneider, and B. Seeger, The R*-tree: An efficient and robust access method for points and rectangles, in *Proc. ACM SIGMOD, 1990.*
14. M. Beigi, A. Benitez, and S.-F. Chang, Metaseek: A content-based meta search engine for images, in *Proc. SPIE Storage and Retrieval for Image and Video Databases, San Jose, CA, 1998*. [demo and document: http://www.ctr.columbia.edu/metaseek]
15. G. Borgefors, Hierarchical chamfer matching: A parametric edge matching algorithm. *IEEE Trans. Patt. Recog. Mach. Intell.*, 1988.
16. C. Buckley and G. Salton, Optimization of relevance feedback weights, in *Proc. of SIGIR'95, 1995*.
17. J. P. Callan, W. B. Croft, and S. M. Harding, The inquery retrieval system, in *Proc. of 3rd Int. Conf. on Database and Expert System Application, Sept. 1992.*

18. C. Carson, S. Belongie, H. Greenspan, and J. Malik, Region-based image querying, in *Proc. of IEEE Workshop on Content-Based Access of Image and Video Libraries, in Conjuction with IEEE CVPR'97, 1997*.
19. S. Chandrasekaran, B. S. Manjunath, Y. F. Wang, J. Winkeler, and H. Zhang, An eigenspace update algorithm for image analysis, *CVGIP: Graphical Models and Image Processing Journal*, 1997.
20. N. S. Chang and K. S. Fu, *A Relational Database System for Images*, Technical Report TR-EE 79-28, Purdue University, May 1979.
21. N. S. Chang and K. S. Fu, Query-by pictorial-example, *IEEE Trans. on Software Engineering* **SE-6**(6), 1980.
22. S.-F. Chang, A. Eleftheriadis, and R. McClintock, Next-generation content representation, creation and searching for new media applications in education, *IEEE Proceedings, 1998*, to appear.
23. S.-F. Chang, J. R. Smith, M. Beigi, and A. Benitez, Visual information retrieval from large distributed online repositories. *Comm. ACM (Special Issue on Visual Information Retrieval) Dec. 1997*, pp. 12–20.
24. S.-K. Chang, Pictorial data-base systems, *IEEE Computer*, 1981.
25. S.-K. Chang and A. Hsu, Image information systems: Where do we go from here? *IEEE Trans. on Knowledge and Data Engineering* **4**(5), 1992.
26. S.-K. Chang, C. W. Yan, D. C. Dimitroff, and T. Arndt, An intelligent image database system, *IEEE Trans. Software Eng.* **14**(5), 1988.
27. S.-F. Chang, Compressed-domain content-based image and video retrieval, in *Proc. Symposium on Multimedia Communications and Video Coding, 1995*.
28. S.-F. Chang, Compressed-domain techniques for image/video indexing and manipulation, in *Proc. ICIP95 Special Session on Digital Library and Video on Demand, 1995.*
29. S.-F. Chang and J. R. Smith, Finding images/video in large archives. *D-Lib. Magazine*, 1997.
30. S.-F. Chang and J. Smith, Extracting mutlidimensional signal features for content-based visual query, in *Proc. SPIE Symposium on Visual Communications and Signal Processing, 1995*.
31. T. Chang and C.-C. J. Kuo, Texture analysis and classification with tree-structured wavelet transform, *IEEE Trans. Image Proc.* **2**(4), 429–441, 1993.
32. M. Charikar, C. Chekur, T. Feder, and R. Motwani, Incremental clustering and dynamic information retrieval, in *Proc. of the 29th Annual ACM Symposium on Theory of Computing, 1997*, pp. 626–635.
33. B. Cheng, Approaches to image retrieval based on compressed data for multimedia database systems, Ph.D. thesis, University of New York at Buffalo, 1996.
34. T. S. Chua, K.-L. Tan, and B. C. Ooi, Fast signiture-based color-spatial image retrieval, in *Proc. IEEE Conf. on Multimedia Computing and Systems, 1997*.
35. G. C.-H. Chuang and C.-C. J. Kuo, Wavelet descriptor of planar curves: Theory and applications, *IEEE Trans. Image Proc.* **5**(1), 56–70, 1996.
36. D. Copper and Z. Lei, On representation and invariant recognition of complex objects based on patches and parts, in *Lecture Notes in Computer Science Series, 3D Object Representation for Computer Vision* (M. Hebert, J. Ponce, T. Boult, and A. Gross, Eds.), pp. 139–153, Springer-Verlag, New York/Berlin, 1995.
37. I. J. Cox, M. L. Miller, T. P. Minka, and P. N. Yianilos, An optimized interaction strategy for bayesian relevance feedback, in *IEEE Conf. CVPR, 1998*.
38. I. J. Cox, M. L. Miller, S. M. Omohundro, and P. N. Yianilos, Pichunter: Bayesian relevance feedback for image retrieval, in *Intl. Conf. on Pattern Recognition*.
39. I. J. Cox, M. L. Miller, S. M. Omohundro, and P. N. Yianilos, Target testing and the pichunter bayesian multimedia retrieval system, in *Advanced Digital Libraries Forum, Washington, DC, May*.
40. G. C. Cross and A. K. Jain, Markov random field texture models, *IEEE Trans. Patt. Recog. and Mach. Intell.* **5**, 25–39, 1983.
41. D. Daneels, D. Campenhout, W. Niblack, W. Equitz, R. Barber, E. Bellon, and F. Fierens, Interactive outlining: An improved approach using active contours, in *Proc. SPIE Storage and Retrieval for Image and Video Databases, 1993*.
42. J. Dowe, Content-based retrieval in multimedia imaging, in *Proc. SPIE Storage and Retrieval for Image and Video Databases, 1993*.
43. R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, Chap. 6, Wiley, New York, 1973.
44. W. Equitz and W. Niblack, *Retrieving Images from a Database using Texture—Alogrithms from the QBIC System*, Technical Report RJ 9805, *Computer Science*, IBM Research Report, May 1994.
45. R. Fagin and E. L. Wimmers, Incorporating user preferences in multimedia queries, in *Proc. of Int. Conf. on Database Theory, 1997*.

46. C. Faloutsos, M. Flickner, W. Niblack, D. Petkovic, W. Equitz, and R. Barber, *Efficient and Effective Querying by Image Content*, Technical Report, IBM Research Report, 1993.
47. C. Faloutsos and K.-I. (David) Lin, Fastmap: A fast alogorithm for indexing, data-mining and visualization of traditional and multimedia datasets, in *Proc. of SIGMOD, 1995*, pp. 163–174.
48. M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafine, D. Lee, D. Petkovic, D. Steele, and P. Yanker, Query by image and video content: The QBIC system, *IEEE Computer*, 1995.
49. C. Frankel, M. J. Swain, and V. Athitsos. Webseer: *An Image Search Engine for the World Wide Web*, Technical Report TR-96-14, Computer Science Department, University of Chicago, 1996.
50. T. Gevers and V. K. Kajcovski, Image segmentation by directed region subdivision, in *Proc. IEEE Int. Conf. on Image Proc., 1994.*
51. Y. Gong, H. J. Zhang, H. C. Chuan, and M. Sakauchi, An image database system with content capturing and fast image indexing abilities, in *Proc. IEEE, 1994.*
52. C. C. Gotlieb and H. E. Kreyszig, Texture descriptors based on co-occurrence matrices, *Computer Vision, Graphics, and Image Processing* **51**, 1990.
53. D. Greene, An implementation and performance analysis of spatial data access, in *Proc. ACM SIGMOD, 1989.*
54. M. H. Gross, R. Koch, L. Lippert, and A. Dreger, Multiscale image texture analysis in wavelet spaces, in *Proc. IEEE Int. Conf. on Image Proc., 1994.*
55. V. N. Gudivada and J. V. Raghavan, Special issue on content-based image retrieval systems, *IEEE Computer Magzine* **28**(9), 1995.
56. A. Gupta and R. Jain, Visual information retrieval, *Communications of the ACM* **40**(5), 1997.
57. A. Guttman, R-tree: A dynamic index structure for spatial searching, in *Proc. ACM SIGMOD, 1984*.
58. M. Hansen and W. Higgins, Watershed-driven relaxation labeling for image segmentation, in *Proc. IEEE Int. Conf. on Image Proc., 1994*.
59. R. M. Haralick, K. Shanmugam, and I. Dinstein, Texture features for image classification, *IEEE Trans. on Sys. Man. and Cyb.* **SMC-3**(6), 1973.
60. H. H. Yu and W. Wolf, Hierarchical, multi-resolution algorithms for dictionary-driven content-based image retrieval, in *Proc. IEEE Int. Conf. on Image Proc., 1997.*
61. K. Hirata and T. Kato, Query by visual example, in *Proc. of 3rd Int. Conf. on Extending Database Technology*.
62. M. K. Hu, Visual pattern recognition by moment invariants, computer methods in image analysis, *IRE Transactions on Information Theory* **8**, 1962.
63. J. Huang, S. Kumar, M. Mitra, W.-J. Zhu, and R. Zabih, Image indexing using color correlogram, in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, 1997.*
64. T. S. Huang, S. Mehrotra, and K. Ramachandran, Multimedia analysis and retrieval system (MARS) project, in *Proc. of 33rd Annual Clinic on Library Application of Data Processing-Digital Image Access and Retrieval, 1996.*
65. IEEE, *Proc. IEEE Int. Conf. on Image Proc., 1994–1997.*
66. IEEE, *Proc. IEEE Int. Conf. on Multimedia Computing and Systems, 1994–1997.*
67. IEEE, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 1995–1997.*
68. M. Ioka, *A Method of Defining the Similarity of Images on the Basis of Color Information*, Technical Report RT-0030, IBM Research, Tokyo Research Laboratory, Nov. 1989.
69. R. Jain, Workshop report: NSF workshop on visual information management systems, in *Proc. SPIE Storage and Retrieval for Image and Video Databases, 1993.*
70. R. Jain, A. Pentland, and D. Petkovic, in *NSF-ARPA Workshop on Visual Information Management Systems, Cambridge, MA, June 1995*.
71. D. Kapur, Y. N. Lakshman, and T. Saxena, Computing invariants using elimination methods, in *Proc. IEEE Int. Conf. on Image Proc., 1995.*
72. A. Kundu and J.-L. Chen, Texture classification using qmf bank-based subband decomposition, *CVGIP: Graphical Models and Image Processing* **54**(5), 1992, 369–384.
73. A. Laine and J. Fan, Texture classification by wavelet packet signatures, *IEEE Trans. Patt. Recog. and Mach. Intell.* **15**(11), 1993, 1186–1191.
74. D. Lee, R. Barber, W. Niblack, M. Flickner, J. Hafner, and D. Petkovic, Indexing for complex queries on a query-by-content image database, in *Proc. IEEE Int. Conf. on Image Proc., 1994.*

75. Z. Lei, D. Keren, and D. B. Cooper, Computationally fast bayesian recognition of complex objects based on mutual algebraic invariants, in *Proc. IEEE Int. Conf. on Image Proc*.
76. M. S. Lew, K. Lempinen, and D. P. Huijsmans, *Webcrawling Using Sketches*, Technical Report, Computer Science Department, Leiden University, The Netherlands, 1997.
77. B. Li and S. D. Ma, On the relation between region and contour representation, in *Proc. IEEE Int. Conf. on Image Proc., 1995*.
78. X. Q. Li, Z. W. Zhao, H. D. Cheng, C. M. Huang, and R. W. Harris, A Fuzzy logic approach to image segmentation, in *Proc. IEEE Int. Conf. on Image Proc., 1994.*
79. F. Liu and R. W. Picard, Periodicity, directionality, and randomness: Wold features for image modeling and retrieval, *IEEE Trans. Patt. Recog. and Mach. Intell.* **18**(7), 1996.
80. H. Lu, B. Ooi, and K. Tan, Efficient image retrieval by color contents, in *Proc. of the 1994 Int. Conf. on Applications of Databases, 1994.*
81. M. Lybanon, S. Lea, and S. Himes, Segmentation of diverse image types using opening and closing, in *Proc. IEEE Int. Conf. on Image Proc., 1994.*
82. W. Y. Ma and B. S. Manjunath, A comparison of wavelet transform features for texture image annotation, in *Proc. IEEE Int. Conf. on Image Proc., 1995.*
83. W. Y. Ma and B. S. Manjunath, *A Pattern Thesaurus for Browsing Large Aerial Photographs*, Technical Report 96-10, Univ. of California at Santa Barbara, 1996.
84. W. Y. Ma and B. S. Manjunath, Texture features and learning similarity, in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 1996*, pp. 425–430.
85. W. Y. Ma and B. S. Manjunath, Edge flow: A framework of boundary detection and image segmentation, in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 1997.*
86. W. Y. Ma and B. S. Manjunath, Netra: A toolbox for navigating large image databases, in *Proc. IEEE Int. Conf. on Image Proc., 1997.*
87. B. S. Manjunath and W. Y. Ma, Image indexing using a texture dictionary, in *Proceedings of SPIE Conference on Image Storage and Archiving System, Vol. 2606*.
88. B. S. Manjunath and W. Y. Ma, *Texture Features for Browsing and Retrieval of Image Data*, Technical Report 95-06, Univ. of California at Santa Barbara, 1995.
89. B. S. Manjunath and W. Y. Ma, Texture features for browsing and retrieval of image data, *IEEE T-PAMI (Special issue on digital libraries)*, Nov. 1996.
90. C. S. McCamy, H. Marcus, and J. G. Davidson, A color-rendition chart, *Journal of Applied Photographic Engineering* **2**(3), 1976.
91. S. Mehrotra, K. Chakrabarti, M. Ortega, Y. Rui, and T. S. Huang, Multimedia analysis and retrieval system, in *Proc. of The 3rd Int. Workshop on Information Retrieval Systems, 1997.*
92. S. Mehrotra, Y. Rui, O.-B. Michael, and T. S. Huang, Supporting content-based queries over images in MARS, in *Proc. of IEEE Int. Conf. on Multimedia Computing and Systems, 1997.*
93. B. M. Mehtre, M. Kankanhalli, and W. F. Lee, Shape measures for content based image retrieval: A comparison, *Information Processing & Management* **33**(3), 1997.
94. T. P. Minka and R. W. Picard, Interactive learning using a "society of models," in *Proc. IEEE CVPR, 1996*, pp. 447–452.
95. M. Miyahara, Mathematical transform of (r,g,b) color data to munsell (h,s,v) color data, *SPIE Visual Commun. Image Process.* **1001**, 1988.
96. D. Murthy and A. Zhang, Webview: A multimedia database resource integration and search system over web, in *WebNet 97: World Conference of the WWW, Internet, and Intranet, Oct. 1997*.
97. A. D. Narasimhalu, Special section on content-based retrieval, *Multimedia Systems*, 1995.
98. R. Ng and A. Sedighian, Evaluating multi-dimensional indexing structures for images transformed by principal component analysis, in *Proc. SPIE Storage and Retrieval for Image and Video Databases, 1996.*
99. W. Niblack, R. Barber, and *et al.*, The QBIC project: Querying images by content using color, texture and shape, in *Proc. SPIE Storage and Retrieval for Image and Video Databases, Feb. 1994.*
100. P. P. Ohanian and R. C. Dubes, Performance evaluation for four classes of texture features. *Pattern Recognition* **25**(8), 1992, 819–833.
101. A. Ono, M. Amano, and M. Hakaridani, A flexible content-based image retrieval system with combined scene description keyword, in *Proc. IEEE Conf. on Multimedia Computing and Systems, 1996.*
102. M. Ortega, Y. Rui, K. Chakrabarti, S. Mehrotra, and T. S. Huang, Supporting similarity queries in MARS, in *Proc. of ACM Conf. on Multimedia, 1997.*

103. T. V. Papathomas, T. E. Conway, I. J. Cox, J. Ghosn, M. L. Miller, T. P. Minka, and P. N. Yianilos, Psychophysical studies of the performance of an image database retrieval system, in *IS&T/SPIE Conf. on Human Vision and Electronic Imaging III, 1998.*
104. G. Pass, R. Zabih, and J. Miller, Comparing images using color coherence vectors, in *Proc. ACM Conf. on Multimedia, 1996.*
105. A. Pentland and R. Picard, Special issue on digital libraries, *IEEE Trans. Patt. Recog. and Mach. Intell.*, 1996.
106. A. Pentland, R. W. Picard, and S. Sclaroff, Photobook: Content-based manipulation of image databases, *International Journal of Computer Vision*, 1996.
107. A. P. Pentland, Fractal-based description of natural scenes, *IEEE Trans. Patt. Recog. and Mach. Intell.* **6**(6), 1984, 661–674.
108. E. Persoon and K. S. Fu, Shape discrimination using fourier descriptors. *IEEE Trans. Sys. Man. Cyb.*, 1977.
109. R. W. Picard, Computer learning of subjectivity, in *Proc. ACM Computing Surveys*.
110. R. W. Picard, Digital libraries: Meeting place for high-level and low-level vision, in *Proc. Asian Conf. on Comp. Vis*.
111. R. W. Picard, *A Society of Models for Video and Image Libraries*, Technical Report, MIT, 1996.
112. R. W. Picard and T. P. Minka, Vision texture for annotation. *Multimedia Systems: Special Issue on Content-based Retrieval*.
113. R. W. Picard, Toward a visual thesaurus, in *Workshops in Computing, MIRO 95, Springer-Verlag, New York/Berlin, 1995*.
114. R. W. Picard, T. P. Minka, and M. Szummer, Modeling user subjectivity in image libraries, in *Proc. IEEE Int. Conf. on Image Proc., Lausanne, Sept. 1996.*
115. L. Rabiner and B.-H. Juang, *Fundamentals of Speech Recognition, Chap. 5, Prentice–Hall, Englewood Cliffs, NJ, 1993.*
116. R. Rickman and J. Stonham, Content-based image retrieval using colour tuple histograms, in *Proc. SPIE Storage and Retrieval for Image and Video Databases, 1996.*
117. B. E. Rogowitz, T. Frese, J. Smith, C. A. Bouman, and E. Kalin, Percpetual image similarity experiments, in *IS&T/SPIE Conf. on Human Vision and Electronic Imageing III, 1998.*
118. Y. Rui, K. Chakrabarti, S. Mehrotra, Y. Zhao, and T. S. Huang, Dynamic clustering for optimal retrieval in high dimensional multimedia databases, in *TR-MARS-10-97, 1997.*
119. Y. Rui, T. S. Huang, and S. Mehrotra, Content-based image retrieval with relevance feedback in MARS, in *Proc. IEEE Int. Conf. on Image Proc., 1997.*
120. Y. Rui, T. S. Huang, and S. Mehrotra, Mars and its applications to MPEG-7, *ISO/IEC JTC1/SC29/WG11 M2290, MPEG97, July 1997.*
121. Y. Rui, T. S. Huang, S. Mehrotra, and M. Ortega, Automatic matching tool selection using relevance feedback in MARS, in *Proc. of 2nd Int. Conf. on Visual Information Systems, 1997.*
122. Y. Rui, T. S. Huang, S. Mehrotra, and M. Ortega, A relevance feedback architecture in content-based multimedia information retrieval systems, in *Proc. of IEEE Workshop on Content-Based Access of Image and Video Libraries, in Conjunction with IEEE CVPR'97, 1997.*
123. Y. Rui, T. S. Huang, M. Ortega, and S. Mehrotra, Relevance feedback: A power tool in interactive content-based image retrieval, *IEEE Trans. on Circuits Systems Video Technol. (Special Issue on Interactive Multi-Media Systems for the Internet)*, Sept. 1998.
124. Y. Rui, A. C. She, and T. S. Huang, Automated shape segmentation using attraction-based grouping in spatial-color-texture space, in *Proc. IEEE Int. Conf. on Image Proc., 1996*.
125. Y. Rui, A. C. She, and T. S. Huang, Modified fourier descriptors for shape representation—a practical approach, in *Proc. of First International Workshop on Image Databases and Multi Media Search, 1996*.
126. G. Salton and C. Buckley, Term-weighting, approaches in automatic text retrieval, *Information Processing and Management*, 1988.
127. G. Salton and M. J. McGill, *Introduction to Modern Information Retrieval*, McGraw-Hill, New York, 1983.
128. R. Samadani and C. Han, Computer-assisted extraction of boundaries from images, in *Proc. SPIE Storage and Retrieval for Image and Video Databases, 1993*.
129. B. Scassellati, S. Alexopoulos, and M. Flickner, Retrieving images by 2d shape: A comparison of computation methods with human perceptual judgments, in *Proc. SPIE Storage and Retrieval for Image and Video Databases, 1994*.
130. B. Schatz and H. Chen, Building largescale digital libraries, *Computer*, 1996.

131. S. Sclaroff, L. Taycher, and M. La Cascia, Imagerover: A content-based image browser for the world wide web, in *Proc. IEEE Workshop on Content-Based Access of Image and Video Libraries, 1997*.
132. T. Sellis, N. Roussopoulos, and C. Faloutsos, The R^+-tree: A dynamic index for multi-dimensional objects, in *Proc. 12th VLDB, 1987*.
133. W. M. Shaw, Term-relevance computations and perfect retrieval performance, *Information Processing and Management*.
134. A. C. She and T. S. Huang, Segmentation of road scenes using color and fractal-based texture classification, in *Proc. ICIP, Austin, Nov. 1994*.
135. J. R. Smith and S.-F. Chang, Querying by color regions using the VisualSEEk content-based visual query system (M. T. Maybury, Ed.), in *Intelligent Multimedia Information Retrieval, 1996.*
136. J. R. Smith and S.-F. Chang, Local color and texture extraction and spatial query, in *Proc. IEEE Int. Conf. on Image Proc., 1996.*
137. J. R. Smith and S.-F. Chang, Visually searching the web for content, *IEEE Multimedia Magazine* **4**(3), 12–20, 1997. [Columbia U. CU/CTR Technical Report 459-96-25]
138. J. R. Smith and S. F. Chang, Transform features for texture classification and discrimination in large image databases, in *Proc. IEEE Int. Conf. on Image Proc., 1994.*
139. J. R. Smith and S.-F. Chang, Single color extraction and image query, in *Proc. IEEE Int. Conf. on Image Proc., 1995*.
140. J. R. Smith and S.-F. Chang, Tools and techniques for color image retrieval, in *IS & T/SPIE Proceedings, Vol. 2670, Storage & Retrieval for Image and Video Databases IV, 1995*.
141. J. R. Smith and S.-F. Chang, Automated binary texture feature sets for image retrieval, in *Proc. ICASSP-96, Atlanta, GA, 1996*.
142. J. R. Smith and S.-F. Chang, Automated binary texture feature sets for image retrieval, in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proc., May 1996*.
143. J. R. Smith and S.-F. Chang, Tools and techniques for color image retrieval, in *Proc. SPIE Sotrage and Retrieval for Image and Video Databases, 1996*.
144. J. R. Smith and S.-F. Chang, Visualseek: A fully automated content-based image query system, in *Proc. ACM Multimedia 96, 1996.*
145. J. R. Smith and S.-F. Chang, Enhancing image search engines in visual information environments, in *IEEE 1st Multimedia Signal Processing Workshop, June 1997*.
146. J. R. Smith and S.-F. Chang, An image and video search engine for the world-wide web, in *Proc. SPIE Storage and Retrieval for Image and Video Databases, 1997*.
147. J. R. Smith and S.-F. Chang, Multi-stage classification of images from features and related text, in *4th Europe EDLOS Workshop, San Miniato, Italy, Aug. 1997*.
148. SPIE, *Proc. Spie Storage and Retrieval for Image and Video Databases, 1993–1997*.
149. R. K. Srihari, Automatic indexing and content-based retrieval of captioned images, *IEEE Computer Magzine* **28**(9), 1995.
150. M. Stricker and A. Dimai, Color indexing with weak spatial constraints, in *Proc. SPIE Storage and Retrieval for Image and Video Databases, 1996*.
151. M. Stricker and M. Orengo, Similarity of color images, in *Proc. SPIE Storage and Retrieval for Image and Video Databases, 1995*.
152. M. Swain and D. Ballard, Color indexing, *International Journal of Computer Vision* **7**(1), 1991.
153. M. J. Swain, Interactive indexing into image databases, in *Proc. SPIE Storage and Retrieval for Image and Video Databases, Vol. 1908, 1993*.
154. D. L. Swets and J. J. Weng, Efficient content-based image retrieval using automatic feature selection, in *Proc. IEEE, 1995*.
155. H. Tagare, Increasing retrieval efficiency by index tree adaption, in *Proc. of IEEE Workshop on Content-based Access of Image and Video Libraries, in Conjunction with IEEE CVPR '97, 1997*.
156. H. Tamura, S. Mori, and T. Yamawaki, Texture features corresponding to visual perception, *IEEE Trans. on Sys., Man. and Cyb.* **SMC-8**(6), 1978.
157. H. Tamura and N. Yokoya, Image database systems: A survey, *Pattern Recognition* **17**(1), 1984.
158. G. Taubin, Recognition and positioning of rigid objects using algebraic moment invariants, in *SPIE Vol. 1570, Geometric Methods in Computer Vision, 1991*.
159. K. S. Thyagarajan, T. Nguyen, and C. Persons, A maximum likelihood aproach to texture classification using wavelet transform, in *Proc. IEEE Int. Conf. on Image Proc., 1994*.

160. I. Wallace and O. Mitchell, Three-dimensional shape analysis using local shape descriptors, *IEEE Trans. Patt. Recog. and Mach. Intell., PAMI-3(3), May 1981*.

161. I. Wallace and P. Wintz, An efficient three-dimensional aircraft recognition algorithm using normalized Fourier descriptors, *Computer Graphics and Image Processing* **13**, 1980.

162. H. Wang and S.-F. Chang, Compressed-domain image search and applications, Technical Report, Columbia Univ., 1995.

163. J. Wang, W.-J. Yang, and R. Acharya, Color clustering techniques for color-content-based image retrieval from image databases, in *Proc. IEEE Conf. on Multimedia Computing and Systems, 1997*.

164. S.Weibel and E. Miller, Image description on the internet: A summary of the cni/oclc image metadata on the internet workshop, Sept. 24–25, 1996, Dublin, Ohio, *D-Lib Magazine*, 1997.

165. J. Weszka, C. Dyer, and A. Rosenfeld, A comparative study of texture measures for terrain classification, *IEEE Trans. on Sys., Man. and Cyb.* **SMC-6**(4), 1976.

166. D. White and R. Jain, Algorithms and strategies for similarity retrieval, in *TR VCL-96-101*, University of California, San Diego, 1996.

167. D. White and R. Jain, Similarity indexing: Algorithms and performance, in *Proc. SPIE Storage and Retrieval for Image and Video Databases, 1996*.

168. J. K. Wu, A. D. Narasimhalu, B. M. Mehtre, C. P. Lam, and Y. J. Gao, Core: A content-based retrieval engine for multimedia information systems, *Multimedia Systems*, 1995.

169. L. Yang and F. Algregtsen, Fast computation of invariant geometric moments: A new method giving correct results, in *Proc. IEEE Int. Conf. on Image Proc., 1994*.

170. C. T. Zahn and R. Z. Roskies, Fourier descriptors for plane closed curves, *IEEE Trans. on Computers*, 1972.

171. H. J. Zhang and D. Zhong, A scheme for visual feature based image retrieval, in *Proc. SPIE Storage and Retrieval for Image and Video Databases, 1995*.

YONG RUI received the B.S. from the Southeast University, People's Republic of China in 1991 and the M.S. from Tsinghua University, People's Republic of China in 1994, both in electrical engineering. Since 1995, he has been with the Department of Electrical and Computer Engineering at the University of Illinois, Urbana-Champaign, pursuing his Ph.D. Since March 1999 he has been a researcher at Microsoft Research, Microsoft Corporation, Redmond, Washington. His research interests include multimedia information retrieval, multimedia signal processing, computer vision, and artificial intelligence. He has published over 30 technical papers in these areas. He received a Huitong University Fellowship in 1989–1990, a Guanghua University Fellowship in 1992–1993, and a CSE College of Engineering Fellowship in 1996–1998.

THOMAS S. HUANG received his B.S. in electrical engineering from National Taiwan University, Taipei, Taiwan, China and his M.S. and Sc.D. in electrical engineering from the Massachusetts Institute of Technology, Cambridge, Massachusetts. He was on the faculty of the Department of Electrical Engineering at MIT from 1963 to 1973. He was on the faculty of the School of Electrical Engineering and was Director of its Laboratory for Information and Signal Processing at Purdue University from 1973 to 1980. In 1980, he joined the University of Illinois, Urbana-Champaign, where he is now the William L. Everitt Distinguished Professor of Electrical and Computer Engineering, research professor at the Coordinated Science Laboratory, and Head of the Image Formation and Processing Group at the Beckman Institute for Advanced Science and Technology. During his sabbatical leaves, Dr. Huang has worked at the MIT Lincoln Laboratory, the IBM Thomas J. Watson Research Center, and the Rheinishes Landes Museum in Bonn, West Germany; he has held visiting professor positions at the Swiss Institutes of Technology in Zurich and Lausanne, the University of Hannover in West Germany, INRS Telecommunications of the University of Quebec in Montreal, Canada, and University of Tokyo, Japan. He has served as a consultant to numerous industrial firms and government agencies both in the U.S. and abroad. Dr. Huang's professional interests lie in the broad area of information technology, especially the transmission and processing of multidimensional signals. He has published 12 books and over 300 papers in network theory, digital filtering, image processing, and computer vision. He is a Fellow of the International Association of Pattern Recognition, IEEE, and the Optical Society of America. He has received a Guggenheim Fellowship, an A. V. Humboldt Foundation Senior U.S. Scientist Award, and a fellowship from the Japan Association for the Promotion of Science. He received the IEEE Acoustics, Speech, and Signal Processing Society's Technical Achievement Award in 1987 and the Society Award in 1991. He is a founding editor of the international journal *Computer Vision, Graphics, and Image Processing* and Editor of the *Information Sciences Series*, published by Springer-Verlag.

SHIH-FU CHANG is currently an associate professor at the Department of Electrical Engineering and New Media Technology Center, Columbia University. He received a Ph.D. in EECS from U. C. Berkeley in 1993.

At Columbia, he also actively participates in Columbia's Digital Library Project. His current research interests include content-based visual query, networked video manipulation, video coding, and communication. He is particularly interested in the application of content-based video processing to image/video retrieval, network resource management, image watermarking, and authentication. His group has developed several large-scale Web-based prototypes of visual information systems, including an MPEG video editing engine, WebClip, and content-based visual search engines, VideoQ and WebSEEk. Professor Chang holds two U.S. patents (with eight more pending) in the area of visual search and compressed video processing. He was the recipient of two best paper awards, an ONR Young Investigator award 1998–2001, an NSF CAREER award 1995–1998, and an IBM UPP Program Faculty Development Award 1995–1998. He actively participates in technical activities for international conferences and publications and has also taught several short courses on visual information.