An Overview of Feature Extraction Methods in CBIR

Dr. Thakur Nileshsingh V.
Shri Ramdeobaba Kamla Nehru Engineering College,
Katol Road, Nagpur, India.
thakurnisvis@rediffmail.com

Dr. Nitin A. Koli
Sant Gadge Baba Amravati University,
Tapovan Road, Amravati, India.
nak187@rediffmail.com

ABSTRACT

In multimedia processing, content based image retrieval plays the important role. This paper presents an overview of the feature extraction methods related to the CBIR. Image segmentation is first key step to have the features of particular image. Image can be indexed by these features in terms of the vectors. Here, we have studied the past and present methods of feature extraction. Retrieval time can be saved by using the efficient feature extraction method. Paper ends with the importance of segmentation and retrieval time. This paper may help the CBIR researchers for identifying or developing the feature extraction strategy or technique.

Categories and Subject Descriptors

H.3 [**Information Storage and Retrieval**]: Information Search and Retrieval – *clustering, information filtering, query formulation, relevance feedback, retrieval models, search process, selection process*

General Terms

Design, Documentation, Management, Theory.

Keywords

Content based image retrieval, Image segmentation, Feature extraction.

1 INTRODUCTION

Image indexing techniques are becoming more and more important and critical to facilitate people searching databases. Even though there have been standard retrieval techniques for text [1], they are not suitable for image data since image annotation is unfeasible for any non-trivial scenario. It is much more difficult to retrieve image data than text due to the subjectivity of human perception, which is difficult to represent formally. The latter problem is called Content-Based Image Retrieval. The research in Image Retrieval began in the 1970s. Initially, a text-based approach was adopted. In this approach, human first manually annotates each image using keywords, and then images are retrieved based on the keywords in the text annotation.

There are two main disadvantages in this approach. One is that it requires a huge amount of human labour in the manual annotation when the image collection is large. The other one is that it is hard to precisely annotate the rich content of an image by humans due to perception subjectivity [2, 3]. The text based approach remained popular until early 1990s when many large-scale image collections emerged and the drawbacks of text-based approach became more and more notorious. In the content-based approach, images are retrieved directly based on their visual content such as color, texture, and shape [2, 3].

In presented paper, we focused on feature methods related to image retrieval research. Feature extraction consists of two steps: image segmentation and feature construction. Which particular features have to be extract so what they can help for retrieval of concerned image from image database? Particular contributions in color, texture, and shape extraction have been discussed in this paper.

This paper is organized as follows: section 2 summarizes the different CBIR system and the development approaches of CBIR system. More stress given on the feature extraction methods, which are discussed in section 3. Section 4 consists of discussion on overview, which specifies the importance of segmentation and feature extraction.

2 CBIR-STATE OF THE ART

A large number of content-based image retrieval systems have been built [2] such as QBIC [4], VisualSEEK [5], and Photobook [6]. In the QBIC system, content-based queries such as query by example image, query by sketch and drawing, and query by selected color and texture patterns are supported. The visual features include color, texture, and shape. Color is represented using a k-bin color histogram. Texture is described by an improved Tamura texture representation. Shape information includes area, circularity, eccentricity, major axis orientation, and moment invariants. KLT is used to reduce the dimension of the feature vectors and R* tree is the indexing structure. The later version integrated text based query [2].

In the VisualSEEK system, both content based query (query by example image and spatial relation pattern) and text-based query are supported. The system uses the following visual features: color represented by color set, texture based on wavelet transform, and spatial relationship between image regions. A binary tree is used to index the feature vectors [2]. The Photobook system is composed of a set of interactive tools for browsing and searching images. It supports query by example. The images are organized in three subbooks from which shape, texture, and face appearance features are extracted respectively [2, 6].

Content-Based Image Retrieval (CBIR) system for retrieval of normal anatomical regions present in Computed Tomography

(CT) studies of the chest and abdomen is proposed in [7]. Invariance to translation, rotation, and scale is required by a good shape representation. Sustaining deformation contour matching is an important issue at the matching process. An efficient and robust shape-based image retrieval system is proposed [8].

A new approach is proposed in [9] to retrieve images by content using a composition of relevant features regarding texture, shape and brightness distribution. New Content-Based Image Retrieval (CBIR) approach, for image databases, based on cluster analysis are presented in [10]. In order to represent the intricate composition of images, the grid-based approach [11] partitions each image into blocks from which a feature representation is derived from the local low-level content.

In the newly emerging multimedia applications such as MPEG-4 and MPEG-7, shape plays an important role in supporting the so called content based functionalities. Many shape representation have been proposed for various purposes. These methods can generally be grouped into contour-based and region based. Contour-based methods, such as chain code [12], shape signature [13], polygonal approximation [14], autoregressive models [15, 16], FD [17, 18] and CSS [19, 20], exploit shape boundary information which is crucial to human perception in judging shape similarity. Region based methods, such as geometric moments [21], Zernike moments [17, 22], grid representation [23] and area, exploit only shape interior information, therefore can be applied to more general shapes.

Some geometric features such as average scale, skew, kurtosis, etc. reflected in the region based methods (geometric moments) are also important perceptual features. For CBIR purpose, a shape representation should be affine invariant, robust, compact, easy to derive and matching, and perceptually meaningful. In terms of these properties, FD, CSSD, ZMD and GD have been recognized as suitable for CBIR. CSSD and ZMD have been adopted in MPEG-7 as shape descriptors.

In literature, many survey papers [2, 24, 25, 26] exist which focus work prior to the year 2000. Other surveys also exist on closely related topics such as relevance feedback [27], high-dimensional indexing of multimedia data [28], face recognition [29] (useful for face based image retrieval), applications of CBIR to medicine [30], and applications to art and cultural imaging [31]. Multimedia information retrieval, as a broader research area covering video, audio, image, and text analysis has been extensively surveyed in [32, 33].

3 FEATURE EXTRACTION IN CBIR

Attribute based approaches are used to retrieve the images from image database. But feature based approaches [34] are generally used in Content-based image retrieval systems. The important common features of the image: color [35], texture [7], shape [8], boundary, intensity levels, frequency domain feature, spatial domain feature are used as the basis to form the feature database in content based image retrieval. In CBIR, image segmentation is the first step; later the features can be extracted from the image to index the image.

3.1 Image Segmentation

Image segmentation is the first step to CBIR. Reliable segmentation is especially critical for characterizing shapes within images, without which the shape estimates are largely meaningless. One of the most important new advances in segmentation employs the Normalized Cuts criterion [36]. The problem of image segmentation is mapped to a weighted graph partitioning problem where the vertex set of the graph is composed of image pixels and edge weights represent some perceptual similarity between pixel pairs. The normalized cut segmentation method in [36] is also extended to textured image segmentation by using cues of contour and texture differences [37], and to incorporate known partial grouping priors by solving a constrained optimization problem [38]. The latter has potential for incorporating real-world application-specific priors, e.g., location and size cues of organs in pathological images. Table 1 summarizes various segmentation approaches.

Table 1: Summary of Various Segmentation Approaches

Approach
Hidden Markov Random Fields and the Expectation-Maximization (EM) algorithm [39]
Spectral clustering approach [40]
Mean shift procedure [41]
Multi-resolution segmentation of low depth of field images [42]
Bayesian framework based segmentation involving the Markov chain Monte Carlo technique [43]
EM algorithm based segmentation using a Gaussian mixture model [44]
A sequential segmentation approach [45]
An unsupervised approach [46]
Hierarchical extraction [47]
Probabilistic modeling [48]

3.2 Types of Features

In a color layout approach, an image is divided into a small number of sub-images and the average color components, e.g., red, green, and blue intensities, are computed for every sub-image. The overall image is thus represented by a vector of color components where a particular dimension of the vector corresponds to a certain sub-image location. The advantage of global extraction is the high speed for both extracting features and computing similarity. However, as evidenced by the rare use of color layout in recent work, global features are often too rigid to represent an image. Specifically, they can be over sensitive to location and hence fail to identify important visual characteristics.

To increase the robustness to spatial transformation, the second approach to form signatures is by local extraction and an extra step of feature summarization. In local feature extraction, a set of features are computed for every pixel using its neighborhood, e.g., average color values across a small block centered on the pixel. To reduce computation, an image may be divided into small non-overlapping blocks, and features are computed individually for every block. The features are still local because of the small block size, but the amount of computation is only a fraction of that for obtaining features around every pixel.

3.3 Color Features and Color Extraction

The majority of image retrieval systems focus mainly on color distribution. The most popular and well-developed techniques are based on color or brightness histograms [49]. Although the process to obtain the histograms have low computational cost, as it is obtained in linear time regarding the image size, a histogram produces an ambiguous image representation, as two different images can have the same histogram. However, histograms can provide a first sift for image retrieval systems, as they only bring extra images, which can be further isolated using more elaborated and complex techniques. Applying the more complex techniques on fewer images reduces the retrieval total cost.

The color feature is the most widely used visual feature in image retrieval because it is more robust to changes due to scaling, orientation, perspective and occlusion of images [3]. Humans perceive a color as a combination of three stimuli, R (red), G (Green), and B (Blue), which form a color space. Separating chromatic information and luminance information can generate more color spaces. To extract color information, a color space must be chosen first. There exist many color spaces. Examples are RGB, YIQ, YUV, CIE LAB, CIE LUV, HSV and its variants. None of them can be used for all applications [2, 3, 50, 51, 52, 53]. RGB is the most commonly used color space primarily because color image acquisition and recording hardware are designed for this space. However, the problem of this space is the close correlation among the three components, which means that all three components will change as the intensity changes. This is not good for color analysis. YIQ and YUV are used to represent the color information in TV signals in color television broadcasting. Y encodes the luminance information and UV or IQ encodes the chromatic information. CIE LAB and CIE LUV are often used in measuring the distance between two colors because of its perceptual uniformity. That is, the Euclidian Distance between two colors represented in these two spaces matches the human perception.

However, its transformation from the RGB space is computationally intensive and dependent on a reference white. H (Hue) S (Saturation) V (Value) and its variants are perceptual color spaces, while all the previous color spaces are not. By perceptual, we mean that the three components (H, S, and V) represent the color attributes associated with how human eyes perceive colors. Hue, which corresponds to the dominant wavelength of a given perceived color stimulus, represents the type of the color such as red, blue, and green. The strength, purity, or richness of a color is represented by Saturation. The color is perceived to be less saturated as more white light is added to it. Value (or intensity) is the amount of light perceived from a given color sensation. White and black are perceived as the maximum and minimum intensity, respectively [50].

The HSV color space can be chosen for two reasons. First, it is perceptual, which makes HSV a proven color space particularly amenable to color image analysis [50, 51, 52]. Second, the benchmark results in [3] show that the color histogram in the HSV color space performs the best. Many schemes, such as color histogram, color moments, color coherence vector, and color autocorrelogram, can be used to describe the color information in an image. Color histogram is the most widely used method since it is more robust to changes due to scaling, orientation, perspective, and occlusion of images [3]. Color histogram represents the joint distribution of three color channels in an image.

Therefore, it characterizes the global color information in an image. Color moments are the first few low-order moments of each color channel. It is a compact representation of the color distribution of an image. Color coherence vector is designed to take into account of the spatial distribution of color in an image. It is obtained by partitioning each histogram bin into two: one with coherent pixels and the other with incoherent pixels. Color autocorrelogram represents the probability of finding a pixel of some color at some distance from a pixel of the same color in an image.

It characterizes both the global and spatial distribution of the color. In the performance evaluation experiments in [3], it is shown that the color histogram runs much faster than the color coherence vector and color autocorrelogram, performs almost as good as the color coherence vector, and does not perform much worse than the best color autocorrelogram. Therefore, color histogram is used in [2][3]. Because there are many different colors, to reduce the complexity in histogram computation, the color space needs to be quantized [3].

The color space can be quantized through color categorization. All possible colors of the pixels are first classified into thirteen categories based on the H, S, and V value ranges. Each category is identified by an ID, and then each pixel is labelled as the ID of the category to which it belongs. Next, a color label histogram is built. The resulting color label histogram is computationally efficient and effective to obtain objects with similar colors. In addition, it reduces the dimension of the color feature vector.

The author in [50] used twelve categories, which are obtained from the experimental result based on the H, S, and V value ranges, to represent the dominant colors of color regions in an image. These twelve categories are black, white, red, bright red, yellow, bright yellow, green, bright green, blue, bright blue, purple, and bright purple. The Hue is partitioned into 10 color slices with 5 main slices (red, yellow, green, blue, and purple) and five transition slices. Each transition slice is counted in both adjacent main slices. Some modifications can be made to compute the color histogram. Firstly, the difference between the bright chromatic pixels and the chromatic pixels is ignored to reduce the total number of bins.

Therefore, bright red and dark red are considered to be in the same color category. Secondly, the transition color slices are considered as separate categories for histogram computation. Thirdly, a new category “gray” is added to consider all possible value ranges since some images in our image database contain the gray color. Hence, there are totally thirteen color categories, which are white, black, gray, red, red-yellow, yellow, yellow-green green, green-blue, blue, blue-purple, purple, and purple-red.

3.4 Texture features

Texture features are intended to capture the granularity and repetitive patterns of surfaces within in a picture. For instance, grass land, brick walls, teddy bears, and flower petals differ in texture by smoothness as well as patterns. Texture features have been studied for long in image processing, computer vision, and computer graphics [54], such as multi-orientation filter banks [55] and wavelet transforms [56]. In image processing, a popular way to form texture features is by using the coefficients of a certain transform on the original pixel values or more sophisticatedly, statistics computed from those coefficients. Examples of texture features using the *wavelet transform* and the discrete cosine transform can be found in [57, 58]. Among the earliest work on the use of texture features for image retrieval are [59]. Texture descriptors, apt for inclusion in the MPEG-7, were broadly discussed in [60]. Such descriptors encode significant, general visual characteristics into standard numerical formats that can use for various higher-level tasks. A *thesaurus* for texture, geared toward aerial image retrieval, has been proposed in [61]. The texture extraction part of this thesaurus building process involves the application of a bank of Gabor filters [62] to the images, to encode statistics of the filtered outputs as texture features. Advances in textured region descriptors have been made, such as affine and photometric transformation invariant features that are also robust to the shape of the region in question [63]. Texture features at a point in the image are meaningful only as a function of its neighborhood, and the (effective) size of this neighborhood can be thought of as a *scale* at which these features are computed. Because a choice of scale is critical to the meaningfulness of such features, it has been explored as an automatic scale selection problem in [44], specifically to aid image retrieval.

3.5 Shape Features

Shape is a key attribute of segmented image regions and its efficient and robust representation plays an important role in retrieval. Synonymous with shape representation is the way such representations are matched with each other. A new shape descriptor for similarity matching, referred to as shape *context*, is proposed which is fairly compacting yet robust to a number of geometric transformations [64].

In [65], curves are represented by a set of segments or *tokens*, whose feature representations (curvature and orientation) are arranged into a metric tree [66] for efficient shape matching and shape-based image retrieval.

A dynamic programming (DP) approach to shape matching is proposed in [67], where shapes are approximated as sequences of concave and convex segments. One problem with this approach is that computation of Fourier descriptors and moments is slow, although pre-computation may help produce real-time results. Continuing with Fourier descriptors, exploitation of both the amplitude and phase, and the use of Dynamic Time Warping (DTW) distance instead of Euclidean distance are shown to be an accurate shape matching technique in [68]. The rotational and starting point invariance otherwise obtained by discarding the phase information is maintained here by adding compensation terms to the original phase, thus allowing its exploitation for better discrimination.

Closely associated are approaches that model spatial relations among local image entities for retrieval. Much of the approaches to spatial modeling and matching have been influenced by earlier work on *iconic indexing* [69, 70] based on the theory of symbolic projections. Here, images are represented based on orthogonal projections of constituent entities, by encoding the corresponding bi-directional arrangement on the two axes as a 2D string of entities and relationships. This way, image matching is effectively converted from a spatial matching problem to a one-dimensional matching one. Many variants of the 2D string model have been proposed since. In recent years, extensions such as 2D Be-string [71] have been proposed. Another work on iconic indexing can be found in [72], where a symbolic representation of real images, termed virtual image is proposed. Spatial modeling of 3D objects is presented in [73].

3.6 Construction of Features Vectors

Distributions extracted from a collection of local feature vectors can be of other forms, for instance, a continuous density function [57], or even a spatial stochastic model [74]. A continuous density in general is more precise to describe a collection of local feature vectors than a discrete distribution with finitely many support vectors. A stochastic model moves beyond a continuous density by taking into account spatial dependence among local feature vectors. For special kinds of images, we may need these sophisticated statistical models to characterize them.

In [74], it is noted that spatial relationship among pixels is crucial for capturing Chinese ink painting styles. On the other hand, more sophisticated statistical models are computationally costly and less intuitive, a probable reason why their usage is limited.

4 CONCLUSION

In this paper, we have discussed the available CBIR systems with their specific features and the feature extraction methods. We have specifically focused on the feature extraction techniques which were used in past where the strong segmentation was the key issue. But presently, without strong segmentation, region formation algorithms are used for feature extraction. This paper may help the CBIR researchers for identifying or developing the feature extraction strategy or technique.

5 REFERENCES

[1] B. Ribeiro-Neto, R. Baeza-Yates, Modern Information Retrieval, Addison Wesley, 1999.

[2] Y. Rui and T.S. Huang, "Image Retrieval: Current Techniques, Promising Directions and Open Issues", Journal of Visual Communication and Image Representation, 1999.

[3] W.-Y. Ma and H. Zhang, "Content-Based Image Indexing and Retrieval", Handbook of Multimedia Computing, CRC Press, 1999.

[4] M. Flicker, H. Sawhney, et. al., "Query by Image and Video Content: The QBIC System", IEEE Computers, Sep. 1995.

[5] J.R. Smith and S.-F. Chang, "VisualSEEK: A Fully Automated Content-Based Image Query System", Proceedings Of ACM Multimedia 96, 1996.

[6] A. Pentland, R.W. Picard, and S. Sclaroff, "Photobook: Content-Based Manipulation of Image Databases", International Journal of Computer Vision, 1996.

[7] Winnie Tsang et. al., " Texture-Based Image Retrieval for Computerized Tomography Databases", Proceedings of the 18th IEEE Symposium on Computer-Based Medical Systems (CBMS'05), 2005.

[8] Hwei-Jen Lin et. al., "A Study of shape-Based Image Retrieval", Proceedings of the 24th International Conference on Distributed Computing Systems Workshops (ICDCSW'04), 2004.

[9] Agma J. M. Traina et. al., "Content-based Image Retrieval Using Approximate Shape of Objects", Proceedings of the 17th IEEE Symposium on Computer-Based Medical Systems (CBMS'04), 2004.

[10] Renato O. Stehling et. al., "An Adaptive and Efficient Clustering-Based Approach for Content-Based Image Retrieval in Image Databases", Proceedings of the International Database Engineering and Applications Symposium (IDEAS'01), 2001.

[11] Charlie Dagli and Thomas S. Huang, "A Framework for Grid-Based Image Retrieval", Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04), 2004.

[12] H. Freeman and A. Saghri, “Generalized chain codes for planar curves”, Proceedings of the 4th International Joint Conference on Pattern Recognition, pages 701-703, Kyoto, Japan, November 7-10 1978.
[13] E. R. Davies. Machine Vision: Theory, Algorithms, Practicalities. Academic Press, 1997.
[14] Chuang Gu., “Multivalued Morphology and Segmentation-based Coding.”, Ph.D thesis of Signal Processing Lab. (LTS) of Swiss Federal Institute of Technology at Lausanne (EPFL) 1995.
[15] Hannu Kauppinen, Tapio Seppanen and Matti Pietikainen, “An Experimental Comparison of Autoregressive and Fourier-Based Descriptors in 2D Shape Classification”, IEEE Trans. PAMI-17(2):201-207.
[16] Jesse Bennett and Alireza Khotanzad, “Multispectral Random Field Models for Synthesis and Analysis of Color Images”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 3, march 1998.
[17] Eric Persoon and King-sun Fu, “Shape Discrimination Using Fourier Descriptors”, IEEE Trans. On Systems, Man and Cybernetics, Vol.SMC-7(3):170-179, 1977.
[18] Charles T. Zahn and Ralph Z. Roskies, “Fourier Descriptors for Plane closed Curves”, IEEE Trans. On Computer, c-21(3):269-281, 1972.
[19] Farzin Mokhtarian and Alan Mackworth, “Scale-Based Description and Recognition of Planar Curves and Two-Dimensional Shapes”, IEEE PAMI-8(1), 1986.
[20] F. Mokhtarian, S. Abbasi and J. Kittler, “Efficient and Robust Retrieval by Shape Content through Curvature Scale Space”, Int. Workshop on Image DataBases and Multimedia Search, pp. 35-42, Amsterdam, The Netherlands, 1996.
[21] Ming-Kuei Hu, “Visual pattern Recognition by Moment Invariants”, IRE Transactions on Information Theory, IT-8:179- 187, 1962.
[22] B.M.Mehtre, Mohan S. Kankanhalli and Wing Foon Lee, “Shape Measures for Content based Image Retrieval: A Comparison”, Information Processing & Management, 33(3): 319-337, 1997.
[23] G. J. Lu and A. Sajjanhar, “Region-based shape representation and similarity measure suitable for content-based image retrieval”, Multimedia System 7:165-174, 1999.
[24] Aigrain, P., Zhang, H., and Petkovic, D. , “Content-based representation and retrieval of visual media: A review of the state-of-the-art.”, Multimedia Tools and Applications 3, 3, 179–202, 1996.
[25] Smeulders, A. W., Worring, M., Santini, S., Gupta, A., and Jain, R., “Content-based image retrieval at the end of the early years”, IEEE Trans. Pattern Analysis and Machine Intelligence 22, 12, 1349–1380, 2000.
[26] Snoek, C. G. M. and Worring, M., “Multimodal video indexing: A review of the state-ofthe- art”, Multimedia Tools and Applications 25, 1, 5–35, 2005.
[27] Zhou, X. S. and Huang, T. S., “Relevance feedback in image retrieval: A comprehensive review”, Multimedia Systems 8, 536–544, 2003.
[28] Bohm, C., Berchtold, S., and Keim, D. A., “Searching in high-dimensional space index structures for improving the performance of multimedia databases”, ACM Computing Surveys 33, 3, 322–373, 2001.
[29] Zhao, W., Chellappa, R., Phillips, P. J., and Rosenfeld, A., “Face recognition: A literature survey”, ACM Computing Surveys 35, 4, 399–458, 2003.
[30] Muller, H., Pun, T., and Squire, D., “Learning from user behavior in image retrieval: Application of market basket analysis”, Intl. Journal of Computer Vision 56, 1/2, 65–77, 2004.
[31] Chen, Y., Wang, J. Z., and Krovetz, R., “Clue: Cluster-based retrieval of images by unsupervised learning”, IEEE Trans. Image Processing 14, 8, 1187–1201, 2005.
[32] Sebe, N., Lew, M. S., Zhou, X., Huang, T. S., and Bakker, E., “The state of the art in image and video retrieval”, Proceedings CIVR, 2003.
[33] Lew, M., Sebe, N., Djeraba, C., and Jain, R., “Content-based multimedia information retrieval: State-of-the-art and challenges”, ACM Trans. Multimedia Computing, Communication, and Applications 2, 1, 1–19, 2006.
[34] Xiuqi Li, et. al. , “An Effective Content-based Visual Image Retrieval System,” Proceedings of the 26 th Annual International Computer Software and Applications Conference (COMPSAC’02), 2002.
[35] Fan Liu, et. al., “Natural Image Retrieval based on Features of Homogeneous Color Regions”.
[36] Shi, J. and Malik, J., “Normalized cuts and image segmentation”, IEEE Trans. Pattern Analysis and Machine Intelligence 22, 8, 888–905, 2000.
[37] Malik, J., Belongie, S., Leung, T. K., and Shi, J., “Contour and texture analysis for image segmentation”, Intl. Journal of Computer Vision 43, 1, 7–27, 2001.
[38] Yu, S. X. and Shi, J., “Segmentation given partial grouping constraints”, IEEE Trans. Pattern Analysis and Machine Intelligence 26, 2, 173–183, 2004.
[39] Zhang, Y., Brady, M., and Smith, S., “Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm”, IEEE Trans. Medical Imaging 20, 1, 45–57, 2001.
[40] Carballido-Gamio, J., Belongie, S., and Majumdar, S., “Normalized cuts in 3-d for spinal mri segmentation”, IEEE Trans. Medical Imaging 23, 1, 36–44, 2004.
[41] Comaniciu, D. and Meer, P., “Mean shift: A robust approach toward feature space analysis”, IEEE Trans. Pattern Analysis and Machine Intelligence 24, 5, 603–619, 2002.
[42] Wang, J. Z., Li, J., Gray, R. M., and Wiederhold, G., “Unsupervised multiresolution segmentation for images with low depth of field”, IEEE Trans. Pattern Analysis and Machine Intelligence 23, 1, 85–90, 2001.
[43] Tu, Z. and Zhu, S.-C., “Image segmentation by data-driven markov chain monte carlo”, IEEE Trans. Pattern Analysis and Machine Intelligence 24, 5, 657–673, 2002.
[44] Carson, C., Belongie, S., Greenspan, H., and Malik, J., “Blobworld: Image segmentation using expectation-maximization and its application to image querying”, IEEE Trans. Pattern Analysis and Machine Intelligence 24, 8, 1026–1038, 2002.
[45] Chen, J., Pappas, T., Mojsilovic, A., and Rogowitz, B., “Adaptive image segmentation based on color and texture”, Proceedings IEEE ICIP, 2002.
[46] Deng, Y. and Manjunath, B., “Unsupervised segmentation of color-texture regions in images and video”, IEEE Trans. Pattern Analysis and Machine Intelligence 23, 8, 800–810, 2001.
[47] Iqbal, Q. and Aggarwal, J. K., “Retrieval by classification of images containing large manmade objects using perceptual grouping”, Pattern Recognition J. 35, 7, 1463–1479, 2002.

[48] Datta, R., Joshi, D., Li, J., and Wang, J. Z., "Tagging over time: Real-world image annotation by lightweight meta-learning", Proceedings ACM Multimedia, 2007.
[49] R. Brunelli and O. Mich, "On the Use of Histograms for Image Retrieval," Proceedings of IEEE Intl. Conf. on Multimedia Computing and Systems (ICMCS), Florence, Italy, June 7-11, pp. 143-147.
[50] D. Androutsos, "Efficient Indexing and Retrieval of Color Image Data Using a Vector-Based Approach", Ph.D. Dissertation, University of Toronto, 1999.
[51] H.D. Cheng, X.H. Jiang, etc. "Color Image Segmentation: Advances and Prospects," Pattern Recognition, 2001.
[52] H.D. Cheng and Y. Sun, "A Hierarchical Approach to Color Image Segmentation Using Homogeneity," IEEE Transactions on Image Processing, Dec. 2001.
[53] Y. Alp Aslandogan and C.T. Yu, "Techniques and Systems for Image and Video Retrieval," IEEE Trans. On Knowledge and Data Engineering, Vol.11, No.1, Jan/Feb 1999.
[54] Haralick, R., "Statistical and structural approaches to texture", Proceedings IEEE 67, 5, 786–804, 1979.
[55] Malik, J. and Perona, P., "Preattentive texture discrimination with early vision mechanisms", Journal of Optical Society of America A 7, 5, 923–932, 1990.
[56] Unser, M., "Texture classification and segmentation using wavelet frames", IEEE Trans. Image Processing 4, 11, 1549–1560, 1995.
[57] Do, M. N. and Vetterli, M., "Wavelet-based texture retrieval using generalized Gaussian density and kullback-leibler distance", IEEE Trans. Image Processing 11, 2, 146–158, 2002.
[58] Li, J., Najmi, A., and Gray, R. M., "Image classification by a two dimensional hidden markov model", IEEE Trans. Signal Processing 48, 2, 527–533, 2000.
[59] Manjunath, B. and Ma, W.-Y., "Texture features for browsing and retrieval of image data", IEEE Trans. Pattern Analysis and Machine Intelligence 18, 8, 837–842, 1996.
[60] Manjunath, B. S., Ohm, J.-R., Vasudevan, V. V., and Yamada, A., "Color and texture descriptors", IEEE Trans. Circuits and Systems for Video Technology 11, 6, 703–715, 2001.
[61] Ma, W.-Y. and Manjunath, B., "Texture thesaurus for browsing large aerial photographs", Journal of American Society for Information Science 49, 7, 633–648, 1998.
[62] Jain, A. and Farrokhnia, F., "Unsupervised texture segmentation using gabor filters", Proceedings Int. Conf. Systems, Man and Cybernetics, 1990.
[63] Schaffalitzky, F. and Zisserman, A., "Viewpoint invariant texture matching andwide baseline stereo", Proceedings IEEE ICCV, 2001.
[64] Belongie, S., Malik, J., and Puzicha, J., "Shape matching and object recognition using shape contexts", IEEE Trans. Pattern Analysis and Machine Intelligence 24, 4, 509–522, 2002.
[65] Berretti, S., Del Bimbo, A., and Pala, P., "Retrieval by shape similarity with perceptual distance and effective indexing", IEEE Trans. Multimedia 2, 4, 225–239, 2000.
[66] Ciaccia, P., Patella, M., and Zezula, P., "M-tree: An efficient access method for similarity search in metric spaces", Proceedings VLDB, 1997.
[67] Petrakis, E. G. M., Diplaros, A., and Milios, E., "Matching and retrieval of distorted and occluded shapes using dynamic programming", IEEE Trans. Pattern Analysis and Machine Intelligence 24, 4, 509–522, 2002.
[68] Bertini, E., Cali, A., Catarci, T., Gabrielli, S., and Kimani, S. 2005. Interaction-based adaptation for small screen devices. Lecture Notes in Computer Science 3538, 277–281.
[69] Chang, S., Shi, Q., and Yan, C., "Iconic indexing by 2-d strings", IEEE Trans. Pattern Analysis and Machine Intelligence 9, 3, 413–427, 1987.
[70] Chang, S., Yan, C., Dimitroff, D., and Arndt, T., "An intelligent image database system", IEEE Trans. Software Engineering 14, 5, 681–688, 1988.
[71] Wang, Y. H., "Image indexing and similarity retrieval based on spatial relationship model", Information Sciences - Informatics and Computer Science 154, 1-2, 39–58, 2003.
[72] Petraglia, G., Sebillo, M., Tucci, M., and Tortora, G., "Virtual images for similarity retrieval in image databases", IEEE Trans. Knowledge and Data Engineering 13, 6, 951–967, 2001.
[73] Berretti, S. and Del Bimbo, A., "Modeling spatial relationships between 3d objects. In Proceedings IEEE ICPR, 2006.
[74] Li, J. and Wang, J. Z., "Studying digital imagery of ancient paintings by mixtures of stochastic models", IEEE Trans. Image Processing 13, 3, 340–353, 2004.