CHAPTER 2

Literature Review

The number of publications on CBIR increased dramatically over the past decade. The comprehensive reviews of Rehman et al., (2012); Thomee and Lew, (2012); Wang et al., (2010); Datta et al., (2008); Liu et al., (2007), Eakins et al., (1999); Rui et al., (1999); Sumeulders et al., (2000) have provided a good insight into the more recent advances in the content-based image retrieval. In this chapter we are presenting the review with respected to some key feature like color feature, texture feature etc.

a) Color Feature

The color feature of an image can be a powerful feature for the purpose of CBIR, if extracted in a perceptually oriented way and kept semantically intact (Kiranyaz et al., 2012). Color feature is comparatively robust to background complication and independent of image size and orientation (Rui et al., 1999). Color space transform is important for color feature extraction and data redundancy reduction. The identification of suitable color space is essential and useful in many color image processing application, such as image display, processing, retrieval, recognition, and compression (Chen and Hao, 2004). Different studies on color spaces can be found in Wang et al., (1997); Kasson and Plouffe, (1992).

In prior work on color descriptors, Kato et al., (1991) used the color of every corresponding pixel in two images for comparison and the number of consequent pixels having the same color determines the similarity between them. In same time Swain and Ballard (1991), proposed an image indexing scheme based on color histogram. In their scheme, a color image is transformed into the gray-scale image. Then, they generated a histogram with 256 bins to record the total number of pixel values in the gray-scale image as image feature to search the similar images from the database. In order to reduce the memory space for storing the histogram, they reduce the number of bins in the histogram from 256 to 64. In addition, the color distributions of each bin were used to improve the retrieve accuracy of the histogram-based retrieval scheme. And histogram intersection technique is utilized for comparing two images. Histograms are
invariant to translation and rotation approximately the viewing axis and vary slowly with changes of view angle, scale, and occlusion. Hafner et al., (1995) proposed efficient color histogram using quadratic distance. In a color-content-based image retrieval system, the dimensions of the color histograms is reduce in order to minimize the storage necessity and to enhance the speed of similarity measurement (Wang et al., 1997).

Smith and Chang, (1997a) proposed color sets, as an approximation to the color histogram, to make possible fast search over large scale image collection. 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 collection of colors from the quantized color space. Because color set feature vectors were binary, a binary search tree was constructed to allow a quick search.

A multiresolution color clustering procedure is suggested by Wan and Kuo, (1998b) to reduce the computational complexity in distance computation of histogram. Feature of such histogram-based color descriptors based on color space such as (RGB, CIE-Lab, CIE-Luv, HSV and YCrCb) is that they cluster the pixels into fixed color bins, which are quantizing the whole color space using a pre-defined color palette, clustering all the pixels having similar color and reducing the color levels from millions to (usually) thousands or even hundreds via quantization scheme. Yang and Lin, (2003) introduce an image retrieval system by incorporating geometry-enhanced information in the color histogram. A good review and an efficient representation of color histograms based on Karhunen–Loeve transform (KLT) can be found in Tran and Lenz, (2005).

Jeong et al., (2004) proposed image retrieval using color histogram generated by Gauss Mixture Vector Quantization (GMVQ). The GMVQ is known to be robust for quantizer mismatch. Sural, et al., (2002); Singha and Hemachandran, (2011 a, b); Yue et al., (2011) proposed color feature extraction using color histogram in HSV color space. Malik and Baharudin, (2012) proposed images retrieval based on the statistical quantized color histogram features where the images enhancement using Laplacian filter. The statistical quantized color histogram features are extracted from sharpened grayscale images using different number of quantization bins. Liu and Yang (2012) proposed image retrieval using Color Difference Histogram (CDH), describe the image feature by edge orientation in L*a*b* color space.
Besides the color histogram, several other color feature representations have been applied in image retrieval, including color moments, color correlogram, Color Coherence Vector (CCV), color sets etc. To overcome the quantization effects, as in the color histogram, Stricker and Orengo, (1995) used three statistic moments, the average, the variance, and the skewness to represent the image. In their scheme, a color image is transformed to three spectrums, H, S and V. The scheme, then, calculated the three statistic moments for each spectrum. The nine moments obtained from the three spectrums are the features of the image. The mathematical foundation of this approach is that any color distribution can be characterized by its moments. Furthermore, since most of the information is determined 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. One drawback of the moment descriptor is that the average of all the colors might be quite different from any of the original colors. A study by Ma and Zhang, (1998) on image retrieval shows that the high-dimensional color histogram descriptor is better than the color moment descriptor. One drawback of the moment descriptor is that the average of all the colors might be quite different from any of the original colors. Given a color moment feature description, it is hard to recover the actual colors in the image (Deng et al., 2001). Recently Singh and Hemachandran, (2012) have used color moment as a color descriptor for image retrieval.

Deng et al., (2001) proposed a set of dominant colors, for indexing. The target application is similarity retrieval in large image databases using color. Colors in a given region are clustered into a small number of representative colors. The feature descriptor consists of the representative colors, their percentages in the region, spatial coherency of the dominant colors, and color variances for each dominant color.

Pass and Zabih, (1996) proposed a split histogram, called Color Coherence Vector (CCV), a way of incorporating spatial information into color histograms through histogram refinement used to distinguished images whose color histograms are indistinguishable. Pass, et al., (1997) reported a better retrieval system than the traditional histogram-based methods, using CCV. Colombo et al., (1998) have also
used CCV to split histogram into two parts: coherent one and non-coherent one depending on the size of their connected component.

Another image retrieval technique based on color correlogram is reported by Huang et al., (1997a). This technique outperforms not only the traditional color histogram method but also the CCV methods for image indexing and retrieval. Ojala et al., (2001) also used the Correlogram based image retrieval in HSV color domain.

Han and Ma, (2002) proposed Fuzzy Color Histogram (FCH), considers the color similarity information by spreading each pixel’s total membership value to all the histogram bins using fuzzy C-means clustering algorithm. Experimental outcome shows that the obtained FCH is less sensitive to noisy interference such as lighting intensity changes and quantization errors than conventional color histogram. Again Konstantinidis et al., (2005) proposed image retrieval based on fuzzy color histogram processing, in that they presented fuzzy linking method of color histogram using L*a*b* color space. Nachtgeael et al. (2007) proposed color image retrieval using fuzzy partition in HIS color space. And Chamorro-Martinez et al., (2007) proposed fuzzy approach for image retrieval using color feature.

Lu and Chang, (2007) proposed the uses of the color distributions, the mean value and the standard deviation, to represent the global characteristics of the image. In addition, the image bitmap is used to represent the local characteristics of the image for increasing the accuracy of the retrieval system.

Moghaddam and Saadatmand-Tarzstan, (2006) proposed another approach, called Gabor Wavelet Correlogram for image indexing and retrieval and further improved the retrieval performance. But it has certain drawbacks like serious computational difficulty and huge memory requirement as identified by Kiranyaz et al., (2010). Lee et al., (2008) proposed Wavelet Correlograms and it performs slightly better than the Correlogram and surpasses other color descriptors such as color histograms and scalable color descriptor.

Chun et al., (2008) conducted comprehensive performance evaluations among several global/spatial color descriptors for CBIR and reported that Auto- Correlogram achieves the best retrieval performance among the others, such as color histograms, CCV, color moments, etc. Color correlogram offer more effectiveness in comparison with color
histogram methods with little efficiency reduction. Despite the achieved advantage, they suffer from sensitivity to scale, color and illumination changes.

Silakari, (2009) proposed color feature extraction using Block Truncation Coding (BTC). In BTC they split the image into R, G, B component and find the average of each component R, G, B and then apply color moment to each splitted component. In color moment they calculate Mean, Standard Deviation, and Skewness. And then they apply k-means clustering to find the cluster.

b) Texture Feature

Texture analysis can be useful when objects in an image are more characterized by their texture than by intensity and conventional thresholding techniques cannot be used efficiently. While there is no proper definition for texture, intuitively this descriptor provides measures of properties such as smoothness, coarseness, and regularity. The three principal approaches used in image processing to explain the texture of a region are statistical, structural and spectral. Statistical approach characterizes texture by the statistical properties of the gray-levels of the pixels in an image. The structural approach deal with the arrangement of image primitives, such as the description of texture based on regularly spaced parallel lines and the spectral approach is based on properties of the frequency domain (Gonzalez and Woods, 2002). Due to its effectiveness and significance in computer vision and pattern recognition, there are many research outcomes in past three decade. Many researchers have given a compressive review on CBIR based on texture feature like Rui et al., (1999); Liu et al., (2007).

In the year 1971, Haralick et al., proposed the co-occurrence matrix for texture feature. They used the spatial relationships of gray levels in texture discrimination. Gotlieb and Kreyszig, (1990) enhanced the method proposed by Haralick et al., (1971) they performed tests using small and homogeneous texture images and found that the variance feature performed the best, followed by the inverse difference moment and the entropy features.

Tamura et al., (1978) explored the texture representation from a different angle and among the six Tamura features: coarseness, directionality, regularity, contrast, line-likeness, contrast and roughness, the first three are more significant. One major
difference 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 (Faloutsos et al., 1994) and the MARS system (Huang et al., 1997b; Ortega et al., 1997) further improved this texture representation. But Liu et al., (2007) in his review paper explore that the limitation of Tamura features is that there was no work at multiple resolutions to account for scale.

Liu and Picard, (1996) treated images as 2-D homogeneous random fields and used the Wold theory to decompose them into three mutually orthogonal components. These components correspond to the perceptually important ‘periodicity’, ‘directionality’ and ‘randomness’ properties. They compared the features that they compute from the 2-D Wold model to other models, namely the shift-invariant principal component analysis, the multiresolution simultaneous autoregressive model, the tree-structured wavelet transform and Tamura. The Wold-based features performed better than others in terms of average recall for a Brodatz texture dataset.

Among the various texture features, Gabor features and wavelet features are widely used for image retrieval and have been reported to sound match the results of human vision study (Ma and Manjunath, 1997a; Wang et al., 2001). Gabor filter (or Gabor wavelet) is widely adopted to extract texture features from the images for image retrieval (Manjunath and Ma, 1996; Dimai, 1999), and has been shown to be very efficient. Manjunath and Ma, (1996) have shown that image retrieval using Gabor features outperforms the features based on Pyramid-structured Wavelet Transform (PWT), Tree-structured Wavelet Transform (TWT) and Multiresolution Simultaneous Autoregressive Model (MR-SAR). Though Gabor wavelet based features give better retrieval performance, but they suffer from the following two main drawbacks: 1) A Gabor basis function is not orthogonal, which increases redundancy and memory requirement. 2) Time required for feature extraction is quite high, which limits the retrieval speed (Kokare et al., 2005). Han and Ma, (2007) applied a simple modification of the conventional Gabor filter to achieving rotation invariance and scale invariance, individually.
In the last decade, wavelet theory has emerged and became a mathematical tool which provides a more proper, solid, and unified framework for multiscale image analysis (Mallat, 1989; Daubechies, 1992). Typically, the wavelet transform decomposes an image on a low resolution image and a series of detail images. The low resolution image is obtained by iteratively blurring the image; the detail images contain the information lost during this operation. The energy or mean deviation of the detail images are the most usually used features for texture classification and segmentation problems (Kundu and Chen, 1992; Chang and Kuo, 1993; Smith and Chang, 1994; Laine and Fan, 1996). Kokare et al., (2002), have used the decomposition scheme based on M-band wavelets, resulting an improvement in retrieval performance. Unlike the standard wavelet decomposition, which gives a logarithmic frequency resolution, the M-band decomposition gives a mixture of a logarithmic and linear frequency resolution. Further as an additional advantage, wavelet decomposition yields a large number of subbands, which improves the retrieval accuracy. One of the drawbacks with M-band wavelet in content-based image retrieval is that computational complexity and hence the retrieval time increases with the number of bands.

Smith and Chang, (1994 and 1996a) have used the statistical parameters (mean and variance) extracted from the wavelet subbands for image retrieval. This approach achieved over 90% accuracy on the 112 Brodatz texture images. Chang and Kuo in 1993 explored the middle-band characteristics and a tree-structured wavelet transform was used to further improve the classification accuracy. The wavelet transform also can be combined with other techniques to achieve a better performance like Gross et al., (1994) using the wavelet transformation together with KL expansion and Kohonen maps. Thyagarajan et al., (1994) and Kundu et al., (1992) combined the wavelet transform with a co-occurrence matrix to take advantage of both statistics-based and transform-based texture analyses. Many other texture analysis methods based on discrete wavelet representation can be found in the references Chang and Kuo, (1993); Liang and Kuo, (1999). The Lifting scheme (Daubechies and Sweldens, 1998; Sweldens, 1996; Uytterhoeven, 1997), a novel approach for constructing the so-called second-generation wavelet transform, provides reasonable alternative for the problems faced by the traditional first-generation wavelet transform in image applications.
The curvelet transform based multiresolution technique was first purposed by Candes and Donoho in (2000). Many authors have attempted to use curvelet transform in their study (Starck et al., 2002; Ni and Leng 2003; Candes et al, 2006; Sumana et al., 2008). Curvelet was formerly proposed for image denoising and has shown promising performance. Ni and Leng, (2003) attempted an initial application of curvelet on color image retrieval, but it was not implemented correctly and insignificant result was reported. In their work, no standard image database was used, no retrieval accuracy was reported and there was no comparison with other techniques. (Sumana et al., 2008) reported that the image retrieval using curvelet transform gives better result as compare to Gabor filter and Wavelet transform. Shen and Yin, (2009) applied curvelet transform on a set of texture based images. Das et al., (2012) also compared wavelet transform with curvelet transform and found that curvelet give better result.

Huang and Dai, (2003) proposed image retrieval system based on two features, the composite sub-band gradient vector and the energy distribution pattern string. Both features are generated from the sub-images of a wavelet decomposition of the original image. A fuzzy matching mechanism based on energy distribution pattern strings serves as a filter to rapidly remove undesired images in the database for further consideration. The images passing the filter will be compared with the query image based on composite sub-band gradient vectors which are very powerful for discriminating detailed textures.

Jafari-Khouzani and Soltanian-Zadeh, (2005) proposed a rotation-invariant texture-analysis technique using radon and wavelet transforms. This technique utilizes the radon transform to convert the rotation to translation and then applies a translation-invariant wavelet transform to the result to extract texture features. A nearest neighbor’s classifier is employed to classify texture patterns. It was shown that the extracted features generate an efficient orthogonal feature space. It was also shown that the proposed features extract both of the local and directional information of the texture patterns.

c) Shape Feature

Like color and texture features, the shape also plays a key role in CBIR. Zhang and Lu, (2004) presented in their review paper the shape can be classified into two categories,
region based methods and contour based methods. The classification is based on whether shape features are extracted from the contour only or are extracted from the whole shape region. The counter based shape methods consist of global shape descriptors (Niblack *et al.*, 1993), shape signatures (Otterloo 1991; El-Ghazal *et al.*, 2007), autoregressive model (Kauppinen *et al.*, 1995), spectral descriptors (Huang and Huang, 1998; Mehrotra and Gary, 1995) and Curvature Scale Space (CSS) methods (Abbasi *et al.*, 1999; Abbasi *et al.*, 2000; Zhang and Lu, 2003). The common global contour shape descriptors are area, eccentricity, circularity, major axis orientation and bending energy (Young *et al.*, 1974) and Peura and Liverinen, (1997) proposed others simple global contour shape descriptor such as convexity, ratio of principle axis, circular variance and elliptic variance. Shape signature represents a shape by a one dimensional function derived from shape boundary points. Shape signature includes centroidal profile, centroid distance, complex coordinates, cumulative angle, curvature, chord length and area. Shape signatures are normalized into being translation and scale invariant (Zhang and Lu, 2004). Asada and Brandy, (1986) have developed the concept of “curvature primal sketch” which has been further adopted by Mokhtarian and Mackworth (1986 and 1992). They have shown that curvature inflection points extracted using a Gaussian scale space can be used to recognize curved objects. One difficulty with this approach is that curves without inflection points fall into the same equivalence class. Dudek and Tsotsos, (1997) presented a technique for shape representation and recognition of objects based on multiscale curvature information. The curvature scale space descriptor treat shape boundary as a 1D signal, and analyses this signal in scale space (Jalba *et al.*, 2006). Spectral descriptor includes Fourier Descriptor (FD) and Wavelet Descriptor (WD), one of the most usually used shape descriptor techniques is Fourier Descriptor (Otterloo, 1991). Mitchell and Grogan, (1984) used Fourier descriptor to describe partial shapes. Fourier descriptor plays a significant role as shape feature in CBIR (Mehrotra and Gary, 1995; Huang and Huang 1998; Lu and Sajjanhar, 1999; Zhang and Lu, 2003; Zhang and Lu, 2005; Yadav *et al.*, 2007; Shahabi and Safar, 2007; Ekombo *et al.*, 2009; Amanatiadis *et al.*, 2011). Eichmann *et al.*, (1990) proposed the Short-time Fourier Descriptor (SFD) for shape description, and Zhang and Lu, (2001a) have found that SFD is out performed by conventional FD methods in shape retrieval. Recently Wavelet Descriptor (WD) has
been used for shape based image retrieval. Chuang and Kuo, (1996) have found that WD has the advantage over FD.

In region based methods, the shape descriptors utilize information from the internal regions of the shape. The region based methods contain geometric moments Hu, (1962), Zernike moments (Teague, 1980; Mehtre et al., 1997), and grid representation Lu and Sajjanhar, (1999). The primary approach used in region-based techniques is the use of invariant moments. Hu in 1962 presents the first paper on the image moment invariants for two-dimensional pattern recognition application. Kim et al., (2000) proposed Zernike moments that are invariant to rotation, translation, and scale by calculating moments based on the unit disk of the image in polar coordinates. The drawbacks with Zernike moments are the inherent redundancy in the features and the examination of shape information along the radial directions (Li et al., 2005). A critical issue with this approach is the inordinate amount of computation required to extract the Zernike moments. Zhang and Lu, (2002 a, b) proposed a Generic Fourier descriptor (GFD) for shapes by applying 2D Fourier Transform (FT) on polar shape images. The FT magnitude coefficients are taken as the image descriptors. Compared with Zernike moments, the method has no redundant features and allows multiresolution analysis in both radial and angular directions. Their experiments showed that the GFD outperformed Zernike moments in the test for rotation, scale, and translation invariance. Li et al., (2005) proposed compound image descriptor (CID) and compared with GFD and found that the CID is more robust to changes caused by image scaling, translation, and/or rotation; the CID also describes the image more accurately. Li et al., (2009) proposed Invariant Zernike Moments Descriptor (IZMD), by combining the Zernike moments magnitude and the corrected phase. The scale and translation invariance of IZMD could be obtained by pre-normalizing the image using the geometrical moments. From their experiment they shows that IZMD is more robust than ZM descriptor to changes caused by image scaling and rotation; it could represent images more accurately and is also more robust to image noise. Bai et al., (2010) proposed the group contextual information on different shapes to improve the efficiency of shape retrieval on several standard data sets (Latecki et al., 2000; Sebastian et al., 2004). The basic idea was to use shapes as each other’s contexts in propagation to reduce the distances between intra class objects. The implementation was done by a graph-based transduction approach, named Label Propagation (LP) (Zhu et al., 2005). Egozi et al.,
(2010) proposed a contextual similarity function, named Meta similarity, which characterizes a given object by its similarity to its K-nearest neighbor (K-NN) objects. An interesting distance learning method called contextual dissimilarity measure (CDM) (Jegou et al., 2010) is motivated by an observation that a good ranking is usually not symmetrical in image search, which is mainly designed for the image search problem. Bai et al., (2012) again proposed Co-Transduction for shape/object retrieval algorithm. The algorithm used to fuse different similarity measures for robust shape retrieval through a semi supervised learning frame work. Their approach is inspired by the co-training algorithm (Blum and Mitchell, 1998) and they showed that Co-Transduction/tri-transduction gives better result from their earlier technique (Bai et al., 2010) on same database.

2.1 Some of the commercial/test based image retrieval systems are explain in the following section.

I. QBIC (Query By Image Content)

QBIC (Daneels et al., 1993; Faloutsos, et al. 1994, Equitz and Niblack, 1994; Lee et al., 1994; Niblack et al., 1993; Flickner et al., 1995) is the first proposed commercial content-based image retrieval system. It used query based on example images, user-constructed sketches and drawings, and selected color and texture patterns. Color features calculated using the 3D average color vector of an object or RGB, YIQ, LAB, and Munsell color space in entire image using 256 bin color histogram (Faloutsos et al., 1994). The texture features are computed in QBIC (Equitz and Niblack, 1994) are the coarseness, contrast, and directionality features proposed by Tamura et al., (1978). The shape feature consists of circularity, shape area, major axis orientation, eccentricity, and an algebraic moment invariants (Scassellati et al., 1994; Faloutsos et al., 1994). 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/.

II. CANDID (Comparison Algorithm for Navigating Digital Image Databases)

CANDID was proposed by Kelly et al., (1995) and developed in Computer Research and Applications Group, Los Alamos National Laboratory, USA. Each image is represented by a signature consisting of a weighted sum of Gaussian functions. The
color features used are the clusters of spectral bands. The texture features used are the mean, skewness, and kurtosis of each pixel. The feature vectors of all pixels together form a point set in higher-dimensional space. On the basis of the k-means algorithms, clusters are formed. A mean vector and covariance matrix are computed for each cluster, and the associated Gaussian distribution is weighted by the number of elements in the corresponding cluster. The distribution of the feature vectors is now approximated by the weighted sum of the Gaussian distributions. The dissimilarity between two image signatures is based on the normalized Euclidean distance or the inner product of two signatures. CANDID is used in the retrieval of pulmonary CT images, and multispectral Landsat satellite images. On line demo found in URL http://public.lanl.gov/kelly/CANDID/index.shtml.

III. Photobook

Photobook was proposed by Pentland et al., in 1996, and developed at the MIT Media Laboratory, Cambridge, MA. It is a set of interactive tools for browsing and searching images. Photobook consists of three subbooks from which 2D shapes and texture, faces images features are extracted, respectively. Based on the corresponding features in each of the three subbooks, users can query. In its next version of Photobook, Four Eyes, including human in the image annotation and retrieval loop by Picard, (1996); Picard et al., (1996); Liu and Picard, (1996). They proposed a ‘society of model’ approach to incorporate the human factor. Experimental results show that this approach is efficient in interactive image annotation (Minka and Picard, 1997; Picard et al., 1996). The face recognition technology of Photobook has been used by Viisage Technology in a FaceID package, which is used in several US police departments. The on line demo found in URL http://vismod.www.media.mit.edu/vismod/demos/photobook/index.html.

IV. VisualSEEk

VisualSEEk developed at Image and Advanced Television Lab, Columbia University, (Smith and Chang, 1997a; Smith and Chang. 1997b) is a visual feature search engine and also the WebSEEk (Smith and Chang, 1997c) is a World Wide Web oriented text/image search engine. Main research features are spatial relationship query of image regions and visual feature extraction from compressed domain (Wang and Chang, 1995; Chang, 1995). They used color set and wavelet transform based texture feature in
their systems for visual features (Smith and Chang, 1994; Smith and Chang, 1995). They also developed binary tree based indexing algorithms, to speed up the retrieval process (Smith and Chang, 1996b). The color set similarity is computed by using quadratic distance. The on-line demos are available at http://www.ee.columbia.edu/sfchang/demos.html.

V. VIR Image Engine

The VIR Image Engine, developed by Gupta and Jain (1997), is an extensible framework for building content based image retrieval systems. Virage supports visual queries based on color, composition (color layout), texture, and structure (object boundary information). The users can regulate the weights associated with the atomic features according to their own importance. The VIR Image Engine provides a set of GUI tools essential for the development of a user interface. These include facilities for image insertion, image query, weight adjustment for re-query, inclusion of keywords, and support for several popular image file formats. Another existing component, the query canvas, allows queries-by-sketch; it consists of a bitmap editor where the user can sketch a picture with drawing tools and color it using the colors from existing collection and modify it using the same drawing tools. Queries can be performed on various user-defined combinations of primitives. The on-line demos of Virage are at http://www.virage.com/products/vir-irw.html.

VI. Netra

Netra is a prototype image retrieval system developed in the UCSB Alexandria Digital Library (ADL) by Ma and Manjunath in 1997b. Netra uses color, texture, shape, and spatial location information in the segmented image regions to search and retrieve similar regions from the database. For color feature they used color codebook which is represented by a sub set of colors in each of image region, and the codebook is constructed using the Generalized Lloyd Algorithm (GLA) (Gersho and Gray, 1992) to vector quantize colors in the RGB color space. The texture features are based on Gabor decomposition (Alexandrov et al., 1995; Ma and Manjunath, 1995), neural net-based image thesaurus construction (Ma and Manjunath, 1996a; Ma and Manjunath, 1996b) and edge flow-based region segmentation (Ma and Manjunath, 1997a). For shape they
used contour representations, which contain curvature function, centroid distance, and complex coordinate function. The on-line demo is at http://vivaldi.ece.ucsb.edu/Netra.

VII. ImageRover

ImageRover system was proposed by Sclaroff et al., (1997), and developed in Boston University. It combines textual and visual statistics in a single index for content-based search of a web image database. Textual statistics are captured using Latent Semantic Indexing (LSI) based on text in the containing HTML document. Visual statistics are captured using color and texture orientation histograms. To initiate a search of the ImageRover index, the user specifies few keywords describing the desired images. Later the user can refine his query through relevance feedback. Both visual and textual cues are used in the relevance feedback loop. Based on relevance feedback from the user, the system selects the appropriate normalized Minkowski metric each time a query is made. The online demo is at http://www.cs.bu.edu/groups/ivc/ImageRover/demo.html.

VIII. FOCUS (Fast Object Color-based Query System)

FOCUS was proposed by Das et al., (1997) and developed at Department of Computer Science, University of Massachusetts, Amherst, MA. In the proposed method, each image is divided into cells of 100X100 pixels and for each cell a color histogram in the HSV space, roughly quantized along the saturation and value axes (64X10X10), is computed. The peaks of all local histograms are determined and combined in a list of unique peaks for the whole image by merging multiple copies of the same peak. Also, a frequency table is constructed which, for each color in the HSV space, gives the number of images that have a peak of that color. The spatial relationships between color regions are represented by means of a Spatial Proximity Graph (SPG) constructed in two phases. First an intermediate SPG is generated, with one node corresponding to each color peak computed for the image cells. Two nodes in this graph are connected if their corresponding peaks are located in the same cell or are located in neighboring cells and have the same color. This graph is then simplified, by unifying all connected nodes of the same color in a single node, and stored using an adjacency matrix representation. For the query image, a global color histogram is computed and color
region relationships are determined at pixel level. The on-line demo is at http://cowarie.cs.umass.edu/colordemo/mdas/demo1/phase0.html.

IX. FIR (Formula Image Retrieval)

FIR system was proposed by Volmer (1997) and developed by Fraunhofer Institute for Computer Graphics, Darmstadt, Germany, in association with Txt Ingegneria Informatica S.P.A. (Italy), Giunti Multimedia Srl (Italy), Epsilon Software (Greece), and Kino TV & Movie Productions S.A. (Greece), as part of the Esprit IV project FORMULA. Using multiresolution wavelet transformation decomposition, shape and color information of an image are presented. They convert RGB color space to Luv color space. Then two-dimensional Haar wavelet transform is performed on each color channel individually, followed by a truncation of the resulting coefficients. Only the coefficients larger than a threshold are retained and after rounding to integer values they are stored in a feature vector. The distance measure between two feature vectors is a weighted Euclidean distance, with different weights for each of the resolution levels in the wavelet decomposition. The on-line demo is at http://www.igd.fhg.de/igd-a7/projects/formula/formula_e.html.

X. MARS (Multimedia Analysis and Retrieval System)

MARS was developed at Department of Computer Science, University of Illinois at Urbana-Champaign, further developed at Department of Information and Computer Science, University of California at Irvine, CA. (Huang et al., 1997b; Mehrotra et al., 1997a; Mehrotra et al., 1997b; Rui et al., 1997a). The Mars system supports queries by example, on combinations of low-level features such as color, texture, shape and textual descriptions. Color is represented using a 2D histogram over the HS coordinates of the HSV color space. Texture is represented by two histograms, one measuring the coarseness and the other one the directionality of the image, and one scalar defining the contrast. 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 combination of DBMS and IR (exact match with ranked retrieval) (Huang et al., 1997a; Ortega et al.,
1997), integration of indexing and retrieval (how the retrieval algorithm can take benefit of the underline indexing structure) (Rui et al., 1997b), and integration of computer and human. The focus of MARS is not only finding a single “best” feature representation, but also how to organize various visual features into a meaningful retrieval architecture which can dynamically adapt to different applications and different users. The similarity distance between two color histograms is computed by histogram intersection. The similarity between two textures of the whole image is determined by a weighted sum of the Euclidean distance between contrasts and the histogram intersection distances of the other two components, after a normalization of the three similarities. For computing the texture similarity between two corresponding subimages, the Euclidean distance between the vector representations is used. MARS also proposed relevance feedback architecture in image retrieval (Rui et al., 1998) and integrates such a technique at various levels during retrieval, including query vector refinement (Rui et al., 1997c), automatic matching tool selection (Rui et al., 1997a), and automatic feature adaption (Rui et al., 1997d; Rui et al., 1998). The on-line demo is at http://jadzia.ifp.uiuc.edu:8000.

XI. C-bird (Content-Based Image Retrieval from Digital libraries)

C-bird (Li et al., 1998; Li, et al., 1999) developed at School of Computing Science, Simon Fraser University, Burnaby, B. C. Canada. For each collected image, a feature descriptor and a layout descriptor are computed. A feature descriptor is a set of four vectors: a color vector, a Most Frequent Color (MFC) vector, a most frequent orientation (MFO) vector, and a chromaticity vector. A 512-bin RGB histogram is stored in the color vector. The centroids of the regions associated with the 5 most frequent colors form the MFC vector and the centroids of regions of the 5 most frequent edge orientations form the MFO vector. The 36-dimensional chromaticity vector is computed as follows: first, a normalization of each RGB channel is made to obtain illumination invariance, and then the 3D color histogram is replaced by a 2D chromaticity histogram. Treating this chromaticity histogram as an image, first a wavelet-based image reduction is applied, and then the Discrete Cosine Transform coefficient matrix is built. The chromaticity vector is made of the 36 values of the upper left corner of the DCT matrix. For search by object model, some geometric data such as the area, the centroid and the eccentricity are computed from color regions...
associated with each of the MFCs. The layout descriptor contains a color layout vector and an edge layout vector. To construct these vectors the image is divided into 64 cells, and for each cell the most frequent colors and the number of edges for each orientation are determined. Also, for images at half and quarter resolution, a feature descriptor like the one described above is stored. The distance between two chromaticity vectors in an illumination invariant color query is the L2 distance. Texture orientation histograms, as well as color histograms for the full image, are matched by histogram intersection. On line demo is present at http://www.cs.sfu.ca/cbird/.

XII.  FIDS (Flexible Image Database System)

FIDS system is proposed by Berman and Shapiro, (1998), developed at Department of Computer Science and Engineering, University of Washington, Seattle, WA, USA. FIDS uses the following features, such as color histogram and the fraction of flesh colored pixels; the histogram of values is used after applying the Sobel edge filter, also the coefficients of the Haar wavelet decomposition. These features are also taken of subimages in a grid, in rows, and columns. The distances between the histograms are the L1-distance. The distance between wavelet coefficients, is some weighted difference. An overall distance can be composed by taking the weighted sum, maximum, or minimum of the individual feature distances, which preserves metric properties. The on-line demo is at http://www.cs.washington.edu/research/imagedatabase/.

XIII.  VIPER (Visual Information Processing for Enhanced Retrieval)

The VIPER network based at the University of Geneva. These include a web-based standard for Query By Example (QBE)-based CBIR systems, image browser benchmarking (Muller et al., 2000, 2004). The group is also behind the development of the Multimedia Retrieval Mark Up Language (MRML) that aims to provide a unified interface for multimedia retrieval and management software. The web-accessible standard for QBE-based CBIR systems was designed to allow developers of image retrieval systems the opportunity to benchmark their systems online at any time. The Benchathlon began with BIRDS-I and was an initial step towards a standardized benchmark for CBIR systems. BIRDS-I was an image retrieval benchmark that was presented as a contest during EI 2001. Participants were required to implement an
image retrieval server to be tested against a ground-truth via a set of defined metrics. The Benchathlon aimed to develop a networked system benchmark for CBIR, along the lines of existing benchmarks for text retrieval and relational database management. (Muller et al., 2005) reported that while the Benchathlon initiated discussion amongst participants no systematic comparison between systems was started. The on-line demo is at http://viper.unige.ch/.

XIV. **iPURE (Perceptual and User-friendly Retrieval of Images)**

iPURE was proposed by Aggarwal et al., (2000) developed IBM India Research Lab, New Delhi, India. Images are segmented using the Mean Shift Algorithm. For each region, color, texture, shape and spatial lay-out features are extracted. The color features are the average Luv color space. Texture features that are used are the coefficients of the Wold decomposition of the image viewed as a random field. The shape features are size, orientation axes, and Fourier descriptor. The spatial lay-out is captured by the centroid, the minimum bounding box, and contiguity. Individual segments are matched by computing a weighted Euclidean distance on the feature vectors. Then the contiguity amongst the retrieved segments is matched against those of the query segments. The on-line demo is at URL http://www.research.ibm.com/irl/projects/mediamining/index.html.

XV. **PicHunter**

PicHunter system was proposed by Cox et al., (2000), developed by NEC Research Institute, Princeton, NJ, USA. The image features used in PicHunter system are color, spatial distribution along with hidden textual annotations. HSV color histogram, HSV color auto-correlogram and RGB Color-Coherence Vector (CCV) are describing the color content of an image. The distance between individual features (color vectors or annotation lists represented as binary vectors) is the L1 distance. These distances are scaled and combined in a global distance. The scaling factors are computed by maximizing the probability of a training set.

XVI. **FRIP (Finding Regions in the Pictures)**

FRIP system was proposed by Ko et al., (2001) and developed at Yonsei University, Korea. In FRIP system, images are segmented into regions using some circular filter.
From each region, color, texture, and shape features are extracted. For color feature, the average color in RGB space is used. Biorthogonal wavelet frame is used for the texture feature. And the shape features are eccentricity, and centroid-boundary distances. In addition, the normalized area, centroid, and length of the major and minor axis are stored. The distance between two color feature vectors is the city block (L1) distance. The distances between two texture ratios, between two normalized areas, and between two eccentricities, are the absolute difference of these two values. The distance between two centroids is the Euclidean (L2) distance. The on-line demo is at URL http://vip.yonsei.ac.kr/frip/.

XVII. MPEG-7

The ISO MPEG Group proposed the ‘MPEG-7 Multimedia Description Language’. It defines standardized descriptions and description systems that allow users to search, identify, filter, and browse visual content (Sikora, 2001). Images are retrieved by defining and the combination of description including color patches or textures and the color and shape of the object, object size, etc. Color descriptors used are Scalable color descriptor which is based on color histogram encoded by Haar transform, dominant color descriptors based on clustering color regions, color layout descriptors, CSD, and GOF/GOP color descriptors. Shape descriptors supported in MPEG-7 are Shape spectrum, Region-based descriptor based on moment invariant based, Contour-based shape descriptor based on curvature scale space based, and 2D/3D shape descriptors. To allow for maximum flexibility for design of applications, MPEG-7 specifies only minor parts of the descriptor extraction methods and not how similarity between content should be measured (Sikora, 2001). On line demo available at http://give-lab.cs.uu.nl/mpeg7-ce-shape1b/.

XVIII. Simplicity

Simplicity system was proposed by Wang et al., (2000), developed a pre-processing phase, before extracting its feature vector of an image they first classified image as texture vs. non-texture, or graph vs. photograph.. They also segment the image into regions. Some of the features are extracted directly from the color components and some from the wavelet transform of the L component, whereas, the shape/non-texture feature is derived from the normalized inertia. The advantage of Simplicity is its ability
to limit the search space by grouping images that belong to the same category together. However, the disadvantage is the time needed to perform the classification. This is done by calculating a feature vector for every 4X4 pixel block and then clustering them to find out which category the image belongs to.

XIX. Blobworld

Blobworld is an image retrieval system proposed by Carson et al., (2002) developed at UC Berkeley. It uses the Expectation Maximization (EM) algorithm to segment the images into regions of uniform color and texture (blobs). The color is described by a histogram of 218 bins of the color coordinates in Lab-space. Texture is represented by mean contrast and anisotropy over the region. Shape is represented by approximate area, eccentricity, and orientation. Query-by-example is performed based on a region from one of the images presented to the user. Blobworld allows the user to view the internal representation of the submitted image and the query results; facilitating the understanding of the retrieval results. A demo of Blobworld is available at http://elib.cs.berkeley.edu/photos/blobworld/start.html

XX. SMURF (Similarity-based Multimedia Retrieval Framework)

SMURF is proposed by Vleugels and Veltkamp, (2002) developed at Center for Geometry, Imaging, and Virtual Environments, Institute of Information and Computing Sciences, Utrecht University, The Netherlands. It is a framework that incorporates color, texture, and shape. In particular there are demos for polyline matching, and weighted point set matching. Matching of two polylines is done by computing the area between the two corresponding turning angle functions, and minimizing over rotations (vertical shift of the function), and starting position (horizontal shift). The weighted point set matching is done with a pseudo metric, the Proportional Transportation Distance. Unlike the Earth Movers Distance, it obeys the triangle inequality for sets of unequal weight, which is essential for the indexing method used. On line demo available at URL http://give-lab.cs.uu.nl

XXI. PicSOM

PicSOM (Koskela, 2003; Laaksonen et al., 2002) is an image browsing system based on the Self-Organizing Map (SOM), developed at the Laboratory of Computer and
Information Science at Helsinki University of Technology, Finland. The SOM is used to organize images into map units in a two-dimensional grid so that similar images are located near each other. PicSOM uses the average RBG color feature, texture feature, Fourier-based shape features and MPEG-7 features. PicSOM applies a tree-structured version of the SOM algorithm (Tree Structured Self-Organizing Map, (TS-SOM)) to create a hierarchical representation of the image database. During the queries, the TS-SOMs are used to retrieve images similar to a given set of reference images. As a basis for retrieving images, the PicSOM system uses a combination of several types of statistical features, which are computed from the image content. Separate feature vectors have been formed for describing colors, textures, and shapes found in the images. A distinct TS-SOM is constructed for each feature vector set and these maps are used in parallel to select the retrieved images. The image queries are performed through the Web interface and the queries are iteratively refined as the system exposes more images to the user. Image retrieval with PicSOM is an iterative process utilizing the relevance feedback approach. The image query begins with an initial set of images uniformly elected from the database. On subsequent rounds, the query focuses more on the user’s need. This is achieved as the system learns the user’s preferences from the selections made during previous rounds. It is worth noting that the Euclidean distance, inherently used in the PicSOM system, may be inappropriate for certain types of features.

XXII. ImageCLEF (Cross Language Evaluation Forum)

ImageCLEF is the image retrieval track of the Cross Language Evaluation Forum (CLEF) (Clough et al., 2004; Muller et al., 2004). It is not strictly a CBIR benchmarking event as it allows the use of meta-data, i.e., text that appears around the image or in the image title. The primary purpose of the CLEF campaign is to promote the multi-lingual searching of such data; however, a secondary goal is to investigate combining text and CBIR. ImageCLEF offers two main image retrieval tasks, one over collections of photographic images and one over collections of medical images. Since its inception in 2003, ImageCLEF has drawn interest from both academics and commercial research organizations from the areas of CBIR, Cross-Language Information Retrieval, and user interaction.
XXIII. WALRUS

In WALRUS (Natsev et al., 1999), while moving a window with proper size throughout a whole image, wavelet transform is applied to the region in the moving window, then, coefficients of the lowest frequency band is used as features. Since wavelet transform is applied for each of the many overlapping windows, the computational cost is prohibitively high. Although they proposed a dynamic-programming algorithm to decrease the cost, it is still high. In addition, in this method texture features cannot be extracted, since it uses coefficients of limited frequency subbands. To determine a proper window size is also difficult. Natsev et al., (2001) in WALRUS tackle the problem of finding objects of different sizes and locations within an image. They do that by calculating the wavelet transform of sliding windows or variable sizes in an image. A feature vector consisting of the sum wavelet coefficients is calculated for each sliding window. Then, the feature vectors are clustered to determine the regions in an image belonging to the same object. In the image matching phase, the regions of the images are compared and the one with the highest percentage of matching regions with the query image is returned as a query match. The advantage of WALRUS is its ability to retrieve matching images of different object size and location within images. The disadvantage of WALRUS is the high computational complexity due to wavelet transforms of the sliding windows and the large search space from all the sub-images generated.

XXIV. WINDSURF (Wavelet-Based Indexing of Images Using Region Fragmentation)

In WINDSURF (Ardizzoni et al., 1999), the method uses the wavelet transform to extract color and texture features from an image and applies a clustering technique to partition the image into a set of “homogeneous” regions. Similarity between images is assessed by using the Bhattacharyya distance to compare region descriptors, and then combining the results at image level. To fragment each image into regions, we use a clustering algorithm on the coefficients obtained through the DWT. They use a simple k-means algorithm with a validity function which is slightly different from that proposed for the fuzzy k-means (Xie and Beni, 1991).
XXV. Integrated Region Matching (IRM)

Li et al., (2000) proposed Integrated Region Matching (IRM) allows for matching a region of one image to several regions of another image. That is, the region mapping between any two images is a many-to-many relationship. As a result the similarity between two images is defined as the weighted sum of distances in the feature space, between all regions from different images. Compared with retrieval systems based on individual regions, such as Blobworld, the IRM approach decreases the impact of inaccurate segmentation by smoothing over the imprecision in distances. IRM incorporates the properties of all the segmented regions so that information about an image can be fully used. Region-based matching is a difficult problem because of inaccurate segmentation. To define the similarity measure, first regions in two images are matched. Being aware that segmentation cannot be perfect, the matching is “softened” by allowing one region of an image to be matched to several regions of another image. Here, a region-region match is obtained when the regions are relatively similar to each other in terms of the features extracted. IRM first segments the image into blocks of 4X4 pixels and extracts a feature vector for each block. The k-means algorithm is used to cluster the feature vectors into several classes with every class corresponding to one region in the segmented image. Six features are used for segmentation. Three of them are color components (L*u*v color space), and the other three represent energy in high frequency bands of the wavelet transform (Daubechies-4 wavelet transform to the L component of the image).

XXVI. CIRES

CIRES system for content-based retrieval in digital image libraries is developed by Iqbal and Aggarwal, (2002). CIRES used image structure in addition to color and texture for feature extraction. In image structure they used edge by line segments, longer linear lines, retained lines, coterminations, “L” junctions, “U” junctions, parallel lines, parallel groups, “significant” parallel groups and polygons. For color, CIE LAB space is used in histogram and for texture Gabor filters have been used. And Cries was updated by using multi image query (Iqbal and Aggarwal, (2003) in that the integration of features extracted from three different techniques, and the proposed mechanism of multi-image queries, and two different techniques for relevance feedback, on the system performance. Structure, color, and texture features were extracted from an
image. They have also weighted of different features by a user based upon images content to improved retrieval system. On line demo available at http://amazon.ece.utexas.edu/~qasim/research.htm.

XXVII. Fuzzy Club

Fuzzy Club image retrieval is proposed by Zhang and Zhang (2002). Fuzzy Club addresses the issue of effective and efficient content based image retrieval by presenting an indexing and retrieval system that integrates color, texture and shape information for the indexing and retrieval, and applies these features regions obtained through unsupervised segmentation, as opposed to applying them to the whole image domain. Fuzzy Club emphasizes improving on a color feature “inaccuracy” problem in the region based literature that color histogram bins are not independent. For instance, if the color spectrum is divided into 10 bins, these bins are not independent some are closer or farther away from each other in the original color space. Fuzzy logic is applied to the traditional color histogram to solve this problem to some degree. Fuzzy Club first segments an image is segmented into regions of 4X4 blocks and extract color and texture features on each block. The k-means algorithm is used to cluster similar pixels together to form a region. The LAB color space is used to extract color features and Haar wavelet transform is used to extract three texture features.

XXVIII. Cortina

Cortina is a large-scale image retrieval system for the World Wide Web proposed by Quack in 2004. The system retrieves images based on visual features and collateral text. Methods are introduced to investigate these multi-modal characteristics of the data and to gain insights into the semantics within the data. They showed that a search process which consists of an initial query-by-keyword and followed by relevance feed-back on the visual appearance of the results is possible for large-scale data sets. They also show that it is superior to the pure text retrieval commonly used in large-scale systems. The precision is shown to be increased by exploiting the semantic relationships within the data and by including multiple feature spaces into the search process. System has been enhanced by Gelasca et al., (2007), at the systems level, the components of Cortina include building image collections using a Web crawler, collecting category information and keywords, and processing images to compute content descriptors.
Functionalities of Cortina include duplicate image detection, category and image content based search, face detection and relevance feedback. A MySQL database is used for storing textual annotations and keywords, whereas the image features are stored in flat file structures. This combination appears to be effective and scalable for large collection of image/video data and is easily parallelizable. On line demo available at http://vision.ece.ucsb.edu/multimedia/cortina.shtml

XXIX. FReBIR (Fuzzy Region-Based Image Retrieval system)

FReBIR (Philipp-Foliguet et al., 2006) is a region-based image retrieval system, in which images are represented as adjacency graphs of fuzzy regions. The goal of this retrieval module is to retrieve images containing a specific object. An algorithm to match sub-graphs of fuzzy regions is then applied in order to retrieve images from partial queries, taking into account the image composition. Images are represented by a set of fuzzy regions, with their features and the composition of the image is stored into an attributed relational graph (ARG) of regions, aiming at representing the relative positions of regions. Which has been further improved in by (Philipp-Foliguet et al., 2009) they present a method of image indexing and retrieval which takes into account the relative positions of the regions within the image. Indexing is based on a segmentation of the image into fuzzy regions. The image retrieval is based on inexact graph matching, taking into account both the similarity between regions and the spatial relation between them. Graph matching, and on the other hand, a measure of similarity between graphs allowing the result. The method is adapted to partial queries, aiming for example at retrieving images containing a specific type of object. Applications may be of two types, firstly an on-line search from a partial query, with a relevance feedback aiming at interactively leading the search, and secondly an off-line learning of categories from a set of examples of the object.

XXX. RETIN

The RETIN (Gony et al., 2007) search engine, developed in ETIS lab, addresses the problem of category search. The system uses an active learning scheme based on binary classification in order to interact with a user looking or documents concepts in databases. In category search, each document has to be classified as belonging or not to the category. The objective of the statistical methods is to update a relevance function
or a binary classification of document using the user labels. For binary classification use a SVM with a Gaussian $\gamma^2$ kernel (Gosselin and Cord, 2006). For color and texture feature extraction, HSV space is used color and for texture analysis, Gabor filters are used with twelve different scales and orientations. Signatures are statistical distributions of colors and textures resulting from a dynamic quantization of the feature spaces. That means that we use color and texture space clustering to compute the image histograms. Both spaces are clustered using an enhanced LBG algorithm. A web graphical interface of RETIN is already available at http://dupont.ensea.fr/~ruven/new.php.

XXXI. CEDD (Color and Edge Directivity Descriptor)

In CEDD (Chatzichristofis and Boutalis, 2008) developed an image retrieval system by extracting low level feature such as “Color and Edge Directivity Descriptor” and incorporates color and texture information in a histogram. CEDD size is limited to 54 bytes per image, rendering this descriptor suitable for use in large image databases. To extract the color feature from histogram using the HSV color space, a set of fuzzy rules undertake the extraction of a Fuzzy-Linking histogram that was proposed by (Chatzichristofis and Boutalis, 2007). Edge Histogram Descriptor (Won et al., 2002) is also used for exporting the texture feature of the image and the Gustafson Kessel fuzzy classifier (Gustafson and Kessel, 1978) for extracting shaped region. On line demo available at: http://orpheus.ee.duth.gr/image_retrieval.

XXXII. Img(Anaktisi)

img(Anaktisi) is a web content based image retrieval system proposed by Zagoris et al., (2009). img(Anaktisi) is developed in C#/NET. It provides efficient retrieval services for various image databases using as a query a sample image, an image sketched by the user and keywords. The image retrieval engine is powered by novel compact and effective descriptors. Also, an Auto Relevance Feedback (ARF) technique is provided to the user. This technique readjusts the initial retrieval results based on user preferences improving the retrieval score significantly. On line demo can be found at http://www.anaktisi.net.
XXXIII. MammoSys

In MammoSys (De Oliveira et al., 2010) present a content-based image retrieval system designed to retrieve mammographies from large medical image database. The system is developed based on density; according to the four categories defined by the American College of Radiology integrated to the database of the Image Retrieval in Medical Applications (IRMA) project provides images with classification ground truth. Two-Dimensional Principal Component Analysis (2DPCA) (Yang et al., 2004) is used in breast density texture characterization, in order to efficiently represent texture and allow for dimensionality reduction. A support vector machine (Vapnik, 1999) is used to solve a variety of learning, classification, prediction problems and the retrieval process.

XXXIV. JPSearch

The JPSearch system was proposed by Temmermans et al., (2012). System contained composed of six parts, in Part 1: System framework and components, in Part 2: Registration, identification and management of schema and ontology, in Part 3: Query format, in Part 4: File format for metadata embedded in image data (JPEG and JPEG 2000), Part 5: Data interchange format between image repositories, and in Part 6: Reference software. The Part 1 provides motivation and overview of the JPSearch framework. The Part 2 supports interoperability among various metadata specifications and social tagging using the specification of core metadata and translation rule description language. The Part 3 provides powerful image query language for interoperability among multiple image databases. The Part 4 specifies a file format based on JPEG and JPEG 2000 file format in which arbitrary number of metadata description of an image can be embedded. The Part 5 supports interoperability in interchanging images with metadata between image repositories. Finally, the Part 6 provides reference and utility software based on which the JPSearch compliant systems.