A Survey on Content Based Image Retrieval

I. Felci Rajam[1] and S. Valli[2]

[1]Department of MCA, St. Joseph's College of Engineering, Jeppiaar Nagar, Chennai-600119, India and Research Scholar, Department of CSE, College of Engineering, Guindy, Anna University Chennai, Chennai-600025, Tamil Nadu, India

[2]Department of CSE, College of Engineering, Guindy, Anna University Chennai,Chennai, 600025, Tamil Nadu, India

Email: felcirajam@gmail.com , valli@annauniv.edu

Abstract: Content Based Image Retrieval (CBIR) is a very important research area in the field of image processing, and comprises of low level feature extraction such as color, texture and shape and similarity measures for the comparison of images. Recently, the research focus in CBIR has been in reducing the semantic gap, between the low level visual features and the high level image semantics. This paper provides a comprehensive survey of all these aspects. This survey covers approaches used for extracting low level features; various distance measures for measuring the similarity of images, the mechanisms for reducing the semantic gap and about invariant image retrieval. In addition to these, various data sets used in CBIR and the performance measures, are also addressed. Finally, future research directions are also suggested.

[I. Felci Rajam, S. Valli. **A Survey on Content Based Image Retrieval.** *Life Sci J* 2013; 10(2): 2475-2487]. (ISSN: 1097-8135). http://www.lifesciencesite.com 343

Keywords: content-based image retrieval (CBIR), affinity matrix, semantic cluster matrix (SCM), invariant moments, relevance feedback (RF), invariant retrieval

1. Introduction

Digital images are currently widely used in medicine, fashion, architecture, face recognition, finger print recognition and bio-metrics etc. Hence, efficient image searching and retrieval are important. The earlier image retrieval systems were text based. Images were represented by using keywords. The keyword for the image was created by human operators. Manually entering keywords for images in a large database can be inefficient, expensive and may not capture every keyword that describes the image. Therefore, Content Based Image Retrieval (CBIR), based on the image content came into existence. Ying Liu et al., 2007 surveyed the CBIR system focusing on high level semantic concepts. Kekre H.B. et al., 2011 discussed feature extraction, distance measures, classifier techniques such as neural network classifiers, K-nearest neighbor algorithm and the performance measures.

The CBIR system relies on color, texture and shape which are low level image features. Section 2 discusses the various possible low level features available in the literature. The low level features are extracted from the database images and stored in a feature database. Similarly, the low level features are extracted from the query image and the query image features are compared with the database image features using the distance measure. Images having the least distance with the query image are displayed as the result. The various popular distance measures reported in the literature are presented in Section 3. The main drawback of the CBIR system is that the images with similar low level features may vary from the query image in terms of the semantics perceived by the user. This problem is known as the 'Semantic gap' problem (Rahman. M. M. et al., 2007, Hui Hui Wang et al., 2010). Hence, reducing the semantic gap between the low level image features and the high level image concepts, became a very interesting and challenging area of research. It is called as Semantic Content Based Image Retrieval (SCBIR) and the various techniques available in SCBIR are presented in Section 4.

The CBIR system must be invariant to the geometric transformations. Hence, this paper discusses invariant image retrieval in Section 5. Finally, the different datasets used in the various CBIR systems are discussed in section 6, and the performance measures to analyze the CBIR systems are presented in section 7. Section 8 gives the comparison of the various CBIR components used in recent papers. Section 9 concludes the paper.

2. Low Level image Features used in CBIR

In CBIR systems, a feature is a characteristic that can capture a certain visual property of an image either globally for the entire image or locally for regions or objects (Selvarajah S. et al., 2011, Kodituwakku S.R., 2010). The low level features commonly used in CBIR are color, texture, shape and edge. The following subsections address the color, texture, shape and edge features used in CBIR.

2.1 Color Features

Color features are extracted using color moments, color histogram, invariant color histogram, and dominant color. These methods of extracting color features are explained in the following section.

2.1.1 Color Moments

The color distribution of the image is characterized by its moments. The first, second and third central moment of each of the color channels is stored as a color feature. If the value of the i^{th} color channel at the j^{th} pixel is p_{ij}, then the first moment mean (μ_i) is given by equation (1). The second moment standard deviation (σ_i) is given by equation (2). The third moment skewness (γ_i) is given by equation (3) (Rahman et al., 2007, Rahman et al., 2009, Hatice Cinar Akakin et al., 2012).

$$\mu_i = \frac{1}{N}\sum_{j=1}^{N} P_{ij} \qquad (1)$$

$$\sigma_i = \left(\frac{1}{N}\sum_{j=1}^{N}\left(P_{ij} - \mu_i\right)^2\right)^{\frac{1}{2}} \qquad (2)$$

$$\gamma_i = \left(\frac{1}{N}\sum_{j=1}^{N}\left(P_{ij} - \mu_i\right)^3\right)^{\frac{1}{3}} \qquad (3)$$

N in equation (1) to (3) is the total number of images.

2.1.2 Color Histogram

The histogram of an image is a graph which contains the occurrence of each intensity value found in that image, obtained by counting all image pixels having that intensity value. For an 8-bit grayscale image there are 256 different possible intensities. So, the histogram will graphically display 256 grayscale values showing the distribution of pixels amongst those numbers. Histograms can also be taken of color images. A color histogram is the representation of the distribution of colors in an image. It is a standard statistical description of the color distribution in terms of the occurrence frequencies of the different regions in a color space (Imtnan-Ul-Haque et al., 2011). To create a color histogram, the color space has to be partitioned into regions. The 24 bit RGB color space has 2^{24} different color regions. A histogram containing 2^{24} bins is too large to be dealt. Hence the color space is quantized into a number of bins, where each bin represents a range of color values. The number of pixels in the image that falls in each of these ranges is counted to get the color histogram. The number of bins is decided based on the loss of precision tolerated and the memory requirement. Color histograms can be built in various color spaces (Deselaers T. et al., 2007, Neetu Sharma et al., 2011, Javad Kangarani Farahani et al., 2012).

The authors Lining Zhang et al., 2012 utilized the color histogram to represent the color information. They found the color histogram in HSV color space. The hue and saturation are quantized into eight bins and value into four bins. Imtnan-Ul-Haque et al. 2011, used three dimensional (3D) histogram as a color feature. The number of bins for each of the color channel is kept common. They have used three color spaces namely RGB, Improved Hue, Luminance and Saturation (IHLS) and L*a*b.

2.1.3 Invariant Color Histogram

Theo Gevers et al., (2004) proposed a robust histogram from photometric color invariants (invariant to illumination, shading, highlights and inter reflections) for object recognition. The histograms are computed by the variable kernel density estimators. The variable kernel density estimator is given in equation (4)

$$\hat{f}(x) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{\alpha(X_i)}K\left(\frac{x-X_i}{\alpha(X_i)}\right) \qquad (4)$$

Here, kernel K is a function satisfying $\int K(x)dx = 1$. The kernel centered on Xi, has its own scale parameter $\alpha(X_i), i = 1, \ldots, n$. For color images, the scale parameter is a function of the RGB-values and the color space transform. This histogram is invariant to illumination, shading, highlights and reflections.

2.1.4 Dominant Color

In region based image retrieval, the regions are segmented and the features are extracted for the regions. Due to the inaccuracy of the segmentation, the average color of a segmented region may be different from that of the original region. To obtain the dominant color of the image, first the histogram is obtained and then the bin with the maximum size is taken as the dominant color of the region. When the segmented region does not have a homogeneous color, then, the average color will not be a good choice for the color feature (Ying Liu, et al., 2008).

2.2 Texture Features

Texture features are extracted using Gray Level Co-occurrence matrix (GLCM), Gabor Transform and Tamura Features. These methods of extracting texture features are explained in the following section.

2.2.1 Gray Level Co-occurrence matrix (GLCM)

The GLCM is created from a gray-scale image. The GLCM finds how often a pixel with a gray-level value i occurs either horizontally, vertically, or diagonally to adjacent pixels with the value j. It is given by the relative frequency of the occurrences of two gray-level pixels i & j, separated by d pixels in the θ orientation, where d is the displacement and θ is the direction. The 'd' can take values 1, 2, 3, etc., and θ can take values 0 °(horizontal), 90 °(vertical), 45 °and 135 ° (diagonal) (Rahman et al., 2007). The construction of the GLCM is shown in Figure 1. Several statistical texture properties like contrast, correlation, energy, homogeneity and entropy can be derived from the GLCM and the formulae are given in equations (5) through (9) (Haralick et al., 1973, Rahman et al., 2009, Najlae Idrissi, 2009).

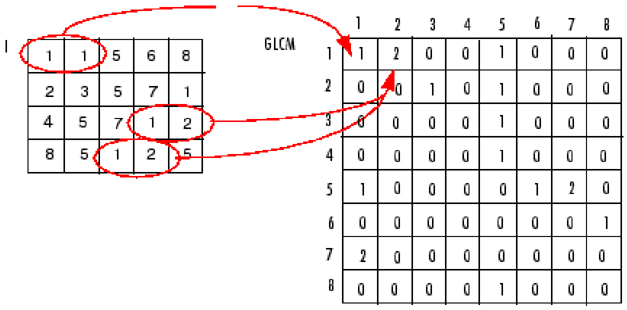

Figure 1. Construction of GLCM

$$Contrast = \sum_i \sum_j |i-j|^2 \, p(i,j) \quad (5)$$

$$Correlation = \sum_i \sum_j \frac{(i-\mu_i)(j-\mu_j)p(i,j)}{\sigma_i \, \sigma_j} \quad (6)$$

$$Energy = \sum_i \sum_j p(i,j)^2 \quad (7)$$

$$Homogeneity = \sum_i \sum_j \frac{p(i,j)}{1+|i-j|} \quad (8)$$

$$Entropy = \sum_i \sum_j p(i,j) \log\big(p(i,j)\big) \quad (9)$$

2.2.2 Gabor Transform

For an image I(x, y) of size P × Q, its discrete Gabor wavelet transform is given by equation (10).

$$g(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right) + 2\pi jWx\right] \quad (10)$$

where W is the modulation frequency σ_x, and σ_y characterizes the spatial extent and frequency bandwidth of the Gabor filter.

If g(x,y) is the mother wavelet transform, then the Gabor filter $g_{mn}(x,y)$ is obtained by equation (11).

$$g_{mn}(x,y) = a^{-2m} g(x',y'), \ a > 1 \quad (11)$$

where $x' = a^{-m}(x\cos\theta + y\sin\theta)$, $y' = a^{-n}(x\sin\theta + y\cos\theta)$, m and n are the scale and orientation of the wavelet with m = 0, 1, …., M-1; n=0, 1, …., N-1, θ =nπ / N and $a = \left(U_n/U_l\right)^{\frac{1}{M-1}}$. U_n and U_l represent the higher and lower frequencies of interest.

For an image I(x, y) with size P × Q, its Gabor wavelet transform is given by equation (12) and g_{mn}^* is the complex conjugate. The mean and standard deviation of the magnitude are used to represent the homogeneous texture feature of the region. The mean and standard deviation are calculated using equations (13) and (14) respectively.

$$G_{mn}(x,y) = \sum_s \sum_t I(x,y) g_{mn}^*(x-s, y-t) \quad (12)$$

$$\mu_{mn} = \frac{\sum_x \sum_y |G_{mn}(x,y)|}{P \times Q} \quad (13)$$

$$\sigma_{mn} = \sqrt{\frac{\sum_x \sum_y (|G_{mn}(x,y)| - \mu_{mn})^2}{PQ}} \quad (14)$$

The texture feature vector f_t is calculated using μ_{mn} and σ_{mn} as feature components. For M scales and N orientations, the texture feature vector f_t is given by equation (15) (Rahman. M. H. et al, 2011, Rahman. M. H. et al, 2012, Selvarajah S.et al., 2011).

$$f_t = \left[\mu_{00}, \sigma_{00}, \mu_{01}, \dots \mu_{(M-1)(N-1)}, \sigma_{(M-1)(N-1)}\right] \quad (15)$$

2.2.3 Tamura Features

Coarseness, contrast, directionality, line-likeness, regularity and roughness are the six Tamura features. Coarseness, contrast and directionality correlate strongly with the human perception, and hence they are very important.

2.3 Shape Features

Shape features are extracted using many approaches. They are one-dimensional functions for shape representation, polygonal approximation, spatial interrelation feature, moments, scale-space methods and shape transform domains (Yang Mingqiang, et al., 2008). There is no general feature which works best for every kind of image. An appropriate shape feature has to be chosen depending upon the situation and the nature of the image (Zhang D. et al., 2004). Some of the shape features are discussed in this section.

2.3.1 Histogram of Edge Directions

The edge histogram captures the general shape information in the image. The edge information contained in the image is obtained, using edge detection algorithms like canny, sobel, etc. The edge directions are quantized into a number of bins (Rahman et al., 2007, Felci Rajam. I et al., 2011b, 2012). For achieving scale invariance, the histogram is normalized with respect to the number of pixels in the image. The histogram is smoothed to make it robust to rotation.

2.3.2 Region Moments

Among region-based descriptors, moments are very popular. These include invariant moments, Zernike moments and Legendre moments.

2.3.2.1 Invariant Moments

Invariant moments or geometric moments are the simplest moment functions with the basis $\Psi_{pq}(x,y) = x^p y^q$. The geometric moment function m_{pq} of order (p+q) is defined by the equation (16).

$$m_{pq} = \sum_x \sum_y x^p y^q f(x,y) \ \ , p,q = 0,1,2,\dots \quad (16)$$

The geometric central moments that are invariant to translation are defined by equation (17)

$$\mu_{pq} = \sum_x \sum_y (x-\bar{x})^p (y-\bar{y})^q f(x,y), \quad p,q = 0,1,2,\dots \quad (17)$$

where $\bar{x} = m_{10}/m_{00}$ and $\bar{y} = m_{01}/m_{00}$

The seven invariant moments are given by the following equations (18) to (24).

$$I_1 = \eta_{20} + \eta_{02} \quad (18)$$

$$I_2 = (\eta_{20} - \eta_{02})^2 + (2\eta_{11})^2 \quad (19)$$

$$I_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (20)$$

$$I_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (21)$$

$$I_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (22)$$

$$I_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \quad (23)$$

$$I_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \quad (24)$$

where $\eta_{p,q} = \frac{\mu_{p,q}}{\mu_{0,0}^{\gamma}}, \gamma = 1 + \frac{p+q}{2}$, for p + q=2, 3, …

Invariant moments are invariant to translation, rotation and scaling (Yang Mingqiang, et al., 2008, Dudani S.A et al., 1977, Asmatullah Chaudhry et al., 2012).

2.3.2.2 Zernike moments

Zernike moments are derived from the orthogonal Zernike polynomials. Hence, it is an orthogonal moment. The Zernike moments are given by equation (25).

$$V_{nm}(x, y) = V_{nm}(r\cos\theta, r\sin\theta) = R_{mn}(r)\exp(jm\theta) \quad (25)$$

where $R_{mn}(r)$ is the orthogonal radial polynomial, and is given by equation (26).

$$R_{mn}(r) = \sum_{s=0}^{(n-|m|)/2} (-1)^s \frac{(n-s)!}{s! X \left(\frac{n-2s+|m|}{2}\right)! \left(\frac{n-2s-|m|}{2}\right)!} r^{n-2s} \quad (26)$$

n=0, 1, 2,…; 0≤$|m|$≤n; and n-$|m|$ is even.

The Zernike moments for the image f(x,y) are defined by equation (27).

$$Z_{nm} = \frac{n+1}{\pi} \sum_r \sum_\theta f(r\cos\theta, r\sin\theta).R_{nm}(r).\exp(jm\theta) \quad r \le 1 \quad (27)$$

Zernike moments are invariant to rotation, translation and scaling. Also they are robust to noise and minor variations in shape. But the computational complexity of Zernike moments is high (Yang Mingqiang, et al., 2008, Saptadi Nugroho, 2011).

2.3.2.3 Legendre Moments

Legendre moments use Legendre polynomials as the kernel function. The two-dimensional Legendre moments of order (p + q) for an image f(x,y) are defined by equation (28).

$$L_{pq} = \frac{(2p+1)(2q+1)}{4} \int_{-1}^{1} \int_{-1}^{1} P_p(x) \times P_q(y) f(x, y) dx dy \quad x, y \in [-1, 1] \quad (28)$$

where the Legendre polynomial $P_p(x)$ of order p is given by equation (29).

$$P_p(x) = \sum_{k=0}^{p} \left((-1)^{\frac{p-k}{2}} \frac{1}{2^p} \frac{(p+k)! x^k}{\left[\frac{p-k}{2}\right]! \left[\frac{p+k}{2}\right]! k!} \right) \quad (29)$$

The Legendre moments described in equation (28) can be expressed in discrete form by equation (30).

$$L_{pq} = \lambda_{pq} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P_p(x_i) P_q(y_j) f(i, j) \quad (30)$$

where $\lambda_{pq} = \frac{(2p+1)(2q+1)}{N^2}$. x_i and y_i are the normalized pixel coordinates, and are given by equation (31) (Ch. Srinivasa Rao et al., 2010).

$$x_i = \frac{2i}{N-1} - 1 \text{ and } y_i = \frac{2j}{N-1} - 1 \quad (31)$$

3. Distance Measures

Distance measures are used for comparing the similarity of two images. There are different kinds of similarity measurements like Euclidean distance, histogram intersection, Bhattacharya distance and Mahalanobis distance for CBIR applications.

3.1 Euclidean Distance

Let *q* be the query image and *t* be the target image and let $p_q(z_q)$ and $p_t(z_t)$ be their respective probability densities. The Euclidean distance between the query image and the target image is given by equation (32).

$$D_{Euc}(q_i, t_i) = \sum_{i=1}^{n} (q_i - t_i)^2 \quad (32)$$

In Euclidean distance, the least value of distance measure indicates the similarity (Rahman et al., 2007, Felci Rajam. I et al., 2011a, 2012).

3.2 Histogram Intersection

It is a distance measure for comparing histograms. It calculates the common part of the two histograms, and neglects the features occurring in a single histogram. The histogram intersection of two histograms H and H' is calculated using equation (33) (Deselaers T et al., 2007).

$$d_\cap(H, H') = \sum_{m=1}^{M} min(H_m, H'_m) \quad (33)$$

3.3 Bhattacharya Distance

The Bhattacharya Distance measures the similarity between two discrete or continuous probability distributions. A popular distance of similarity between two Gaussian distributions is the Bhattacharya distance. The Bhattacharya distance (D_{Bhatt}) between the query image q and the target image t in the database is given by equation (34).

$$D_{Bhat}(q, t) = \frac{1}{8}(\mu_q - \mu_t)^T \left[\frac{\Sigma_q - \Sigma_t}{2}\right]^{-1} (\mu_q - \mu_t) + \frac{1}{2} \ln \frac{|(\Sigma_q + \Sigma_t)/2|}{\sqrt{|\Sigma_q||\Sigma_t|}} \quad (34)$$

where μ_q and μ_t are the mean vectors, and Σ_q and Σ_t are the covariance matrices of the query image *q* and the target image *t*, respectively (Rahman et al., 2007, 2009, Felci Rajam. I et al., 2011a, 2012).

3.4 Mahalanobis Distance

The Mahalanobis Distance is based on the correlations between variables, and is used to analyze various patterns. It is useful in determining the similarity between an unknown sample set and a known one. The unknown sample set is the query image, and the known set is the images in the database. The Mahalanobis distance between the query image q and the target image t is given by equation (35) (Prabir Bhattacharya et al., 2006).

$$D_{Maha}(q,t) = (\mu_q - \mu_t)^T \Sigma^{-1} (\mu_q - \mu_t) \qquad (35)$$

4. Semantic Content Based Image Retrieval

Many techniques have been evolved for reducing the 'semantic gap' between the low level image features and high level semantics. The CBIR system which reduces this semantic gap is called as the semantic CBIR. This section describes the various techniques that are developed in reducing the semantic gap.

4.1 Machine Learning Techniques

The high level semantic features are derived from the image DB with the help of machine learning techniques. Supervised and unsupervised machine learning techniques are available.

4.1.1 Supervised Machine Learning techniques

Neural networks, Decision trees, and Support Vector Machines (SVMs) are some of the supervised machine learning techniques, which learn the high level concepts from low-level image features. The supervised machine learning techniques perform the classification process with the help of the already categorized training data. For the training data, the input (low level image features) and the desired output (category of the image (ie) the high level semantic concept) is already known. When the supervised learning algorithms are trained with the known training data, it is able to generalize the new unseen data. Hence, given a query image, the low level features are extracted and it is given as input to any one of the machine learning algorithms which is already trained with the training data. The machine learning algorithm predicts the category of the query image which is nothing but the semantic concept of the query image. Hence instead of finding similarity between the query image and all the images in DB, it is found between the query image and only the images belonging to the query image category. Also when the entire DB is searched, the retrieval result contains images of various categories. But when the machine learning techniques are used, since the query image's category (semantic concept) is predicted, the retrieval results will contain the images belonging to that category alone.

4.1.1.1 Neural Network

Neural networks are also useful in concept learning. The low level features of the segmented regions of the training set images are fed into the neural network classifiers, to establish the link between the low level image features and high level semantics. The disadvantage of this method is that it requires a large amount of training data, and is computationally intensive (Liu, Y et al., 2007, Aymn E.Khedr et al., 2012, Thenmozhi et al., 2012). When the query image feature vector is presented to the neural network, it gives its semantic concept.

4.1.1.2 Decision Tree Learning Methods

Decision tree (DT) learning methods such as ID3, C4.5 and CART perform well in data classification. The ID3 algorithm requires the value of the input attributes to be discrete. The ID3 finds the most useful attribute in classifying the given set. The attribute with the maximum information gain is the most useful attribute. The C4.5 algorithm can handle continuous attributes. At each node of the tree, C4.5 chooses one attribute of the data that most effectively splits the set of samples into two subsets. The attribute with the highest normalized information gain is chosen to make the decision. The Classification and Regression Tree (CART) is a binary decision tree. It is constructed by splitting a node into two child nodes, repeatedly beginning with the root node that contains the whole learning sample. The CART algorithm will itself identify the most significant variables and eliminate the non-significant variables (Ying Liu et al., 2008). Decision Tree – Semantic Template (DT-ST) method was proposed by Ying Liu et al., 2008. The sample regions are extracted from the training set images. Then, the color and texture features are extracted from the sample regions, and the color and texture semantic templates are constructed for each concept. The training samples with discrete attribute values are used to construct the DT. The DT-ST method outperforms the ID3 and C4.5 algorithms (Liu, Y et al., 2007). The constructed decision tree is used to find the semantic concept of the query image.

4.1.1.3 Support Vector Machine (SVM)

SVM was first designed for binary classification. Later, the SVM multi class classifiers were developed using the SVM binary classifiers. In the One-against-all (OVA) multi class classifier, N binary classifiers are constructed for each of the N classes. Each binary SVM takes one class sample as the positive example, and all the remaining class samples as negative example. A data point is classified under a certain class, if and only if, the corresponding class SVM accepts it, and all the other SVMs reject it. The disadvantage is that more than one SVM may accept or all SVMs may reject. In such a case, the data point cannot be classified. The One-against-one (or) Pair Wise Coupling (PWC) constructs $N*(N-1)/2$ SVM binary classifiers. For each pair of classes, there is one classifier. The class which gets the highest vote is the class of the data point. Hence, when a query image feature vector is given to the SVM, it predicts its class (semantic concept) (Madzarov G et al., 2009).

4.1.1.4 SVM-Binary Decision Tree

The SVM-BDT takes advantage of the efficient computation of the binary tree architecture, and the high classification accuracy of the SVMs. Here, (N-1) SVMs are needed to train an N class problem. For the construction of the SVM-BDT, first, the semantic template of each of the categories or classes is found. Construction of semantic template is explained in section 4.4. The semantic template of the i^{th} class is given by $(\bar{f}_1^i, \bar{f}_2^i, \ldots, \bar{f}_m^i)$. The Euclidean distance between the semantic templates of each of the N classes is the $N \times N$ distance matrix. Two classes that have the largest Euclidean distance are assigned to each of the two clustering groups. The semantic template of these two classes is the cluster center of the two groups. The class closest to each of the cluster centers is found, and assigned to the corresponding group. The cluster center is updated to the semantic template of the newly added class. All the classes are assigned to one of the two possible groups of classes. The process continues, until there is only one class per group. The SVM binary classifier is used to train the samples in each non leaf nodes of the decision tree (Felci Rajam I. et al., 2011a).

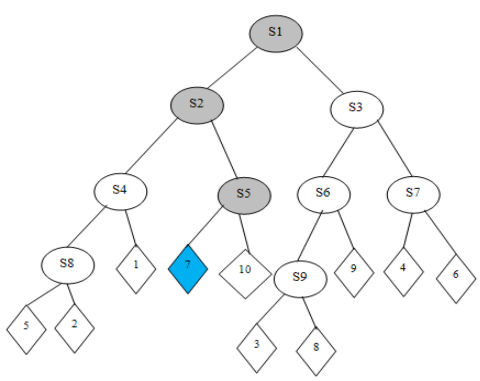

Figure 2. SVM-BDT for 10 category image set

During the testing time, the feature vector of the query image is given as input to the SVM-BDT, and only $\lceil log_2 N \rceil$ classifiers are consulted during the testing time. The SVM-BDT predicts the label (semantic category) of the query image. Hence, the similarity distance between the query image and the predicted category images are computed and the images with least distance are retrieved. An example of a 10-class SVM-BDT is shown in Figure 2. For SVM-based image classification, recent work shows that the radial basis kernel function (RBF) works well, when the relation between the class labels and attributes is nonlinear (Rahman et al., 2007).

4.1.2 Unsupervised Learning

Unsupervised learning refers to the problem of trying to find hidden structure in the unlabeled data. It has no measurements of outcome, to guide the learning process. Image clustering is a typical unsupervised learning technique. It groups the sets of image data in such a way, that the similarity within a cluster should be maximized, and the similarity between different clusters must be minimized (Liu. Y et al., 2007).

K-means clustering aims to partition the given n observations into k clusters. The mean of each cluster is found and the image is placed in a cluster, whose mean has the least Euclidean distance with the image feature vector. Due to the complex distribution of the image data, the k-means clustering often cannot separate images with different concepts well enough (Liu, Y et al., 2007).

NCut clustering is used to cluster the database images into different semantic classes. A set of n images is represented by a weighted undirected graph G = (V, E). V = { 1, 2, …, n} represents images; the edges E= {(i, j) / i, j Є V}are formed between every pair of nodes. The weight w_{ij} of an edge (i, j) is a function of the distance between those two nodes (images) d(i, j). The system displays the image clusters and adjusts the model of similarity measure according to user feedbacks. (Zakariya S. M., 2010).

Fuzzy clustering models provide a promising solution to the clustering problem. The Fuzzy c-means (FCM) clustering is the most widely used fuzzy clustering algorithm. This algorithm is based on an iterative optimization of a fuzzy objective function. The degree of membership of a data item to a cluster is between [0, 1]. For a given query image, the output of the FCM is the membership value of the image with each of the K classes. The query image belongs to the class for which the membership value is high (Rahman et al., 2007, Felci Rajam I. et al., 2011b, Celia B et al., 2010).

Hence the unsupervised learning algorithms takes the uncategorized image data as input and clusters those data into a number of categories by finding the hidden structure in the unlabelled data. The clustering algorithms divide the given data into n clusters and give cluster centers of each cluster. When a query image features are given to the clustering algorithm, it finds the distance between the query image and all the cluster centers. The query image belongs to the cluster for which the distance is a minimum. Hence, here also the similarity distance is found between the query image and the images belonging to the predicted cluster alone.

4.2 Relevance Feedback (RF)

The relevance feedback mechanism was initially used in retrieving documents, and later was used for retrieving images. This mechanism brings the user in the retrieval loop to reduce the semantic gap between the low level features, and what the user thinks. In RF, initially the system provides the retrieval results with the help of distance measures or some of the machine learning techniques or the affinity matrix etc. The user states whether the retrieval results are relevant (positive results) or irrelevant (negative results) to the query. Such relevance feedback is used to perform a new search with the modified search parameters, (ie), the position of the query point, the similarity metric and other tuning parameters. The machine learning algorithms can also be refined to learn the user's feedback. The feedback can be got from the user again and again, till the user is satisfied with the results. By this continuous learning process through interaction with the end user, the performance of the CBIR system is boosted (Giacinto, G. et al., 2005, Das. G. et al., 2007).

Query reweighting (QR) and Query Point Movement (QPM) are some of the relevance feedback mechanisms, which are discussed in Liu. Y. et al., 2007. Query Expansion (QEX) is another relevance Feedback technique. The RF strategies QR and QPM did not completely cover the user's interest in the broad feature space. QEX groups the similar relevant points into several clusters, and selects good representative points from these clusters to construct the multipoint query. QEX is more effective than QPM and QR (Ja-Hwung Su et al., 2011). Samuel Rota Bulò et al., 2011 proposed a novel approach to content-based image retrieval with relevance feedback, which is based on the random walker algorithm. Ja-Hwung Su et al., 2011, proposed a Navigation-Pattern-based Relevance Feedback (NPRF), to achieve high efficiency in CBIR. The NPRF search makes use of the discovered navigation patterns and three kinds of query refinement strategies, namely QPM, QR, and QEX to converge the search space toward the user's intention effectively.

4.3 Affinity Matrix

The affinity matrix is useful in capturing the high level semantic concepts in the image. The rows of the affinity matrix contain the query regions, and the columns contain all the images in the database. Entries in the affinity matrix are the accumulated scores acquired through the Relevance Feedback (RF). Initially all the entries of the affinity matrix are set to zero. During the testing phase, if the user marks a certain image as 'positive' to the query image, the score of the corresponding image column is increased by 1. If the image is marked as negative, then the corresponding entry in the matrix is decreased by 1. The affinity matrix is filled by integers that represent the semantic closeness of the image region with the database images (Liu et al., 2009). Given a query image, the user selects the region of interest. The distance between the query region and all other semantic regions are found. The query region belongs to the cluster for which the distance is less. The database images which have the positive affinity values constitute the search space. The similarity distance between the query image and the images in this reduced search space is found and the top 'k' images with least distances are displayed and the feedback is updated in the affinity matrix.

4.3.1 Markov Model Mediator (MMM)

Shyu et al. (2007) proposed a Markov model mediator (MMM) to facilitate the capturing of high level image concepts in the CBIR. It learns the high level concepts of the images from the history of the user access patterns, and the access frequencies of the images in the database. The affinity relationship among the images in the database is contained in the relative affinity matrix A. The training process was done off-line. The user selects one query image. The query message is sent to the server. The server retrieves the related images from the database and sends the query results to the client. Upon receiving the results, user selects the relevant images. This feedback is sent back to the server. When the server receives this feedback message, it updates the user access patterns and access frequencies accordingly (Shyu et al., 2007).

4.3.2 Semantic Clustering Scheme

Liu et al. (2009) proposed a semantic clustering scheme for region based image retrieval. The semantic similarities between the image regions are obtained with the help of the user query and feedback history. The affinity matrix is used to store the users' feedback about the image. Initially, the query region is entered as a new semantic cluster and its feature vector is treated as the centroid of the semantic cluster. For every image that is labeled as positive, the system finds out which region of the positive image has the shortest Euclidean distance from the query region, and puts this positive region in the same cluster of the query region. All the other regions in that image have the label 'unknown'. All the regions in the images marked as negative are marked as 'negative' in the corresponding query region. Given a query image, the user has to select the region of interest from the query image. The Euclidean distance between the query region and all the other semantic regions in the affinity matrix is used to find the cluster to which the query image belongs. Then the positive images in that cluster are matched with the query image and the top images with high similarity are displayed (Liu et al., 2009).

4.4 Semantic Template (ST)

ST is defined as the 'representative' feature of a concept, calculated from a collection of sample images. For a concept, the ST is defined as the centroid of the low level features of all the sample regions of the concept (Ying Liu et al., 2008). If the color and texture features are given by $\{c_1, c_2, \ldots, c_{n1}\}$ and $\{t_1, t_2, \ldots, t_{n2}\}$, and suppose there are n sample regions for each concept, then, the centroid of the first dimension of the color and texture features can be calculated by the following equation (36) and (37) respectively.

$$\bar{c_1} = \frac{1}{n}\sum_{i=0}^{n} c_1^i \quad (36)$$

$$\bar{t_1} = \frac{1}{n}\sum_{i=0}^{n} t_1^i \quad (37)$$

The color semantic template and texture semantic template are given by equation (38) and (39) (Ying Liu et al., 2008).

$$C_i = (\bar{c_1}, \bar{c_2}, \dots, \overline{c_{n1}}) \quad (38)$$

$$T_i = (\bar{t_1}, \bar{t_2}, \dots, \overline{t_{n2}}) \quad (39)$$

The color threshold (texture threshold) for class i is the maximum distance between any 2 color (texture) feature vectors in the cluster and is defined by the equation (40) and (41) respectively.

$$R_i^c = max_{j=0}^{n} d_{c(j,i)}^i \quad (40)$$

$$R_i^t = max_{j=0}^{n} d_{t(j,i)}^i \quad (41)$$

The semantic template is a feature which collectively represents a semantic concept. Given a query image, the color and texture features are extracted from it. The Euclidean distance between the color (texture) feature of the query region and the color(texture) semantic template of each concept is obtained. The label of the query image is found by the equation (42).

$$CTlabel = \begin{cases} m, & if\ d_{c(j,i)}^m \le R_m^c\ and\ d_{t(j,i)}^m \le R_m^t \\ -1, & Otherwise \end{cases} \quad (42)$$

Hence the label of the query image (ie) the semantic concept of the query image is found and the searching is restricted to that category images alone. Hence semantically relevant images will be retrieved in lesser amount of time.

4.5 Semantic Cluster Matrix (SCM)

Machine learning methods provide good results, when the query image belongs to an already trained category. But if the query image is an unknown image, then the retrieval performance is poor. There is a mismatch in even the top level semantic category. This is called as 'Category mismatch'. Felci Rajam I. et al., 2012 designed a matrix called the Semantic Cluster Matrix (SCM) for training the system adaptively at the testing time, to avoid the problem of 'category mismatch'. The SCM combines the power of the affinity matrix and the relevance feedback mechanism to provide efficient retrieval, not only for the already trained category images, but also for the untrained category images. The SCM contains semantic information, such as the semantic template for the semantic cluster, semantic threshold values, label of the cluster, number of regions in the cluster, category of the semantic cluster, the affinity matrix for the cluster and the feature vectors of all the regions belonging to this semantic cluster. Each row contains the details of one semantic cluster. The system is initially trained using the SVM-BDT. For a known category query image, the SVM-BDT provides good results. But for an unknown category image, the SVM-BDT fails. In such case, the SCM is searched. The distance between the query image feature vector and the semantic template of each cluster in the SCM is found. Then, the distances are compared with the semantic thresholds of the clusters and the distances which are within the semantic thresholds are alone considered. The categories of these clusters are analyzed and the category which gets the highest vote is taken and among these, the label of the least distance cluster is picked as the closest semantic cluster to the query image. The positive images in the affinity matrix of that cluster are alone compared with the query image to provide the target results. If the SCM produces correct results then, the feedback is obtained from the user and the corresponding semantic cluster is updated with the query image information. The SCM is also initially trained with the known category images. Hence, at the beginning, the SCM also fails to provide the correct results. In such a case, the entire DB is searched and the results are displayed. The user gives the feedback about the positive images; then, that category of images is searched and the results are displayed. Once again the user gives feedback and the query image information becomes a new cluster in the SCM (Felci Rajam I. et al., 2012).

4.6 Object Ontology

Object ontology is another means of reducing the semantic gap in image retrieval. The semantics of an image is nothing but the description of the image. For example, the sea can be described as 'lower, uniform and blue region', and the sky can be described as 'upper, uniform, and blue region'. The low level features are associated with different intervals of the intermediate level descriptors, such as 'light blue, medium blue, dark blue'. The object ontology provides a qualitative definition of the high level query concepts. The database images are classified into different categories, by associating such descriptors with the images based on knowledge (Liu, Y. et al., 2007). In this method, the ontology is developed to allow the user to query an image database, using the semantically meaningful concepts. An example of object ontology used by V. Mezaris et al., is shown in Figure 3. In this system, each region of an image is described by its average color, its position in vertical and horizontal axis, its size and shape. The query will be given using key words. An example query is "Find images with upper, large, light blue region". The images of the database are also described with these keywords. Since the DB images are categorized based on the object ontology, when a query is given, the images that have the matching descriptors to the query are retrieved. Hence the system reduces the semantic gap and support query by keywords.

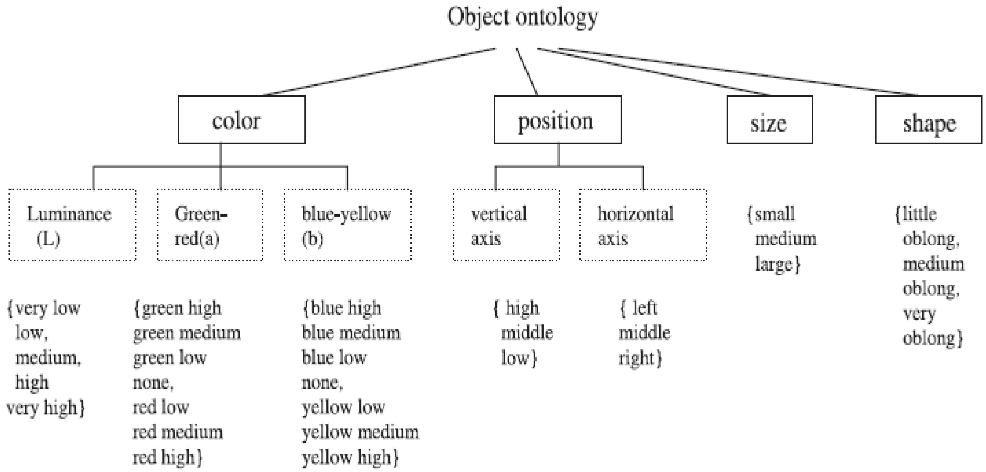

Figure 3. An example of Object Ontology

5. Invariant Image Retrieval

The existing image retrieval methods utilize the features of an image to describe and retrieve similar images. However, many features lack invariance when geometrical transformations are applied on an image. This will result in the mismatch of the retrieval process when the image's orientation, and position or scales are altered. This section discusses some of the content-based image retrieval methods, when the query image is subjected to geometric transformations.

For a given query image, the retrieval method (Cheng-Hao Yao et al., 2003) retrieves the images that are subjected to translation, rotation and/or scaling. This retrieval method is based on the similarity measures on feature distributions, and retrieves texture images from a database of color textures. The similarity measure is the histogram intersection technique, and this similarity measure is extended to retrieve the texture regions from a database of natural images. This method is invariant to translation, scaling and rotation, but it is only suitable for texture based retrieval. Zhe-Ming Lu et al., (2006) proposed an image retrieval system based on rotation, scaling and translation invariant features, by performing the Log-Polar transformation on the image.

Saptadi Nugroho et al., (2011) proposed a rotation invariant indexing for images, using Zernike moments and the R–tree. The Zernike moments has a rotation invariant characteristic. The R-tree algorithm stores and searches the magnitude of the Zernike moments. Theo Gevers et al., (2004) proposed a robust histogram from photometric color invariants (invariant to illumination, shading, highlights and inter reflections) for object recognition. This method is robust under Gaussian noise.

6. Datasets Used

For performing experiments in the CBIR, many datasets are available. Many of the researchers used the COREL dataset (Yixin Chen, 2003, Ja-Hwung Su, 2011, Yu Ma et al., 2010), which contains natural scenery images. The Corel image database contains a large number of images of various contents, ranging from animals and outdoor sports to natural scenes. These images are pre-classified into different categories of size 100 by domain professionals. The Caltech dataset (Samuel Rota Bul et al., 2011, Felci Rajam. I et al., 2012) is a benchmark dataset, which contains natural images and it has 101 different object categories. The Oliva Dataset, Flickr18 (Nikhil Rasiwasia, 2007), and IAPR TC-12 benchmark also contain natural image datasets. The SIVAL (Spatially Independent, Variable Area, and Lighting) benchmark consists of 25 different objects, and 10 scenes. SIVAL emphasizes the task of localized CBIR through nearly identical scenes, which only vary by the localized target objects (Feng et al., 2010). The Vistex texture database (Raghu Krishnapuram et al., 2004, Imtnan-Ul-Haque et al., 2011), Outex database (Imtnan-Ul-Haque et al., 2011), Brodatz texture database (Yu Ma et al., 2010), and Meastex (Najlae Idrissi et al., 2009) are the texture databases. ImageCLEFmed contains medical images. The MNIST, Pendigit, Optdigit and Statlog are databases containing handwritten characters and digits (Gjorgji Madzarov et al., 2009).

7. Performance measures

The performance of the CBIR system can be analyzed, using several performance metrics. Many of the CBIR systems use precision and recall as the performance metric for analyzing the performance(Rahman et al., 2007, Rahman et al., 2009, Hatice Cinar Akakin, 2012, Lining Zhang et al., 2012, Timothy Tian-Ming Zheng, 2010, Neetu Sharma, 2011, Ja-Hwung Su, 2011, WangXing-yuan et al., 2012, Mann-Jung Hsiao et al., 2010). The precision calculation is given in equation (43) and recall can be calculated using equation (44).

$$Precision = \frac{No.of\ relevant\ images\ retrieved}{Total\ No.of\ images\ retrieved} \quad (43)$$

$$Recall = \frac{No.of\ relevant\ images\ retrieved}{Total\ No.of\ relevant\ images\ in\ DB} \quad (44)$$

Accuracy percentage is also another metric to analyze the CBIR system. The accuracy of the system is calculated using equation (45).

$$Accuracy\ \% = \frac{\sum_{i=1}^{T} q_i}{T} \times 100 \quad (45)$$

where q_i is 1 in equation (45), if the i^{th} query image has the correct resultant images; q_i = 0, otherwise. T is the total number of query images in testing (Hatice Cinar Akakin, 2012, Neetu Sharma, 2011).

8. Comparison of the various works

This section summarizes the above mentioned components of the CBIR used in various recent papers. The main features of the CBIR that are used in the recent IEEE and other papers are summarized in the Table 1.

Table 1. Comparision of the CBIR components used in recent papers

Paper	Low level features				Data Set	Distance measures	Machine Learning	Accuracy	Invariance			
	Color Feature	Texture Feature	Shape Feature	Others					T	S	R	N
"Scale and Rotation Invariant Gabor Features for Texture Retrieval", IEEE Conference DICTA, [Rahman M.H. et al., 2011]	--	Invariant Gabor Descriptor	--	--	Brodatz database	--	--	99.7%	-	√	√	-
"Rotation Invariant Indexing For Image Using Zernike Moments and R-Tree", TELKOMNIKA, 2011 [Saptadi et al, 2011]	--	--	Zernike moments	--	MRI images	Error tolerance distance	R-Tree Data Structure	95.2%	-	-	√	-
"SRBIR", Journal of Computer Science, [Felci Rajam et al., 2011a]	Color moments	GLCM	--	--	COREL dataset	Euclidean, Bhattacharya, Mahalanobis	SVM Binary Decision Tree	95.4%	-	-	-	-
"QSVMBDT", Springer CCIS 205, [Felci Rajam et al., 2011b]	Color moments	GLCM	Histogram of edge direction &invariant moments	--	COREL dataset	Euclidean, Bhattacharya, Mahalanobis	SVM Binary Decision Tree	97.6%	-	-	-	-
"Generalized Biased Discriminant Analysis for CBIR", IEEE Trans. on Systems, Man, and Cybernetics, [Lining et al., 2012]	Color histogram	Weber's law descriptor	Edge directional histogram.	--	COREL dataset	Mahalanobis distance	Generalized Biased Discriminant Analysis (GBDA)	83.5%	-	-	-	-
"Choice of a pertinent color space for color texture characterization", Pattern Recognition, [Imtnan et al., 2011]	3D Color histogram	--	--	--	Vistex database, Outex database	Kullback–Leibler(KL) divergence	Parametric Spectral Analysis	95.8	-	-	-	-
"Content-Based Microscopic Image Retrieval System for Multi-Image queries", IEEE Trans on Info. Tech. in Biomedicine, [Hatice et al., 2012]	Color moments	Co-occurrence Histogram	--	--	Follicular Lymphoma, Neuroblastoma	correlation distance measure	SVM, Nearest neighbor search	93% 86%	-	-	-	-
"An effective method for color image retrieval based on texture", Elsevier Computer Standards &	--	Color co-occurrence matrix	--	--	Color images collected from	Normalized Euclidean distance	--	82.94%	-	-	-	-

Paper	Low level features				Data Set	Distance measures	Machine Learning	Accuracy	Invariance			
	Color Feature	Texture Feature	Shape Feature	Others					T	S	R	N
Interfaces, [WangXing -yuan, 2012].					public sources							
"CBIR with Relevance Feedback using Random Walks", Pattern Recognition, [Samuel et al., 2011].	Color histogram, Color histogram layout, color moments	GLCM		Global Scene (GIST)	Oliva dataset, Caltech dataset	l^1- norm	Random walk with relevance feedback	95% at 6th RF round	-	-	-	-
"Efficient Relevance Feedback for CBIR by Mining User Navigation Patterns", IEEE Trans. on Knowledge and Data Engineering, [Ja-Hwung et al., 2011]	Color Layout, Color Structure	Homogeneous Texture	Edge Histogram, Region Shape	--	COREL dataset	Weighted KNN search	Navigation Pattern based Relevance Feedback	Precision= 0.91 at 6th iteration	-	-	-	-
"Texture feature extraction method for scale and rotation invariant image retrieval", Electronics Letters, [Rahman M.H. et al., 2012].	--	invariant Gabor Descriptor	--	--	Brodatz texture database, Food database	Canberra distance	--	92.18% 89.33%	-	√	√	-
"Fast Semantic Diffusion for Large-Scale Context-Based Image and Video Annotation", IEEE Trans. on Image Processing, [Yu-Gang et al., 2012].	Grid based color moments	Wavelet Texture	--	Bag of visual words	NUS-WIDE TRECVID	--	Semantic graph	Precision= 0.459	-	-	-	-
"Parallel Implementation of Shape based Image Retrieval approach", IJCA Special Issue on "Novel Aspects of DIA", [Kuldeep et al., 2011].	--	--	DCT method to extract DC and AC coefficients	--	Data set of MRI, CT-scan and X-ray	Euclidean Distance	Parallel implementation feature extraction and feature matching.	Average accuracy=66.88%	-	-	-	-
"Region-based image retrieval using the semantic cluster matrix and adaptive learning", IJCSE, [Felci et al., 2012].	Color moments	GLCM	7 invariant moments	--	Caltech	Euclidean, Bhattacharya, Mahalanobis	SVM-BDT, SCM	91.8%	-	-	-	-

9. Conclusion

This paper discusses the various methodologies used for extracting the salient low level features and various distance measures to find the similarity between images. Also, this paper provides a detailed review of the works carried out, in reducing the semantic gap between the low level features and the high level semantic concepts. A discussion of invariant image retrieval under various geometric transformations is also made. The various datasets and the performance measures used in analyzing the efficiency of CBIR systems are also presented. From the above survey, it is concluded that for efficient and invariant image retrieval, low level features invariant to geometric transformations, and robust under various types of noises are needed. Based on the current trends, the open research issues include algorithms for feature extraction, that are robust under various geometric transformations, noise and various lighting conditions, recognizing the semantic concepts present in the image, automatic segmentation of the different semantic concepts in the image, effective learning of the high level semantic concepts, etc. To obtain an efficient CBIR framework, one must choose the components of the CBIR in a balanced manner, and this paper helps in analyzing all the components of the CBIR framework.

References

[1] Kodituwakku S. R, Selvarajah S. Comparison of Color Features for Image Retrieval. Indian Journal of Computer Science and Engineering, Vol. 1, No. 3, pp: 207-211, 2010.

[2] Selvarajah S and Kodituwakku S. R. Analysis and Comparison of Texture Features for Content Based

Image Retrieval. International Journal of Latest Trends in Computing, Volume 2, Issue 1, pp. 108 – 113, 2011.
[3] Deselaers T, Keysers D, Ney H. Features for image retrieval: an experimental comparison. Inf. Retr. 11(2), pp. 77–107, 2007.
[4] Rahman M.M., Bhattacharya M.P. and Desai B.C. A framework for medical image retrieval using machine learning and statistical similarity matching techniques with relevance feedback. IEEE Trans. Inform. Technol. Biomed., Vol. 11, No. 1, pp.58–69, 2007.
[5] Haralick R.M., Shanmugam K. and Dinstein I., Textural features for image classification. IEEE Trans. Syst., Man, Cybern. SMC-3(6),pp. 610–621, 1973.
[6] Rahman M.H., Pickering M.R., Frater M.R. Scale and Rotation Invariant Gabor Features for Texture Retrieval. IEEE International Conference on Digital Image Computing Techniques and Applications (DICTA), pp. 602-607, 2011.
[7] Cheng-Hao Yao and Shu-Yuan Chen. Retrieval of translated, rotated and scaled color textures. Pattern Recognition, Vol. 36, pp. 913 – 929, 2003.
[8] Yang Mingqiang, Kpalma Kidiyo, Ronsin Joseph. A survey of shape feature extraction techniques. Pattern Recognition, Peng-Yeng Yin (Ed.) pp. 43-90, 2008.
[9] Zhang D, Lu G. Review of shape representation and description techniques. Pattern Recognition 37, pp. 1-19, 2004.
[10] Saptadi Nugroho and Darmawan Utomo. Rotation Invariant Indexing For Image Using Zernike Moments and R–Tree. TELKOMNIKA, Vol.9, No.2, pp. 335-340, 2011.
[11] Dudani S.A., Breeding K.J. and McGhee R.B. Aircraft identification by moment invariants. IEEE Trans. on Computers C-26(1), pp. 39–46, 1977.
[12] H.B. Kekre, Dhirendra Mishra, Anirudh Kariwala. A Survey of CBIR Techniques and Semantics. International Journal of Engineering Science and Technology (IJEST), Vol. 3, No. 5, PP. 4510-4517, 2011.
[13] Yixin Chen, James Z. Wang, Robert Krovetz. Content-Based Image Retrieval by Clustering. Proceedings of the 5th ACM SIGMM international workshop on Multimedia information retrieval MIR '03, pp. 193-200, 2003.
[14] Raghu Krishnapuram, Swarup Medasani, Sung-Hwan Jung, Young-Sik Choi, Rajesh Bala subramaniam. Content-Based Image Retrieval Based on a Fuzzy Approach. IEEE Transactions on Knowledge and Data Engineering, vol.16, no.10, pp.1185–1199, 2004.
[15] Madzarov G. and Gjorgjevikj D. Multi-class classification using support vector machines in decision tree architecture. Proceeding of the IEEE EUROCON 2009, pp.288–295, 2009a.
[16] Madzarov G., Gjorgjevikj D. and Chorbev I. A multi-class SVM classifier utilizing binary decision tree. Informatica, Vol. 33, No. 2, pp.233–241, 2009b.
[17] Md. Mahmudur Rahman A, Prabir Bhattacharya B, Bipin C. Desai. A unified image retrieval framework on local visual and semantic concept-based feature spaces. Journal of Visual Communication and Image Representation, Vol. 20, issue 7, pp. 450–462, 2009.
[18] Felci Rajam I. and Valli S. SRBIR: semantic region based image retrieval by extracting the dominant region and semantic learning. Journal of Computer Science, Vol. 7, No. 3, pp.400–408, 2011a.
[19] Felci Rajam I and Valli S. Content-Based Image Retrieval Using a Quick SVM-Binary Decision Tree – QSVMBDT. Springer Communications in Computer and Information Science 205, pp: 11-22, 2011b.
[20] Felci Rajam I. and Valli S. Region-based image retrieval using the semantic cluster matrix and adaptive learning. International Journal of Computational Science and Engineering, Vol. 7, No. 3, pp.239–252, 2012.
[21] Celia B, Felci Rajam I. An efficient content based image retrieval framework using machine learning techniques. Proceedings of the Second international conference on Data Engineering and Management (ICDEM '10), Springer LNCS, Vol. 6411, pp 162-169, 2010.
[22] Prabir Bhattacharya, Md. Mahmudur Rahman, Bipin C. Desai. Image Representation and Retrieval Using Support Vector Machine and Fuzzy C-means Clustering Based Semantical Spaces. Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06), 2006
[23] Ying Liu, Dengsheng Zhang, Guojun Lu. Region-based image retrieval with high-level semantics using decision tree learning. Pattern Recognition 41, pp. 2554 – 2570, 2008.
[24] Najlae Idrissi, Jos ́e Martinez, Driss Aboutajdine. Bridging the Semantic Gap for Texture-based Image Retrieval and Navigation. Journal of Multimedia, Vol. 4, No. 5, pp: 277-283, 2009.
[25] Nikhil Rasiwasia, Pedro J. Moreno and Nuno Vasconcelos. Bridging the Gap: Query by Semantic Example. IEEE Transactions on Multimedia, Vol. 9, No. 5, pp: 923-938, 2007.
[26] Lining Zhang, Lipo Wang andWeisi Lin. Generalized Biased Discriminant Analysis for Content-Based Image Retrieval. IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics, Vol. 42, No. 1, pp. 282-290, 2012.
[27] Imtnan-Ul-Haque Qazi, OlivierAlata, Jean-ChristopheBurie, Ahmed Moussa, ChristineFernandez-Maloigne. Choice of a pertinent color space for color texture characterization using parametric spectral analysis. Pattern Recognition 44, pp. 16–31, 2011.
[28] Hatice Cinar Akakin and Metin N. Gurcan. Content-Based Microscopic Image Retrieval System for Multi-Image queries. IEEE Transactions on Information Technology in Biomedicine, Vol. 16, No. 4, pp. 758 – 769, 2012.
[29] WangXing-yuan, ChenZhi-feng, YunJiao-jiao. An effective method for color image retrieval based on texture. Computer Standards & Interfaces 34, pp. 31–35, 2012.
[30] Timothy Tian-Ming Zheng, Wing W. Y. NG, XU-Sheng Huang, Shi-Ting Yang, Patrick P.K. Chan, Weiwei Lai, Daniel S. Yeung. Shape-Based Image Retrieval of Chinese Paper Cutting Using RBFNN

with Invariant Moment. IEEE Ninth International Conference on Machine Learning and Cybernetics, Qingdao, pp. 808-814, 2010.

[31] Samuel Rota Bul_o, Massimo Rabbi and Marcello Pelillo. Content-Based Image Retrieval with Relevance Feedback using Random Walks. Pattern Recognition, Vol. 44, No. 9, pp. 2109-2122, 2011.

[32] Neetu Sharma S, Paresh Rawat S and jaikaran Singh S. Efficient CBIR Using Color Histogram Processing. Signal & Image Processing : An International Journal(SIPIJ), Vol.2, No.1, pp. 94-112, March 2011.

[33] Ja-Hwung Su, Wei-Jyun Huang, Philip S. Yu, Fellow, and Vincent S. Tseng. Efficient Relevance Feedback for Content-Based Image Retrieval by Mining User Navigation Patterns. IEEE Transactions on Knowledge and Data Engineering, Vol. 23, No. 3, pp. 360-372, 2011.

[34] Theo Gevers and Harro Stokman. Robust Histogram Construction from Color Invariants for Object Recognition. IEEE Transactions on Pattern Analysis And Machine Intelligence, Vol. 26, No. 1, pp. 113-118, 2004.

[35] Yu Ma, Xiaodong Gu , Yuanyuan Wang. Histogram similarity measure using variable bin size distance. Computer Vision and Image Understanding, Vol. 114, pp. 981–989, 2010.

[36] Liu Y., Zhang D., Lu G. and Ma W.Y. A survey of content-based image retrieval with high-level semantics. Patt. Recog., 40, pp. 262-282, 2007.

[37] Shyu M-L., Chen S-C., Chen M., Zhang C. and Sarinnapakorn K. Capturing high-level image concepts via affinity relationships in image database retrieval. Springer Multimedia Tools and Applications, Vol. 32, No. 1, pp.73–92, 2007.

[38] Liu Y., Chen X., Zhang C. and Sprague A. Semantic clustering for region-based image retrieval. Journal of Visual Communication and Image Representation, Vol. 20, No. 2, pp.157–166, 2009.

[39] Zakariya S. M., Rashid Ali and Nesar Ahmad. Content Based Image Retrieval by Combining Visual Features of an Image With A Threshold. Special Issue of IJCCT Vol. 2, No. 2, 3, 4, pp. 204-209, 2010.

[40] Feng S., Lang C. and Xu D. Localized content-based image retrieval using saliency-based graph learning framework. ICSP 2010 Proceedings, pp.1029–1032, 2010.

[41] Giacinto G., Roli F. Instance-Based Relevance Feedback for Image Retrieval. Proceedings of Advances in Neural Information Processing Systems", Vol. 17, pp. 489–496, 2005.

[42] Das G., Ray S. A Comparison of Relevance Feedback Strategies in CBIR. 6th IEEE/ACIS International Conference on Computer and Information Science, pp.100–105, 2007.

[43] Rahman M.H. , Pickering M.R. , Frater M.R. , Kerr D. Texture feature extraction method for scale and rotation invariant image retrieval. Electronics Letters, Volume: 48 , Issue: 11 , Page(s): 626 – 627, 2012.

[44] Zhe-Ming Lu, Dan-Ni Li and Hans Burkhardt. Image retrieval based on RST-invariant features. IJCSNS International Journal of Computer Science and Network Security, Vol. 6, No. 2A, pp. 169-174, 2006.

[45] Yu-Gang Jiang, Qi Dai, Jun Wang, Chong-Wah Ngo, Xiangyang Xue and Shih-Fu Chang. Fast Semantic Diffusion for Large-Scale Context-Based Image and Video Annotation. IEEE Transactions on Image Processing , Vol. 21, No. 6, pp. 3080 – 3091, 2012.

[46] Hui Hui Wang, Dzulkifli Mohamad & N.A. Ismail. Semantic Gap in CBIR: Automatic Objects Spatial Relationships Semantic Extraction and Representation. International Journal of Image Processing (IJIP), Vol. 4, Issue 3, pp. 192-204, 2010.

[47] Kuldeep Yadav, Avi Srivastava and Ankush Mittal, M.A Ansari. Parallel Implementation of Shape based Image Retrieval Approach on CUDA in Compressed Domain. IJCA Special Issue on "Novel Aspects of Digital Imaging Applications", pp: 15 – 22, DIA, 2011.

[48] Ch. Srinivasa Rao , Srinivas Kumar S and Chandra Mohan B. Content Based Image Retrieval Using Exact Legendre Moments and Support Vector Machine. The International Journal of Multimedia and its applications (IJMA), Vol. 2, No. 2, pp. 69-79, 2010.

[49] Md. Mahmudur Rahman, Bipin C. Desai, and Prabir Bhattacharya. Supervised Machine Learning Based Medical Image Annotation and Retrieval in ImageCLEFmed 2005. proceedings of the CLEF 2005, LNCS 4022, pp. 692–701, 2006.

[50] V. Mezaris, I. Kompatsians, M.G. Strintzis. An ontology approach to object-based image retrieval. Proceedings of the ICIP, vol II, pp. 511-514, 2003.

[51] Javad Kangarani Farahani, Reza Ahmadi, Zahra Askari, Mohammad Hosein Bayat. Improve image contrast using the histogram of the matrix obtained in a uniform method of histogram and without noise histogram overlay. Life Sci J, vol. 9, no. 4, pp. 3460-3463, 2012.

[52] Aymn E.Khedr and Abd El-Ghany A. M. Mohmed. A proposed image processing framework to support Early liver Cancer Diagnosis. Life Sci J, vol. 9, no. 4, pp. 3808-3813, 2012.

[53] Thenmozhi. S, Balasubramanie. P, Venkatesh. J, Aarthy. C. Detection of reflection in iris images using back propagation. Life Sci J, vol. 9, no. 3, pp. 2446-2450, 2012.

[54] Asmatullah Chaudhry, Javed Ullah, M. Arfan Jaffar, Jin Young Kim, Tran Anh Tuan. Human Activity Recognition System: Using Improved Crossbreed Features and Artificial Neural Network. Life Sci J, vol. 9, no. 4, pp. 5351-5356, 2012.

[55] Mann-Jung Hsiao, Yo-Ping Huang, Tienwei Tsai, Te-Wei Chiang. A General and Effective Two-Stage Approach for Region-Based Image Retrieval, Life Sci J, vol. 7, no. 3, pp. 73-80, 2010.

6/18/2013