Low-level Features Extraction of an Image for CBIR: Techniques and Trends

Kumari Puja
Department of EC, TIET
Jabalpur, India
kmpuja@gmail.com

Dr. Pravin Dakhole (Head of the Dept.)
Department of EC, YCCE
Nagpur, India
pravin_dakhole@yahoo.com

Abstract - Content-based Image Retrieval (CBIR) has gained much attention in the past decades. CBIR is a technique to retrieve images from an image database such that the retrieved images are semantically relevant to a query image provided by a user. It is based on representing images by using low-level visual features, which can be extracted from images such as color, texture and shape. Each of the features is represented using one or many feature descriptor. In this paper, Dominant colors in HSV and YCbCr and Color Coherent Vector (CCV)are implemented to describe the color feature of an image. For texture feature extraction, Gabor wavelet is best method. Texture features are found by calculating the mean and variation of the Gabor filtered image. Rotation normalization is realized by a circular shift of the feature elements so that all images have the same dominant direction. Another feature descriptor of texture is GLCM (gray level cooccurrence matrix). The features extracted from GLCM are its energy, entropy, contrast and inverse difference moment.

Keywords: low level features, Dominant color, Color Coherent Vector, Gabor wavelet, Gray level Cooccurrence matrix.

I. INTRODUCTION

IMAGE retrieval is the field of study concerned withsearching and retrieving digital images from a collection ofdatabase. This research has been explored since the 1970s[1, 2] Image retrieval attracts interest among researchers in thefields of image processing,multimedia, digital libraries,remote sensing, astronomy, database applications and othersrelated area. An effective image retrieval system is able tooperate on the collection of images to retrieve the relevantimages based on the query image which conforms as closelyas possible to human perception.A typical image retrieval system includes three major components: 1) feature extraction (usually in conjunction with feature selection), 2) high dimensional indexing and 3) system design [3]. In this work, we have consider the first component; that of low-level feature extraction, and we attempt to answer the following question: What are the color and texture features that need to be extractedfrom an image, in order to achieve the highest retrieval performance, at a relatively lowcomputational cost? In this paper we are illustrating four color descriptor and two two texture descriptor that can consider being the best low level features to describe an image for the retrieval. In Section II, we discuss the three color feature extraction techniques: 1) the Dominant Colors in HSV, 2) the Dominant Colors in YCbCr, 3) the Color Coherent Vector (CCV). In Section III, we discuss the two texture feature extraction techniques: 1) the Gabor wavelet transform, and 4) the GLCM (gray level cooccurrence matrix).
From the experiments it has found that by taking only these low level features into consideration we can give good result in image retrieval.

II. COLOR FEATURE EXRACTION MODELS

A. Dominant colors

Color is extremely subjective and personal. To try to attribute numbers to the brains reaction to visual stimuli is very difficult. The aim of color spaces is to aid the process of describing color, either between people or between machines or programs. There are a large number of color spaces in use in the world today, they are: CIE, RGB, YUV- YCbCr, HSL/HSV, and CMYK. For the dominant color we have concentrated on

HSV and YCbCr color space as these are closer to the human visual system.
Dominant color region in an image can be represented as a connected fragment of homogeneous color pixels which is perceived by human vision [4]. Image Indexing [2] is based on this concept of dominant color regions present in the image. The segmented out dominant regions along with their features are used as an aid in the retrieval of similar images from the image database. Image path, number of regions found, region information like color, normalized area and location of each region are stored in file for further processing.

The main issue is "how to calculate dominant colors in a image"? In this we needs to count colors in a image. The count should not take all the different colors in the image, that means it should merge similar colors and count it as one. For e.g if there is a Red and light red then the count should be one, i.e Red. Just like human eye counts colors. We have solved this issue with the help of Euclidean distance algorithm and K-Means clustering algorithm [5].

The steps involved for extracting Dominant Color feature are given below:

Step 1: Scan through the image and get all the pixels.Group similar pixels and also increase the count. We have used a hash table for grouping. After the scanning we got a hash table with all the colors and the count of each colors.

Step 2: Find out the dominant colors using the pixel count and remove the nearest pixels. To find the nearest color in the same domain we used Euclidian Distance algorithm. I meant by dominant color is the color that has more pixel count. While removing the color, we should not remove the dominant color.

Step 3: The above step still will not give accurate results, this step 2 result will just a give a starting point for color count. We used the result from Step 2 as the cluster for applying K-Means clusteringalgorithm. For clusters take only the top n higher pixel counted colors from Step 2. Apply the clustering on the pixel data we got from Step 1.

Step 4: Apply the Euclidean distance on the result we got from Step 3. The result will be closer to the count of colors in the image. We can tweak to get closer result by increasing or decreasing the distance cutoff value.

B. Color Coherent Vector

Another color description is Color Coherent Vector (CCV). A color histogram provides no spatialinformation; it merely describes which colorsare present in the image, and in what quantities. We are considering CCV which is similar to color histograms, but which also takes spatial information into account.

Figure 3: Two images with similar color histograms

the images shown in Figure 1have similar color histograms, despite their rather different appearances. The color redappears in both images in approximately thesame quantities. In the left image the redpixels (from the flowers) are widely scattered,while in the right image the red pixels (fromthe golfer's shirt) form a single coherent region.

Our coherence measure classifies pixels aseither coherent or incoherent. Coherent pixelsare a part of some sizable contiguous region,while incoherent pixels are not. Colorcoherencevectorrepresents this classification foreach color in the image. CCV's prevent coherentpixels in one image from matching incoherentpixels in another. This allows finedistinctions that cannot be made with colorhistograms.

Computing CCV's

We first blur the image slightly by replacing pixel values with the average value in a smalllocal neighborhood. We thendiscretize the color space, such that there areonly n distinct colors in the image. The next step is to classify the pixels withina given color bucket as either coherent or incoherent.

A coherent pixel is part of a largegroup of pixels of the same color. We determine the pixel groups by computing connected components.A connected component *C* is a maximal set ofpixels such that for any two pixels *p, p′ Є C*,there is a path in *C* between p and *p′*. (Formally,
a path in *C* is a sequence of pixels

$$p = p1,p2,\ldots, pn = p'$$

such that each pixel *pi*is in *C* and any two sequential pixels *pi*, *pi+1*are adjacent to each other (eight neighbors).When thisis complete, each pixel will belong to exactlyone connected component. We classify pixelsas either coherent or incoherent depending onthe size in pixels of its connected component.A pixel is coherent if the size of its connectedcomponent exceeds a fixed value ζ, otherwise,the pixel is incoherent.
For a given discretized color, some of thepixels with that color will be coherent andsome will be incoherent. Let us call the numberof coherent pixels of the j'th discretizedcolor αj and the number of incoherent pixelsβj . Clearly, the total number of pixels withthat color is αj +βj, and so a color histogramwould summarize an image as

(α1 +β1, …….., αn +βn)

Instead, for each color we compute the pair

(α, βj)

which we will call the *coherence pair* for the j'th color. The *color coherence vector* for the image consists of

((α1, β1), ……..,(αn, βn))

This is a vector of coherence pairs, one for each discretized color.

In our experiments, all images were scaledto contain M = 38;976 pixels, and we haveused ζ = 300 pixels (so a region is classifiedas coherent if its area is about 1% of the image).With this value of ζ , an average imagein our database consists of approximately 75%coherent pixels.

III. TEXTURE FEATURE EXTRACTION MODELS

Under this we have considered Gabor filter method and GLCM (gray level cooccurrence matrix) for texture feature extraction.

A. Gabor Filter

Gabor filter (or Gabor wavelet) is widely adopted to extracttexture features from the images for image retrieval [6, 7]. Basically, Gabor filters are a group of wavelets, with eachwavelet capturing energy at a specific frequency and a specific direction.Basically, Gabor filters are a group of wavelets, with eachwaveletcapturing energy at a specific frequency and a specificdirection.Currently, most techniques make anassumption that all the images are captured under the sameorientations. In many applications like Content base Image Retrieval such an assumption isunrealistic. In this paper we illustrateda rotation normalization method that achieves rotation invariance by a circular shift of the feature elements so that all images have the same dominant direction.

For a given image *I*(*x*, *y*) with size *P*X*Q*, its discrete Gaborwavelet transform is given by a convolution:

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

where, *s* and *t* are the filter mask size variables, and Ψ_{mn}* is the complex conjugate of Ψ_{mn} which is a class of self-similar functions generated from dilation and rotation of the followingmother wavelet: Ψ_{mn} =

$$\frac{1}{2\pi\sigma_x\sigma_y}\exp[-\frac{1}{2}(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2})]\cdot\exp(j2\pi Wx)$$

where *W* is called the modulation frequency. The self-similar Gabor wavelets are obtained through the generating function:

$$\psi_{mn}(x,y) = a^{-m}\psi(\tilde{x},\tilde{y})$$

where *m* and *n* specify the *scale* and *orientation* of the wavelet respectively, with *m* = 0, 1, …*M*-1, *n* = 0, 1, …, *N-1*, and

$$\tilde{x} = a^{-m}(x\cos\theta + y\sin\theta)$$

$$\tilde{y} = a^{-m}(-x\sin\theta + y\cos\theta)$$

where *a* >1 and θ = *n* π/ *N*.

The variables in the above equations are defined as follows:

$$a = (U_h/U_l)^{\frac{1}{M-1}},$$
$$W_{m,n} = a^m U_l$$
$$\sigma_{x,m,n} = \frac{(a+1)\sqrt{2\ln 2}}{2\pi a^m (a-1)U_l},$$
$$\sigma_{y,m,n} = \frac{1}{2\pi \tan(\frac{\pi}{2N})\sqrt{\frac{U_h^2}{2\ln 2} - (\frac{1}{2\pi\sigma_{x,m,n}})^2}}$$

In our implementation, we used

$U_l = 0.05$, $U_h = 0.4$,

s and t range from 0 to 60, i.e., filter mask size is 60x60.

After applying Gabor filters on the image with differentorientation at different scale, we obtain an array of magnitudes:

$$E(m,n) = \sum_x \sum_y |G_{mn}(x,y)|,$$

$m = 0, 1, \ldots, M-1$; $n = 0, 1, \ldots, N-1$

These magnitudes represent the energy content at differentscale and orientation of the image. For the homogenous texture feature of the region we use mean $\mu_{mn}$and standard deviation $\sigma_{mn}$of the magnitude of the transformed coefficients.

$$\mu_{mn} = \frac{E(m,n)}{P \times Q}$$
$$\sigma_{mn} = \frac{\sqrt{\sum_x \sum_y (|G_{mn}(x,y)| - \mu_{mn})^2}}{P \times Q}$$

A feature vector **f** (texture representation) is created using $\mu_{mn}$and σ_{mn}as the feature components.

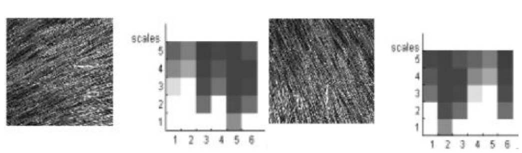
$$\mathbf{f} = (\mu_{00}, \sigma_{00}, \mu_{01}, \sigma_{01}, \ldots, \mu_{45}, \sigma_{45}).$$

The texture similarity measurement of a query image Q and a target image T in the database is defined by:

$$D(Q,T) = \sum_m \sum_n d_{mn}(Q,T)$$

where

$$d_{mn} = \sqrt{(\mu_{mn}^Q - \mu_{mn}^T)^2 + (\sigma_{mn}^Q - \sigma_{mn}^T)^2}$$



(a) (b) (c) (d)

Figure 2. (a) *an image.* (b) *Energy map of (a).* (c) *Rotated image of (a).* (d) *Energy map of (c).*

Since this similarity measurement is not rotation invariant,similar texture images with different direction may be missed out from the retrieval or get a low rank.In this paper we illustrated a simple circular shift on the feature map to solve therotation variant problem associate with Gabor texture features.Specifically, we calculate total energy for each orientation. Theorientation with the highest total energy is called the dominant orientation/direction. We then move the feature elements in thedominant direction to be the first elements in ***f***[8].The other elements are circularly shifted accordingly.For example, if theoriginal feature vector is "*abcdef*" and "*d*" is at the dominantdirection, then the normalized feature vector will be "*defabc*".This normalization method is based on the assumption that tocompare similarity between two images/textures they should berotated so that their dominant directions are the same.

B. Gray Level Cooccurrence Matrix (GLCM).

Gray level co-occurrence matrix (GLCM)[9], one of the most known texture analysis methods, estimates image properties related to second-order statistics. Each entry (i,j) in GLCM corresponds to the number of occurrences of the pair of gray levels i and j which are a distance **d** apart in original image.In order to estimate the similarity between differentgray level co-occurrence matrices, 14 statistical features extracted from them. To reduce thecomputational complexity, only some of these features were selected. The description of 4 most relevant features that are widely used in [10, 11] is given in Table1.

Energy	$\sum_i \sum_j P_d^2(i,j)$
Entropy	$-\sum_i \sum_j P_d(i,j) \log P_d(i,j)$
Contrast	$\sum_i \sum_j (i-j)^2 P_d(i,j)$
Inverse Difference Moment	$\sum_i \sum_j \frac{P_d(i,j)}{\lvert i-j \rvert^2}, i \neq j$

Table 1. Some texture features extracted from GLCM

ENERGY is a measure of textural uniformity of an image. Energy reaches its highest value when gray level distribution has either a constant or a periodic form. A homogenous image contains very few dominant gray tone transitions, and therefore the P matrix for this image will have fewer entries of larger magnitude resulting in large value for energy feature.
ENTROPY measures the disorder of an image and itachieves its largest value when all elements in P matrix are equal.
CONTRAST is a difference moment of the P and itmeasures the amount of local variations in an image.
INVERSE DIFFERENCE MOMENT measures imagehomogeneity. This parameter achieves its largest valuewhen most of the occurrences in GLCM are concentratednear the main diagonal.

IV. EXPERIMENTAL RESULTS

The simulations were performed in MATLAB. For color feature extraction, the HSV space and YCbCr Space wasquantized to 128 color bins. For texture feature extraction, the Gabor transform parameters were set toperform an eight-orientation of the image.
The color and texture features were extracted from the images in the Corel. The Corel dataset is a database of 10 classes, each containing 100 images.
The features are extracted from eachimage were represented as a vector, and Euclidean distance was used to measure thedistance from thefeature vector of the query to the feature vector of every image in the database.
A retrieval score was computed according to the following evaluation criterion: for each query,the system returned the 10 closest images to the query, including the query image itself (as thedistance from the query image to itself is zero). The number of mismatches was computed as the
number of images returned that belong to a class different than that of the query image, inaddition to the number of images that belong to the query image class, but that have not beenreturned by the system. The retrieval score for one class was then computed as

$$100 \times \left[1 - \left(mismatches/_{10}\right)\right]\%$$

Finally, the average retrieval score for all classes was computed asthe average of the retrieval scores obtained for each class.
Table 2 & 3 display the obtained retrieval results by considering the color and texture features individually.

	Dominat Color (HSV)	Dominat Color (YCbCr)	Color Coherent Vector
Average Retrieval Score	49%	43%	55%

Table 2

	Gabor Filter	GLCM
Average Retrieval Score	60%	68%

Table 3

If we consider all the above features together then the retrieval rate increases to 86%.

V. CONCLUSION AND FUTURE WORK

The main contribution of this work is compare the different color and texture feature extraction techniques for CBIR.A combination of these color, and texture features provides a robust feature set for image retrieval.
In future by including Feedback Algorithm we can increase the retrieval rate.

VI. REFRENCES

[1]. A. Rosenfeld. “Picture processing bycomputer.”ACM Computing Surveys, 1(3):147-176, 1969.
[2]. H. Tamura and S. Mori. “A data management system for manipulating large images.” In Proceedings of Workshop on Picture Data Description and Management, pages 45-54, Chicago, Illinois, USA, April 1977
[3]. Y. Rui, T. S. Huang and S. Chang, “Image Retrieval: Current Techniques, Promising Directions and Open Issues ”, *Journal of Visual Communication and Image Representation,* vol. 10, pp. 39-62, March 1999.
[4]. T.F.Syeda_Mahmood. “Data and Model driven selection using color regions.”

International Journal of ComputerVision,21(1/2):9-36,1997.
[5]. Tapas Kanungo, Nathan S. Netanyahu and Angela Y. Wu, “An Efficient k-Means Clustering Algorithm: Analysis and Implementation”, In IEEE Transactions On Pattern Analysis And Machine Intelligence, VOL. 24, NO. 7, JULY 2002
[6]. B. S. Manjunath and W. Y. Ma. “Texture features forbrowsing and retrieval of large image data” IEEETransactions on Pattern Analysis and Machine Intelligence,(Special Issue on Digital Libraries), Vol. 18 (8), August
1996, pp. 837-842.
[7]. Alexander Dimai. “Rotation Invariant Texture Descriptionusing General Moment Invariants and Gabor Filters”. InProc. Of the 11th Scandinavian Conf. on Image Analysis.Vol I. June, 1999, pp.391-398.
[8]. Sylvie Jeannin (ed.), “ISO/IEC JTC1/SC29/WG11/N3321:MPEG-7 Visual Part of eXperimentation Model Version5.0”. March 2000.
[9]. R. M. Haralick, K. Shanmugam, and I.Dinstein, “Textural Features for ImageClassification”, IEEE Trans. on Systems, Man,and Cybernetics, Vol. SMC-3, No.6, November1973, pp.610-621.
[10]. R. C. Gonzalez, and R. E. Woods, “Digital Image Processing”, 1993.

[11] A. Baraldi and F. Parmiggiani, “An investigationof the Textural Characteristics ssociatedwith Gray Level Cooccurrence MatrixStatistical Parameters ”, IEEE Trans. OnGeoscience and Remote Sensing, vol. 33, no. 2, pp. 293-304, 1995.