International Journal of Research in Science and Technology http://www.ijrst.com/
(IJRST) 2011, Vol. No. 1, Issue No. III, July-Sept **ISSN: 2249-0604**

CONTENT BASED IMAGE RETRIEVAL

Dnyanda Namdeo Hire., S. S. Katariya.
Amrutvahini College of Engineering,Sangamner,
Tal-Sangamner, Dist- Ahmadnagar.

ABSTRACT

Feature extraction is a key issue in content based image retrieval (CBIR). In the past, a number of texture features have been proposed in literature, including statistic methods and spectral methods. However, most of them are not able to accurately capture the edge information which is the most important texture feature in an image. Recent researches on multi-scale analysis, especially the curvelet research, provide good opportunity to extract more accurate texture feature for image retrieval. Curvelet was originally proposed for image denoising and has shown promising performance. In this paper, image retrieval using various spectral methods is discussed.

INTRODUCTION

Nowadays, due to the availability of large storage spaces, huge number of images have been produced and stored around the world. With this huge image database, people want to search it and make use of the images in it. Here comes the challenge of image retrieval and researchers try to find out accurate ways of searching images. Basically, images can be retrieved in two ways, firstly, text based and secondly, content-based or query by example based. Text based retrieval approach is very well known and widely used. In this process users are provided a text area to enter the key words on the basis of which image searching is done. It is widely used in Yahoo web based image searching technique. Though the concept is very familiar to us, this approach has two notable drawbacks. (1) The images in the database are annotated manually by annotators using key words, which is a very time consuming process [1] for a large database. (2) The retrieval solely depends on the human perception based text annotation. The consequence is that there is significant inconsistency in understanding of image content by different annotators, so the key words vary a lot and retrieval results are usually very poor.

To avoid the above mentioned problem, the second approach, Content-based image retrieval has been proposed by the researchers. The term CBIR seems to have originated in the early 1990's [1]. Since then it is an ongoing process. Different from text based image search, CBIR techniques use low-level features like texture, color, and shape to represent images and retrieves images relevant to the query image from the image database. Among those low level image features, texture feature has been shown very effective and subjective. A variety of techniques have been developed for extracting texture features. These techniques can be broadly classified into spatial methods and spectral methods. In spatial approach, most techniques rely on computing values of what are known as low order statistics from query and stored images [2]. These methods compute texture features such as the degree of contrast, coarseness, directionality and regularity [1, 2, 3]; or periodicity, directionality and randomness [4]. Alternative methods of texture analysis for image retrieval include the use of Gabor filters [5, 6, 7], wavelet [8, 9, 10] and DCT [11]. Statistic techniques suffer from insufficient number of features and sensitive to image noise. The spectral methods in literature, however, do not capture edge information accurately.

In this paper, spectral methods for image retrieval are discussed. There are following spectral methods for retrieving relevant images: Fourier Transform, Wavelet Transform, Gabor filter Transform, Curvelet Transform.

The rest of the paper is organized as following. In section 2, we describe the Fourier transform, Wavelet transform, Gabor filter transform and curvelet transform. In section 3, concludes the paper.

DESCRIPTION OF TRANSFORMS

FOURIER TRANSFORM

The purpose of Fourier transform (FT) is to convert a time-domain signal into the frequency-domain. FT uses Fourier analysis to measure the frequency components of the signal. The discrete Fourier transform for a 2-D image f (x,y) can be represented as:

$$F(u,v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) e^{-j2\pi(ux/M + vy/N)}$$

where x=0,1, 2,...,M−1 and y= 0,1, 2,...,N−1 denote an M × N image. Fourier transform provides the pattern information of an image that is collected from its frequency components as shown in fig.1.

Fig. 1. Original Lena image (512× 512) (left) and its Fourier transform (right).

The frequency components at specific locations of an image are used to represent the texture features of that image. Texture features computed from high frequency components are the main distinguishing factors between images which are used in CBIR. Therefore, frequency information at specific locations is required to distinguish images in CBIR. However, the disadvantage of Fourier transform is it captures only the global spectral features but does not provide any information about the exact location of these features.

DISCRETE COSINE TRANSFORMS (DCT)

It is similar to the discrete Fourier transform, it transforms a signal or image in the spatial domain to the frequency domain and obtains DCT coefficients which can be used in various image processing purpose. Given an image f (x,y) of size n×n, its 2-D discrete cosine transform can be defined as:

$$C(u,v)=\alpha(u,v)\sum_{x=0}^{n-1}\sum_{y=0}^{n-1}f(x,y)\cos[\frac{(2x+1)u\pi}{2n}]\cos[\frac{(2y+1)v\pi}{2n}]$$

where $u,v = 0,1,2,\ldots ,n-1$.

Discrete cosine transform (DCT) has been adopted as an effective technique for image and video compression. It possesses the property of preserving most of the image energy in the low-frequency DCT coefficients which makes it so popular for data compression. Different approaches for shape, texture, and color feature extraction and indexing using DCT can be found in [12]. Similar to FT, DCT texture features can only capture global features while ignoring local details. Therefore, it is not suitable for CBIR.

SHORT TIME FOURIER TRANSFORM (STFT)

Short time Fourier transform (STFT) provides a time-frequency representation that is not possible in FT. In STFT, a window function is chosen in such a way that the portion of a non-stationary signal which is covered by the window function seems stationary. This window function is then convolved with the original signal so that the signal or data part covered by the window is selected only. FT is then applied to the newly generated stationary signal. The window is then moved to the next slot of signal and the previously mentioned steps are applied repeatedly until the whole signal is completely analyzed. This time-frequency representation is necessary for CBIR. For an image f (x,y), its STFT can be defined as:

$$X(\tau_1,\tau_2,\omega_1,\omega_2)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(x,y)W^*(x-\tau_1,y-\tau_2)e^{-j(\omega_1 x+\omega_2 y)}dxdy$$

τ_1,τ_2

where W(x,y) is the selected 2-D window to convolve the 2-D signal, represent the spatial positions of the window and ω1 ,ω2 represents the spatial frequency parameters.

GABOR FILTERS TRANSFORM

Gabor filters transform is a good multiresolution approach that represents the edges of image in an effective way using multiple orientations and scales. Gabor filters have a spatial property that is similar to mammalian perceptual vision, thereby providing researchers a good opportunity to use it in image processing. Gabor filters transform creates a filter bank consisting of Gabor filters with various scales and orientations. Then the filters are convolved with the image. The Gabor filters transform can be represented as:

$$G_{m,n}(x,y)=\sum_{s}\sum_{t}f(x_1,y_1)g_{m,n}{}^{*}(x-x_1,y-y_1)$$

where s and t are the filter mask size variables, x1 =x−s and y1 =y−t. m and n represent the scale and orientation of a Gabor wavelet, respectively. The mother Gabor wavelet g(x,y) and its Fourier transform G(u,v) are defined as:

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

$$G(u,v)=\exp\{-\frac{1}{2}[\frac{(u-W)^2}{\sigma_u{}^2}+\frac{v^2}{\sigma_v{}^2}]\}$$

where $\sigma_u = 1/2\pi\sigma_x$, $\sigma_v = 1/2\pi\sigma_y$ and W is the modulation frequency. Similar Gabor filters are then obtained from the dilation and rotation of the mother Gabor wavelet:

$$g_{m,n}(x,y)=a^{-m}G(x',y'),$$

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

where a >1, θ = nπ /K, K is the total number of orientations. The values of the parameters a, σ_u and σ_v are given as:

$$a=(U_h/U_l)^{-\frac{1}{S-1}},$$

$$\sigma_u=\frac{(a-1)U_h}{(a+1)\sqrt{2\ln 2}},$$

$$\sigma_v=\tan(\frac{\pi}{2k})[U_h-2\ln(\frac{\sigma_u^2}{U_h})][2\ln 2-\frac{(2\ln 2)^2\sigma_u^2}{U_h{}^2}].$$

Here S is the number of scales, W=U h and U h and U l specify the upper and lower centre of frequencies. Gabor filtered images for the Lena picture are shown in fig.2.

Though Gabor filters texture features have been found to be the best over all, they have some limitations as they use wavelets as a filter bank. Wavelets are not so effective in representing edge discontinuities in images. Moreover, Gabor spectral cover is not complete, to avoid overlap in the spectral domain, Gabor filters only use half peak magnitudes in the frequency domain. As a consequence, information loss results in the spectral domain (Fig. 3). Gabor filters tiling of frequency ranges between Ul = 0.05 and Uh = 0.4 is shown in the Fig. 3 where the half peak-magnitude filters only touch each other. Consequently, the high frequency components, which are considered to be the most important in characterizing image textures, are not effectively captured. Redundant information also exists in transformed images as Gabor filters do not involve image down sampling. These limitations need to be improved in texture based CBIR.

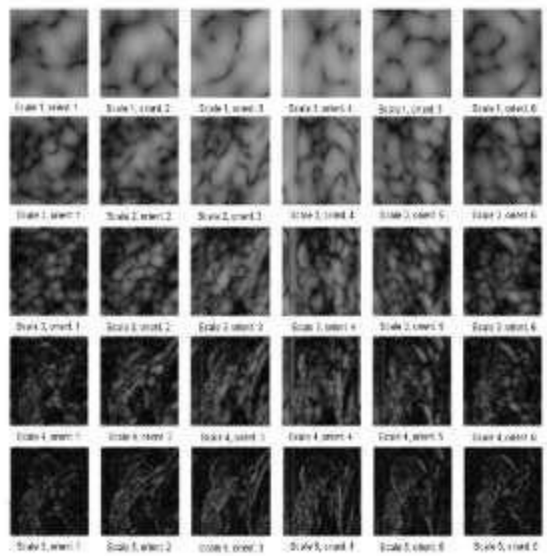

Fig. 2. Gabor filtered subbands for a 512× 512 Lena image.

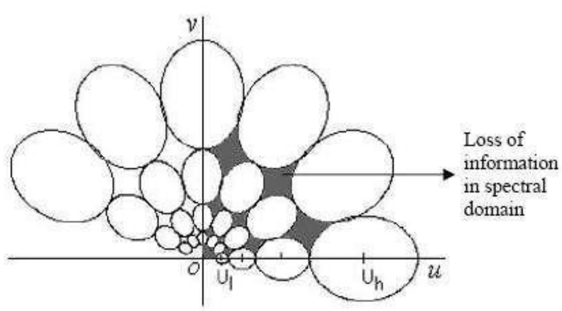

Fig. 3. Frequency tiling of half frequency plan by Gabor filters, the ovals are the covered spectrum . Uh and Ul specify the upper and lower centre of frequencies.

Discrete wavelet transforms use down sampling of images when extracts texture features. Next we describe how this transform addresses the problems found in FT and STFT.

DISCRETE WAVELET TRANSFORM (DWT)

Wavelet transform is introduced with the advancement in multiresolution transform research. Discrete wavelet transform is one of the most promising multiresolution approaches used in CBIR. It has the advantage of a time-frequency representation of signals where Fourier transform is only frequency localized. The location, at which a frequency component of an image exists, is important as it draws the discrimination line between images. Given an image f (x,y), its continuous wavelet transform is given by:

$$WT_{\psi}(a_1,a_2,b_1,b_2)=\int_R\int_R f(x,y)\overline{\psi_{a_1,a_2,b_1,b_2}(x,y)}dxdy$$

where a wavelet with scale parameter a1, a2 and position parameter b1,b2 can be described as follows:

$$\psi_{a_1,a_2,b_1,b_2}(x,y)=a_1^{-\frac{1}{2}}a_2^{-\frac{1}{2}}\psi(\frac{x-b_1}{a_1})\psi(\frac{y-b_2}{a_2})$$

Unlike the FT and STFT, the window size varies at each resolution level when the wavelet transform is applied to an image. In discrete wavelet transform, the original image is high-pass filtered yielding three detail images, describing the local changes in horizontal, vertical and diagonal direction of the original image.

DISCRETE CURVELET TRANSFORM

Basically, curvelet transform extends the ridgelet transform to multiple scale analysis. Therefore, we start from the definition of ridgelet transform. Given an image f (x,y), the continuous ridgelet coefficients are expressed as:

$$\Re_f(a,b,\theta)=\iint \psi_{a,b,\theta}(x,y)f(x,y)dxdy.$$

Here, a is the scale parameter where a > 0, b∈ R is the translation parameter and θ ∈ [0,2π) is the orientation parameter. Exact reconstruction is possible from these coefficients. A ridgelet can be defined as:

$$\psi_{a,b,\theta}(x,y)=a^{-\frac{1}{2}}\psi(\frac{x\cos\theta+y\sin\theta-b}{a})$$

where θ is the orientation of the ridgelet. Ridgelets are constant along the lines x cosθ + ysinθ=const and transverse to these ridges are wavelets.

Curvelet transform based on wrapping of Fourier samples takes a 2-D image as input in the form of a Cartesian array f[m,n] such that 0≤m<M,0≤n<N and generates a number of curvelet coefficients indexed by a scale j , an orientation l and two spatial location parameters (k1, k2) as output. To form the curvelet texture descriptor, statistical operations are applied to these coefficients. Discrete curvelet coefficients can be defined by:

$$C^D(j,l,k_1,k_2)=\sum_{\substack{0\le m<M\\ 0\le n<N}} f[m,n]\,\varphi^D_{j,l,k_1,k_2}[m,n].$$

Here, each $\varphi^D_{j,l,k_1,k_2}[m,n]$ is a digital curvelet waveform. The digital curvelet transform is implemented using the fast discrete curvelet transform. Basically, it is computed in the spectral domain to employ the advantage of FFT. Given an image, both the image and the curvelet are transformed into Fourier domain, then the convolution of the curvelet with the image in spatial domain becomes the product in Fourier domain. Finally the curvelet coefficients are obtained by applying inverse Fourier transform on the spectral product. But due to the frequency response of a curvelet is a nonrectangular wedge; the wedge needs to be wrapped into a rectangle to perform the inverse Fourier transform. The wrapping is done by periodic tiling of the spectrum using the wedge, and then collecting the rectangular coefficient area in the centre. Through this periodic tiling, the rectangular region collects the wedge's corresponding portions from the surrounding periodic wedges. The complete feature extraction process using a single curvelet is illustrated in Fig. 4.

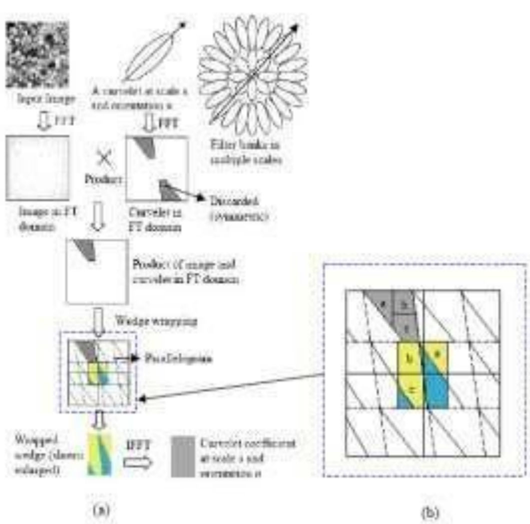

Fig. 4. Fast discrete curvelet transform to generate curvelet coefficients.

Once the curvelet coefficients have been obtained, the mean and standard deviation are computed as the texture features for the curvelet. That is, for each curvelet, two texture features are obtained. If n curvelets are used for the transform, 2n texture features are obtained. A 2n dimension texture feature vector is used to represent each image in the database for image retrieval.

This feature extraction is applied to each of the images in the database. At the end, each image in the database is represented and indexed using its curvelet feature vector. Given a query, its curvelet features are computed using the process shown in Figure 4. The system then compares the query feature vector with all the feature vectors in the database using L2 distance:

$$d(\mathbf{Q},\mathbf{T}) = \left(\sum_{i=1}^{2n}(Q_i - T_i)^2\right)^{\frac{1}{2}}.$$

Here, Q = {Q1 ,Q2 ,……..Q2n }is the feature vector of the query image, and T = {T1 ,T2 ,…......T2n } is the feature vector of the target image in the database. Finally, the images in the database are ranked according to their distance d to the query image, and the ranked list of images is returned to the user. This retrieval process is shown in Fig. 5.

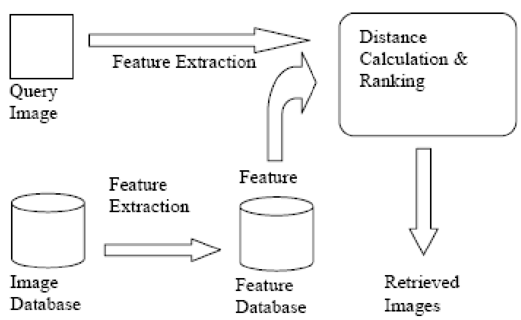

Fig. 5. Image retrieval mechanism.

CONCLUSIONS

Content based image retrieval is a challenging method of capturing relevant images from a large storage space. Although this area has been explored for decades, no technique has achieved the accuracy of human visual perception in distinguishing images. Whatever the size and content of the image database is, a human being can easily recognize images of same category.
From the very beginning of CBIR research, texture is considered to be a primitive visual cue like the color and shape of an image. Though image retrieval using texture features is not a brand new approach, there are still scopes to enhance the retrieval accuracy with a proper representation of texture features. In this research, we aimed to obtain a high image retrieval accuracy using the multiresolution discrete curvelet texture features.

REFERENCES

[1] F. Long, H. J. Zhang and D. D. Feng, Fundamentals of Content-based Image Retrieval, In Multimedia Information Retrieval and Management, D. Feng Eds, Springer, 2003.

[2] H. Tamura, S. Mori and T.Yamawaki, Texture Features Corresponding to Visual Perception, IEEE Trans. on Systems, Man and Cybernetics, 8(6): 460-473, 1978.

[3] B. S. Manjunath et al, Color and Texture Descriptors, IEEE Transactions CSVT, 11(6):703-715, 2001.

[4] F. Liu and R.W.Picard, Periodicity, Directionality, and Randomness: Wold Features for Image Modeling and Retrieval, IEEE Trans. On Pattern Analysis and Machine Intelligence, 18(7):722-733, 1996.

[5] L. Chen, G. Lu, and D. S. Zhang, Effects of Different Gabor Filter Parameters on Image Retrieval by Texture, In Proc. of IEEE 10th International Conference on Multi-Media Modelling, pp.273-278, Australia, 2004.

[6] B. S. Manjunath and W. Y. Ma, Texture Features for Browsing and Retrieval of Large Image Data, IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(8):837-842, 1996.

[7] B. S. Manjunath et al, Introduction to MPEG-7, John Wiley & Son Ltd., 2002.

[8] J. Z. Wang, J. Li, and G. Wiederhold, SIMPLIcity: Semantics-sensitive Integrated Matching for Picture Libraries, IEEE Trans. Pattern and Machine Intelligence, 23(9):947-963, 2001.

[9] S. Bhagavathy and K. Chhabra, A Wavelet-based Image Retrieval System, Technical Report—ECE278A, Vision Research Laboratory, University of California, Santa Barbara, 2007.

[10] N. Suematsu et al, Region-Based Image Retrieval using Wavelet Transform, Proc. 15th International Conf. on Vision Interface, May, 2002.

[11] Z. Lu, S. Li, and H. Burkhardt, A Content-Based Image Retrieval Scheme in JPEG Compressed Domain, International Journal of Innovative Computing, Information and Control, 2(4): 831-839, 2006.

[12] C. W. Ngo, T. C. Pong, and R. T. Chin, "Exploiting Image Indexing Techniques in DCT Domain," Pattern Recognition, vol. 34(9), pp. 1841-1851, September 2001.