Low-Level Color and Texture Feature Extraction for Content-Based Image Retrieval

Final Project Report, May 09, 2008

EE 381K: Multi-Dimensional Digital Signal Processing
Michele Saad

Abstract

An up-to-date comparison of state-of-the-art low-level color and texture feature extraction methods, for the purpose of content-based image retrieval (CBIR) is presented in this report. CBIR is a technique that utilizes the visual content of an image, to search for similar images in large-scale image databases, according to a user's interest. The CBIR problem is motivated by the need to search the exponentially increasing space of image and video databases efficiently and effectively. The visual content of an image is analyzed in terms of low-level features extracted from the image. These primarily constitute color and texture features. We implement and compare four color feature extraction algorithms and four texture feature extraction algorithms in this work. For color feature extraction, the conventional color histogram, the fuzzy color histogram, the color correlogram, and a color/shape-based method were implemented and compared. For texture feature extraction, the steerable pyramid, the contourlet transform, the Gabor wavelet transform, and the complex directional filter bank were implemented and compared. The fuzzy color histogram and the Gabor wavelet transform were shown to yield the highest color and texture retrieval results respectively, at the expense of more computation relative to the other proposed methods.

1. Introduction

The increase in computing power and electronic storage capacity has lead to an exponential increase in the amount of digital content available to users in the form of images and video, which form the bases of many educational, entertainment and commercial applications [1].

Consequently, the search for relevant information in the large space of image and video databases has become more challenging. The main challenge lies in the reduction of the semantic gap between low-level features extracted from the image and high-level user semantics. How to achieve accurate retrieval results, is still a challenging and an unsolved research problem. 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 [2]. In this work, we study 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 extracted from an image, in order to achieve the highest retrieval performance, at a relatively low computational cost? The main contribution of this work is a comprehensive comparison of four color feature extraction approaches and four texture feature extraction approaches for CBIR. In Section 2, we discuss the four color feature extraction techniques: 1) the conventional color histogram, 2) the fuzzy color histogram, 3) the color correlogram, and 4) a color/shape-based method. In Section 3, we discuss the four texture feature extraction techniques: 1) the steerable pyramid, 2) the contourlet transform, 3) the Gabor wavelet transform, and 4) the complex directional filter bank. In Section 4, we present a comparison of the color and texture methods. In Section 5, we present our experimental procedure and results, and we conclude in Section 6.

2. Color Feature Extraction Models

The extraction of the color features for each of the four methods is performed in the HSV (hue, saturation and value) perceptual color space, where Euclidean distance corresponds to the human visual system's notion of distance or similarity between colors.

2.1 The Conventional Color Histogram

The conventional color histogram (CCH) of an image indicates the frequency of occurrence of every color in the image. From a probabilistic perspective, it refers to the probability mass function of the image intensities. It captures the joint probabilities of the intensities of the color channels. The CCH can be represented as $h_{A,B,C}(a,b,c) = N.Prob(A=a, B=b, C=c)$, where A, B and C are the three color channels and N is the number of pixels in the image [3] (key paper #1). Computationally, it is constructed by counting the number of pixels of each color (in the quantized color space).

2.2 The Fuzzy Color Histogram

In the fuzzy color histogram (FCH) approach, a pixel belongs to all histogram bins with different degrees of membership to each bin. More formally, given a color space with K color bins, the FCH of an image I is defined as $F(I)=[f_1,f_2,\ldots f_k]$ where $f_i = \frac{1}{N}\sum_{j=1}^{N}\mu_{ij}$, where N is the number of pixels in the image and μ_{ij} is the membership value of the j^{th} pixel to the i^{th} color bin, and it is given by $\mu_{ij} = \frac{1}{1 + d_{ij}/\varsigma}$, where d_{ij} is the Euclidean distance between the color of pixel j (a 3-dimensional vector of the H, S and V components), and the i^{th} color bin, and ς is the average distance between the colors in the quantized color space [4] (key paper #2).

2.3 The Color Correlogram

The color correlogram (CC) expresses how the spatial correlation of pairs of colors changes with distance. A CC for an image is defined as a table indexed by color pairs, where the d^{th} entry at location (i,j) is computed by counting number of pixels of color j at a distance d from a pixel of color i in the image, divided by the total number of pixels in the image [5].

2.4 The Color/Shape-Based Method

Howe *et al.* have proposed in [6] (key paper #3) a color-shape based method (CSBM) in which a quantized color image I' is obtained from the original image I by quantizing pixel colors in the original image. A connected region having pixels of identical color is regarded as an object. The area of each object is encoded as the number of pixels in the object.

Further, the shape of an object is characterized by 'perimeter intercepted lengths' (PILs), obtained by intercepting the object perimeter with eight line segments having eight different orientations and passing through the object center.

3. Texture Feature Extraction Models

The notion of texture generally refers to the presence of a spatial pattern that has some properties of homogeneity [4]. Directional features are extracted to capture image texture information. The four texture feature extraction methods presented in this section generate a multi-scale, multi-directional representation of an image.

3.1 The Steerable Pyramid

The steerable pyramid recursively splits an image into a set of oriented sub-bands and a lowpass residual. The image is decomposed into one decimated lowpass sub-band and a set of undecimated directional sub-bands. Analytically, the bandpass filter in polar coordinates, at each orientation *l,* is composed of a radial part *H(r)* and an angular part $G_l(\theta)$. These are defined as: $H(r)=\begin{cases}\cos\left(\frac{\pi}{2}\log_2(\frac{2r}{\pi})\right), & \frac{\pi}{4}<r<\frac{\pi}{2}\\ 1, & r\geq\frac{\pi}{2}\\ 0, & r\leq\frac{\pi}{4}\end{cases}$, and

$G_l(\theta)=\begin{cases}\alpha_L[\cos(\theta-\frac{\pi l}{L})]^{L-1}, & \left|\theta-\frac{\pi l}{L}\right|<\frac{\pi}{2}\\ 0, & else\end{cases}$, where $\alpha_L=2^{L-1}\frac{(L-1)!}{\sqrt{L[2(L-1)]!}}$, and L is the total number of orientations [7] (key paper #4).

3.2 The Contourlet Transform

The contourlet transform is a combination of a Laplacian pyramid (LP) and a directional filter bank (DFB). The LP provides the multi-scale decomposition, and the DFB provides the multi-directional decomposition. The LP is a decomposition of the original image into a hierarchy of images, such that each level corresponds to a different band of image frequencies. This is done by taking the difference of the original image and the Gaussian-lowpass-filtered version of the image, (at the appropriate scale σ). The Gaussian lowpass kernel is defined as: $H(w_1,w_2) = \exp\{-2(\pi\sigma)^2(w_1^2 + w_2^2)\}$, where w_1 and w_2 are the horizontal and vertical frequencies respectively. The bandpass images from the Laplacian pyramid are fed into the DFB so that directional information can be captured. The DFB realizes a division of the spectrum into 2^L wedge-shaped slices, as shown in Figure 1. A detailed description of the DFB is provided in [7]. The low frequency components are separated from the directional components. After decimation, the decomposition is iterated using the same DFB.

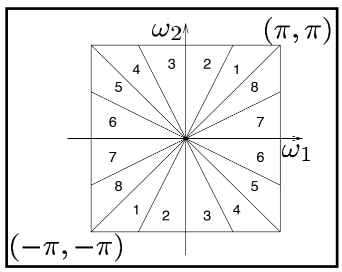

Fig. 1: DFB decomposition into 2^3 wedges

3.3 The Gabor Wavelet Transform

The Gabor function, in the Fourier domain, is given by $G(u,v) = \exp\{-\frac{1}{2}\left(\frac{u^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2}\right)\}$, where σ_u and σ_v are the bandwidths of the filter. The Gabor wavelet transform dilates and rotates the two-dimensional Gabor function. The image is then convolved with each of the obtained Gabor functions. To obtain a Gabor filter bank with *L* orientations and *S* scales, the Gabor function is rotated and dilated as follows: $G_{mn}(x,y) = a^{-m}G(x,y)$, where $x = a^{-m}(x\cos\theta + y\sin\theta)$, $y = a^{-m}(-x\sin\theta + y\cos\theta)$, and $\theta = n\pi/L$, n = 1, 2, …, L, and m = 0, 1, …, S-1 [7].

3.4 The Complex Directional Filter Bank

The complex directional filter bank (CDFB) consists of a Laplacian pyramid and a pair of DFBs, designated as primal and dual filter banks. The filters of these filter banks are designed to have special phase functions, so that the overall filter is the Hilbert transform of the primal filter bank. A multi-resolution representation is obtained by reiterating the decomposition in the lowpass branch [7]. The block *P* in Figure 2 shows one level of the CDFB, where $L_o(w)$, $G(w)$ and $F(w)$ are lowpass filters. A more detailed explanation of the CDFB is provided in [7].

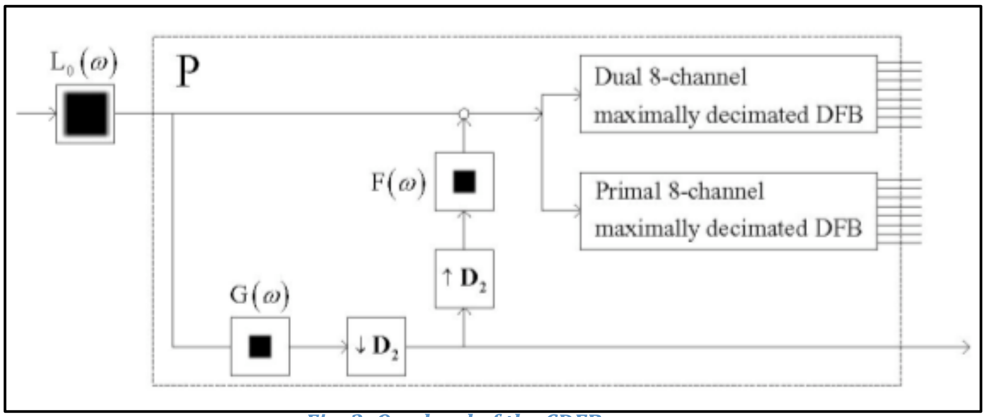

Fig. 2: One level of the CDFB

A plot of the frequency response of the four methods is shown in Figure 3.

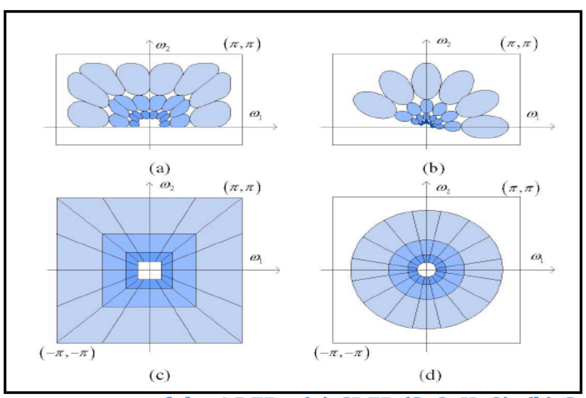

Fig. 3: Frequency response of the 4 DFBs. (a) CDFB (S=3, K=8), (b) Gabor Wavelet (S=4, K=6), (c) Contourlet Transform (S=3, K=8), (d) Steerable Pyramid (S=3, K=8)

4.Comparison of the Color and Texture Features

Table 1 lists the pros and cons of the color features described in section 2.

Color Feature	Pros	Cons
Conventional Color Histogram	-Simple -Fast computation	-High dimensionality -No color similarity -No spatial info
Fuzzy Color Histogram	-Fast computation -Encodes color similarity -Robust to quantization noise -Robust to change in contrast	-High dimensionality -More computation -Appropriate choice of membership weights needed
Color Correlogram	-Encodes spatial info	-Very slow computation -High dimensionality -Does not encode color similarity
Color/Shape Method	-Encodes spatial info -Encodes area -Encodes shape	-More computation -Sensitive to clutter -Choice of appropriate color quantization thresholds needed

Table 1: Pros and cons of the four-color features

The Table 2 lists the pros and cons of the texture features described in section 3.

Texture Feature	Pros	Cons
Steerable Pyramid	-Supports any number of orientation	-Sub-bands undecimated, hence more computation and storage
Contourlet Transform	-Lower sub-bands decimated	-Number of orientations supported needs to be power of 2
Gabor Wavelet Transform	-Achieves highest retrieval results	-Results in over-complete representation of image -Computationally intensive
Complex Directional Filter Bank	-Competitive retrieval results	-Computationally intensive

Table 2: Pros and cons of the four texture features

5. Experiments and Results

The simulations were performed in MATLAB. For color feature extraction, the HSV space was quantized to 128 color bins. For texture feature extraction, the transform parameters were set to perform an eight-orientation decomposition of the image at three levels of resolution. In other words, the scale parameter S was set to three, and the orientation parameter L was set to eight.

5.1 The Datasets

The color and texture features were extracted from the images in the Corel and the Brodatz image datasets respectively. The Corel dataset is a database of 10 classes, each containing 100 images [8]. In general, images within the same class have a similar color distribution. An example image from each class is shown in Figure 4. The Brodatz dataset is a database of 13 classes, each containing 10 images of one texture rotated at different angles [9]. An example image from each class is shown in Figure 5.

Fig. 4: An example image from each of the 10 classes in the Corel database

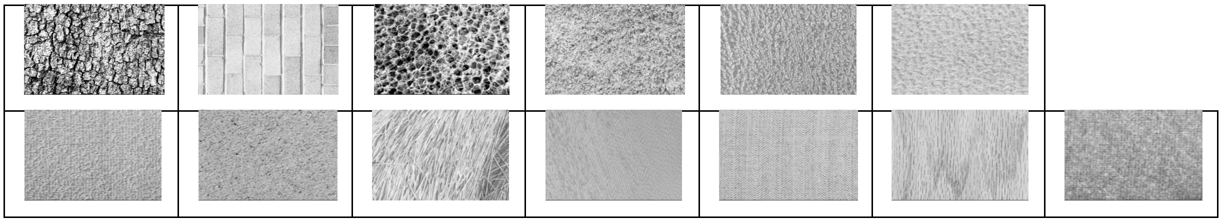

Fig. 5: An example image from each of the 13 classes in the Brodatz database

5.2 Distance Measure and Retrieval Score

One image from each class was chosen as a query image. The color (or texture) features were then extracted from the query image and from all the images in the database, (Corel in the case of color features, and Brodatz in the case of texture features). The features extracted from each image were represented as a vector in R^D, and Euclidean distance was used to measure the distance form the feature 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 the distance 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, in addition to the number of images that belong to the query image class, but that have not been returned by the system. The retrieval score for one class was then computed as $100 \times \left[1 - (mismatches/10)\right]\%$. Finally, the average retrieval score for all classes was computed as the average of the retrieval scores obtained for each class.

5.3 Color and Texture Retrieval Results

Tables 3 and 4 display the obtained color and texture retrieval results respectively.

	CCH	FCH	Correlogram	Color/Shape
Average Retrieval Score	80.12%	82.05%	69.48%	70.03%

Table 3: Color retrieval scores

	Steerable Pyramid	Contourlet Transform	Gabor	Complex Directional Filter Bank
Average Retrieval Score	63.02%	63.67%	81.48%	76%

Table 4: Texture retrieval scores

6. Conclusion and Future Work

The main contribution of this work is a comprehensive comparison of state-of-the-art color and texture feature extraction techniques for CBIR. The FCH and the Gabor wavelet transform were found to yield the highest color and texture retrieval results, respectively, at the cost of higher computational complexity. In future work, we will explore methods for combining color and texture features, in addition to incorporating user-feedback into the system. Another issue that will need to be addressed, is the issue of distance measures between feature vectors. Euclidean distance was used in this report because of its simplicity and interpretability, but it would be valuable to evaluate other distance measures and their effect on retrieval performance.

References

[1]. A. Bovik, *Handbook of Image and Video Processing*, 2nd Edition, Elsevier Academic Press, ISBN 0-12-119792-1, pp. 993-1013, 2005.

[2]. 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.

[3] J. R. Smith and S.-F. Chang, " Automated image retrieval using color and texture", Technical Report CU/CTR 408-95-14, Columbia University, July 1995.

[4]. J. Han and K. Ma, "Fuzzy Color Histogram and Its Use in Color Image Retrieval",*IEEE Trans. On Image Processing,* vol. 11, pp. 944 – 952, Aug. 2002.

[5]. J. Huang, S. R. Kumar, M. Mitra, W. J. Zhu and R. Zabih, "Image Indexing Using Color Correlograms", *Proc. IEEE Conf. on Computer Vision and Pattern Recognition,* pp. 762 – 768, June 1997.

[6]. N. R. Howe and D. P. Huttenlocher, "Integrating Color, Texture and Geometry for Image Retrieval", *Proc. IEEE Conf. on Computer Vision and Pattern Recognition,* vol. II, pp. 239-246, June 2000.

[7]. S. Oraintara and T. T. Nguyen, "Using Phase and Magnitude Information of the Complex directional Filter Bank for Texture Image Retrieval", *Proc. IEEE Int. Conf. on Image Processing,* vol. 4, pp. 61-64,Oct. 2007.

[8]. http://wang.ist.psu.edu/docs/related/

[9]. http://www.ux.uis.no/~tranden/brodatz.html

[10]. P Liu, K. Jia, Z. Wang and Z. Lv, "A New and Effective Image Retrieval Method Based on Combined Features", *Proc. IEEE Int. Conf. on Image and Graphics,* vol. I, pp. 786-790, August 2007.