COLOR TEXTURED IMAGE RETRIEVAL BY COMBINING TEXTURE AND COLOR FEATURES

Cong BAI, Kidiyo KPALMA and Joseph RONSIN

Université Européenne de Bretagne, France
INSA de Rennes, IETR, UMR 6164, F-35708, RENNES
{cong.bai, kidiyo.kpalma, joseph.ronsin}@ insa-rennes.fr

ABSTRACT

A new approach for color textured image retrieval based on the combination of color and texture features is proposed. The features are extracted in DCT domain. For texture featuring, Texture-Pattern is proposed to be constructed by using three groups of AC coefficients of each DCT block from the luminance component. And for color featuring, LumaColor-Pattern is constructed by using the DC coefficients from the luminance and chroma components. The histograms of dominant components of these two patterns are constructed and their combination is used for image retrieval. Experimental results on VisTex database have shown that the proposed method yields higher performance than referred approaches which are reported in recently published literature.

***Index Terms*—** Content-based image retrieval, DCT, color, texture

1. INTRODUCTION

Color and texture are two important features in Content-Based Image Retrieval (CBIR). A combined use of color and texture would provide better performance than that of color or texture alone. In general, two kinds of approaches of using the combination of color and texture features can be used for image retrieval: jointly and separately [1].

Under the jointly aspect, N pseudo gray-level images are derived from different spectral bands of color image. Traditional methods of extracting texture features in gray-level domain are implemented on these pseudo gray-level images. For example, in [2], red, green, and blue (RGB) channels from color iamge are seen as three pseudo gray-level images and Gabor filters are applied to them. In [3], a Chromatic Statistical Landscape Features (CSLF) method is proposed by applying Statistical Landscape Features (SLF) on the three pseudo gray-level images that are extracted from the three chromatic components of the HSI color model. And a recent approach is presented in [4], in which, multivariate Laplace distributions and Student-t distribution are considered to model the wavelet coefficients in RGB channels.

Another aspect of analyzing color images is to process them separately, that means to transform the color image into luminance and chrominance components and color and texture feature are extracted separately. For example, in [5], RGB image was firstly transformed into HSV image. The autocorrelogram of wavelets coefficients extracted from Hue and Saturation components is used as color feature, and the first and second moments of the BDIP (block difference of inverse probabilities) and BVLC (block variation of local correlation coefficients) for each subband of Value component is used as texture feature. In [6], color feature is represented by 2D histogram of *CIE Lab* chromaticity coordinates and texture features are extracted using Discrete Wavelet Frames (DWF) analysis.

In this paper, we propose a new approach using the combination of texture and color features to do color textured image retrieval. Color image is firstly converted from RGB to YCbCr. And then, the image is decomposed into 4x4 blocks which are transformed by Discrete Cosine Transform (DCT). Texture-Pattern is constructed by using the statistical information of directional group of AC coefficients in each DCT block from Y component and LumaColor-Pattern is constructed by using DC coefficients of each DCT block from Y, Cb and Cr components. Feature descriptors are represented by the histograms of dominant components of these two patterns and χ^2 distance is used to measure the similarity of feature descriptors between query and images in the database.

The rest of the paper is organized as follows: the general description of our proposal is presented in section 2. Section 3 gives the the procedure of forming texture and color features. Similarity measurement and combination of color and texture features are given in section 4. Experimental results are shown in section 5 and conclusion is given in the last section.

This work is partly supported by CAIMAN project within the French ANR (Agence Nationale de Recherche) program.

2. GENERAL DESCRIPTION

Figure 1 shows the block diagram of the proposed approach. The color image is firstly converted to a YCbCr color space which is adopted in JPEG standard. In this color space, there are three components: one is luminance component I_Y, and the other two are chroma components I_{Cb} and I_{Cr}. Then each component is decomposed into 4x4 blocks which are transformed by DCT. So we get one DC coefficient and fifteen AC coefficients for each DCT block. Both the non-overlapping and 50% overlapping DCT blocks were tested, and we found that 50% overlapping performs better. So we adopt this overlapping. Furthermore, luminance normalization should be applied to the DCT coeffcients to eliminate the effect of luminance variation [7]. And then the DCT coefficients are quantized with a quantization coefficient QP after normalization. In our approach, $QP = 30$. D_Y, D_{Cb} and D_{Cr} represent the DCT coefficients after normalization and quantization from each component. As the I_Y component can be seen as a gray-level copy of the original color image, and the texture feature is considered as mainly appearring in the luminance component of the image, the texture feature is extracted from this component. We select 9 AC coefficients of every block in D_Y to construct Texture-Pattern. As the DC coefficient is proportional to the average value of each block, LumaColor-Pattern is constructed by DC coefficients of each block from each component, D_Y, D_{Cb} and D_{Cr}. The histogram of patterns are generated as the number of occurrences of patterns in the DCT domain. Finally, we use the histogram of dominant Texture-Pattern H_T and histogram of dominant LumaColor-Pattern H_C to do retrieval.

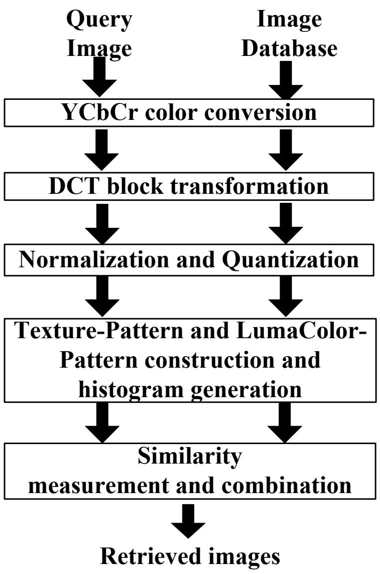

Fig. 1. Block diagram of proposal

3. PATTERN CONSTRUCTION AND HISTOGRAM GENERATION

3.1. Texture-Pattern construction

DCT has the capability to compact the energy, it means that much of the energy lies in low frequency coefficients, so that high frequences can be discarded without visible distortion. Furthermore, the AC coefficients of some regions represent some directional information.

From the point of view above, we select 9 coefficients out of all 15 AC coefficients in each block and categorize them into 3 groups: horizontal (Group H), vertical (Group V) and diagonal (Group D). For each group, the sum of the coefficients is calculated firstly and then the squared-differences between each coefficient and the sum of this group are calculated. Finally, the sums of squared-differences of each group are used to construct Texture-Patterns. The procedure of forming Texture-Pattern is shown in Figure 2.

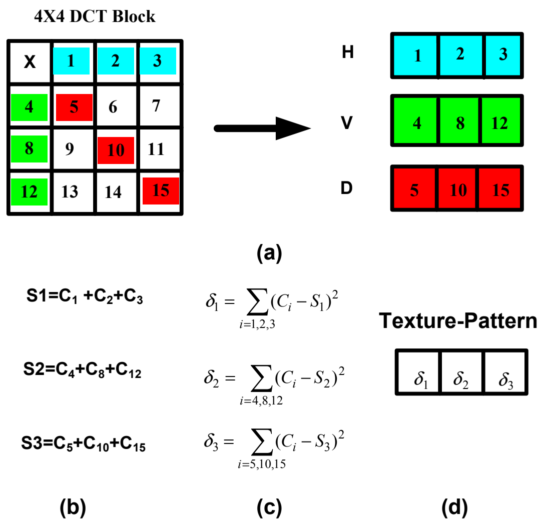

Fig. 2. Texture-Pattern:
(a) Three groups of AC coefficients extracted from DCT block (b) Sums of each group (c) Sums of squared-differences (d) Texture-Pattern

3.2. LumaColor-Pattern construction

As the DC coefficient is proportional to the average value of each block, DC coefficients in D_Y can represent the average luminance of each DCT block and DC coefficients in D_{Cb} and D_{Cr} can be seen as the average chrominance of each block.

From the above observation, the LumaColor-Pattern is constructed by the DC coefficients from the 3 components of each block in the image. The procedure of forming LumaColor-Pattern is shown in Figure 3.

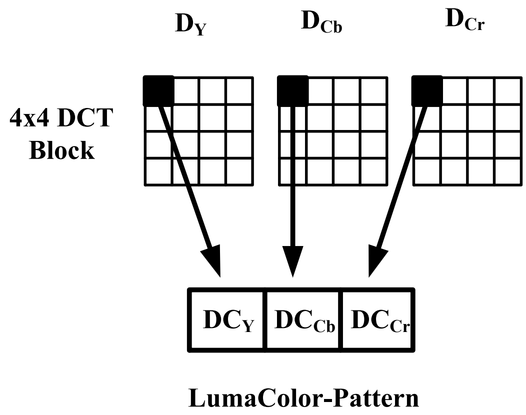

Fig. 3. LumaColor-Pattern

3.3. Histogram generation

As mentioned before, the histogram of Texture-Pattern and LumaColor-Pattern are defined as the number of occurrences of patterns in DCT domain. A disadvantage of the histogram method is that it requires a large number of histogram bins, typically several hundreds, to capture information of feature vector accurately. Thus it leads to complexity in both storage of image features and retrieval timing. To overcome this drawback, we adopt two improvements as demonstrated in [7][8]. Here we show the modifications in the histogram of Texture-Pattern only, but in the aspect of the histogram of LumaColor-Pattern, same amelioration can also be adopted.

The original histogram of Texture-Pattern is shown in Figure 4. From this histogram, we can get two conclusions and two improvements can be adopted respectively: the first is that there is only a few part of Texture-Patterns which appear in large quantities and a large number of Texture-Patterns that

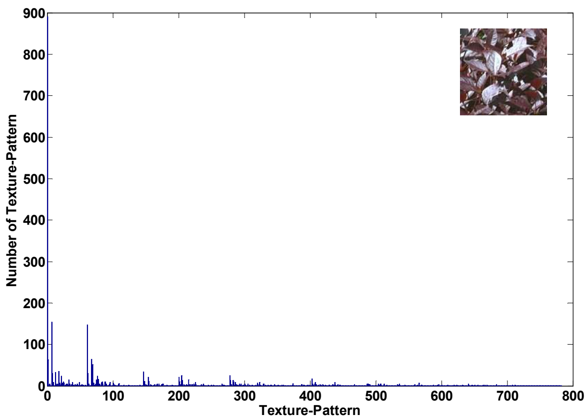

Fig. 4. Histogram of original Texture-Pattern

appear rarely. So in consideration of time-consuming and efficiency, we select those of Texture-Patterns which have higher frequency to construct the descriptor H_T. We use parameter $Tbins$ to represent the number of bins that are selected in the histogram of Texture-Patterns. The second one is that the first Texture-Pattern inside the histogram is very dominant. This Texture-Pattern mainly corresponds to uniform blocks of image and will not be considered as a representative pattern in the Texture-Pattern histogram.

So the histogram of Texture-Patterns that we will use for image retrieval is as shown in Figure 5. This histogram is obtained by selecting the first 300 ($Tbins = 300$) high frequencies Texture-Patterns from the histogram of Figure 4.

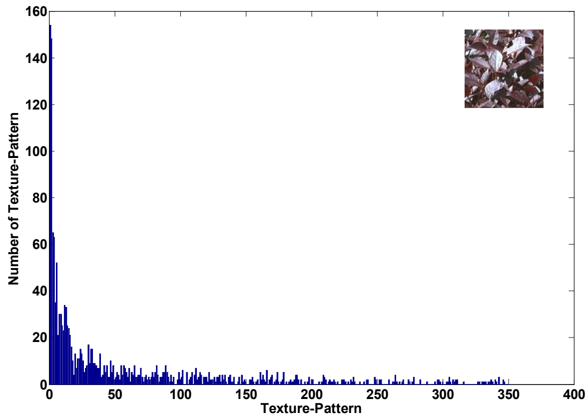

Fig. 5. Histogram of selected Texture-Pattern

Finally, we get H_T from Texture-Pattern and H_C from LumaColor-Pattern as the feature descriptors which are used for image retrieval.

4. SIMILARITY MEASUREMENT AND COMBINATION

4.1. Similarity measurement

The similarity between query and images in the database is measured by the distances between the descriptor of Texture-Pattern and the descriptor of LumaColor-Pattern. A similarity measure assigns a lower distance to the most similar images.

Assuming H_Q and H_D are feature descriptors (descriptor of Texture-Pattern and descriptor of LumaColor-Pattern in our study) of the query Q and images D in the database respectively, many classical methods can be used to compute the distance in metric space. In our approach, we choose χ^2 distance as the similarity measurement. The χ^2 distance is defined as follows:

$$Dis(Q, D) = \sum_{k=1}^{m} \frac{(H_Q(k) - H_D(k))^2}{H_Q(k) + H_D(k)} \quad (1)$$

in which m indicates the dimension of the histograms.

4.2. Distance normalization

Since each feature descriptor has its own physical meanings, and the value ranges of them are totally different, before the distance of color descriptor and that of texture descriptor are fused, they should be normalized.

Distances can be normalized through the following ways: given a query image, by calculating distances of texture descriptors between this query and all images in database, and those of color descriptors between this query and all images in database, two sets of distance $\{Dis_T(i)\}$ and $\{Dis_C(i)\}$ are obtained, where $i = 1, \cdots, N$. N is the number of images in the database. $Dis_T(i)$ is the distance of texture descriptor between query image Q and ith image in the database, and $Dis_C(i)$ is the distance of color descriptor between query image Q and ith image in the database. Thus the distance normalization can be implemented as:

$$
\begin{aligned}
Dis_{NT}(Q, D_i) &= \frac{Dis_T(i) - min\{Dis_T(i)\}}{max\{Dis_T(i)\} - min\{Dis_T(i)\}} \\
Dis_{NC}(Q, D_i) &= \frac{Dis_C(i) - min\{Dis_C(i)\}}{max\{Dis_C(i)\} - min\{Dis_C(i)\}} \qquad (2)
\end{aligned}
$$

where $Dis_{NT}(Q, D_i)$ and $Dis_{NC}(Q, D_i)$ are the normalized distances between query image Q and ith images in the database of texture and color descriptors respectively. Both type of distances now range from 0 to 1.

The global distance that is used to evaluate the similarity between the query and images in the database is then given by:

$$
Dis_G(Q, D_i) = (1-\alpha) \times Dis_{NT}(Q, D_i) + \alpha \times Dis_{NC}(Q, D_i) \qquad (3)
$$

where $\alpha \in \{0, 1\}$ is a weight parameter that can control the impact of color feature and texture feature in the procedure of image retrieval.

5. EXPERIMENTAL RESULTS

We have performed experiments of image retrieval on VisTex texture database [9]. The whole VisTex texture database has 167 natural textured images. Each image in the VisTex collection is of size 512×512 and for our experiments all of them are divided into sixteen non-overlapping 128×128 subimages. To compare with the other approaches, we evaluate our proposal on two sets of VisTex: one is the classical selection of 40 classes of textures that are used by many literatures about texture retrieval [10] [11]. This selection is displayed in Figure 6. And the other is the whole collection of VisTex, that means the selection of 167 classes of texture.

In the retrieval experiments, for both data sets, each subimage in the database is used once as a query. For comparison purpose, retrieved images are the first 16 most similar images for each query in our experiments. The relevant images for each query consists of all the subimages from the same original texture. We use the average retrieval rate (ARR) to evaluate the performance. For a given query image, and a given set of retrieved images, the retrieval rate is defined as the percentage of the number of relevant images retrieved on the total number of retrieved images. So the retrieval rate (RR) is defined as:

$$
RR = \frac{\sharp(relevant \quad images \quad retrieved)}{\sharp(retrieved \quad images)} \qquad (4)
$$

$\sharp(a)$ denotes the number of a.

ARR is defined as the mean value of the set of retrieval rate of each query. We also should emphasize that for different α in Equation 3, various ARR can be got because of different impact of color and texture feature in the process of retrieval. All the results presented below are the ARR when $\alpha = 0.35$. This weight parameter assures the best ARR that we can get.

Table 1 presents the comparative experimental results on the data set of 40 texture classes with referred methods. In this table, GCG represents Gaussian Copula with Gamma distributed margins [12], GFP represents Gaussian distribution with Fixed Point covariance matrix estimators [4] and Student-t is also the proposal presented in [4]. The comparison shows that our proposal performs better.

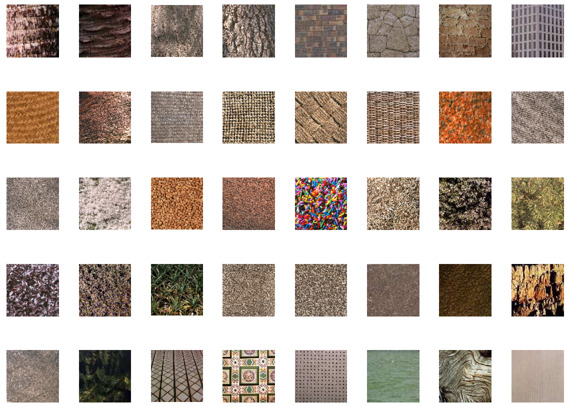

Fig. 6. Selected textures from VisTex database

Table 1. ARR in the selection of VisTex

Method	GCG	GFP	Student-t	Our Proposal
ARR(%)	85.83	88.23	89.65	**90.16**

Table 2 presents the retrieval performance on the whole VisTex database. CSLF indicates Chromatic Statistical Landscape Features [3]. Gabor indicates the Gabor filters used to do color texture retrieval in [2]. From this table, we can see that as many classes of texture in VisTex are not homogeneous, the retrieval rate is much lower than that of 40 classes. But our proposal still outperforms.

Table 2. ARR in the whole VisTex

Method	CSLF	Gabor	Our Proposal
ARR(%)	56.2	52.0	**58.09**

6. CONCLUSION

In this paper, we have proposed a new approach for image retrieval based on the combination of color and texture features. We presented a new method to extract texture feature and color feature from color image: this method uses the statistical information of directional group of AC coefficients of DCT from luminance component to construct Texture-Pattern and uses DC coefficients from luminance and chroma components to construct LumaColor-Pattern. Experiments that are excuted both on the 40 selected classes and the whole of the VisTex database have shown that our proposal outperforms the referred methods which are reported in recently published papers.

7. REFERENCES

[1] Maenpaa T.; Pietikainen M.;, "Classification with color and texture: jointly or separately?," Pattern Recognition, vol.37, no.8, pp.1629-1640, 2004.

[2] Jain, A.; Healey, G.;, "A multiscale representation including opponent color features for texture recognition," IEEE Transactions on Image Processing, vol.7, no.1, pp.124-128, Jan 1998

[3] Cun Lu Xu; Xian Tong Zhen;, "Chromatic Statistical Landscape Features for Retrieval of Color Textured Images," Internet Computing for Science and Engineering (ICICSE), 2009 Fourth International Conference on, On page(s): 98 - 101, Volume: Issue: , 21-22 Dec. 2009

[4] Bombrun, Lionel; Berthoumieu, Yannick; Lasmar, Nour-Eddine; Verdoolaege, Geert;, "Multivariate texture retrieval using the geodesic distance between elliptically distributed random variables," Image Processing (ICIP), 2011 18th IEEE International Conference on, pp.3637-3640, 11-14 Sept. 2011

[5] Young Deok Chun; Nam Chul Kim; Ick Hoon Jang;, "Content-Based Image Retrieval Using Multiresolution Color and Texture Features," IEEE Transaction on Multimedia, vol.10, pp. 1073-1084, 2008.

[6] Liapis, S.; Tziritas, G.;, "Color and texture image retrieval using chromaticity histograms and wavelet frames," IEEE Transactions on Multimedia, vol.6, no.5, pp. 676- 686, Oct. 2004

[7] Zhong Daidi; Defée I.;, "DCT histogram optimization for image database retrieval," Pattern Recognition Letters, vol. 26, pp.2272-2281, 2005

[8] Cong Bai; Kidiyo Kpalma; Joseph Ronsin;, "A New Descripor Based On 2D DCT For Image Retrieval," Proceedings of the International Conferenceon Computer Vision Theory and Applications (VISAPP), pages714-717, 2012

[9] Media Laboratory, MIT. VisTex Database of textures. `http://vismod.media.mit.edu/vismod/imagery/VisionTexture/`, Accessed Dec. 2010

[10] Kwitt, R.; Uhl, A., "Lightweight Probabilistic Texture Retrieval," IEEE Transactions on Image Processing, vol. 19, issue:1, pp.241253, Jan. 2010

[11] Kokare M.;Biswas P.K. and Chatterji, B.N.;, "Texture image retrieval using new rotated complex wavelet filters," IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 35, issue: 6, pp. 1168-1178, Dec. 2005

[12] Stitou, Y.; Lasmar, N.; Berthoumieu, Y.;, "Copulas based multivariate gamma modeling for texture classification," Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on, pp.1045-1048, 19-24 April 2009