



Efficient colour texture image retrieval by combination of colour and texture features in wavelet domain

Cong Bai, Wenbin Zou, Kidiyo Kpalma, Joseph Ronsin

► **To cite this version:**

Cong Bai, Wenbin Zou, Kidiyo Kpalma, Joseph Ronsin. Efficient colour texture image retrieval by combination of colour and texture features in wavelet domain. *Electronic Letters*, 2012, 48 (23), pp.1463-1465. <hal-00776192>

HAL Id: hal-00776192

<https://hal.archives-ouvertes.fr/hal-00776192>

Submitted on 15 Jan 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Efficient color texture image retrieval by combination of color and texture features in wavelet domain

Cong BAI, Wenbin ZOU, Kidiyo KPALMA and Joseph RONSIN

A new approach for image retrieval by combination of color and texture features is proposed. This approach uses the histogram of feature vectors which are constructed from the coefficients of some subbands of wavelet transform and chosen according to their intrinsic characters. K-means algorithm is used to quantize feature vectors. The experimental results both on small size database (40 classes of textures) and large size database (167 classes of textures) show that, compared with the state-of-the-art approaches, the proposed approach can get better retrieval performance.

Introduction: Content-based image retrieval (CBIR) is an active research field in pattern recognition and computer vision. Color and texture are two important features that are used in CBIR. Using the combination of both features provides better performance than that of color or texture alone. For example, in [1], red, green and blue (RGB) channels of color images are treated as three respective pseudo gray-level images and Gabor filters are applied on these three images to extract features. In our previous work [5], texture features are constructed from the AC coefficients of Discrete Cosine Transform (DCT) and color features are constructed from the DC coefficients. As wavelet is widely used as an efficient tool for extracting features, some researchers have presented image retrieval methods based on wavelet in recent years. So in [2], color features are represented by 2D histogram of *CIE Lab* chromaticity coordinates and texture features are extracted by using Discrete Wavelet Frames (DWF) analysis. In [3], RGB images were firstly transformed into HSV model. The color feature is represented by the autocorrelogram of wavelet coefficients extracted from Hue and Saturation components, 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. A recent approach is presented in [4], in which, the wavelet coefficients in RGB color channels are modeled by multivariate Laplace distribution and Student-t distribution. Other than mentioned methods, this paper presents a new method for color texture image retrieval combining color and texture features in wavelet domain. This method constructs the color and texture features from the coefficients of some subbands of wavelet transform.

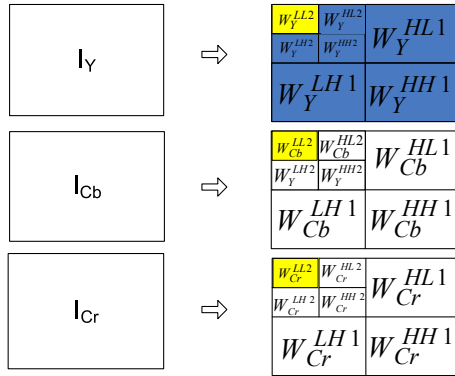


Fig. 1. Wavelet decomposition

Featuring principle : In our approach, color image is firstly converted to YCbCr whose components are I_Y , I_{Cb} and I_{Cr} . And then, each component is decomposed by 2-level subsampled dyadic Discrete Wavelet Transform (DWT). Results are represented by W_C^{mn} , where $C \in \{Y, Cb, Cr\}$ denotes the components and $m \in \{LL, HL, LH, HH\}$ denotes the subband orientation and $n \in \{1, 2\}$ the wavelet decomposition level. In our approach, we choose CDF 9/7 wavelets used in JPEG2000. 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 appearing in the luminance component of the image, the multiresolution texture-vectors

are constructed from wavelet coefficients of high-frequency subbands of Y channel in both decomposition levels, that's to say W_Y^{HLn} , W_Y^{LHn} and W_Y^{HHn} , where $n \in 1, 2$, as marked blue in Figure 1. On the other hand, color-vector is constructed from the coefficients of the lowest subbands of each component, that's to say W_Y^{LL2} , W_{Cb}^{LL2} and W_{Cr}^{LL2} , as marked yellow in Figure 1. By this way, only 41% wavelet coefficients are used for constructing feature vector.

Multiresolution texture-vectors and color-vector: Multiresolution texture-vectors contain two levels of resolution, respectively high and low, and referred as: HiRes texture-vector (HiResV) and LoRes texture-vector (LoResV). They are constructed by the coefficients at the same position from high-frequency subbands of two decomposition levels in Y component. HiResV contains three coefficients from three subbands W_Y^{HL1} , W_Y^{LH1} and W_Y^{HH1} respectively in first decomposition level. LoResV is constructed in the same way, but in second decomposition level from three different subbands W_Y^{HL2} , W_Y^{LH2} and W_Y^{HH2} respectively. Color-vector (ColV) is built in the similar way. Each color-vector includes three coefficients at the same position from the lowest-frequency subbands of Y, Cb and Cr components, W_Y^{LL2} , W_{Cb}^{LL2} and W_{Cr}^{LL2} respectively, and defined as follows:

$$\begin{aligned} HiResV &= [W_Y^{HL1}(x_H, y_H), W_Y^{LH1}(x_H, y_H), W_Y^{HH1}(x_H, y_H)] \\ LoResV &= [W_Y^{HL2}(x_L, y_L), W_Y^{LH2}(x_L, y_L), W_Y^{HH2}(x_L, y_L)] \\ ColV &= [W_Y^{LL2}(x_C, y_C), W_{Cb}^{LL2}(x_C, y_C), W_{Cr}^{LL2}(x_C, y_C)] \end{aligned} \quad (1)$$

where (x_H, y_H) , (x_L, y_L) and (x_C, y_C) indicate the coordinates of the coefficients in each subband. If the size of image is $M \times N$, then $x_H = (1, \dots, \frac{M}{2})$, $y_H = (1, \dots, \frac{N}{2})$, $x_L = (1, \dots, \frac{M}{4})$, $y_L = (1, \dots, \frac{N}{4})$ and $x_C = (1, \dots, \frac{M}{4})$, $y_C = (1, \dots, \frac{N}{4})$. These three vectors are used to construct feature descriptors.

Descriptor construction: We use the histogram of these vectors as the descriptor of the image. With the objective of reducing dimensions of descriptors, before generating the histogram, K-means algorithm is used to partition the color-vectors and texture-vectors of all the images in the database into clusters. The number of clusters K for different kinds of vectors that assure best retrieval performance are found experimentally. So for HiRes texture-vectors and LowRes texture-vectors, they are respectively clustered into 400 centers (K=400), and for color-vectors, K=2000. The histogram of vectors of an image is defined as the number of vectors that are the nearest to each clusters issued from K-means partition. Therefore the histogram of HiResV and that of LoResV both have 400 bins respectively. The histogram of ColV has 2000 bins..

Similarity measurement and combination: χ^2 distance is used to measure the similarity between feature descriptors of query and images from the database. This distance is defined as follows:

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

in which H_Q and H_D are feature descriptors of the query image Q and image D in the database, m indicates the dimension of the descriptors.

Since each feature descriptor has its own physical meanings, and its ranges of values are totally different, before using the combination of the distances of different descriptors, they should be normalized.

Distances can be normalized through the following ways: given a query image, by calculating distances of feature descriptors from this query and all images from database, one set of distances $\{Dis(i)\}$ is obtained, where $i = 1, \dots, N$. N is the number of images in the database. $Dis(i)$ is the distance between feature descriptors of query image Q and i^{th} image in the database. Thus the distance normalization can be implemented as:

$$Dis_N(Q, D_i) = \frac{Dis(i) - \min\{Dis(i)\}}{\max\{Dis(i)\} - \min\{Dis(i)\}} \quad (3)$$

where $Dis_N(Q, D_i)$ are the normalized distances between descriptors of query image Q and i^{th} image in the database. With this normalization, all type of distances now range from 0 to 1.

Let $Dis_{NT}(Q, D_i)$ and $Dis_{NC}(Q, D_i)$ be the normalized distances of texture and color descriptors respectively. And $Dis_{NT}(Q, D_i) = Dis_{NTH}(Q, D_i) + Dis_{NTL}(Q, D_i)$, where $Dis_{NTH}(Q, D_i)$ and $Dis_{NTL}(Q, D_i)$ are the normalized distances of HiRes texture-vector and LowRes texture-vector respectively. The global distance that is used to evaluate the similarity between the query and images in the database is then given by:

$$Dis_G(Q, D_i) = (1 - Wp) \times Dis_{NT}(Q, D_i) + Wp \times Dis_{NC}(Q, D_i) \quad (4)$$

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

Experimental results: We have evaluated our method on VisTex texture database [7]. The whole VisTex texture database has 167 natural texture images. Each image is of size 512×512 and all of them are divided into sixteen non-overlapping 128×128 subimages [4][5][8]. So we have a texture database with 16×167 subimages. We evaluate our proposal by two experiments: one is on the classical selection of 40 classes of textures that are used by many literatures about texture retrieval [4]. And the other is on the whole collection of VisTex, that means the selection of 167 classes of texture [1][5][8].

In the retrieval experiments, for both data sets, each subimage in the database is used once as a query. Average retrieval rate (ARR) is then used 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 over the total number of retrieved images. So the retrieval rate (RR) is defined as:

$$RR = \frac{\#(\text{relevant images retrieved})}{\#(\text{retrieved images})} \quad (5)$$

where $\#(a)$ denotes the number of a . For comparison purpose, the first 16 most similar images for each query are retrieved. The relevant images for each query consists of all the subimages from the same original texture. ARR is defined as the mean value of the set of retrieval rate of all query.

We also should emphasize that for different Wp in Equation (4), 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 for $Wp = 0.35$: it 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 including state of the art methods. In this table, GCG represents Gaussian Copula with Gamma distributed margins [6], GFP represents Gaussian distribution with Fixed Point covariance matrix estimators [4] and Student-t is also the proposal presented in [4]. DCT represents the method presented in [5]. The comparison shows that the new proposal performs better.

Table 1: ARR on the selection of VisTex

Method	GCG	GFP	Student-t	DCT	Proposal
ARR(%)	85.83	88.23	89.65	90.16	91.68

Table 2 presents the retrieval performance on the whole VisTex database. CSLF indicates Chromatic Statistical Landscape Features [8]. Gabor indicates the Gabor filters used to do color texture retrieval in [1]. Like previously, DCT represents the method presented in [5]. 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 on the case of 40 classes. But the new proposal provides significant improvement: more than 9% is obtained.

Table 2: ARR on the whole VisTex

Method	Gabor	CSLF	DCT	Proposal
ARR(%)	52.0	56.2	58.09	67.15

Conclusion: A simple but efficient approach for image retrieval by combination of color and texture features extracted from wavelet domain is presented. A new way for using some subbands of wavelet to construct the feature vectors is proposed. With CDF 9/7 wavelets, this approach can be applied easily on the JPEG2000 compressed images. Experimental

results on two data sets show that this new approach obtains better retrieval performance than the state-of-the-art approaches. Future work will concentrate on the improvement of scale and rotation invariance.

Cong BAI, Wenbin ZOU, Kidiyo KPALMA and Joseph RONSIN (*Université Européenne de Bretagne, INSA de Rennes, IETR UMR-6164, France*)

E-mail: {cong.bai, wenbin.zou, kidiyo.kpalma, joseph.ronsin}@insa-rennes.fr

References

- Jain, A.; Healey, G.: 'A multiscale representation including opponent color features for texture recognition', *IEEE Transactions on Image Processing*, 1998, **7**, pp.124-128
- Liapis, S.; Tziritis, G.: 'Color and texture image retrieval using chromaticity histograms and wavelet frames', *IEEE Transactions on Multimedia*, 2004, textbf6, pp. 676- 686
- Young Deok Chun; Nam Chul Kim; Ick Hoon Jang: 'Content-Based Image Retrieval Using Multiresolution Color and Texture Features', *IEEE Transaction on Multimedia*, **10**, pp. 1073-1084
- Bombrun, L.; Berthoumieu, Y.; Lasmar, N.; Verdoolaege, G.: 'Multivariate texture retrieval using the geodesic distance between elliptically distributed random variables', *Image Processing (ICIP)*, 2011 IEEE International Conference on, 2011, pp.3637-3640
- Cong, B.; Kidiyo K.; Joseph R.: 'Color textured image retrieval by combining texture and color features', To be appeared on Proceedings of 20th European Signal Processing Conference (EUSIPCO), 2012
- Stitou, Y.; Lasmar, N.; Berthoumieu, Y.: 'Copulas based multivariate gamma modeling for texture classification', *Acoustics, Speech and Signal Processing, IEEE International Conference on*, 2009, pp.1045-1048
- Media Laboratory, MIT. VisTex Database of textures. <http://vismod.media.mit.edu/vismod/imagery/VisionTexture/>, Accessed Dec. 2010
- Cunlu X.; Xiantong Zh.: 'Chromatic Statistical Landscape Features for Retrieval of Color Textured Images', *Internet Computing for Science and Engineering (ICICSE)*, International Conference on, 2009 pp.98 - 101