WAVELET-BASED TEXTURE RETRIEVAL USING INDEPENDENT COMPONENT ANALYSIS

Rui Zhang, Xiao-Ping Zhang, Ling Guan

Department of Electrical and Computer Engineering
Ryerson University
350 Victoria Street, Toronto, ON, Canada, M5B 2K3

ABSTRACT

In this paper, a novel approach to texture retrieval using independent component analysis (ICA) in wavelet domain is proposed. It is well recognized that the wavelet coefficients in different subbands are statistically correlated, resulting in the fact that the product of the marginal distributions of wavelet coefficients is not accurate enough to characterize the stochastic properties of texture images. To tackle this problem, we employ (ICA) in feature extraction to decorrelate the analysis coefficients in different subbands, followed by modeling the marginal distributions of the separated sources using generalized Gaussian density (GGD), and perform similarity measure based on the maximum likelihood criterion. It is demonstrated by simulation results on a database consisting of 1776 texture images that the proposed method improve the accuracy of texture image retrieval in terms of average retrieval rate, compared with the traditional method using GGD for feature extraction and Kullback-Leibler divergence for similarity measure.

Index Terms— Content-based image retrieval, texture retrieval, independent component analysis, generalized gaussian density, mutual information

1. INTRODUCTION

With the availability of large volume storage equipments and easy access to the Internet, image databases in various areas, such as academic research, medical applications, and personal photo album management, have been unprecedentedly enriched. Content-based image retrieval (CBIR), intended to provide users with effective and efficient tools for searching and browsing such databases still remains a challenging problem due to its limited performance based on low level features, such as color, shape, and texture.

Texture information receives much attention throughout the development of CBIR techniques, in which the class of statistical approaches to texture analysis in transform domain plays an important role. An approach to texture retrieval based on Gabor wavelet features and norm-based distance function was proposed in [1]. Wavelet-based texture retrieval using a more flexible statistical model, i.e. generalized Gaussian density (GGD), was studied in [2]. In addition, the similarity measure (SM) and feature extraction (FE) were jointly considered as the estimation and detection in a maximum likelihood (ML) framework, providing a justifiable definition on the SM using Kullback-Leibler divergence (KLD). The KLD-based distance measure was proved to be asymptotically equivalent to the maximum likelihood criterion yet with reduced computational complexity. In the above methods, the SM is performed by applying the respective distance functions in different wavelet subbands, followed by an additive combination of the resulting distances. According to the chain rule of KLD, the afore-mentioned methodology of SM can only be used to approximate the distance between the probability density functions (PDF's) of two texture images, taking into consideration that the marginal distributions of the coefficients in different subbands are correlated to each other. To compensate for this drawback, a vector wavelet domain hidden Markov model (WD-HMM) was developed for texture characterization and retrieval problems in [3].

We tackle the problem of inter-subband dependence of wavelet-based texture characterization from a different perspective, i.e. independent component analysis (ICA) is employed to decorrelate the marginal distributions of the coefficients in different subbands such that the modeling of the joint distribution as the product over the marginal distributions is justifiably accurate. The application of the proposed method to wavelet-based texture retrieval is studied. The proposed method is theoretically simpler than those modeling the inter-subband dependence using complex models, such as WD-HMM. The application of ICA to texture analysis was first considered in [4], in which the problem of texture classification and synthesis were addressed. However, the separated sources were modeled using channel histograms and only 8 images were used in the experiments, by which the effectiveness of the method using ICA with application to texture retrieval is not predictable. To demonstrate the potential of our

The work presented is jointly supported by Canadian Foundation of Innovation (CFI) and Nature Sciences and Engineering Research Council (NSERC).

method through its ability to reduce the cross-subband dependence, we compared the mutual information (MI) before and after using ICA. As a filter-based method, we employ steerable pyramid decomposition. We study the application of ICA to wavelet-based texture retrieval through the combination of ICA with GGD and make further investigation on the effectiveness of the proposed method using an image database consisting of 1776 texture images. In addition, we explain why KLD-based SM is not applicable to our method. As an initial investigation of the proposed method and its application to CBIR, we compared it with the wavelet-based texture retrieval using GGD. It is shown by simulation results that the new method outperforms the one using only GGD in terms of average retrieval rate.

2. DESCRIPTION OF THE PROPOSED METHOD

2.1. A Maximum Likelihood Framework of CBIR

An ML framework of CBIR can be described as follows. Denote the image data by $\boldsymbol{X}$, which in our case is the image signatures with respect to the filtering functions of the multi-scale multi-orientation decomposition. Assuming that $\boldsymbol{X}$ consist of L independently and identically distributed (i.i.d.) random vectors, i.e. $\boldsymbol{X} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_L\}$, drawn from $p(\boldsymbol{x}|\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is the parameter set of the PDF of the image. Considering a database of M images, the FE can be regarded as a problem of ML estimation, i.e.

$$\hat{\boldsymbol{\theta}}_i = \underset{\boldsymbol{\theta}_i \in \boldsymbol{\Theta}}{\operatorname{argmax}} \sum_{l=1}^{L} \log p(\boldsymbol{x}_l^i|\boldsymbol{\theta}_i), i = 1, 2, \ldots, M, \quad (1)$$

where $\hat{\boldsymbol{\theta}}_i$ is the feature set of the ith image in the database with image data represented as $\boldsymbol{x}_l^i$, and $\boldsymbol{\Theta}$ is the set of possible values on which the PDF is defined. Jointly considered with the FE, the SM can be thought of as a problem of ML detection, in which the likelihood in (2) is maximized,

$$\log p(\boldsymbol{X}^q|\boldsymbol{\theta}_i) = \sum_{l=1}^{L} \log p(\boldsymbol{x}_l^q|\boldsymbol{\theta}_i), \quad (2)$$

by choosing $\boldsymbol{\theta}_i$, the feature of the nth top matched image, from the set of features of all images in the database, given the query image data $\boldsymbol{X}^q$. Asymptotically equivalent to the ML rule, a KLD-based approach to SM can be applied to the above problem.

It can be observed that the parameter estimation of a multivariate distribution is required in the above problem. The complexity, however, can be considerably reduced if independent marginal distributions are found, while avoiding the inaccuracy caused by the approximation used in [2].

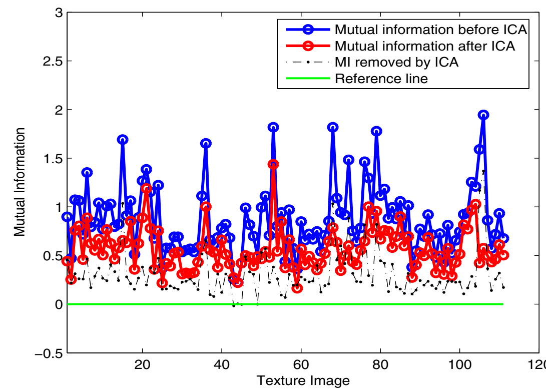


Fig. 1. The comparison between the MI's with and without ICA.

2.2. Decorrelation of the Marginal Distributions of Subband Coefficients Using Independent Component Analysis

As a useful statistical signal processing method, ICA characterizes the nongaussian structure of the data being analyzed. It transforms the data using a set of basis functions adapted to the data such that the signatures of the data in different channels are statistically independent. The ICA algorithm using the extended infomax principle [5] is employed in the proposed method. To illustrate the potential of ICA to remove the cross-subband dependence, we calculated the MI's, before and after ICA, of the response of a filter bank to 111 representative texture images, of which each one is from a different texture class in the database. As described in the first section, a steerable pyramid decomposition with 2 scales and 4 orientations was used. To reduce the computational intensity, we selected the analysis coefficients from two subbands in each scale. As illustrated in Fig. 1, the marginal distributions are indeed decorrelated which is indicated by the reduction of MI. It should be noted that in practice ICA can not completely separate the multivariate random data into statistically independent sources, i.e. residual dependence always remains.

2.3. Texture Retrieval Using ICA and GGD

The CBIR framework under consideration is composed of two parts, i.e. FE and SM. Shown in Fig. 2 is the block diagram of the proposed texture retrieval framework.

In the FE step, an image is first decomposed using a multi-resolution analysis, which in our case is the steerable pyramid decomposition. Then the analysis coefficients are transformed using the basis vectors adapted to the texture image, followed by the statistical modeling using GGD defined as

$$p(x|\alpha, \beta) = \frac{\beta}{2\alpha\Gamma(1/\beta)} e^{-(|x|/\alpha)^\beta}, \quad (3)$$

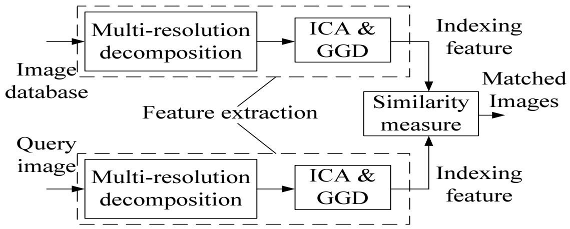


Fig. 2. The block diagram of the proposed framework of texture retrieval.

where x is a scalar random variable, and α and β are scale and shape parameters of the GGD, controlling the width and the decreasing rate of the peak of the PDF. In our study, texture images within the same class are assumed to have the same ICA basis functions with the purpose of reducing computational complexity. Therefore, the feature of an image is composed of the parameters of the GGD and the filtering matrix of the ICA. In the SM step, the likelihood in (2) can be expressed with appropriate normalization as

$$\frac{1}{L}\log p(\boldsymbol{X}^q|\boldsymbol{\theta}_i)=\frac{1}{L}\sum_{j=1}^{J}\sum_{l=1}^{L}\log p(s^q_{k,j,l}|\psi^i_j) - \log|\det(\boldsymbol{W}_k^{-1})|, \tag{4}$$

where $s^q_{j,l,k}$ is the lth IC of the query image in the jth channel filtered using the basis functions $\boldsymbol{W}_k$ adapted to the kth texture class to which the ith candidate image belongs, and $\psi^i_j = \{\alpha^i_j, \beta^i_j\}$ is the parameter of the source distribution.

It should be noted that scheme of SM in [2] is not applicable to the proposed method. Based on the ML criterion as expressed in (4), the subband coefficients of a given query image have to be filtered using the basis functions adapted to candidate texture image to be compared, which depend on the texture class to which the to-be-compared image belongs. The marginal distributions of the projection coefficients after the transforms by different $\boldsymbol{W}_k$'s can not always be considered as being independent, unless the to-be-compared texture image and the query image are in the same class, which is not known prior to the retrieval.

3. SIMULATION RESULTS

The database of texture images used in our experiments were constructed by dividing a texture image with the resolution of 640×640 into 16 texture images with the resolution of 160×160, which constitute a class of images with the same type of texture. Images within the same class were obtained from 16 non-overlapping areas of each of the 111 texture images with higher resolution. A few sample texture images were shown in Fig. 3. In our experiments, query images were taken from the image database, which means the ideal retrieval result is that all of the 16 images in the same class as the query image are returned as the top 16 images with the first one being the query image itself. Two steerable pyramid filters were employed with 2 scales of decomposition, i.e. sp3Filter and sp5Filter, which are characterized by 4 orientations/scale and 6 orientations/scale, respectively. Following [1], the performance of the proposed method is evaluated using the average retrieval rate which is defined as the average percentage of the images in the same class as the query among the top 16 matched images. Through the comparison in Table 1, it can be shown that our method outperforms the wavelet-based texture retrieval using GGD. In addition, we compared the average retrieval rates of the two methods as a function of the number of top matched images considered. As illustrated in Fig. 4, to achieve the same average retrieval rate, the number of matched images which need to be considered to reach a certain retrieval accuracy is considerably reduced by the proposed method. To compare the effectiveness of the two methods to different type of texures, they were also compared based on the average retrieval rates with respect to different classes of textures. We randomly selected the results of 35 classes out of 111 classes. As demonstrated in Fig. 5, it can be observed that the proposed method consistently provides better retrieval accuracy than that of the method using GGD. We also developed a graphic user interface for the purpose of subjective evaluation, in which the query image is not considered as part of the retrieval result. As illustrated in Fig. 6, 4 and Fig. 4, all of the 15 images in the same class as the query image are returned as the top matched images, while three irrelevant images in a different class are retrieved using the method employing GGD.

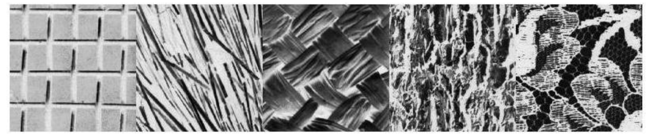


Fig. 3. Sample texture images with resolution of 640×640.

	GGD+KLD	GGD+ML	GGD+ICA+ML
sp3Filter	62.94%	62.24%	70.54%
sp5Filter	61.64%	61.31%	62.62%

Table 1. Performance comparison by average retrieval rate.

4. CONCLUSION

The wavelet-based texture retrieval using ICA is considered, in which the ICA is used to decorrelate the marginal distributions of the analysis coefficients resulting from a multi-resolution decomposition of a texture image. Based on the mutual information, we demonstrate that the cross-subband dependence is considerably removed, resulting in the fact that the marginal distributions of the separated sources modeled by GGD and the estimated ICA transformation matrix provides a more accurate statistical modeling of texture images, which in turn improves the performance of texture retrieval in terms of average retrieval rate. Simulation results on a database consisting 1776 texture images indicate the superior performance of the proposed approach to the one using GGD. As an initial study, we leave as part of the future work the comparison between the new method and those characterizing cross-scale cross-orientation dependence by complicated models, such as WD-HMM. In addition, the method of SM suitable to the proposed method with lower computational intensity comparable to that of KLD will be investigated as well.

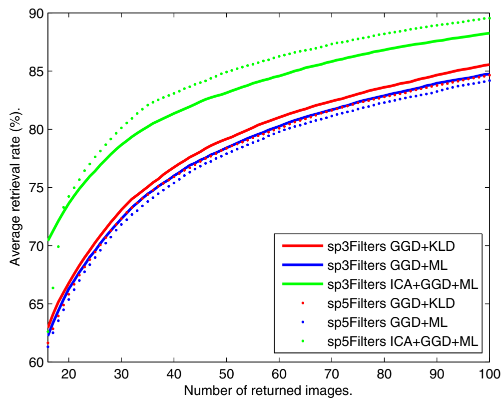


Fig. 4. The average retrieval rate with respect to different numbers of returned images considered.

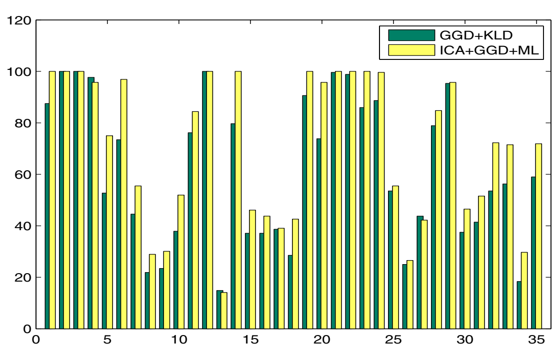
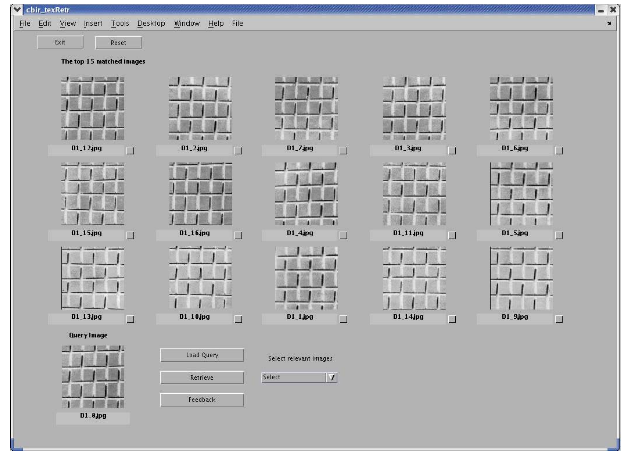
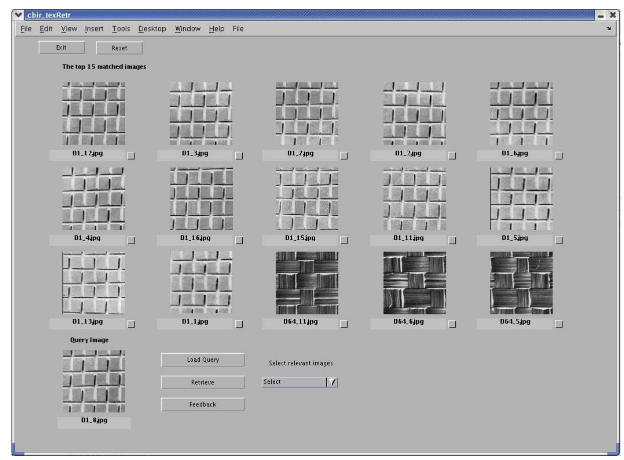


Fig. 5. Performance comparison by average retrieval rate with respect to different texture classes using sp3Filter.



(a) Result using ICA and GGD.



(b) Result using GGD.

Fig. 6. Subjective evaluation of the retrieval results.

5. REFERENCES

[1] B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18, no. 8, pp. 837–842, Aug. 1996.

[2] M. N. Do and M. Vetterli, "Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance," *IEEE Trans. Image Processing*, vol. 11, no. 2, pp. 146–158, Feb. 2002.

[3] M. Do and M. Vetterli, "Rotation invariant texture characterization and retrieval using steerable wavelet-domain hidden markov models," *IEEE Trans. Multimedia*, , no. 4, pp. 517–527, 2002.

[4] R. Manduchi and J. Portilla, "Independent component analysis of textures," in *Proc. the 7th IEEE International Conference on Computer Vision (ICCV 1999)*, Kerkyra, Greece, 1999, vol. 2, pp. 1054–1060.

[5] T. Lee, M. Girolami, and T. J. Sejnowski, "Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources," *Neural Computation*, pp. 417–441, 1999.