Multivariate statistical modeling for texture analysis using wavelet transforms

Nour-Eddine Lasmar, Yannick Berthoumieu

MULTIVARIATE STATISTICAL MODELING FOR TEXTURE RETRIEVAL USING WAVELET TRANSFORMS

Nour-Eddine LASMAR, Yannick BERTHOUMIEU

IMS- Groupe Signal - UMR 5218 CNRS, ENSEIRB – Université de Bordeaux, France
Emails : {nour-eddine.lasmar, yannick.berthoumieu}@ims-bordeaux.fr

ABSTRACT

This paper deals with texture modeling for classification or retrieval systems using multivariate statistical features. We propose to model neighborhoods of wavelet coefficients using Spherically Invariant Random Vectors (SIRVs). Under this multivariate model we provide a closed form of Kullback-Leibler divergence between joint distributions to measure similarity. The performances of the proposed model in retrieval are conducted on the VisTex image database aiming to compare the recognition rates with conventional approach of using univariate models such as the Generalized Gaussian distribution and with a recent multivariate model of wavelet coefficients called Multivariate Bessel K forms (MBKF).

Index Terms— image texture analysis, multivariate model, Kullback-Leibler Divergence, wavelet transforms, information retrieval

1. INTRODUCTION

The accurate characterization of texture is fundamental in various image processing applications, ranging from retrieval in large image databases to segmentation and texture synthesis.

Some of the most popular texture extraction methods for retrieval are based on filtering or wavelet-like approaches. The conventional scheme of multiscale texture analysis consists of modeling subband coefficients and uses the model hyperparameters as a signature for a specific texture class.

Many univariate prior models such as the Generalized Gaussian distribution (GGD) [1] and the Student t-distribution model [2] have been used to successfully characterize the marginal probability distribution of wavelet coefficients of textured and natural images in subbands. In order to model coefficients magnitude, Gamma and Weibull distributions have been introduced to achieve good retrieval rates [3] [4].This kind of representation leads to a simple and attractive approach but univariate modeling does not provide a complete statistical description of the images.

Latterly, researchers started to study the joint statistics of the wavelet coefficients of both textured and natural images. Some models and methods were formulated to explore statistical dependences existing across scale, orientation and position. Portilla and Simoncelli presented a statistical model based on joint statistics of steerable pyramid coefficients [5]. In their work, efficient algorithm of texture synthesis was developed giving increased synthesis quality. However this model is not tractable for classification applications due to the largeness of the signature. Tzagkarakis *et al.* [6] proposed a computationally complex Gaussianization procedure of the filter banks output in order to model wavelet coefficients with a multivariate normal distribution. Powerful statistical algorithms have been developed for image denoising using a Multivariate Generalized Gaussian Distribution (MGGD) [7] and Elliptically Contoured Distribution (ECD) [8], but no closed expression exists for the Kullback-Leibler Divergence (KLD) between these joint distributions to measure similarity in a retrieval or classification context. Recently, Boubchir *et al.* [9] introduced Multivariate Bessel K Forms density (MBKF) to characterize joint statistics in wavelet domain. Oppositely to MGGD and ECD, we can develop a closed solution for KLD between MBKF densities, so we can introduce MBKF model in a retrieval scheme.

In this paper, we propose to model the joint probability distribution of wavelet coefficients in a neighborhood using SIRV based representation: the product of the square root of a positive scalar quantity and a d-dimensional zero mean Gaussian vector. Our contribution is to propose an accurate parametric model characterizing the joint statistics of wavelet detail subband coefficients and to provide closed form solution to the KLD to measure similarity.

The remainder of this paper is as follows. We review in section 2 the statistical retrieval framework. In section 3 SIRV based multivariate model is introduced and the related KLD is calculated. Finally, in section 4 experimental results are given to evaluate retrieval performance and section 5 concludes the paper and proposes an outlook on further research.

2. PROBABILISTIC IMAGE RETRIEVAL

We establish the formal framework of probabilistic image retrieval. Consider an image database with M images $I_i, 1 \leq i \leq M$. Each image is represented by a data matrix $D_i = [\vec{x}_{i1}, \ldots, \vec{x}_{in}]$. From the probabilistic point of view, each data matrix contains n realizations of *i.i.d* random vectors $\vec{X}_1, \ldots \vec{X}_n$, which follow a parametric joint distribution with probability density function (PDF) $p_X(\vec{x}; \theta_i)$.

The retrieval task is to search the N most similar images to a given query image I_q. It is natural to select the most similar image I_r to I_q as the one whose parameter θ_s leads to a maximization of the log-likelihood function, i.e.

$$s = \arg\max_i \frac{1}{n} \sum_{j=1}^{n} \log\left(p(\vec{x}_{ij}; \theta_i)\right) \tag{1}$$

Using the weak law of large number, we have

$$s \overset{n \to \infty}{=} \arg\max_i E_{\theta_q}\left[\log\left(p_X(\vec{x}; \theta_i)\right)\right] \tag{2}$$

$$= \arg\max_i \int \ldots \int p_X(\vec{x}; \theta_q) \log\left(p_X(\vec{x}; \theta_i)\right) d\vec{x} \tag{3}$$

where the term E_{θ_q} denotes the expectation with respect to $p_X(\vec{x}; \theta_q)$. Equation (3) can be rewritten as the following minimizing problem

$$s = \arg\min_i -\int \ldots \int p_X(\vec{x}; \theta_q) \log\left(p_X(\vec{x}; \theta_i)\right) d\vec{x} \tag{4}$$

$$= \arg\min_i \int \ldots \int p_X(\vec{x}; \theta_q) \log\left(\frac{p_X(\vec{x}; \theta_q)}{p_X(\vec{x}; \theta_i)}\right) d\vec{x} \tag{5}$$

This can be seen as equivalent to minimizing the KLD between $p_X(\vec{x}; \theta_q)$ and $p_X(\vec{x}; \theta_i)$ noted $KLD(p_q \| p_i)$. So, to select the N top matches to the query image I_q we retrieve the set of images $\{I_{k_1}, I_{k_2}, \ldots, I_{k_N}\}$ where

$$KLD(p_q \| p_{k_1}) \leq KLD(p_q \| p_{k_2}) \leq \cdots \leq KLD(p_q \| p_{k_N}) \tag{6}$$

3. SIRV BASED MODELING

Spherically Invariant Random Vectors (SIRVs) has been appropriately used in modeling non-Gaussian problems. This is for instance, the case for radar clutter returns [10], radio fading analysis [11], or sonar interferences [12]. The joint statistics of wavelet coefficients also exhibit the evident non-Gaussianity and SIRV model is suitable to characterize these statistics.

Consider an image decomposed into oriented subbands at multiple scales. We denote as $x^{s,o}(n, m)$ the wavelet coefficient at scale s, orientation o and centered at spatial location $(2^s n, 2^s m)$. We denote as $\vec{x}$ a neighborhood of coefficients clustered around this reference coefficient.

We assume the coefficients within each local neighborhood around a reference coefficient of a subband are characterized by a SIRV model. Formally, a random vector $\vec{x}$ is a Spherically Invariant Random Vector [13] if it is the product of the square root of positive random variable τ called the *texture* and a d-dimensional independent zero-mean Gaussian vector $\vec{g}$ with covariance $M = E(\vec{g}^t \vec{g})$ verifying $Tr(M) = d$:

$$\vec{x} = \sqrt{\tau} \times \vec{g} \tag{7}$$

The joint density of vector $\vec{x}$ is determined by the covariance matrix M and the mixing density $p_\tau(\tau)$:

$$p_{\vec{X}}(\vec{x}) = \int_0^{+\infty} \frac{p_\tau(\tau)}{(2\pi)^{d/2}|\tau M|^{1/2}} \exp\left(-\frac{\vec{x}^t M^{-1} \vec{x}}{2\tau}\right) d\tau \tag{8}$$

To complete the model, we need to specify the probability density $p_\tau(\tau)$. We propose Weibull distribution as an appropriate description of the *texture* τ, given by:

$$p_\tau(\tau; a, b) = \frac{a}{b}\left(\frac{\tau}{b}\right)^{a-1} \exp\left(-\left(\frac{\tau}{b}\right)^a\right), \quad \tau \succ 0 \tag{9}$$

where $a \succ 0$ is the shape parameter and $b \succ 0$ is the scale parameter.

By inserting Eq. (9) to Eq. (8), the joint density which models the vector of wavelet neighbors is:

$$p_{\vec{X}}(\vec{x}) = \int_0^{+\infty} \frac{a\tau^{a-1}}{b^a(2\pi)^{d/2}|\tau M|^{1/2}} \exp\left(-\left(\frac{\vec{x}^t M^{-1} \vec{x}}{2\tau} + \left(\frac{\tau}{b}\right)^a\right)\right) d\tau \tag{10}$$

In this case, the hyperparameters of the correspondent joint distribution $p_{\vec{X}}(\vec{x})$ are noted (M, a, b).

To estimate these parameters we will first use a fixed point (FP) estimate for M. We note $x_i, i = 1, \ldots, N$ the realizations of the d-dimensional vector $\vec{x}$ (in our case, d is the size of the neighborhood). In their recent work [14], Pascal *et al* provide the proof of existence and uniqueness of $\hat{M}$ as the solution of the following:

$$\hat{M} = f\left(\hat{M}\right) \tag{11}$$

where f is given by

$$f\left(\hat{M}\right) = \frac{d}{N} \sum_{i=1}^{N} \frac{\vec{x}_i \vec{x}_i^t}{\vec{x}_i^t \hat{M}^{-1} \vec{x}_i} \tag{12}$$
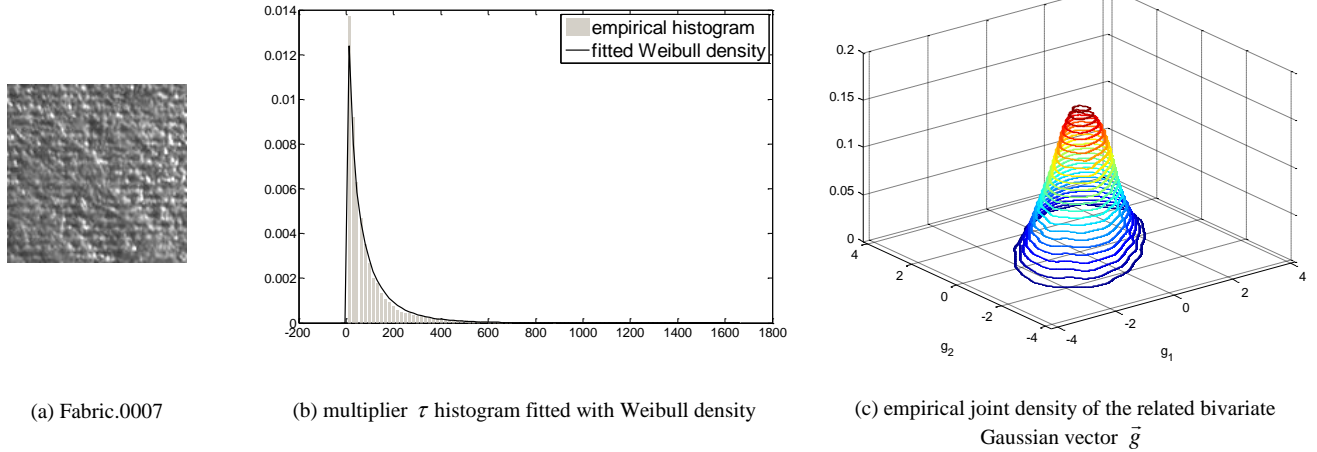
Equation (11) is solved using an iterative procedure with the initial guess from the Maximum Likelihood (ML) estimate

$$\hat{M}_0 = \frac{1}{N} \sum_{i=1}^{N} \vec{x}_i \vec{x}_i^t$$

Experiments show that typically only around five iteration steps are required to obtain convergence. We note that FP estimation of the covariance matrix M does not depend on the *texture* τ but only on the vectors x_i.

Then, the ML estimate of *texture* τ is given by:

$$\hat{\tau}_i = \frac{\vec{x}_i^t \hat{M}^{-1} \vec{x}_i}{d}, i = 1, \ldots, N \tag{13}$$

(a) Fabric.0007   (b) multiplier τ histogram fitted with Weibull density   (c) empirical joint density of the related bivariate Gaussian vector $\vec{g}$

Fig. 1.Example texture and SIRV representation in the case of bivariate modeling for the subband B_{11} resulting from steerable pyramid decomposition (B_{so} denotes the subband at scale s and orientation o)

Once we estimate $\hat{\tau}_i, i = 1,\dots,N$, the Weibull parameters (a,b) can be estimated using ML which we can found in [15]:

$$\frac{1}{\hat{a}} = \frac{\sum_{i=1}^{N} \tau_i^{\hat{a}} \ln(\tau_i)}{\sum_{i=1}^{N} \tau_i^{\hat{a}} - \ln(\tau_i)} \quad ; \hat{b} = \left(\frac{1}{N}\sum_{i=1}^{N} \tau_i^{\hat{a}}\right)^{1/\hat{a}} \tag{14}$$

Using SIRV modeling, joint distribution of wavelet coefficient is represented by a univariate Weibull distribution and a multivariate Gaussian distribution. Taking a neighborhood of dimension equal to 2, this is illustrated in Fig.1 (b) that shows the good fitting of Weibull density with the normalized histogram of the multiplier τ estimated from a detail subband obtained by steerable pyramid decomposition; the empirical density of the correspondent bivariate Gaussian vectors $\vec{g}$ is presented in Fig.1 (c).

In the best of our knowledge, the proposed joint distribution in Eq. (10) doesn't have a closed analytical form. However, we can derive a closed form solution for the KLD thanks to SIRV representation.

Consider two joint distributions $f_1(\vec{x}; M_1, a_1, b_1)$ and $f_2(\vec{x}; M_2, a_2, b_2)$. In a SIRV representation, the *texture τ* and the Gaussian vector $\vec{g}$ are independent so the KLD between the two joint distributions is the sum of the KLD between the KLD of the two Weibull distributions and the KLD between the two multivariate Gaussian densities:

$$KLD(f_1(\vec{x}; M_1, a_1, b_1) \| f_1(\vec{x}; M_2, a_2, b_2)) =$$
$$\underset{Weibull}{KLD}(p_1(\tau; a_1, b_1) \| p_2(\tau; a_2, b_2)) + \underset{Gaussian}{KLD}(p_1(\vec{g}; M_1) \| p_2(\vec{g}; M_2)) \tag{15}$$

In other hand

$$\underset{Weibull}{KLD}(p_1(\tau; a_1, b_1) \| p_2(\tau; a_2, b_2)) = \Gamma(\frac{a_2}{a_1} + 1)(b_1/b_2)^{a_2} + \ln(b_1^{-a_1} a_1)$$
$$- \ln(b_2^{-a_2} a_2) + \ln(b_1)(a_1 - a_2) + \frac{\gamma a_2}{a_1} - \gamma - 1$$

where γ denotes the Euler-Mascheroni constant ($\gamma \approx 0.57721$) and $\Gamma(.)$ the gamma function.
and the KLD for the d -dimension Gaussian case is:

$$\underset{Gaussian}{KLD}(p_1(\vec{g}; M_1) \| p_2(\vec{g}; M_2)) = 0.5\left(tr(M_2^{-1} M_1) + \ln\left(\frac{|M_2|}{|M_1|}\right) - d\right)$$

4. EXPERIMENTAL RESULTS

The experiments give an evaluation of the proposed model in the framework of texture retrieval. We use the same experimental setup presented in [1] and [4]. We work with 40 texture classes from VisTex database [16]. From each of these texture images of size 640x640 pixels, 16 subimages of 160x160 are created. A test database of 640 texture images is thus obtained. A query image is any one of these images in the database. The relevant images for each query are the other 15 images obtained from the same original 640x640 image.

$$recall = \frac{number\ of\ relevant\ retrieved\ images}{number\ of\ relevant\ images}$$

$$precision = \frac{number\ of\ relevant\ retrieved\ images}{number\ of\ retrieved\ images}$$

We employ the steerable pyramid decomposition proposed in [5], we note *Nsc* the number of scales and *Nor* the number of orientations.Nevertheless, orthonormal or biorthogonal wavelet representations can be used. We use the conventional criterion of *precision/ recall* to compare the performance of the proposed SIRV model with the retrieval approach using GGD presented in [1] and with the MBKF distribution [9].

For multivariate modeling, the neighborhood may include wavelet coefficients from other subbands (i.e. corresponding to nearby scales and orientations) as well as from the same subband. In our experiments, we used a neighborhood drawn from the same subband.
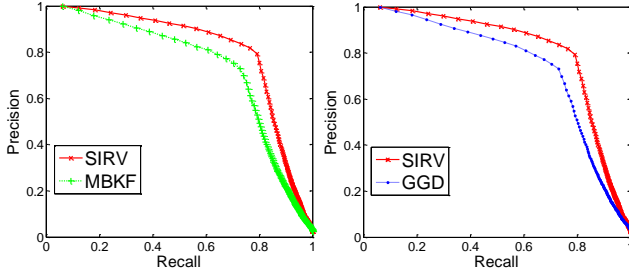
Fig.2. Recall-Precision curves showing the impact improvement of using SIRV model compared to GGD and MBKF models

($Nsc = 2$, $Nor = 6$)

Nsc	Nor	GGD	MBKF	SIRV
1	5	63.4473	63.9941	**77.6270**
1	6	64.5996	65.6152	**78.0957**
2	5	72.5098	71.8262	**78.3984**
2	6	73.1152	72.8711	**79.3164**

Table 1: Average retrieval rate (%) comparison
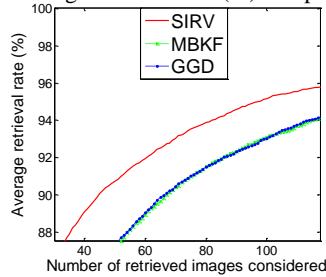
Fig.3. models convergence comparison

We use KLD as image similarity for the tree compared model. For MBKF distributions, developing thesolution leads to:

$$\underset{MBKF}{KLD}(p_X(\vec{x};\alpha_1,\Sigma_1)\|p_X(\vec{x};\alpha_2,\Sigma_2)) = (\Psi(\alpha_1)-1)(\alpha_1-\alpha_2)$$

$$+\ln\left(\frac{\Gamma(\alpha_2)}{\Gamma(\alpha_1)}\right)+\alpha_2\ln\left(\frac{\alpha_1}{\alpha_2}\right)+0.5(tr(\Sigma_2^{-1}\Sigma_1)+\ln\left(\frac{|\Sigma_2|}{|\Sigma_1|}\right)-d)$$

where $\Psi(.)$ is the digamma function.

The recall/precision curves presented in Fig.2 show the improvement obtained by using SIRV model.

The results of average retrieval rates according to different scales and orientations are summarized in Table 1. We can see that SIRV modeling significantly improves recognition rates, e.g from 73% to 79%, compared with GGD and MBKF.

Furthermore, we observe also that the proposed method converge faster than the two others. For example we retrieve 90% of the relevant images with a query of size 45 when we must consider a query of size 75 to retrieve the same percentage if we employ the two others models

5. CONCLUSION

In this work we have shown that image retrieval improve considerably when wavelet coefficients are jointly modeled using SIRV model. In a statistical retrieval framework, we have proposed a closed form for the Kullback-Leibler Divergence as a similarity measure. The model is validated in the retrieval system and achieves better recognition rates

compared to GGD and MBKF distributions. Future work includes a novel wavelet-based Bayesian denoiser based on the presented model.

REFERENCES

[1] M. Do and M. Vetterli, "Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance," *Image Processing, IEEE Transactions on,* vol. 11, 2002, pp. 146-158.

[2] Jinggang Huang and D. Mumford, "Statistics of natural images and models," *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on., 1999*, p. 547 Vol. 1.

[3] J. Mathiassen, A. Skavhaug, and K. Bø, "Texture Similarity Measure Using Kullback-Leibler Divergence between Gamma Distributions," *European Conference on Computer Vision 2002*, 2002, pp. 19-49.

[4] R. Kwitt and A. Uhl, "Image similarity measurement by Kullback-Leibler divergences between complex wavelet subband statistics for texture retrieval," *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, 2008, pp. 933-936.

[5] J. Portilla and E.P. Simoncelli, "A Parametric Texture Model Based on Joint Statistics of Complex Wavelet Coefficients," *International Journal of Computer Vision,* vol. 40, 2000, pp. 49-70.

[6] G. Tzagkarakis, B. Beferull-Lozano, and P. Tsakalides, "Rotation-invariant texture retrieval with gaussianized steerable pyramids," *Image Processing, IEEE Transactions on*, vol. 15, 2006, pp. 2702-2718.

[7] D. Cho and T.D. Bui, "Multivariate statistical modeling for image denoising using wavelet transforms," *Signal Processing: Image Communication*, vol. 20, Jan. 2005, pp. 77-89.

[8] S. Tan and L. Jiao, "Multivariate Statistical Models for Image Denoising in the Wavelet Domain," *International Journal of Computer Vision*, vol. 75, Nov. 2007, pp. 209-230.

[9] L. Boubchir, R. Boumaza and B. Pumo, "Multivariate Statistical Modeling of Images in Wavelet and Curvelet Domain using the Bessel K Form Densities," *Image Processing, 2009. ICIP 2009. 16th IEEE International Conference on*, 2009, accepted.

[10] E. Conte, A. De Maio, and G. Ricci, "Recursive estimation of the covariance matrix of a compound-Gaussian process and its application to adaptive CFAR detection ," *Signal Processing, IEEE Transactions on*, vol. 50, 2002, pp. 1908-1915.

[11] K. Yao, M. Simon, and E. Bigiieri, "Unified theory on wireless communication fading statistics based on SIRP," *Signal Processing Advances in Wireless Communications, 2004 IEEE 5th Workshop on*, 2004, pp. 135-139.

[12] T. Barnard and F. Khan, "Statistical normalization of spherically invariant non-Gaussian clutter," *Oceanic Engineering, IEEE Journal of*, vol. 29, 2004, pp. 303-309.

[13] Kung Yao, "A representation theorem and its applications to spherically-invariant random processes," *Information Theory, IEEE Transactions on,* vol. 19, 1973, pp. 600-608.

[14] F. Pascal, Y. Chitour, J. Ovarlez, P. Forster, and P. Larzabal, "Covariance Structure Maximum-Likelihood Estimates in Compound Gaussian Noise: Existence and Algorithm Analysis," *Signal Processing, IEEE Transactions on*, vol. 56, 2008, pp. 34-48.

[15] K. Krishnamoorthy, *Handbook of Statistical Distributions with Applications,* Chapman & Hall, 2006.

[16] MIT Vision and modeling group. Vision texture. [online].
Available from: http://vismod.www.media.mit.edu