MULTIVARIATE STATISTICAL MODELING OF IMAGES WITH THE CURVELET TRANSFORM

Larbi Boubchir and Jalal M. Fadili

Image processing group, GREYC UMR CNRS 6072
ENSICAEN, 6 Bd du Maréchal Juin 14050 Caen, France
E-mail: {boubchir,jfadili}@greyc.ensicaen.fr

ABSTRACT

This paper proposes a theoretical and statistical study to characterize the dependencies of the curvelet coefficients of images across position, scale and orientation. Our study was based on estimated histograms of the marginal and joint distributions to study the statistical properties of curvelet coefficients, and on the mutual information to measure the level of dependence between these coefficients. Finally, a novel multivariate statistical model, namely the anisotropic multivariate generalized gaussian (AMGGD), was proposed to characterize these dependencies.

1. INTRODUCTION

During the last ten years, wavelets had an immense success in image processing, and were used for many problems such as image restoration and compression [1]. These problems often involve parcimonious image representation through modern adaptive sparse representations borrowed from harmonic analysis (e.g wavelets and beyond).

However, wavelets are not optimal for detecting anisotropic objects in image (e.g. lines, contours), but are efficient for detecting isotropic objects at different scales. Recently, new X-let multiscale transforms have been developed -such as the curvelets, contourlets and bandlets- which integrate the concept of directionality and detect objects in an optimal way. Their effectiveness in image processing still remains to confirm.

In this paper, we present a statistical analysis of the coefficients' dependencies of the images across position, scale and orientation (directions) in the curvelet domain [2, 3]. To characterize these dependencies, we study the marginal and joint statistics of the curvelet coefficients and assess the dependency levels via mutual information[4]. In order to capture these dependencies, we have proposed a novel analytical multivariate model, namely the anisotropic multivariate generalized gaussian (AMGGD).

2. BACKGROUND

2.1. Curvelets

Curvelets as proposed by E. Candès and D. Donoho [2], constitute a relatively new family of frames that are designed to represent edges and other singularities along curves much more efficiently than the traditional wavelet-based transforms. For example, as shown in Fig.1(a), it would take many wavelet coefficients to accurately represent such a curve. Compared with wavelets, curvelets can represent a smooth contour with much fewer coefficients for the same precision (Fig.1(b)).

The curvelet transform is a multiscale transform with frame elements indexed by location, scale and orientation parameters, and have time-frequency localization properties of wavelets but also shows a very high degree of directionality and anisotropy. More precisely, we here use a new tight frame of curvelets recently developed in [5]. Further details about curvelets can be found in [2, 5].

Fig.2(a) shows a typical set of curvelet coefficients of the image "Boat" divided into three scales and six orientations using the Digital Curvelet Transform (DCT) of [6]. Fig.2(b) shows the corresponding spectral partitioning of the curvelet transform [5, 3, 6]. A curvelet is shown in Fig.2(c).
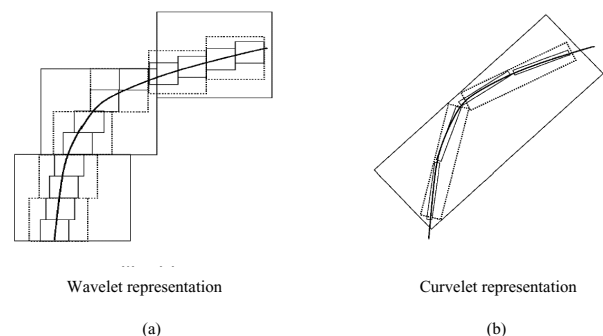

Wavelet representation   Curvelet representation

(a)   (b)

Fig. 1. Comparison of non-linear approximation performance for curvelets and wavelets.

2.2. Coefficients Relationships

For each curvelet coefficient X, we define its *neighbors* (NX) in the same orientation as its eight adjacent coefficients. Next, the coefficient in the same spatial location
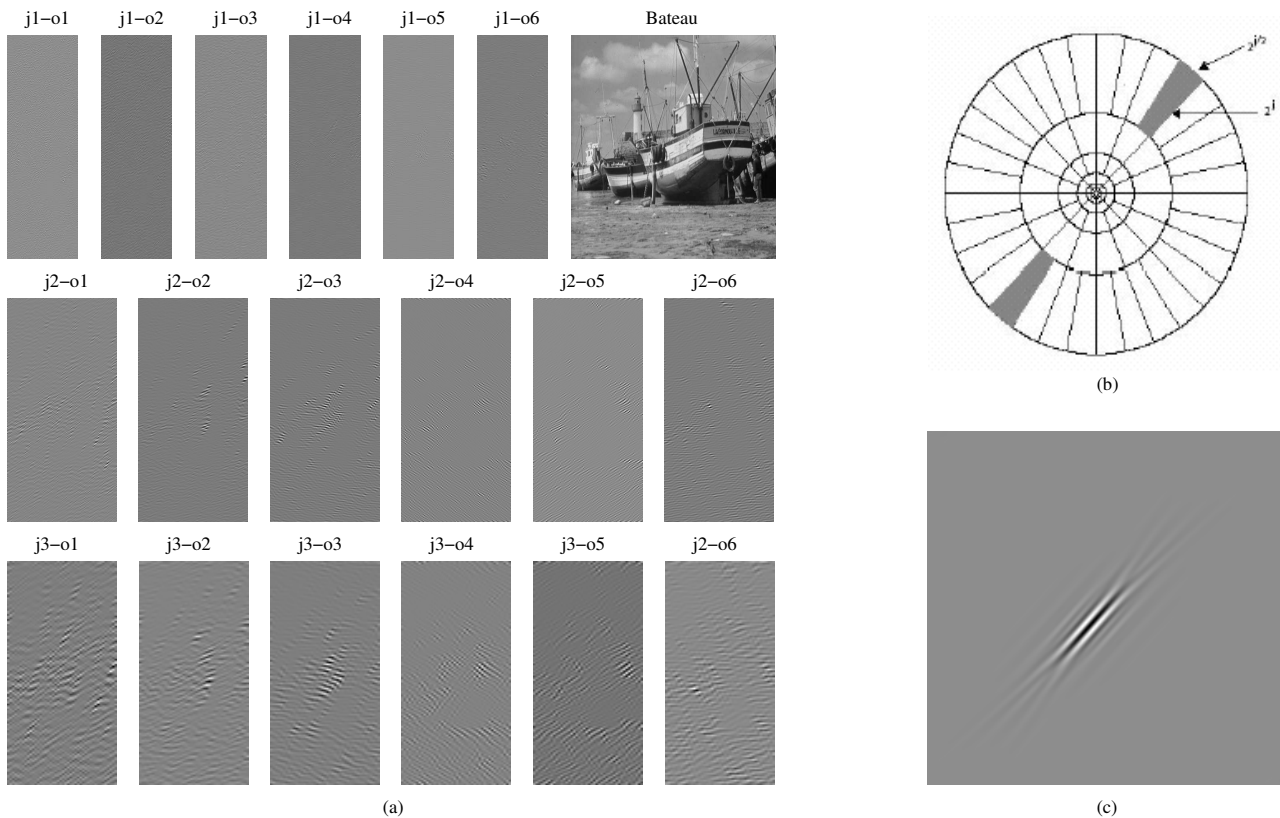
Fig. 2. (a) An example of digital curvelet transform of the image "Boat". Only three scales and six orientations are shown (j: is the scale, o: is the orientation). (b) Curvelet partitioning of the spectral plane. (c) A curvelet.

in the immediately coarser scale corresponds to its *parent* (PX), and the coefficients in the same scale and same spatial location but at different orientation are referred to as *cousins* (CX) of each other. There are more orientations in the curvelet representation compared to the wavelet representation where there are only three cardinal directions. Fig.3 graphically depicts these some important curvelet coefficient relationships.
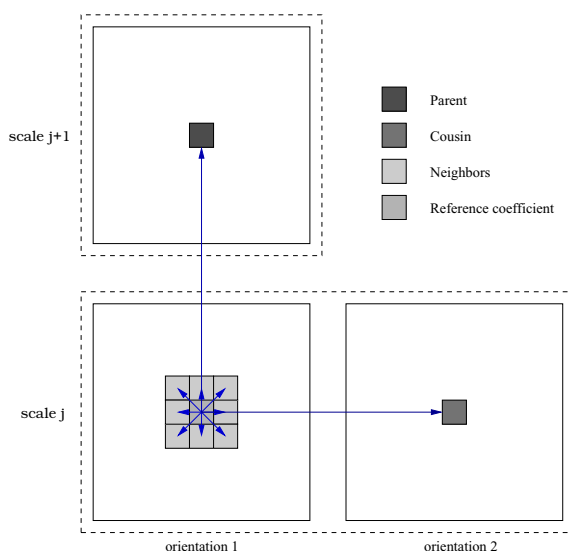

Fig. 3. Curvelet coefficient relationships.

3. CURVELET STATISTICS

Here, we present a statistical study to characterize the inter- and intra-scale dependencies of the curvelet coefficients of images.

3.1. Marginal statistics

We first study the marginal statistics of the curvelet coefficients of images. Fig.4(a) shows the histograms of two orientations at two successive scales of the image "Boat" (coarser and next-to-coarser scale). The distributions are characterized by a very sharp peak at zero amplitude and extended tails to both sides of the peak (leptokurtic). This leptokurtic behavior is observed on all histograms of all orientations and scales of all images in our test set. This implies that the curvelet transform is very sparse, as the majority of coefficients have amplitudes close to zero. The kurtosis of these distributions are measured and are higher than the gaussian value of 3. Thus, the marginal distributions of images in curvelet domain are highly non-gaussian.

3.2. Joint Statistics

Next, we study the joint statistics of curvelet coefficients at different positions, scales and orientation. In Fig.2(a), it can be seen that large coefficients tend to cluster spatially around the edges of objects in the image "Boat", and also persist to other scales and other orientations.
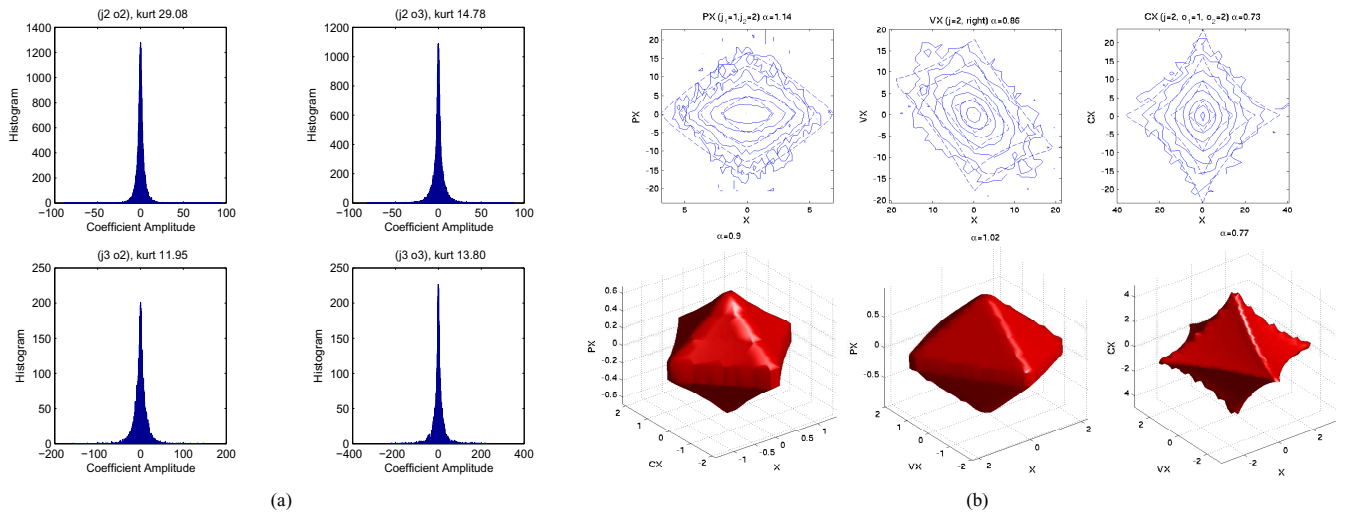
Fig. 4. (a) Marginal distributions of two orientations at two successive scales of the image "Boat". (b) Joint distributions of curvelet coefficients of the image "Boat" in the bivariate and trivariate case, $P(X, .)$ and $P(X, ., .)$, where $(.)$ refers to the parent PX, neighbors VX and cousin CX.

Fig.4(b) plots the joint distributions of curvelet coefficients of the image "Boat" in the bivariate and trivariate case, $P(X, .)$ and $P(X, ., .)$, where $(.)$ refers to the parent, neighbor or cousin.

3.3. Mutual Information

Here, we propose a quantitative study of the joint statistics of curvelet coefficients to numerically assess the dependencies between coefficients. The mutual information [4] is used as a quantitative measure of dependencies [7, 8].

Table 1 shows the average mutual information estimate over a database of 100 test images [9]. It can be observed that the position dependency is the most prominent, followed by the scale and orientation coefficients dependencies. This confirms our expectations from Fig.4(b). Furthermore, The inter-orientation dependencies are less important because of the curvelet partitioning illustrated in Fig.1(a). However, the curvelets system don't form an orthogonal base but a frame. Therefore, dependencies between orientations still subsist.

Table 1. The average mutual information estimate over a database of 100 test images.

	Coarser scale	Next-to-coarser scale
$MI(X; PX)$	0.164	0.194
$MI(X; VX)$	0.374	0.555
$MI(X; CX)$	0.142	0.151

4. MULTIVARIATE STATISTICAL MODEL

Here, we introduce a class of joint PDFs which are able to model the non-gaussian heavy tailed behavior and the dependencies in a multivariate setting.

4.1. Proposed model

Definition 1
Let $\mathbf{X} = (X_1, X_2, \ldots, X_d)^T$ a vector of random variables defines in $\mathbb{R}^d$. If $\mathbf{X}$ is distributed according to an **anisotropic** multivariate generalized gaussian, then its PDF can be written as follows:

$$p_{\mathbf{X}}(\mathbf{x}; \alpha, \Sigma) = \frac{\det \Sigma^{-1/2}}{(Z(\alpha)A(\alpha))^d} \exp\left(-\left\|\frac{\Sigma^{-1/2}\mathbf{x}}{A(\alpha)}\right\|_\alpha^\alpha\right) \tag{1}$$

where

$$Z(\alpha) = \frac{2}{\alpha}\Gamma\left(\frac{1}{\alpha}\right), \alpha > 0 \tag{2}$$

and

$$A(\alpha) = \sqrt{\frac{\Gamma\left(\frac{1}{\alpha}\right)}{\Gamma\left(\frac{3}{\alpha}\right)}} \tag{3}$$

α is the shape parameter and Σ is a symmetric positive definite matrix. $\|\mathbf{x}\|_\alpha = \sum_{i=1}^d |x_i|^\alpha$ stands for the l_α norm of vector $\mathbf{x}$.

This definition includes the particular case $d = 1$ (one random variable) and multivariate normal distribution case ($\alpha = 2$). This PDF model enjoys several interesting properties that we have established.

4.2. Hyperparameters estimation

We here propose estimators (moments and maximum likelihood) of the hyperparameters associated to: α and Σ.

An attractive property of parameterization Eq. 1 is that

$$\text{cov}(\mathbf{X}) = \Sigma \tag{4}$$

Therefore, $\text{cov}(\mathbf{X})$ can be easily estimated using the moments estimator. Plugging this in the likelihood equations, we can obtain α using the maximum likelihood es-

timator (ML). The ML estimator of α given Σ can be derived as being:

$$\hat{\alpha}_{ML} = \arg\min_{\alpha>0} LL(\alpha) = \arg\min_{\alpha>0} -\sum_{i=1}^{n} \log p_{\mathbf{X}}(\mathbf{x}_i; \alpha, \Sigma)$$
$$(5)$$

and $\hat{\alpha}_{ML}$ is then the solution of the following equation:

$$
\begin{aligned}
f(\alpha) &= \frac{\alpha \partial LL(\alpha)}{\partial \alpha} \\
&= \sum_{i,j} \left| \frac{y_{i,j}}{A(\alpha)} \right|^{\alpha} . \log \left| \frac{y_{i,j}}{A(\alpha)} \right|^{\alpha} \\
&\quad - \frac{1}{2} \sum_{i,j} \left| \frac{y_{i,j}}{A(\alpha)} \right|^{\alpha} . \left(\Psi\left(\frac{1}{\alpha}\right) - 3\Psi\left(\frac{3}{\alpha}\right) \right) \quad (6) \\
&\quad - nd - \frac{3nd}{2\alpha} \left(\Psi\left(\frac{1}{\alpha}\right) - \Psi\left(\frac{3}{\alpha}\right) \right) \\
&= 0
\end{aligned}
$$

j represents the jth component of vector $\mathbf{y}_i$ where $\mathbf{y}_i = \Sigma^{-1/2}\mathbf{x}_i$, $\mathbf{x}_{i,\{i=1,\dots,n\}}$ are the realizations of $\mathbf{X}$, and Ψ is the Digamma function (reminder, $\Psi(z) = \frac{d\ln(\Gamma(z))}{dz}$).

The following result establishes existence of the ML estimator of α.

Proposition 1

(i) $f(\alpha)$ has at least one root in $\mathbb{R}^{+*}$, then there exists at least one no-unique solution $\hat{\alpha}_{ML}$ as the ML estimator of α given Σ.

(ii) if $M = \max_{i,j\in\mathbb{N}} |\mathbf{y}_{i,j}| \leq \sqrt{3}$, then will always be strictly upper-bonded by $\frac{nd\log(3)}{2}$ at the global minimum $\hat{\alpha}_G$ of $LL(\alpha)$.

The goal of this result (Prop.1-(ii)) is that we can reject adaptatively some local minima of $LL(\alpha)$ (those above $\frac{nd\log(3)}{2}$).

According to Eq. 4 and proposition 1, we suggest the following hyperparameters estimation algorithm:

Algorithm 1 Hyperparameters estimation

1: Estime Σ
$$\hat{\Sigma} = \frac{1}{n} \sum_i \mathbf{x}_i \mathbf{x}_i^T \quad (7)$$

2: Estime α
To obtain $\hat{\alpha}$ we can solve numerically Eq. 6 using a descent method (e.g, gradient) taking advantage of Prop.1-(ii) in order to avoid local minima.

4.3. Application to AMGGD

In Fig.4(b), the PDF model was fitted (dotted) to the observed joint PDF (solid) in the bivariate case. One can appreciate the adequacy of this model to the observed statistics. This procedure has been applied successfully to 100 images database [9]. The proposed PDF model with it estimator behaved very stably

5. CONCLUSION

We have studied the properties of the curvelet coefficients of images. We have also shown that large coefficients tend to cluster spatially around the edges of objects in the image, and persist to other scales and other orientations. This implies that the curvelet coefficients are highly dependent on their parents and neighbors, as well as cousins in different scales and orientations. These dependencies can be quantitatively verified using mutual information, which shows that curvelet coefficients exhibits level of dependence on their neighbors, followed by parents, and then cousins. Based on the above properties, we have proposed an analytical multivariate model of PDF AMGGD in order to capture the dependencies between curvelet coefficients. Our future work is now focused on the application of this PDF model to Bayesian image restoration.

6. REFERENCES

[1] S. G. Mallat, *A Wavelet Tour of Signal Processing*, 2nd Edition, San Diego: Academic Press, 1999.

[2] E. Candès and D. Donoho, "Curvelets: A surprisingly effective nonadaptive representation of objects with edges," *Curves and Surface, Vanderbilt University Press, Nashville, TN*, pp. 123–143, 1999.

[3] J. Starck, E. Candès, and e Donoho, "The curvelet transform for image denoising," *IEEE Transactions on Image Processing*, vol. 11, no. 6, pp. 670–684, June 2002.

[4] T. M. Cover and J. A. Thomas, *Elements of information theory*, John Wiley and Sons, Inc., New York, 1991.

[5] E. Candès and D. Donoho, "New tight frames of curvelets and optimal representations of objects with c^2 singularities," Tech. Rep., Department of Statistics, Stanford University, USA, November 2002.

[6] E. Candès and L. Demanet, "The curvelet representation of wave propagators is optimally sparse," Tech. Rep., Applied and Computational Mathematics, California Institute of Technology, USA, 2004.

[7] J. Liu and P. Moulin, "Information-theoretic analysis of interscale and intrascale dependencies between image wavelet coefficients," *IEEE Transactions on Image Processing*, vol. 10, no. 10, pp. 1647–1658, November 2001.

[8] D. D. Y. Po and M. N. Do, "Directional multiscale statistical modeling of images x," *Proc. of SPIE conference on Wavelet Applications in Signal and Image Processing*, vol. 5207, pp. 69–79, August 2003.

[9] "http://sipi.usc.edu/services/database/database.html," .