HOW TO PERFORM TEXTURE RECOGNITION FROM STOCHASTIC MODELING IN THE WAVELET DOMAIN

Abdourrahmane M. ATTO *and Yannick* BERTHOUMIEU

Université de Bordeaux, UB1, IPB, ENSEIRB-MATMECA,
Laboratoire IMS UMR 5218, Groupe Signal et Image,
351 cours de la libration, 33405 Talence Cedex, FRANCE

ABSTRACT

The paper addresses content-based image retrieval from texture databases, by using stochastic modeling in the wavelet domain. It proposes an analysis of the key parameters involved in such a content-based texture retrieval. These parameters are the *wavelet order* and the *goodness-of-fit measure* used to select the best family of distributions for modeling the subband wavelet coefficients. It is shown that taking suitable parameters into consideration makes it possible to attain high retrieval rates in content-based texture retrieval.

Index Terms— Stochastic modeling ; Wavelets ; Texture ; Similarity.

1. INTRODUCTION

Stochastic modeling in the wavelet domain has proven efficiency in Content Based Image Retrieval (CBIR) on texture databases, see [1], [2], [3] and [4], among others. The contribution of this paper in this research field consists in providing statistical analyzis that makes possible selection of nearly optimal parameters among the key parameters involved in CBIR. These fundamental parameters are 1) the wavelet function and 2) the goodness-of-fit measure used for selecting the best family of stochastic models that will be associated with the database. Looking for relevant parameters is addressed from a non-parametric statistical point of vue.

The organisation of the paper is as follows. First, Section 2 discusses the selection of the best wavelets with respect to desirable statistical properties such as stationarization, decorrelation and higher order dependency reduction. These properties are required in that they make it possible to achieve texture representations that can be more closely approximated by stochastic modeling, under the assumption that the subband wavelet coefficients are approximately *independent and identically distributed, iid*.

Then, Section 3 addresses the selection of a measure that can provide accurate information on how suitable the stochastic modeling is, with respect to the whole database. This measure involves analysing the statistics of the sequence of goodness-of-fit measures from the database under consideration. In contrast with the standard measures based on computing the acceptance rate of the corresponding stochastic models, the approach proposed in this paper is shown to be coherent with CBIR performance measurements.

Section 4 provides CBIR results to highlight the importance of selecting relevant parameters and Section 5 concludes the work.

2. WAVELET OF STOCHASTIC PROCESSES: WHICH BASIS, WHICH WAVELET ORDER?

Textures are inherently non-smooth and stochastic modeling has proven its suitability in their representations. Best representations for a stochastic process relate to transforms that have stationarization and statistical dependency reduction properties between the several random variables describing the temporal or spatial evolution of the stochastic process. Among the transforms that approximately achieve this goal, wavelet decompositions are of interest in this work because wavelets operate unconditionally with respect to the input process and tend to achieve desirable stationarization and decorrelating properties for a large class of stochastic processes. This class contains stationary random processes (see [5], [6], [7], among others). It also contains many non-stationary random processes (see [8], [9], [10], [11], [12], among others).

For non-stationary random processes, stationarization mainly occurs in detail wavelet coefficients (see for instance [8]). In addition, when the coefficients are or become stationary, then dependency reduction is shown to be higher with larger order wavelets [6], [8], [9], the *wavelet order* refering hereafter to the number of vanishing moments of the wavelet function (see [8] for a more formal definition). This follows from that the asymptotic decorrelation and/or higher order cumulant decay are proven in [6], [8], [9], provided that this wavelet order tends to infinity.

From these results, one can assume that, in a more general context involving many stochastic processes, then wavelets with higher orders will lead to a more stationarization effect and better dependency reductions. Proving formally this assumption is not straightforward and is not addressed in this paper since all the theoretical results avalaible from the literature on *wavelets and stochastic processes* confirm its truthfulness (see the above references). In this respect, a wavelet with order $r \gg r_0$ will hereafter sayed to yield a better *iid*-like representation than a wavelet with order r_0.

In practice, order $r = \infty$ is unworkable numerically since it yields infinite sequences of filter coefficients. However, orders $r = 7, 8$ are shown to be reasonable for approximately attaining the desirable statistical properties mentioned above for many stochastic processes, [6], [8]. We will thus consider a symlet wavelet of order 8 as a relevant wavelet and comparison will be held with respect to the representation and performance achieved by the symlet wavelet of order 1, that is the Haar wavelet. In particular, we will see, in Section 3.2, that the "Symlet 8" wavelet achieve a better *iid*-like representation than the Haar wavelet, from stochasticity measurements under *iid* assumption on the wavelet coefficients of VisTeX[1] textures.

At this stage, it is worth emphasizing that the theoretical results cited above are proven in the framework of orthonormal wavelet and wavelet packet transforms. However, when analysis (and not compression) is concerned, then redundancy is often desirable since it provides additional information for the input random process de-

[1]MIT Vision Texture database, available at `http://vismod.www.media.mit.edu`.

scription. It is then convenient to find redundant transforms that behave approximately like the orthonormal wavelet and wavelet packet transforms. The discrete *Stationary Wavelet Transform*, SWT [13] and its wavelet packet based extension, SWPT, are such good candidates: the SWT (resp. SWPT) can be seen as the union of several orthogonal discrete wavelet (resp. wavelet packet) transforms, any of these transforms having the desired statistical properties. For the sake of tractability of the representation, we will consider the SWT in the following. The SWT is a particular basis among the SWPT bases and the above issue is justified by that the SWT has the advantage of presenting lower computational load than a full SWPT.

3. HOW TO PERFORM MODEL VALIDATION WHEN NO ALTERNATIVE HYPOTHESIS EXISTS?

This section addresses the problem of selecting the best family of distribution functions for modeling the subband wavelet coefficients. The selection assumes avalaibility of a suitable Goodness Of Fit (GOF) measure with respect to the distribution model and the (residual) stationarity and correlation structures that may remain in wavelet subbands. We first address the selection of the GOF measure in Section 3.1. Cumulative and uniform criteria based on GOF measurements on whole databases is then proposed in order to select the best distribution family for stochastic modeling, in Section 3.2.

3.1. Selection of a goodness-of-fit measure

A relevant GOF measure should be sensitive to

- a weak correlation characterizing a set of data issued from the observation of a random process (Example: detail wavelet coefficients in presence of a smooth region),
- a few number of large coefficients that decay in a certain order, among a set of *iid* data (Example: exponential decay of the wavelet coefficients in the neighborhood of edges).

In order to find the relevant GOF measure, we analyze in the following, the sensitivity of different GOF tests for detecting such correlation structures on synthetic data. The tests selected after a first comparison from the literature are the Kolmogorov-Smirnov — KS — test [14] (uniform norm for comparing the empirical *cdf* and the model), the chi-square — χ_2 — test [15] (ℓ_2 norm for comparing the model to a corresponding *pdf* obtained from binned data) and the Anderson-Darling — AD — test [16] (ℓ_1 norm applied for comparing the empirical *cdf* to the model).

The sensitivity of these GOF tests with respect to the above correlation structures is tested in the following experimental setup:

(C1) Correlation is inserted by setting $y_k = \nu y_{k-1} + x_k$, with $0 < \nu < 1$, where the sequence $x = \{x_k\}_{k=1,2,\ldots,N}$ is *iid* Gaussian with zero-mean and variance σ_X^2.

(C2) Data are assumed to be of the form $y = \{\theta, x\}$, where $\theta = \{\theta_i\}_{i=1,2,\ldots,M_1}$ and $x = \{x_j\}_{j=1,2,\ldots,M_2}$, with $M_1 + M_2 = N$. Subset θ of y is chosen so as to follow a geometric progression, $\theta_k = r\theta_{k-1}$, $r = 1/\sqrt{2}$, $\theta_1 = 2 \times \max|x|$ and subset x is a realization of some *iid* random variables.

Under **(C1)**, the random variable Y_k has zero-mean and variance $\sigma_{Y_k}^2 \mathbb{E}[Y_k^2] = \nu^2 \mathbb{E}[Y_{k-1}^2] + \sigma_X^2$, with $\sigma_{Y_0}^2 = \sigma_X^2$. When $\nu \cong 0$, it is reasonable to assume as null hypothesis that: $y = \{y_k\}_{k=1,2,\ldots,N}$ is *iid* with the same *cdf* as X. The experiment concerning **(C1)** consists in testing this null hypothesis when ν increases from 0 to $1/2$. A suitable stochasticity measurement should significantly reject this hypothesis when ν is larger than 0. The *cdf* of X will be Gaussian with standard deviation $\sigma_X^2 = 1$ in the experiments under **(C1)**.

Under **(C2)**, the random sequence $Y_k, k = 1, 2, \ldots, N$ is composed of a large number of *iid* random variables $Y_k = X_k$ for $k = 1, 2, \ldots, M_2$ with *cdf* F and some $M_1 = N - M_2$ *non-stochastic subset*. The experiment concerning **(C2)** consists in testing the null hypothesis that is: $Y_k, k = 1, 2, \ldots, N$ is *iid* with *cdf* F when the number M_1 of non-stochastic data increases. F will be the either Gaussian or Weibull *cdfs* concerning experiments under **(C2)**.

Table 1. Tests for the assesment of the deviation from the null hypothesis: Average values in "percentage of reject" for the *iid* assumption under correlation structures **(C1)** and **(C2)**.

		Correlation structure (C1)						
ν:		0.05	0.1	0.15	0.20	0.25	0.30	0.35
Gaussian	AD	05.14	05.23	05.13	05.21	05.20	05.31	05.56
	KS	03.28	07.71	40.00	98.72	100	100	100
	χ_2	14.55	44.21	97.85	100	100	100	100

		Correlation structure (C2)						
$100 \times M_1/N$:		0.1	0.4	0.7	1.0	1.3	1.6	1.9
Gaussian	AD	0	0	0	0	0	0	0
	KS	02.58	04.45	09.94	23.80	58.32	100	100
	χ_2	0	0	0	0	0	0	0
Weibull	AD	100	100	100	100	100	100	100
	KS	02.73	04.46	09.81	23.90	58.78	100	100
	χ_2	0	0	0	0	0	0	0

Experimental results are provided in Table 3.1. We have that:

χ_2 : The χ_2 test is performant for detecting **(C1)** and irrelevant for assesing **(C2)**. The binning and cumulative norm has the effect of masking deviation from the *cdf* specified, when this deviation concerns a small portion of the data. In addition, the χ_2 test requires a sufficient sample size for an accurate binning and its approximation lacks to be precise for small sample sizes. This is why this test is not sufficiently relevant in our context.

AD : The AD test is very relevant when Weibull distribution is concerned under **(C2)**, but is highly irrelevant for Gaussian distribution under **(C1)** and **(C2)**. This is possibly due to that the AD behaviour closely relates on the model distribution in the sense that its critical values are distribution-dependent.

KS : The KS test proves its relevancy through results of Table 3.1. Indeed, the KS test progressively rejects the null hypothesis when the correlation **(C1)** (resp. **(C2)**) induces deviation from the null hypothesis (large value of ν, resp. large portion of non-stochastic data).

3.2. Selection of the best family of stochastic models

From the results of Section 3.1, we have that the KS test is the most suitable GOF test for emphazing the relevance of a model with respect to most commonly encountered correlation structures when the detail wavelet subbands are concerned. In this sense, we consider the Kolmogorov measure associated with this test in the rest of the paper. Let $x = \{x(\ell)\}_{\ell=1,2,\ldots,N}$. Then, the Kolmogorov measure describes how well x can be considered as a realization of *iid* random variables with *cdf* F by measuring the deviation of F from the empirical distribution $F_{x,N}$ of x:

$$\lambda_N(x, F) = \sqrt{N} \sup_t |F_{x,N}(t) - F(t)|. \quad (1)$$

Assume that a given database have to be associated with a family of stochastic models, among a given set of stochastic families. The main question is: how to assess the relevance of a particular family from Kolmogorov measurements computed over the database?

This question is motivated by that the standard criterion consisting in computing the acceptance rate of the null hypothesis (number of suitable models) turns out to be irrelevant for CBIR purpose in real world textures. First, we have that admissible critical values for the KS test yield a very high percentage of reject for the null hypothesis, when modeling the wavelet subbands of standard texture databases. This holds true for all the standard distribution families. In addition, the acceptance rate tends to incoherent with CBIR performance statistics. Indeed, on the basis of the χ_2 test and this criterion, the "Generalised Gaussian" family was shown to be more relevant than "Weibull" family for modeling textures of the VisTeX database, whereas the "Weibull" family has shown better CBIR performance than the "Generalised Gaussian" family on the same database: [2]. This inconsistency is probably due to that for rejected samples, the above criterion does not take into account, information on the closeness of the model to the empirical distribution ; whereas none of the texture samples is rejected when computing the CBIR similarity measurements: all textures in the database are finally associated with stochastic models so that no alternative hypothesis is avalaible in such a CBIR.

Instead of computing the acceptance rate of the models issued from a given stochastic family per texture, the model validation scheme proposed below involves selecting the best model family by analyzing the statistical behaviour of the sequence composed of all GOF measurements characterizing the whole database. More precisely, let $\mathcal{D}$ be a database composed of M elements: $\mathcal{D} = \{\mathbf{z}_k : k = 1, 2, \ldots, M\}$, where $\mathbf{z}_k = \{\mathbf{z}_k(\ell)\}_{\ell=1,2,\ldots,N}$ for every k. Denote by $f_{\mu,\theta}$, the distribution indexed by the family μ and where index $\theta = \theta(\mu)$ refers to the parameters of the distribution. Example: if $\mu =$ "Generalised Gaussian" , then $\theta(\mu) =$ (location, scale, shape).

In what follows, we assume that for a given sample set $\mathbf{z}_k$, the parameters $\theta(\mu)$ are computed from the maximum likelihood over the set of all possible parameters for the distributions indexed by μ. For the sake of simplifying notation, we let $f_{\mu,\theta} \equiv f_\mu$. Let F_μ be the *cdf* associated with f_μ. The sequence of GOF measurements associated with distribution family μ over $\mathcal{D}$ is

$$s_{\mu,N} = \{\lambda_N(\mathbf{z}_k, F_\mu(\mathbf{z}_k)) : k = 1, 2, \ldots, M\}.$$

Among the set of statistics characterizing the sequence $s_{\mu,N}$, we consider its mean value as the main parameter and the maximum value as the second informative parameter for selecting the best family of model. The mean value is a cumulative measure taking into account the above deviations on every element of $\mathcal{D}$. The maximum value highlights the deviation resulting from the worst stochastic modeling on $\mathcal{D}$. Since the parameter λ_N measures the deviation of the emprical *cdf* from the corresponding model, then, for a given sample $\mathbf{z}_k$, the stochastic modeling associated with the smallest λ_N is considered to be the more relevant. Similarly, when both mean and maximum parameters are the smallest parameters for a specific family μ_0 of stochastic models, then this family is decidedly the most relevant between the tested families.

In the following experimental results, μ is either a "Generalised Gaussian" (GG) or a "Weibull" (WBL) distribution. We consider 40 textures from the VisTeX database. We proceed by splitting each image into 16 non-overlapping subimages (128 × 128 pixels per subimage), forming a set of 640 texture samples. Every subimage I is then decomposed by using the SWT where the decomposition level is fixed to $J_0 = 4$. Let $(c_{j,n}[I])_{j,n}$, $j \in \{1, 2 \ldots, J_0\}$, $n \in \{1, 2, 3\}$, be the sequence of SWT subbands of I. Then the test database $\mathcal{D}$ is the set of $M = 640 \times 3 \times J_0$ subband SWT detail coefficients.

Table 2 provides the above reference statistics for sequences $s_\mu = s_{\mu,N}/\sqrt{N}$ for $\mu =$ GG, WBL. The results given in Table 2 highlights that 1) the WBL family performs a more relevant fitting than GG family and 2) "Symlet 8" representation is better than "Haar" representation in terms of the *iid* property. The above items 1) and 2) hold true both in terms of cumulative (mean value) and uniform (max value) deviations of the models versus the empirical distributions on $\mathcal{D}$. Furthermore, we has that item 2) corroborates the analysis performed in Section 2, meaning that stationarization and *iid* properties are better achieved by using wavelets with higher order.

Table 2. Statistics for $s_\mu = s_{\mu,N}/\sqrt{N}$ from $\mathcal{D}$, where $\mu \in \{\text{GG}, \text{WBL}\}$. Smallest parameters refers to a best fitting of the database by the family μ.

Model	$\mu = \text{GG}$	$\mu = \text{WBL}$
Statistics	Mean(s_μ) - Max(s_μ)	Mean(s_μ) - Max(s_μ)
"Haar"	0.1138 - 0.2141	0.0614 - 0.1261
"Symlet 8"	0.1110 - 0.2125	0.0205 - 0.0447

Table 3. GG based texture-specific retrieval from the VisTeX database.

Texture	"Haar"	"Symlet 8"	Texture	"Haar"	"Symlet 8"
Bark.00	63.28	67.58	Food.08	23.05	99.22
Bark.06	68.75	67.97	Grass.01	94.92	98.05
Bark.08	62.11	65.63	Leav.08	66.80	73.83
Bark.09	38.28	42.19	Leav.10	66.80	60.94
Bric.01	99.22	100	Leav.11	63.67	65.23
Bric.04	86.33	88.28	Leav.12	81.25	79.69
Bric.05	84.77	88.67	Leav.16	55.08	63.67
Buil.09	74.61	74.22	Meta.00	84.38	85.94
Fabr.00	99.22	87.5	Meta.02	100	100
Fabr.04	79.69	85.55	Misc.02	95.70	97.27
Fabr.07	90.23	96.88	Sand.00	96.09	94.53
Fabr.09	42.97	100	Ston.01	48.83	75.78
Fabr.11	90.63	85.55	Ston.04	92.97	93.75
Fabr.14	75.78	100	Terr.10	46.09	57.42
Fabr.15	94.53	93.36	Tile.01	55.47	60.55
Fabr.17	97.66	94.92	Tile.04	41.02	98.83
Fabr.18	98.44	98.83	Tile.07	100	97.27
Flow.05	75.78	77.73	Wate.05	100	100
Food.00	100	94.92	Wood.01	60.94	60.55
Food.05	65.63	65.23	Wood.02	100	100

4. EXPERIMENTAL RESULTS

In this section, we present experimental results proving that performing CBIR with respect to a relevant wavelet (in terms of stationarization, decorrelation and higher order dependence reduction) and best joint family of stochastic models (in terms of the statistical properties of the sequence $s_{\mu,N}$) leads to high retrieval rates in texture recognition.

The same database $\mathcal{D}$ introduced in Section 3.2 is considered for the experimental results given below. The stochastic modeling of the subband wavelet coefficients is addressed by using the GG and the WBL distribution functions.

Recall that $\mathcal{D}$ have been obtained from the detail SWT subbands

$$(c_{j,n}[I_\ell])_{j\in\{1,2\ldots,J_0\},n\in\{1,2,3\}}$$

of subimages $I_\ell, \ell = 1, 2, \ldots, 640$, from the VisTeX database. Then, in order to compare a query subimage I_q with an arbitrary subimage I_ℓ, we use the following cumulative similarity measure from the SWT detail coefficients (which assumes that independence is approximately attained in the SWT subbands):

$$\mathcal{K}_{\mathcal{W}}(I_q, I_\ell) = \sum_{\substack{j\in\{1,2\ldots,J_0\}\\ n\in\{1,2,3\}}} \mathcal{K}(c_{j,n}[I_q], c_{j,n}[I_\ell]) \qquad (2)$$

where $c_{j,n}[I_q]$, $c_{j,n}[I_\ell]$ are the subband (j, n) SWT coefficients of I_q, I_ℓ respectively and $\mathcal{K}$ is a similarity measure chosen to be a symmetric version of the Kullback-Leibler divergence. For two random variables X_1 and X_2 having *pdfs* f_{X_1} and f_{X_2}, this symmetric Kullback-Leibler divergence is defined by

$$\mathcal{K}(X_1, X_2) = \mathcal{K}(X_1||X_2) + \mathcal{K}(X_2||X_1), \qquad (3)$$

$$\text{with} \quad \mathcal{K}(X_i||X_j) = \int_{\mathbb{R}} f_{X_i}(x) \log \frac{f_{X_i}(x)}{f_{X_j}(x)} dx, \quad i, j = 1, 2.$$

For Generalized Gaussian distributions, the Kullback-Leibler divergence is given in [1]. We have that the symmetric version this divergence is:

$$\mathcal{K}(X_1, X_2) = \left(\frac{\alpha_1}{\alpha_2}\right)^{\beta_2} \frac{\Gamma\left(\frac{1+\beta_2}{\beta_1}\right)}{\Gamma(1/\beta_1)} + \left(\frac{\alpha_2}{\alpha_1}\right)^{\beta_1} \frac{\Gamma\left(\frac{1+\beta_1}{\beta_2}\right)}{\Gamma(1/\beta_2)} - \frac{\beta_1+\beta_2}{\beta_1\beta_2}.$$

For Weibull distributions, the Kullback-Leibler divergence is given in [2]. We have that the symmetric version of this divergence is

$$\mathcal{K}(X_1, X_2) = \Gamma\left(1 + \frac{k_2}{k_1}\right)\left(\frac{\lambda_1}{\lambda_2}\right)^{k_2} + \Gamma\left(1 + \frac{k_1}{k_2}\right)\left(\frac{\lambda_2}{\lambda_1}\right)^{k_1} + (k_1 - k_2)\log\frac{\lambda_1}{\lambda_2} + e\left(\frac{k_1}{k_2} + \frac{k_2}{k_1} - 2\right) - 2.$$

where e is the Euler-Mascheroni constant.

Tables 3 and 4 provide the average retrieval rates when the GG and WBL modeling are concerned, respectively. Table 5 provides a summarized point of vue of results given in 3 and 4. These results highlight that the analysis performed in terms of relevant wavelet and best probability distribution family leads to more relevant CBIR strategy: WBL based modeling of the "Symlet 8" SWT subbands is with the highest CBIR performance. Finally, for comparison purpose, we have that: in the same context (same database and same experimental setup) but when the transform used is the dual-tree complex wavelet transform, CBIR experiments yield 74.12 % (resp. 81.58 %) of cumulative retrieval rate when stochastic modeling is performed by using the GG (resp. WBL) distribution family [2].

5. CONCLUSION

The paper has addressed content-based image retrieval from texture databases, by using stochastic modeling in the wavelet domain. It has proposed an analysis of the main parameters involved in this content-based image retrieval. These parameters are the wavelet order and the goodness-of-fit measure used for selecting the best family of distribution models. Higher order wavelets are shown to be more relevant in that they allows for a better representation under the *iid* criterion. Under this *iid* criterion, then the mean and maximum values of the sequence of all goodness-of-fit measures from the Kolmogorov parameter are shown to be relevant criteria for the selection of the best family of distribution models. Further prospects could concern 1) modeling the approximation subband in order to increase performance, 2) selecting the optimal decomposition level depending on the database, 3) selecting the best decomposition tree when the SWPT decomposition is concerned and/or 4) information fusion from several decompositions.

Table 4. WBL based texture-specific retrieval from the VisTeX database.

Texture	"Haar"	"Symlet 8"	Texture	"Haar"	"Symlet 8"
Bark.00	66.80	66.80	Food.08	57.81	99.61
Bark.06	72.27	67.97	Grass.01	78.13	94.14
Bark.08	61.72	64.45	Leav.08	67.97	73.44
Bark.09	34.77	41.02	Leav.10	63.28	62.5
Bric.01	98.44	100	Leav.11	62.5	65.63
Bric.04	86.72	88.67	Leav.12	83.98	80.47
Bric.05	82.42	86.72	Leav.16	50.39	64.84
Buil.09	100	95.31	Meta.00	82.03	85.55
Fabr.00	98.83	86.33	Meta.02	97.66	100
Fabr.04	69.53	85.55	Misc.02	95.31	98.05
Fabr.07	90.23	96.48	Sand.00	94.14	94.14
Fabr.09	98.44	100	Ston.01	55.47	75.39
Fabr.11	87.89	83.59	Ston.04	92.97	93.75
Fabr.14	100	100	Terr.10	47.27	55.86
Fabr.15	83.98	93.36	Tile.01	68.75	60.16
Fabr.17	100	99.61	Tile.04	83.2	98.83
Fabr.18	98.44	98.83	Tile.07	100	97.27
Flow.05	72.66	78.52	Wate.05	100	100
Food.00	92.97	94.14	Wood.01	57.81	60.55
Food.05	66.41	64.45	Wood.02	100	100

Table 5. Comparison chart (mean values) that summerizes the results given in Table 3 (GG) and those given in Table 4 (WBL).

GG		WBL	
"Haar"	"Symlet 8"	"Haar"	"Symlet 8"
76.52	83.44	80.03	83.80

6. REFERENCES

[1] M. N. Do and M. Vetterli, "Wavelet-based texture retrieval using generalized gaussian density and kullback-leibler distance," *IEEE Transactions on Image Processing*, vol. 11, no. 2, pp. 146 – 158, Feb. 2002.

[2] R. Kwitt and A. Uhl, "Image similarity measurement by kullback-leibler divergences between complex wavelet subband statistics for texture retrieval," *IEEE International Conference on Image Processing, ICIP,* San Diego, California, USA, 12-15 October, pp. 933–936, 2008.

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

[4] N. Lasmar and Y. Berthoumieu, "Multivariate statistical modeling for texture analysis using wavelet transforms," *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP,* Dallas, Texas, USA, 14 - 19 March, 2010.

[5] A. M. Atto and D. Pastor, "Central limit theorems for wavelet packet decompositions of stationary random processes," *IEEE Transactions on Signal Processing*, vol. 58, no. 2, pp. 896 – 901, Feb. 2010.

[6] A. M. Atto, D. Pastor, and A. Isar, "On the statistical decorrelation of the wavelet packet coefficients of a band-limited wide-sense stationary random process," *Signal Processing*, vol. 87, no. 10, pp. 2320 – 2335, Oct. 2007.

[7] J. Zhang and G. Walter, "A wavelet-based KL-like expansion for wide-sense stationary random processes," *IEEE Transactions on Signal Processing*, vol. 42, no. 7, pp. 1737–1745, July 1994.

[8] A. M. Atto, D. Pastor, and G. Mercier, "Wavelet packets of fractional brownian motion: Asymptotic analysis and spectrum estimation," *IEEE Transactions on Information Theory*, vol. 56, no. 9, Sep. 2010.

[9] P. F. Craigmile and D. B. Percival, "Asymptotic decorrelation of between-scale wavelet coefficients," *IEEE Transactions on Information Theory*, vol. 51, no. 3, pp. 1039 – 1048, Mar. 2005.

[10] T. Kato and E. Masry, "On the spectral density of the wavelet transform of fractional brownian motion," *Journal of Time Series Analysis*, vol. 20, no. 50, pp. 559–563, 1999.

[11] E. Masry, "The wavelet transform of stochastic processes with stationary increments and its application to fractional brownian motion," *IEEE Transactions on Information Theory*, vol. 39, no. 1, pp. 260–264, Jan. 1993.

[12] P. Flandrin, "Wavelet analysis and synthesis of fractional brownian motion," *IEEE Transactions on Information Theory*, vol. 38, no. 2, pp. 910–917, Mar. 1992.

[13] R. R. Coifman and D. L. Donoho, *Translation invariant de-noising*. Lecture Notes in Statistics, 1995, no. 103, pp. 125–150.

[14] J. F. J. Massey, "The kolmogorov-smirnov test for goodness of fit," *Journal of the American Statistical Association*, vol. 253, no. 2, pp. 68 – 78, Mar. 1951.

[15] J. F. Kenney and E. S. Keeping, *Mathematics of Statistics, Pt. 2*, Princeton, Ed. Van Nostrand, 1951.

[16] T. W. Anderson and D. A. Darling, "Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes," *Annals of Mathematical Statistics*, vol. 23, pp. 193 – 212, 1952.