



HAL
open science

Centroid-based texture classification using the SIRV representation

Aurélien Schutz, Lionel Bombrun, Yannick Berthoumieu

► **To cite this version:**

Aurélien Schutz, Lionel Bombrun, Yannick Berthoumieu. Centroid-based texture classification using the SIRV representation. IEEE International Conference on Image Processing, Sep 2013, Melbourne, Australia. pp.3810-3814. hal-00865595

HAL Id: hal-00865595

<https://hal.science/hal-00865595>

Submitted on 24 Sep 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.

CENTROID-BASED TEXTURE CLASSIFICATION USING THE SIRV REPRESENTATION

Aurélien Schutz, Lionel Bombrun and Yannick Berthoumieu

Université de Bordeaux, ENSEIRB-Matmeca, Laboratoire IMS, Groupe Signal et Image
{*aurelien.schutz, lionel.bombrun, yannick.berthoumieu*}@ims-bordeaux.fr

ABSTRACT

This paper introduces a centroid-based (CB) supervised classification algorithm of textured images. In the context of scale/orientation decomposition, we demonstrate the possibility to develop centroid approach based on multivariate stochastic modeling. The main interest of the multivariate modeling comparatively to the univariate case is to consider spatial dependency as additional features for characterizing texture content. The aim of this paper is twofold. Firstly, we introduce the Spherically Invariant Random Vector (SIRV) representation for the modeling of wavelet coefficients. Secondly, from the specific properties of the SIRV process, i.e. the independence between the two sub-processes of the compound model, we derive centroid estimation scheme. Experiments from various conventional texture databases are conducted and demonstrate the interest of the proposed classification algorithm.

Index Terms— textured images, Jeffrey divergence, SIRV model, centroid, supervised classification.

1. INTRODUCTION

Classification of textured images is used in a large field of applications ranging from the classification of orchards from remote sensing images, to quality check of manufactured pieces by comparison of internal structures. Among classification methods, clustering approaches have known an increased interest providing effective and tractable algorithms for various domains. Classification techniques based on clustering such as supervised centroid-based (CB) and unsupervised k -means methods assume that (i) textured images are sorted in k sub-collections of samples, i.e. the clusters, (ii) each cluster can be represented by the most centrally localized object, i.e. the barycenter or centroid. Evaluating a centroid implies to define an adapted measure of similarity/dissimilarity between a set of estimated parameters characterizing each sample in the cluster. For texture clustering, the main purpose is thus to define an effective set of parameters and a dissimilarity measure which can be minimized in order to estimate the centroid coordinates in the parameter space knowing the sub-set of samples associated to the cluster. Over the last decade, numerous

works devoted to texture analysis have shown the interest to use jointly scale-space decomposition and stochastic modeling for characterizing the textural content [1, 2, 3, 4, 5, 6]. The more recent works proposed parametric probability density function (pdf), i.e. prior such as Generalized Gaussian density (GGD) or Weibull density, to fit the empirical histogram of sub-band coefficients [2, 3]. Moreover, some works have pointed out the pertinence to consider multivariate modeling such as multivariate GGD [7], Spherically Invariant Random Vector (SIRV) [8, 9], and copulas based models [10] in order to take into account the spatial dependency [7, 8, 9, 10] rather than simple univariate modeling. Incorporate the dependency in the model enables us to increase the performance of classification methods. However, regardless of the model used, the result is that homogeneous texture samples are summarized by a finite set of pdfs, one for each detail sub-band of the decomposition, characterized respectively by a limited set of parameters. In complement, these works have also proposed to exploit probabilistic dissimilarity such as Kullback-Leibler divergence (KL) [2, 3], geodesic distance [8], optimal transport [11], or Jeffrey divergence (J) [12] following the well-founded Bayesian theory. Concerning the issue of the centroid estimation in the context of texture analysis, Choy and Tong [3] have developed a solution based on the univariate modeling using the GGD model and the Kullback-Leibler divergence. However, no propositions have yet been made to extend this proposal to non-trivial multivariate model, which is the main contribution of the paper.

This paper introduces a CB classification algorithm based on the SIRV distribution for the modeling of wavelet coefficients [9]. By exploiting the independence between the multiplier τ and the Gaussian vector \vec{g} , the Jeffrey divergence of the joint vector $\vec{y} = (\tau, \vec{g})$ admits a close-form expression. Based on this observation, we propose an algorithm to estimate the centroid from a collection of SIRV parameters. The paper is structured as follows. Section 2 introduces the SIRV distribution for the modeling of wavelet coefficients. Section 3 derives the proposed algorithm to compute the centroid. Some classification results are next presented in Section 4 to evaluate the performance of the proposed CB classification algorithm on texture databases. Conclusions and future works are finally reported in Section 5.

Thanks TOTAL for funding.

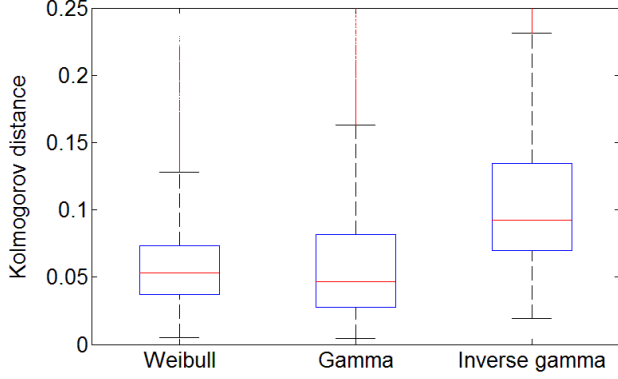


Fig. 1. Box-plots of the Kolmogorov distance on the VisTex database for various multiplier models (Weibull, Gamma and Inverse Gamma).

2. STOCHASTIC MODEL

To model wavelet coefficients, Spherically Invariant Random Vectors (SIRV) have been proposed in [9]. This class of models has been introduced to take into account the non-Gaussianity of the signal. Compared to other multivariate models (copula, MGGD, ...), this family of distributions has found a great interest in modeling the spatial dependency of wavelet coefficients. Let \vec{x} be a d -dimensional vector following a SIRV distribution, it yields that \vec{x} admits the stochastic representation:

$$\vec{x} = \sqrt{\tau} \vec{g}, \quad (1)$$

where τ is a scalar random variable called multiplier ($\tau \in \mathbb{R}^+$) and \vec{g} a real Gaussian vector with zero mean and covariance matrix $\Sigma = \mathbb{E}\{\vec{g}\vec{g}^T\}$. Processes τ and \vec{g} are assumed independent. In the literature [8], various models issued from the Pearson system have been introduced to represent the multiplier τ such as Gamma, Inverse Gamma and Fisher distributions. In this paper, the univariate Weibull distribution is considered to model the multiplier. Its pdf is:

$$p_w(\tau; a, b) = \frac{a}{b} \left(\frac{\tau}{b}\right)^{a-1} \exp\left\{-\left(\frac{\tau}{b}\right)^a\right\} \quad (2)$$

where a and b are respectively the shape and scale parameters. Since the SIRV model is uniquely defined with respect to the covariance matrix parameter up to a multiplicative constant, the multiplier τ is normalized to have an unitary mean, *i.e.* $\mathbb{E}\{\tau\} = 1$. It yields that $b = (\Gamma(1/a + 1))^{-1}$. SIRV parameters are hence extracted according to the SIRV estimation scheme developed in [13].

To evaluate the benefit of the SIRV model for characterizing texture content, the empirical histogram of the multiplier τ has been computed for each sub-band of textured images from the VisTex database. This histogram is then modeled

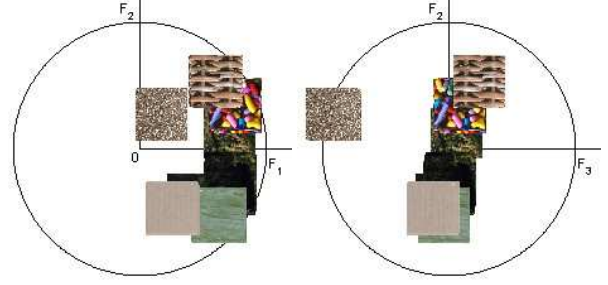


Fig. 2. Principal component analysis on dissimilarity matrix between - textured images: left the 2 principal directions, right the 2nd and 3rd direction.

by three univariate distributions [8] *i.e.* Weibull, Gamma, and Inverse Gamma. The Kolmogorov distance, denoted d_K , is next used to evaluate the goodness-of-fit of the multiplier. The Kolmogorov distance is defined as:

$$d_K = \sup_{\tau} |F(\tau) - F_N(\tau)|. \quad (3)$$

where $F_N(\cdot)$ is the empirical cumulative distribution function (cdf) and $F(\cdot)$ is the theoretical (hypothesized) cdf. Fig. 1 draws three box-plots of the Kolmogorov distance for three multiplier models, namely Weibull, Gamma and Inverse gamma pdfs. In this experiment, the steerable pyramid decomposition has been used and a 3×3 neighborhood has been considered to model the spatial dependency of the wavelet coefficients. As observed, the Weibull model exhibits the best performance. In the following, the Weibull distribution will be considered as prior model for the multiplier.

Unfortunately, no closed-form expression exists for the pdf of the multivariate vector \vec{x} for the case of Weibull distributed multiplier. However, characterizing texture content can be provided considering vector $\vec{y} = (\tau, \vec{g})$ resulting from the SIRV representation. In this case and using the independence between τ and \vec{g} , the joint pdf of vector \vec{y} is

$$p_Y(\vec{y}; \lambda) = p_w(\tau; a) p_G(\vec{g}; \Sigma) \quad (4)$$

where $\lambda = \{a, \Sigma\}$ is the parametric vector associated to the SIRV model. The components λ form a parametric Riemannian manifold. In the sequel of the paper, we call \mathcal{M} the corresponding manifold.

In the general context of classification, a dissimilarity measure is required to compute the similarity of two images based on their respective set of features. In this paper, the Jeffrey divergence (J) is considered to compute the probabilistic distance between two parametric vectors. By working on the vector $\vec{y} = (\tau, \vec{g})$, the Jeffrey divergence is obtained using the chain rule, since the multiplier parameter τ and the Gaussian

process \vec{g} are independent in the SIRV model.

$$J(p(\vec{y}; \lambda), p(\vec{y}; \lambda')) = J(p_G(\vec{g}; \Sigma), p_G(\vec{g}; \Sigma')) + J(p_w(\tau; a), p_w(\tau; a')). \quad (5)$$

The first term in (5) corresponds to J for the multivariate Gaussian process, while the second term corresponds to J between the stochastic model of the multiplier. Both terms admit a closed-form expression recalled below:

$$J(p_G(\vec{g}; \Sigma), p_G(\vec{g}; \Sigma')) = \frac{1}{2} \text{Tr}((\Sigma')^{-1} \Sigma) + \frac{1}{2} \text{Tr}(\Sigma^{-1} \Sigma') - d \quad (6)$$

and

$$J(p_w(\tau; a), p_w(\tau; a')) = \left(2 - \frac{a'}{a} - \frac{a}{a'}\right) \Psi(1) - 2 + A(a, a') + A(a', a) + (a - a') \log \left\{ \frac{\Gamma(1/a' + 1)}{\Gamma(1/a + 1)} \right\} \quad (7)$$

where Γ denotes the gamma function defined by $\Gamma(z) = \int_{\mathbb{R}^+} t^{z-1} e^{-t} dt$, Ψ is the digamma function $\Psi(z) = \Gamma'(z)/\Gamma(z)$ and

$$A(a, a') = \left(\frac{\Gamma(1/a' + 1)}{\Gamma(1/a + 1)} \right)^{a'} \Gamma\left(\frac{a'}{a} + 1\right). \quad (8)$$

To evaluate the potential of the proposed SIRV model for representing texture images, a spectral clustering technique is used. After computing the similarity matrix based on the Jeffrey divergence, a dimensionality reduction algorithm is performed to observe the repartition of the initial images. Fig. 2 displays, for six classes of textured images from the VisTex database, principal component analysis of their respective dissimilarity. As observed, for each class, the cluster can be well approximated by one element: the centroid.

3. CENTROID COMPUTATION

Let $\Lambda = (\lambda_n)_{n=1}^{N_{Tr}}$ be a collection of N_{Tr} parametric vectors from a specific class of textured images. In [3], Choy and Tong have introduced an iterative algorithm to estimate the barycentric sample $\bar{\lambda}$ (also called centroid) from this collection of samples. Let $l(\lambda)$ be the cost function defined by:

$$l(\lambda) = \frac{1}{N_{Tr}} \sum_{n=1}^{N_{Tr}} J(p(\vec{y}; \lambda), p(\vec{y}; \lambda_n)), \quad (9)$$

the centroid $\bar{\lambda}$ is obtained as the solution of the following optimization problem:

$$\bar{\lambda} = \arg \min_{\lambda \in \mathcal{M}} l(\lambda). \quad (10)$$

By combining (5) and (10), the optimization problem can be split into two simpler problems: one for the multivariate Gaussian part and one for the univariate Weibull part. It yields that the centroid $\bar{\lambda} = \{\bar{a}, \bar{\Sigma}\}$ is composed by two centroids:

$$\bar{a} = \arg \min_{a \in \mathcal{M}_w} \frac{1}{N_{Tr}} \sum_{n=1}^{N_{Tr}} J(p_w(\tau; a), p_w(\tau; a_n)); \quad (11)$$

$$\bar{\Sigma} = \arg \min_{\Sigma \in \mathcal{M}_G} \frac{1}{N_{Tr}} \sum_{n=1}^{N_{Tr}} J(p_G(\vec{g}; \Sigma), p_G(\vec{g}; \Sigma_n)). \quad (12)$$

In the following subsection, we sequentially present how those two centroids are computed.

3.1. Centroid for the multivariate Gaussian part

Let $(\Sigma_n)_{n=1}^{N_{Tr}}$ be a collection of covariance matrices. Banerjee *et al.* have considered the Kullback-Leibler divergence as similarity measure and have derived a closed-form expression for the right-sided ($\bar{\Sigma}_R$) and left-sided ($\bar{\Sigma}_L$) centroids [14], it yields:

$$\bar{\Sigma}_R = \left(\frac{1}{N_{Tr}} \sum_{n=1}^{N_{Tr}} \Sigma_n^{-1} \right)^{-1} \quad \text{and} \quad \bar{\Sigma}_L = \frac{1}{N_{Tr}} \sum_{n=1}^{N_{Tr}} \Sigma_n. \quad (13)$$

Nielsen and Nock have further extended this work to compute the Jeffrey centroid $\bar{\Sigma}$. They proved in [15] that $\bar{\Sigma}$ is inside the geodesic existing between $\bar{\Sigma}_R$ and $\bar{\Sigma}_L$:

$$\exists \xi \in [0, 1] \text{ such that } \bar{\Sigma} = (1 - \xi) \bar{\Sigma}_R + \xi \bar{\Sigma}_L, \quad (14)$$

and that $\bar{\Sigma}$ is equidistant to each centroid. It yields:

$$\text{KL}(p_g(\vec{g}; \bar{\Sigma}_R) \| p_g(\vec{g}; \bar{\Sigma})) = \text{KL}(p_g(\vec{g}; \bar{\Sigma}) \| p_g(\vec{g}; \bar{\Sigma}_L)). \quad (15)$$

This first optimization problem is solved by dichotomy on the geodesic linking left-sided centroid $\bar{\Sigma}_L$ to right-sided centroid $\bar{\Sigma}_R$.

3.2. Centroid of Weibull

Let $(a_n)_{n=1}^{N_{Tr}}$ be a collection of Weibull shape parameters. The centroid \bar{a} , solution of (11), is obtained by finding the root of:

$$\frac{dl_w}{da}(a) = A_1 \Psi(1) - A_2 \Psi\left(\frac{1}{a} + 1\right) + A_3 + A_4, \quad (16)$$

where

$$A_1 = \frac{1}{a^2} \sum_{n=1}^{N_{Tr}} \frac{1}{N_{Tr}} a_n - \frac{1}{N_{Tr}} \sum_{n=1}^{N_{Tr}} \frac{1}{a_n}, \quad (17)$$

$$A_2 = \frac{1}{a^2} \sum_{n=1}^{N_{Tr}} \frac{1}{N_{Tr}} a_n (1 - A(a, a_n)) - \frac{1}{a} + \frac{1}{a} \sum_{n=1}^{N_{Tr}} \frac{1}{N_{Tr}} A(a_n, a), \quad (18)$$

Algorithm 1 Pseudo-code to compute the centroid $\bar{\lambda}$

Require: Collection of parametric vector $\Lambda = (a_n, \Sigma_n)_{n=1}^{N_{Tr}}$

Ensure: Centroid $\bar{\lambda} = (\bar{a}, \bar{\Sigma})$

- 1: Compute $\bar{\Sigma}_L$ and $\bar{\Sigma}_R$ with (13)
 - 2: Initialization $\xi_m = 0$ and $\xi_M = 1$
 - 3: **for** $i = 1, \dots, N_{iter}$ **do**
 - 4: $\xi \leftarrow \frac{1}{2}(\xi_m + \xi_M)$
 - 5: $\bar{\Sigma} \leftarrow (1 - \xi)\bar{\Sigma}_R + \xi\bar{\Sigma}_L$
 - 6: $tmp \leftarrow \text{KL}(p_G(\vec{g}; \bar{\Sigma}_R) \| p_G(\vec{g}; \bar{\Sigma}))$
 - 7: $si \leftarrow tmp - \text{KL}(p_G(\vec{g}; \bar{\Sigma}) \| p_G(\vec{g}; \bar{\Sigma}_L))$
 - 8: **if** $si < 0$ **then** $\xi_m \leftarrow \xi$ **else** $\xi_M \leftarrow \xi$ **endif**
 - 9: **end for**
 - 10: Let f be the derivative (16)
 - 11: $\bar{a} \leftarrow \text{root of } f$
-

$$A_3 = -\frac{1}{a^2} \sum_{n=1}^{N_{Tr}} \frac{1}{N_{Tr}} a_n A(a, a_n) \Psi\left(\frac{a_n}{a} + 1\right), \quad (19)$$

and

$$A_4 = \sum_{n=1}^{N_{Tr}} \frac{1}{N_{Tr}} (1 - A(a_n, a)) \log \left\{ \frac{\Gamma(1/a_n + 1)}{\Gamma(1/a + 1)} \right\} + \sum_{n=1}^{N_{Tr}} \frac{1}{N_{Tr}} \frac{A(a_n, a)}{a_n} \Psi\left(\frac{a}{a_n} + 1\right). \quad (20)$$

Practically, a simple dichotomy on \mathbb{R}^+ is considered to find the root of (16). To summarize, Algorithm. 1 is used to compute the centroid $\{\bar{a}, \bar{\Sigma}\}$ from a collection of SIRV parametric vectors.

4. NUMERICAL APPLICATION

Let N_s be the number of sub-bands of a multi-scale decomposition. Let us consider the parametric vector λ_s of the pdf associated to each sub-band. The collection $T = (\lambda_s)_{s=1}^{N_s}$ of those parametric vectors will represent the textured image. Let $(T_{c,n})_{n=1}^{N_{Tr}}$ be N_{Tr} training samples from the same class c . Then, the centroid of this collection of sample is defined as $\bar{T} = (\bar{\lambda}_s)_{s=1}^{N_s}$, where $\bar{\lambda}_s = (\bar{a}_s, \bar{\Sigma}_s)$ is the centroid computed as the solution of (11) and (12) at sub-band s . For each texture class $c = 1, \dots, N_{cl}$, one centroid \bar{T}_c is computed according to the proposed algorithm. Let T_t be a test sample. This sample is labeled to the class \hat{c} , corresponding to the class of the closest centroid, *i.e.*

$$\hat{c} = \arg \min_c J(T_t \| \bar{T}_c), \quad (21)$$

where the dissimilarity measure J between two instances of T is computed as the sum of the dissimilarity measures J between all sub-band distributions at each scale and orientation.

To evaluate the performance of the proposed supervised classification algorithm, the database is split into a training database and a disjoint testing database. Practically, N_{Tr}

Table 1. Kappa index for various stochastic models

	GGD	MG	SIRV		
	[12]		\vec{g}	τ	(τ, \vec{g})
Neigh.	1x1	3x3			
VisTex	73%	88%	88%	59%	90%
OuTex	57%	62%	62%	47%	65%
VisTex C.	41%	57%	58%	26%	60%

training samples are randomly selected for each texture class, the remaining sample are used as testing samples. Three databases are considered here: VisTex [16], 40 classes, 16 sample per class; OuTex (TC 13) [17], 68 classes, 20 sample per class and VisTex Complete [16], 167 classes, 16 sample per class. In the following, 100 Monte Carlo runs have been used to evaluate the performance of the different classifiers (kappa index). The kappa index refers to the proportion of consistent classifications observed beyond that expected by chance alone [18, 19].

Table 1 draws the classification results of the centroid based classifier obtained for various stochastic models. The first column shows the classification results for an univariate model (with GGD assumption) [12]. The next columns are the classification results obtained with a model representing the spatial distribution of wavelet coefficients. Various models are considered: the multivariate Gaussian (MG), and a SIRV model when considering the joint distribution (τ, \vec{g}) , the Gaussian part (\vec{g}) and the multiplier part (τ). As observed, for all texture databases, the classification results obtained with the proposed SIRV model allows a gain of 2 points compared to the multivariate Gaussian model.

5. CONCLUSION

This paper has addressed the problem of centroid-based (CB) classification in the multivariate case. After introducing the Spherically Invariant Random Vector (SIRV) distribution for the modeling of wavelet coefficients, we have proposed an algorithm to compute the centroid from a collection of SIRV parameters. Supervised classification results on various texture databases have shown a gain compared to other classical models.

Further works will concerns the extension of the proposed work to a multi-barycentric classification algorithm in order to handle the intra-class diversity of natural texture images.

6. REFERENCES

- [1] S.G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 11, no. 7, pp. 674–693, 1989.
- [2] M.N. Do and M. Vetterli, "Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance," *Image Processing, IEEE Transactions on*, vol. 11, no. 2, pp. 146–158, Feb. 2002.
- [3] S.-K. Choy and C.-S. Tong, "Supervised texture classification using characteristic generalized Gaussian density," *Journal of Mathematical Imaging and Vision*, vol. 29, pp. 35–47, Aug. 2007.
- [4] K. I. Kim, K. Jung, S. H. Park, and H. J. Kim, "Support vector machines for texture classification," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 11, pp. 1542–1550, Nov. 2002.
- [5] M. Varma and A. Zisserman, "A statistical approach to texture classification from single images," *International Journal of Computer Vision*, vol. 62, pp. 61–81, 2005.
- [6] J. Rabin, G. Peyré, J. Delon, and M. Bernot, "Wasserstein barycenter and its application to texture mixing," in *Proceedings of the Third International Conference on Scale Space and Variational Methods in Computer Vision*, 2011, vol. 8.
- [7] G. Verdoolaege, Y. Rosseel, M. Lambrechts, and P. Scheunders, "Wavelet-based colour texture retrieval using the Kullback-Leibler divergence between bivariate generalized Gaussian models," in *Image Processing, IEEE International Conference on*, Nov. 2009, pp. 265–268.
- [8] L. Bombrun, N.-E. Lasmar, Y. Berthoumieu, and G. Verdoolaege, "Multivariate texture retrieval using the SIRV representation and the geodesic distance," in *Acoustics, Speech and Signal Processing, IEEE International Conference on*, May 2011, pp. 865–868.
- [9] N.-E. Lasmar and Y. Berthoumieu, "Multivariate statistical modeling for texture analysis using wavelet transforms," in *Acoustics Speech and Signal Processing, IEEE International Conference on*, Mar. 2010, pp. 790–793.
- [10] N.-E. Lasmar and Y. Berthoumieu, "Gaussian copula multivariate modeling for texture image retrieval using Kullback-Leibler divergence," 2010.
- [11] S. Ferradans, G.-S. Xia, G. Peyré, and J.-F. Aujol, "Optimal transport mixing of Gaussian texture models," in *Image Processing, IEEE International Conference on*, 2012.
- [12] A. Schutz, Y. Berthoumieu, F. Turcu, C. Naornita, and A. Isar, "Barycentric distribution estimation for texture clustering based on information-geometry tools," in *Electronics and Telecommunications (ISETC), 10th International Symposium on*, Nov. 2012.
- [13] F. Gini and M. Greco, "Covariance matrix estimation for CFAR detection in correlated heavy tailed clutter," *Signal Processing*, vol. 82, no. 12, pp. 1847–1859, 2002.
- [14] A. Banerjee, S. Merugu, I.S. Dhillon, and J. Ghosh, "Clustering with Bregman divergences," *The Journal of Machine Learning Research*, vol. 6, pp. 1705–1749, 2005.
- [15] F. Nielsen and R. Nock, "Sided and symmetrized bregman centroids," *IEEE Transactions on Information Theory*, vol. 55, no. 6, pp. 2048–2059, June 2009.
- [16] R. Picard, C. Graczyk, S. Mann, J. Wachman, L. Picard, L. Campbell, and N. Negroponte, "Vision texture database," the Media Laboratory, MIT, Cambridge, Massachusetts, 1995.
- [17] T. Ojala, T. Mäenpää, M. Pietikäinen, J. Viertola, J. Kyllönen, and S. Huovinen, "Outex - new framework for empirical evaluation of texture analysis algorithms," in *Pattern Recognition, IEEE International Conference on*, 2002, vol. 1, pp. 701–706.
- [18] J. Cohen, "A coefficient of agreement for nominal scales," *Educational and psychological measurement*, vol. 20, no. 1, pp. 37–46, 1960.
- [19] D. Gomez and J. Montero, "Determining the accuracy in image supervised classification problems," *EUSFLAT*, vol. 1, no. 1, pp. 342–349, Jul. 2011.