

A Graph-based approach to derive the geodesic distance on Statistical manifolds: Application to Multimedia Information Retrieval

Zakariae Abbad

LISAC

FSDM, Sidi Mohamed Ben Abdellah University

Fez, Morocco

zakariaeabbad@gmail.com

Ahmed Drissi El Maliani

LRIT Rabat IT Center

Faculty of sciences, Mohammed V University

Rabat, Morocco

maliani.ahmed@gmail.com

Said Ouatik El Alaoui

Laboratory of Engeneering Sciences

National School of Applied Sciences, Ibn Tofail University

Kenitra, Morocco

ouatikelalaoui.said@uit.ac.ma

Mohammed El Hassouni

DESTEC, FLSHR, LRIT IT Center

University of Mohammed V

Rabat, Morocco

mohamed.elhassouni@gmail.com

Abstract—In this paper, we leverage the properties of non-Euclidean Geometry to define the Geodesic distance (GD) on the space of statistical manifolds. The Geodesic distance is a real and intuitive similarity measure that is a good alternative to the purely statistical and extensively used Kullback-Leibler divergence (KLD). Despite the effectiveness of the GD, a closed-form does not exist for many manifolds, since the geodesic equations are hard to solve. This explains that the major studies have been content to use numerical approximations. Nevertheless, most of those do not take account of the manifold properties, which leads to a loss of information and thus to low performances. We propose an approximation of the Geodesic distance through a graph-based method. This latter permits to well represent the structure of the statistical manifold, and respects its geometrical properties. Our main aim is to compare the graph-based approximation to the state of the art approximations. Thus, the proposed approach is evaluated for two statistical manifolds, namely the Weibull manifold and the Gamma manifold, considering the Content-Based Texture Retrieval application on different databases.

Index Terms—Geodesic distance, Statistical manifold, Texture retrieval, Graph theory, Wavelet decomposition

I. INTRODUCTION

Statistical manifolds are geometric representations of smooth families of probability density functions. A class of probability density functions (Pdf) is characterized by a vector θ that forms a coordinate system of continuous parameters and has geometrical properties as a result of local compositions of the distributions. So naturally, a manifold (S) that governs these pdfs is a collection of points (representations of the pdfs) with a coordinate system. This latter makes the one-to-one mapping from S to \mathbb{R}^n [1] [2].

Information geometry is a branch of differential geometry that gives a geometrical perspective to the probabilistic theory and statistics. The research on the geometrical properties of

probability distributions started with the work of Fisher [3] and then was developed in the literature ([1], [4] and [2]). After that, information geometry has been utilized to study the geometry of the statistical models as in [5] where the geometry of the univariate normal model was used in image classification. In [6], the Generalized Gaussian distribution geometry was studied. Also, The study of G_0^1 distributions geometry was applied to region discrimination in [7]. The information geometry was also used to learn finite mixture models, such as the mixture of Gaussians [8] and the mixture of multivariate Gaussians [9].

Among all (curvature, affine connection, ...), the similarity measurement (SM) between probability distributions remains the most prominent example of the advantage of the geometrical understanding. It permits to derive a real distance, namely the geodesic distance (GD) on top of the statistical manifold. The GD is a natural distance that governs the statistical manifolds and is represented by the shortest path that exists between two points on the manifold. In this context, there have been numerous studies to investigate the GD among other metrics such as in [10], [11] where Geert et al. proposed the Geodesic Distance as a similarity measurement between the zero-mean Multivariate Generalized Gaussian distributions (MGGD) in the context of texture retrieval. In [12], the Cauchy-Schwarz divergence was used to estimate the similarity between the Mixtures of Generalized Gaussian distributions in the texture retrieval context.

It is to note that the Kullback-Leibler divergence (KLD) is the most popular in the context of texture and image retrieval. KLD was used in numerous papers like in [13] where it was combined with the Gamma distribution. In [14], the KLD was used in combination with the Weibull distribution in the same context of texture retrieval. However, the KLD is not

considered a real distance on the probability space, because it does not satisfy the triangular inequality and it is not symmetric.

In this work, we use the GD as a similarity measurement instead of the KLD, in order to integrate a geometric thinking in the context of image retrieval and more generally in the context of Multimedia Information retrieval.

We will compute the GD on the Gamma and the Weibull manifolds. Both manifolds have been extensively studied, such as in [15], [2], [16] for the Gamma manifold, and [17], [18] for the Weibull manifold.

Due to the cumbersomeness of the geodesic equations, there is no closed-form but approximations of the GD in both cases. We propose a new approximation, namely the graph-based approximation to compute the GD, in order to provide more precision to the retrieval process. Until the finish date of this work, no endeavor to do so was made. Experiments on two well-known texture datasets show that the GD through the graph-based approximation achieves an obvious higher performance compared with the classic Kullback-Leibler divergence. The structure of this paper is as follows: In the next section, we provide the geometrical properties of the Gamma & Weibull manifolds. In section 3, we provide the graph-based approach that we used to approximate the geodesic distance. In section 4, we present the experimental results of the texture retrieval using the GD and the KLD, before concluding in section 5.

II. THE GEOMETRY OF STATISTICAL MANIFOLDS

A. The Gamma manifold

The Gamma manifold is defined by the parametric family of the probability density function using the scale-shape parametrization:

$$f(x; \alpha, \beta) = \frac{x^{\beta-1}}{\alpha^\beta \Gamma(\beta)} e^{-\left(\frac{x}{\alpha}\right)}. \quad (1)$$

where $\gamma(\cdot)$ is the standard Gamma function.

The Gamma parameters (α and β) can be estimated by the maximum likelihood estimation (MLE), by solving this equation:

$$\hat{\theta} = \arg \max_{\theta} \log \prod_{i=1}^n f(x_i; \alpha, \beta). \quad (2)$$

Which leads to the following system of equations:

$$\hat{\alpha} = \frac{1}{n\hat{\beta}} \sum_{i=1}^n x_i, \quad (3)$$

$$\log(\hat{\beta}) - \frac{\gamma'(\hat{\beta})}{\gamma(\hat{\beta})} = \log\left(\frac{1}{n} \sum_{i=1}^n x_i\right) - \frac{1}{n} \sum_{i=1}^n \log(x_i). \quad (4)$$

The mean and the variance of the Gamma distribution, are respectively given by:

$$E(X) = \alpha\beta, \quad (5)$$

$$Var(X) = \beta\alpha^2. \quad (6)$$

In information geometry [1], a statistical model $\{p_\theta; \theta \in \Theta\}$, where $\Theta \subset \mathbb{R}^r$, can be provided with a Riemannian geometry, that is determined by the Fisher information matrix [19]:

$$g_{ij}(\theta) = E \left\{ \frac{\partial \ln p(X|\theta)}{\partial \theta^i} \frac{\partial \ln p(X|\theta)}{\partial \theta^j} \mid \theta \right\}. \quad (7)$$

with $(i, j = 1, 2, \dots, r)$. This is calculated in the case of the Gamma manifold by the following matrix [16]:

$$g_{ij}(\theta) = \begin{pmatrix} \frac{\beta}{\alpha^2} & 0 \\ 0 & \frac{\partial^2 \ln(\Gamma)}{\partial \beta} - \frac{1}{\beta} \end{pmatrix}.$$

For each $\alpha \in \mathbb{R}$, the α -connection is the torsion-free affine with the components:

$$\Gamma_{ij,k}^{(\alpha)} = \frac{1-\alpha}{2} \partial_i \partial_j \partial_k \varphi(\theta). \quad (8)$$

Where $\varphi(\theta) = \log(\Gamma(\beta)) - \beta \log(\alpha)$ is the corresponding potential function.

So in the case of Gamma manifold it becomes:

$$\begin{aligned} \Gamma_{11,1}^{(\alpha)} &= -\frac{(1-\alpha)\beta}{\alpha^3}, \\ \Gamma_{12,1}^{(\alpha)} &= \Gamma_{12,2}^{(\alpha)} = \frac{(1-\alpha)}{2\alpha^2}, \\ \Gamma_{22,2}^{(\alpha)} &= \frac{(1-\alpha)\psi''(\beta)}{2}. \end{aligned} \quad (9)$$

B. The Weibull manifold

The Weibull manifold is also defined by the parametric family of the probability density function using the scale-shape parametrization:

$$f(x; \lambda, \mu) = \frac{\mu}{\lambda} \left(\frac{x}{\lambda}\right)^{\mu-1} e^{-\left(\frac{x}{\lambda}\right)^\mu}. \quad (10)$$

where $\gamma(\cdot)$ is the standard Gamma function.

Using the MLE algorithm, parameters λ and μ are estimated by solving the following system of equations:

$$\hat{\lambda}^\mu = \frac{1}{n} \sum_{i=1}^n x_i^\mu, \quad (11)$$

$$\hat{\mu}^{-1} = \frac{\sum_{i=1}^n x_i^\mu \ln(x_i)}{\sum_{i=1}^n x_i^\mu} - \frac{1}{n} \sum_{i=1}^n \ln(x_i). \quad (12)$$

The mean and the variance of the Weibull distribution, are respectively given by:

$$E(X) = \lambda \gamma\left(1 + \frac{1}{\mu}\right), \quad (13)$$

$$Var(X) = \lambda^2 \left[\Gamma\left(1 + \frac{2}{\mu}\right) - \left(\Gamma\left(1 + \frac{1}{\mu}\right) \right)^2 \right]. \quad (14)$$

The Fisher information matrix in this case is defined by:

$$g_{ij}(\theta) = \begin{pmatrix} \frac{\mu^2}{\lambda^2} & \frac{\xi - 1}{\lambda} \\ \frac{\xi - 1}{\lambda} & \frac{\xi^2 - 2\xi + \frac{\pi^2}{6} + 1}{\mu^2} \end{pmatrix}.$$

Where ξ is the Euler constant.

The α -connection exists, but it has long analytical expression, so we just mention that the Weibull manifold has a constant Christoffel symbols [17]:

$$\begin{aligned} \Gamma_{11}^1 &= \frac{6(\xi\mu - \mu - \frac{\pi^2}{6})}{\pi^2\lambda}, \Gamma_{11}^2 = -\frac{\mu^3}{\pi^2\lambda^2}, \\ \Gamma_{21}^1 &= \Gamma_{12}^1 = \frac{6(\xi^2 - 2\xi + \frac{\pi^2}{6} + 1)}{\pi^2\mu}, \\ \Gamma_{21}^2 &= \Gamma_{12}^2 = \frac{6\mu(1 - \xi)}{\pi^2\lambda}, \\ \Gamma_{22}^1 &= -\frac{6\lambda(1 - \xi)(\xi^2 - 2\xi + \frac{\pi^2}{6} + 1)}{\pi^2\mu^3}, \\ \Gamma_{22}^2 &= -\frac{6(\xi^2 - 2\xi + \frac{\pi^2}{6} + 1)}{\pi^2\mu}. \end{aligned} \quad (15)$$

III. APPROXIMATION OF THE GEODESIC DISTANCE USING THE GRAPH-BASED APPROACH

A. The geodesic distance

In [20], Rao proposed the Rao-Geodesic distance for computing similarity between distributions of a parametric family, all of whose members satisfy certain conditions [21]. The metric is based on a Riemannian geometry, and is described in terms of the information matrix elements the family. In fact, by considering that $g_{ij}(\theta)$ is strictly positive, for each $\theta \in \Theta$, a Riemannian metric on Θ is defined by:

$$ds^2(\theta) = \sum_{i,j=1}^r g_{ij}(\theta) d\theta^i d\theta^j. \quad (16)$$

Once this metric is introduced, given two probability measures P_{θ_1} and P_{θ_2} which belong to the statistical manifold, the geodesic distance between P_{θ_1} and P_{θ_2} is defined as the Riemannian distance between $\delta(\theta_1, \theta_2) \in \Theta$, and is given by:

$$\delta(\theta_1, \theta_2) = \left| \int_{t_1}^{t_2} \left[\sum_{i,j=1}^r g_{ij}(\theta) \frac{d\theta^i}{dt} \frac{d\theta^j}{dt} \right]^{\frac{1}{2}} dt \right| \quad (17)$$

Particularly, among the curves between θ_1 and θ_1 , we are interested by the one that represents the minimum distance between these two points. It is called the Geodesic, and it is given as a solution to differential equations, called the geodesic equations :

$$\ddot{\theta}^k(t) + \sum_{i,j} \Gamma_{ij}^k[\theta(t)] \dot{\theta}^i(t) \dot{\theta}^j(t) = 0. \quad (18)$$

where the $\Gamma_{\mu\nu}^k$ are the Christoffel symbols of the second kind, defined by:

$$\Gamma_{\mu\nu}^k = \frac{1}{2} \sum_{\rho} g^{k\rho} \left(\frac{\partial g_{\nu\rho}}{\partial \theta^\mu} + \frac{\partial g_{\mu\rho}}{\partial \theta^\nu} - \frac{\partial g_{\mu\nu}}{\partial \theta^\rho} \right). \quad (19)$$

and $g^{\mu\nu}$ denotes the components of the inverse metric.

B. The graph-based approach

In this work, the statistical manifolds are viewed as weighted graphs. It results that the geodesic distance between two points on the manifold is approximated by the shortest path that exists between these two points on the graph. We will use the Floyd-Warshall algorithm [22]. The goal of this algorithm is to look for the shortest paths between all pairs of vertices in a weighted graph, where weights can be positive or negative. The complexity of this algorithm when computing the shortest paths between any 2 vertices is $O(N^3)$, and N is the number of vertices. In the context of image retrieval, an additional step is joined to the SM step, where the Floyd-Warshall algorithm is applied to the matrix of distances (D) to compute the shortest paths between all vertices (images). The input matrix (D) represents the distance between the images existing in the dataset and is calculated using the KLD. It results that all paths that connect the vertices to each other in the weighted graph are initialized by their KLD measure. Then, the GD is approximated by the shortest path method.

IV. EXPERIMENTAL RESULTS

To assess the performance of the proposed approach and its potential, we conduct series of experiments on two popular databases considering the image retrieval application:

- **Dataset1:** The first dataset is a collection of 40 classes from the Vistex database [23]. Each class contains 16 images of 512 x 512 pixels that results in a dataset of 640 images. (Figure. 1).
- **Dataset2:** The second dataset is the Brodatz [24] database which contains 111 gray-level texture images (Figure. 2). Each of those images is divided into 16 of 640 x 640 pixel sub-images which results in a dataset of 1776 images.

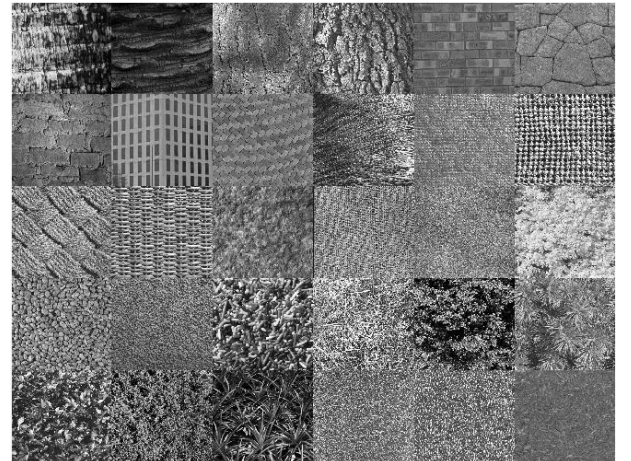


Fig. 1. 30 texture images from the Vistex database.

The texture retrieval framework begins with the Feature extraction (FE) step or in other terms, the construction of

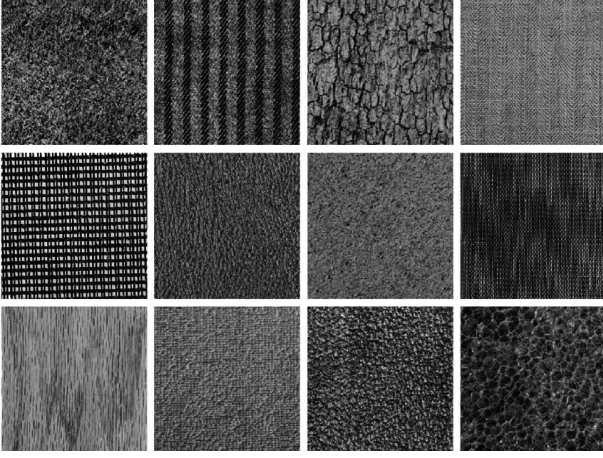


Fig. 2. 12 texture images from the Brodatz database.

the manifold. It starts by converting all the color images that exist in the dataset to grayscale images. Then each image is decomposed into frequency subbands by the DTCWT wavelet decomposition [25] using the Q-Shift (14,14) filters when the decomposition levels are greater or equal than two, and (13,19) near-orthogonal filters for level one. Afterward, we will rely on the Gamma and Weibull distributions to model the histogram of subbands coefficients and estimate their parameters. The estimated parameters for all subbands will construct the signature (point of the manifold) of a single image.

The similarity measurement step, where a distance (here the GD) is calculated between the points of the manifold, especially between signatures of two images I_Q (query image) and I_T (target image) as:

$$GD(I_Q \parallel I_T) = \sum_{j=1}^{6k} GD((\alpha_j^Q, \beta_j^Q, \lambda_j^Q, \Sigma_j^Q) \parallel (\alpha_j^T, \beta_j^T, \lambda_j^T, \Sigma_j^T)). \quad (20)$$

The retrieval accuracy is estimated by presenting every point of the manifold as a query image, for which we retrieve the most similar images. The acquired retrieval rates for all images are averaged to compute the average retrieval rate (ARR) by this formula:

$$ARR(K) = \frac{1}{N_t N_R} \sum_{q=1}^{N_t} n_q(K) \Big|_{K \geq N_R}. \quad (21)$$

where N_t and N_R represent the total number of images in the dataset and the number of relevant images for each query. For each query image q , $n_q(K)$ is the number of correctly retrieved images among the K retrieved ones (i.e K best matches).

To measure the competitiveness of the proposed approaches Gamma+GDFloyd and Weibull+GDFloyd, we compare them with other literature methods. In this regard, we will investigate the following approaches:

- **Weibull+KLD** [14]: It represents the histogram of the wavelet coefficient subbands through the Weibull distribution combined with a closed-form of the KLD,
- **Gamma+KLD** [13]: It models the wavelet coefficient subbands using the Gamma distribution with closed-form of KLD as a similarity measure,
- **GGD+KLD** [26]: In this method the wavelet coefficient subbands are modeled using Generalized gaussian distributions and it uses a closed-form expression of the KLD during the SM step.
- **Gamma+GDSKLD** [27]: In this method, the wavelet subbands are represented through the Gamma distribution and the GD is approximated by the Symmetric KLD,
- **Weibull+GDSKLD**: Same as Gamma+GDSKLD, wavelet subbands are modeled through the Weibull distribution, and the GD is computed approximated by the Symmetric KLD,

Table I shows the ARR of the proposed methods in comparison with the literature methods considering Dataset1 and Dataset2. It is to remark that the proposed approaches lead to higher performances. It means that taking into consideration the extra information given by the geometrical properties of the statistical manifolds improves the retrieval rate. It also confirms the effectiveness of the graph-based approximation of the GD. This is mostly a result of the good representation of the manifold by the graph approach. This latter preserves its structure and geometrical properties.

The performance of our approach is also explained by the use of the GD, which is a real distance compared to the KLD which is not symmetric and does not satisfy the triangular inequality.

Moreover, it is to note that more levels of DTCWT decomposition improves the performance of the ARR. Be that as it may, the underlying two decomposition levels represent by far most of the representation power since the higher improvement is accomplished from 1-Level to 2-Level. The third decomposition level gets little improvement when added together with the underlying two levels. Moving towards another level is not advantageous since it won't improve results.

The difference in performances of all methods considering the two datasets is explained by the number of images (640 vs 1776), and the heterogeneous nature of the images that makes the retrieval harder in case of Dataset2.

Table II compares the proposed method to another GD approximation. In the latter case, the GD is approximated by the square root of the double of the Symmetric KLD (SKLD), since it is proved for distributions that lie infinitesimally close on the probabilistic manifold [28]. We note that the graph-based approach outperforms the SKLD based approximation. This proves the already made conclusion on Table I, and supports that the graph representation conserves the geometrical properties of the manifold.

Figure. 3 shows the Recall/Precision curves obtained using all approaches on Dataset1, considering the third level of decomposition of DTCWT. We note that the proposed

TABLE I
THE AVERAGE RETRIEVAL RATES (%) CONCERNING DATASET1 AND DATASET2 .

	1-Level		2-Level		3-Level	
	Dataset1	Dataset2	Dataset1	Dataset2	Dataset1	Dataset2
Gamma+GDFloyd	75.76	64.92	81.70	73.18	84.62	75.74
Weibull+GDFloyd	76.16	65.11	81.98	73.04	85.16	75.63
Gamma+KLD [13]	71.43	61.74	77.02	69.82	80.77	73.11
Weibull+KLD [14]	71.38	61.93	77.31	69.71	81.01	73.01
GGD+KLD [26]	70.55	62.24	76.17	69.25	79.70	72.35

TABLE II
AVERAGE RETRIEVAL RATE (%) USING DIFFERENT APPROXIMATION METHODS OF THE GEODESIC DISTANCE FOR THREE DTCWT SCALES, ON THE DATASET1 AND DATASET2

	Dataset1	Dataset2
Gamma+GDFloyd	84.62	75.74
Weibull+GDFloyd	85.16	75.63
Gamma+GDSKLD [27]	80.94	73.12
Weibull+GDSKLD [27]	81.64	72.60

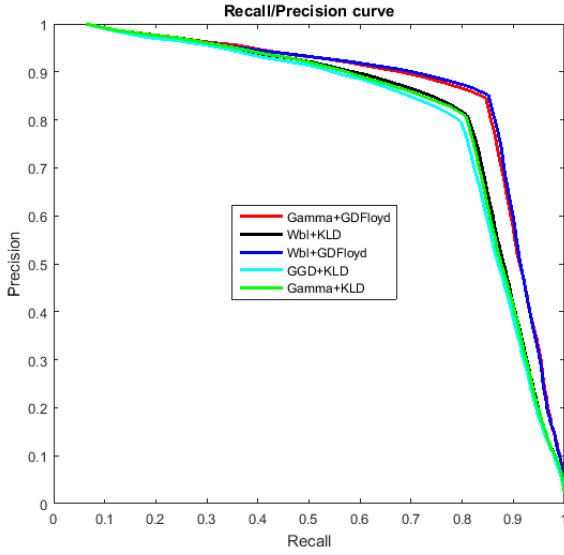


Fig. 3. Recall/Precision curves obtained on Dataset1.

approaches (Gamma+GDFloyd and Weibull+GDFloyd) are clearly outperforming the other literature methods all along the curve. As aforesaid, this is due to the extra information given by the geometrical properties of the statistical manifolds and exploited through the GD, which improves the retrieval rate.

Figure. 4 shows the retrieval precision as a function of the number of retrieved images for Dataset1, considering the

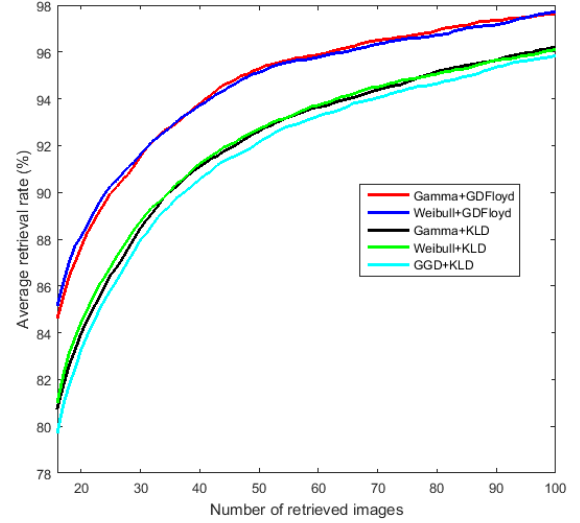


Fig. 4. Retrieval effectiveness with respect to the number of top N matches considered on theDataset1

third level of decomposition of DTCWT. We notice that when considering 16 retrieved images, all the methods gives the same performance as presented in Table I. Nevertheless, the proposed approach shows better performance from the beginning and stays in parallel with the other literature methods all along the curve as the number becomes near to 100 retrieved images. This gives additional proof of the effectiveness of the proposed approaches and show the importance of using the geometrical properties of the statistical manifolds and the accuracy of the graph-based method in exploiting these properties.

V. CONCLUSION

In this work, we proposed the intrinsic geodesic distance as an accurate similarity measurement instead of the statistical Kullback-Leibler divergence. Our main goal was to exploit the properties of the statistical Gamma and Weibull manifolds by using the GD through a graph-based method that preserves most of the geometrical properties. We succeeded to propose a graph-based method to approximate the GD in the case of

Gamma and Weibull manifolds, that outperforms the previous GD approximations. The experimental results indicate that the GD achieves higher performances than the Kullback-Leibler divergence when using the graph-based approximation. Moreover, the advantage with the GD remains important as it fulfills the conditions of a real distance and permits to make geometric interpretations on the studied manifold. Our future works will be devoted to the geometric studies of other manifolds with an application to color texture retrieval.

REFERENCES

- [1] Amari Shun-ichi. *Methods of Information Geometry*. OXFORD University Press, 1993.
- [2] K.Murray Michael and W.Rice John. *Differential Geometry and Statistics*. Springer-Science+Business Media, B.V, 1993.
- [3] Ronald Aylmer Fisher. *Theory of statistical estimation*. *Proceedings of the Cambridge Philosophical Society*, 1925.
- [4] I. Csiszar. *I-Divergence Geometry of Probability Distributions and Minimization Problems*. *The Annals of Probability*, Vol. 3, No. 1, pp. 146-158 (13 pages)., 1975.
- [5] Said Salem, Lionel Bombrun, and Yannick Berthoumieu. *New Riemannian Priors on the Univariate Normal Model*. *The Entropy Journal*, 2014.
- [6] Andai Attila. *On the geometry of generalized Gaussian distributions*. *Journal of Multivariate Analysis*, Volume 100, Issue 4, Pages 777-793, 2009.
- [7] J.N.Torres, J.Gambini, and A.C.Frery. *The Geodesic Distance between $G(1, 0)$ Models and its Application to Region Discrimination*. *Computer Vision and Pattern Recognition*, 2017.
- [8] Chen Yongxin, T. Georgiou Tryphon, and Pavon Michele. *Optimal Steering of a Linear Stochastic System to a Final Probability Distribution - Part III*. *Proceedings of the Cambridge Philosophical Society*, 1925.
- [9] C. T. J. Dodson, Galanis George, and Vassiliou Efstathios. *Geometry in a Fréchet Context A Projective Limit Approach*. London Mathematical Society Lecture Note Series 428. Cambridge University Press, 2015s.
- [10] Verdoolaege Geert and Paul Scheunders. *The Geometry of Multivariate Generalized Gaussian Models— Part I: Metric and Geodesic Equations*. *IEEE TRANSACTIONS ON IMAGE PROCESSING*, 2009.
- [11] Verdoolaege Geert and Paul Scheunders. *The Geometry of Multivariate Generalized Gaussian Models— Part II: Texture Retrieval*. *IEEE TRANSACTIONS ON IMAGE PROCESSING*, 2009.
- [12] F. Nielsen. *Closed-form information-theoretic divergences for statistical mixtures*. *Pattern Recognition (ICPR)*, 21st International Conference on, IEEE, pp. 1723-1726, 2012.
- [13] J.R. Mathiassen, A. Skavhaug, and K. Bo. *Texture Similarity Measure Using Kullback-Leibler Divergence between Gamma Distributions*. in *Proceedings of the 7th European Conference on Computer Vision-Part III (ECCV'02)*, pp. 133–147, Springer., 2002.
- [14] Roland Kwitt and Andreas Uhl. *Image similarity measurement by Kullback-Leibler divergences between complex wavelet subband statistics for texture retrieval*. *15th IEEE International Conference on Image Processing*., 2008.
- [15] C.T.J. DODSON and HIROSHI MATSUZOE. *AN AFFINE EMBEDDING OF THE GAMMA MANIFOLD*. *APPS. Applied Sciences*, Volume: 5, Issue: 1, page 7-12, 2003.
- [16] Arwini Khadiga and C.T.J. Dodson. *Information geometry: Near randomness and near independence*. *Lecture notes in mathematics*, Springer, 2008.
- [17] Arwini Khadiga and C.T.J. Dodson. *Alpha-geometry of the Weibull manifold*. *Second Basic Science Conference*, Al-Fatah University, Tripoli, Libya., 2007.
- [18] Vasileios Zografos, Reiner Lenz, and Michael Felsberg. *The Weibull manifold in low-level image processing: An application to automatic image focusing*. *Image and Vision Computing* 31(5):401-417, 2013.
- [19] Ovidiu Calin and Constantin Udriște. *Geometric Modeling in Probability and Statistics*. Springer, 2014.
- [20] C. Radhakrishna Rao. *Information and the accuracy attainable in the estimation of statistical parameters*. *Calcutta Math*, 1945.
- [21] Ann F.S.Mitchell and W.J Krzanowski. *Rao's Distance Measure*. *The Indian Journal of Statistics*, February 2002.
- [22] Robert W. Floyd. *Algorithm 97: Shortest Path*. *Communications of the ACM*. 5 (6): 345, 1962.
- [23] MIT. Vision texture, jun 2009. <http://vismod.media.mit.edu/vismod/imagery/VisionTexture/>.
- [24] P. Brodatz. Brodatz dataset, 1966. <http://www.ux.uio.no/tranden/brodatz.html>.
- [25] S. Mallat. *A Theory for Multiresolution Signal Decomposition: The Wavelet Representation*. *IEEE TRANSACTIONS PAMI-11*, 1989.
- [26] Minh N.Do and Martin Vetterli. *Wavelet-Based Texture Retrieval Using Generalized Gaussian Density and Kullback–Leibler Distance*. *IEEE TRANSACTIONS ON IMAGE PROCESSING*, feb 2002.
- [27] Z. Abbad, A.E. Drissi, S.E. Ouatik, and M. El Hassouni. *Rao-Geodesic distance on the generalized gamma manifold: Study of three sub-manifolds and application in the Texture Retrieval domain*. *Note Di Matematica journal*, 2017.
- [28] S. Kullback. *Information Theory and Statistics*. Dover Publications, 1968.