

Second order model deviations of local Gabor features for texture classification

David Picard, Inbar Fijalkow

▶ **To cite this version:**

David Picard, Inbar Fijalkow. Second order model deviations of local Gabor features for texture classification. Signals, Systems and Computers, 2014 48th Asilomar Conference on, Nov 2014, Pacific Grove, CA, United States. pp.917-920, 10.1109/ACSSC.2014.7094586 . hal-01151327

HAL Id: hal-01151327
https://hal.science/hal-01151327

Submitted on 12 May 2015

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.

Second Order Model Deviations of Local Gabor Features for Texture Classification

David Picard and Inbar Fijalkow
ETIS UMR 8051
ENSEA, Université Cergy-Pontoise, CNRS
F-95000, Cergy, France

***Abstract*—In this paper, we tackle the problem of texture classification with a local approach based on measuring second order deviations with respect to a dictionary of characteristic patterns. At each pixel, we extract local signal properties thanks to several Gabor filters that are aggregated on a small support region. Then, we compute a dictionary of such features that serves as a universal model. The texture signature is the deviation of second order statistics between its local features and the universal model. Experiments are made on two sets of photographic paper textures, and show the soundness of the approach.**

I. Introduction

In this paper, we are interested in the identification of photographic paper textures sharing physical properties (same sheet, same package, same manufacturer), based solely on visual content [1]. The identification of photographic paper is an important challenge for conservation science, since the aging of old photographs depends strongly upon the material it was printed of. Since these photographs are of high value, non-destructive identification techniques have to be used. In our case, the identification is performed on high resolution pictures of the paper, using digital image processing techniques.

Usually, texture identification systems are designed to discriminate between a wide variety of textures such as wood, marble, cloth, *etc*. In our case, images from different papers are visually very similar, and the distinction is made on tiny variations. As such, usuall texture recognition approaches are not well adapted to this specific problem.

To solve the problem, we propose an approach based on Gabor filters, which have been widely used in texture discrimination [2]. However, such characteristics are often computed on the whole image and are not enough to discriminate between the fine categories of photographic paper textures. We argue these textures are composed of several distinctive patterns (or *textons*), each having its own frequency characteristics. To capture such patterns, we propose to compute a universal dictionary of local responses of a Gabor filter bank on all the available images. Then, we compute the second order moments of the local Gabor responses of an image with respect to the universal dictionary, which amounts to map the image on several Riemannian manifolds (of symmetric positive definite matrices $\mathbb{S}_{++}$). The distance between textures is computed using a geodesic distance on $\mathbb{S}_{++}$, allowing to take into account the specific topology of the proposed representation space.

The remaining of this paper is as follows: First we present a quick overview of texture recognition methods as well as related work. Then, we present our method in section III. We perform experiments in section IV before we conclude.

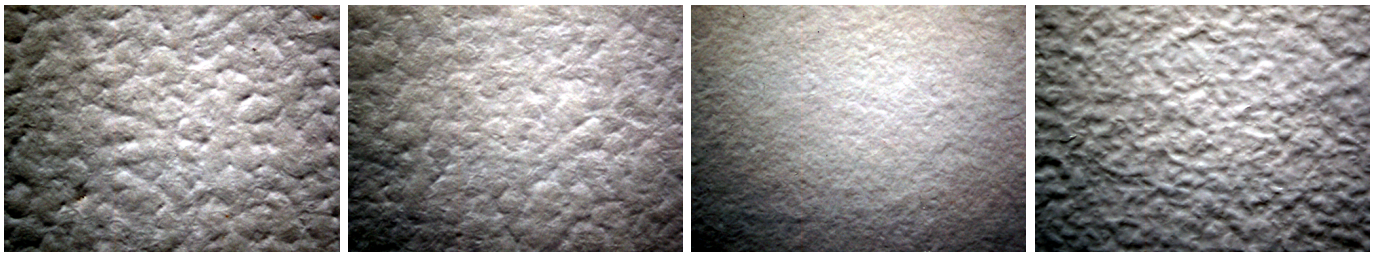
II. Related work

Texture recognition has been a major topic of interest in the image processing community for several decades. Early methods focused on characterizing the statistical distributions of specific frequency band of the Fourier domain [2], [3]. in [4], the authors used polarograms to obtain rotation invariant statistical features. In [5], the authors proposed to used Hough transforms to describe the line structure of the texture.

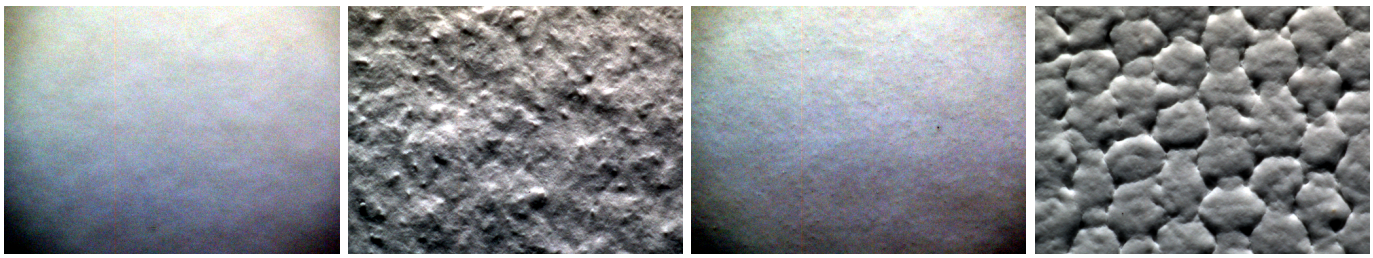
All these global approaches have been outperformed by local approaches that extract characteristic information on a small neighborhood of the image. In [6], the authors proposed for example to use local image patches quantized into a texton codebook. In [7], the difference of intensities in the neighborhood of a pixel is encoded into a binary code that can be rotation invariant. the resulting Local Binary Pattern (LBP) is shown to have very good performances in texture recognition, although it encodes image properties at a small local scale. This descriptor has been extended to encode more information, like in Complete LBP [8] or in the Pattern of Oriented Edge Magnitude [9].

Even more recently, visual dictionaries have been use to perform texture recognition [10]. In these approaches, a set of local features is extracted from the images and quantized so as to produce a dictionary of prototypical descriptors. The image is represented by computing moments of occurrences of the dictionary entries. This approach has been very popular in computer vision to perform object recognition [11]. Given a set of local descriptors (such as LBP), the Bag of Word method consist in assigning each descriptor to its nearest entry in the dictinnary, and computing the histogram of such assignments. This roughly corresponds to computing the zero order moment of the image with respect to the dictionary. Several authors extended this approach by computing higher order moments [12], [13], [14]. In particular, second order moments have shown very good performances in both image retrieval [15]and image classification [16].

In this paper, with take advantage of these second order dictionary approaches, while also taking inspiration from the filter banks methods. In particular, we propose to compute local descriptors using Gabor filters similar to that of the VZ approach [6], while computing second order moments not unlike [13] and [14].



(a) Images from the *bw* dataset.



(b) Images from the *inkjet* dataset.

Fig. 1. Images extracted from the texture datasets. The first row contains images from the *bw* dataset, while the second row contains images from the *inkjet* dataset. Each image is from a different category.

III. Proposed method

Our method is composed of three steps: Local Descriptors extraction, feature computation, similarity measure. The first step of local descriptors extraction is performed using Gabor filters, while the features are computed by evaluating the second order moments of a Gaussian mixture model. Finally, since our features are covariance matrices, we compute the similarity between textures by tacking into account the specific topology of the space of positive definite matrices.

A. Local descriptors

We describe each pixel of the image by the responses to a set of Gabor filters $g(\theta, f)$ [17], taken at m different orientations θ and n different frequencies f:

$$g(\theta, f) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x_\theta^2 + \gamma^2 y_\theta^2}{2\sigma^2}\right) cos\left(2\pi f x_\theta + \frac{\pi}{2}\right) \quad (1)$$

with

$$x_\theta = x\cos(\theta) + y\sin(\theta), \quad y_\theta = -x\sin(\theta) + y\cos(\theta),$$

and $\gamma = \pi/m$. To allow better robustness to translation, we compute the modulus of such responses at any pixel p:

$$r_{\theta,f}(p) = |g_{\theta,f} \star I(p)| \quad (2)$$

To better account for the local properties, the responses are then aggregated in a neighborhood around the pixel p by convolution with a uniform filter u to form the descriptor $d(p)$:

$$d(p) = [u \star r_{\theta,f}(p)]_{\theta,f} \quad (3)$$

Each pixel of the image is thus described by a feature vector of dimension mn, with m being the number of orientations and n the number of frequencies. The pixels are then sorted by decreasing norm of their feature vectors d and subsequently normalized to have unit ℓ_2 norm.

B. Texture features

Starting from a sampling of the feature vectors of all the images, we compute a dictionary of characteristic patterns. The feature space is modeled as a Gaussian Mixture Model trained on the subset of the feature vectors with highest norm (before normalization), using the EM algorithm.

The images are represented by computing second order moments with respect to the model obtained previously. For each Gaussian of the model, we compute the covariance matrix $\Gamma_{I,c}$ of all feature vectors $d(p), p \in I$ of image I, weighted by their respective likelihood $h_c(x)$ to model component c:

$$\Gamma_{I,c} = \frac{\sum_{p\in I} h_c(d(p))(d(p) - \mu_{I,c})(d(p) - \mu_{I,c})^\top}{\sum_{p\in I} h_c(d(p))} \quad (4)$$

with

$$\mu_{I,c} = \frac{\sum_{p\in I} h_c(d(p))d(p)}{\sum_x h_c(d(p))}. \quad (5)$$

Each image is thus represented by an ordered set of positive definite matrices. These matrices correspond to the specificity (or deviation) of the distribution of local Gabor features with respect to a universal model, as measured by their second order moments.

C. Similarity measure

The computation of the covariance matrix for each component induces a mapping of the images into the space of symmetric positive definite matrices $\mathbb{S}_{++}$:

$$\psi_c : \mathbb{R}^{M\times N} \to \mathbb{S}_{++} \quad (6)$$

$$I \mapsto \Gamma_{I,c} \quad (7)$$

To compute the distance between any 2 images I_1 and I_2, we consider the concatenation of the spaces induced by each component of the Gaussian mixture. The square distance in

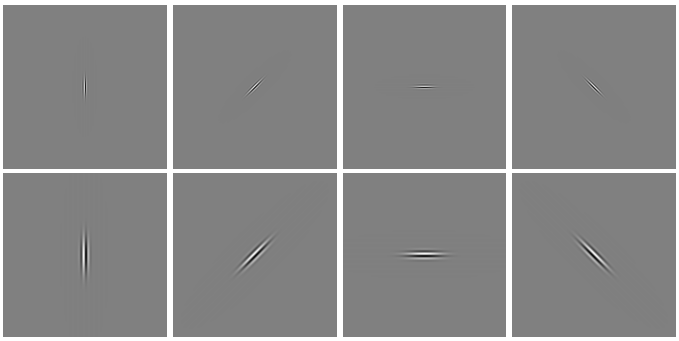


Fig. 2. Gabor filter bank used in this paper.

the concatenated space is then simply the sum of the square distances in the spaces induced by each component:

$$dist(I_1, I_2) = \sum_c dist_{\mathbb{S}_{++}}(\Gamma_{I_1,c}, \Gamma_{I_2,c}) \tag{8}$$

To take into account the topology of $\mathbb{S}_{++}$, we propose to use the Log-Euclidean distance instead of the Euclidean (Frobenius) distance:

$$d(I_1, I_2)^2 = \sum_c \| \log(\Gamma_{I_1,c}) - \log(\Gamma_{I_2,c}) \|_F^2 \tag{9}$$

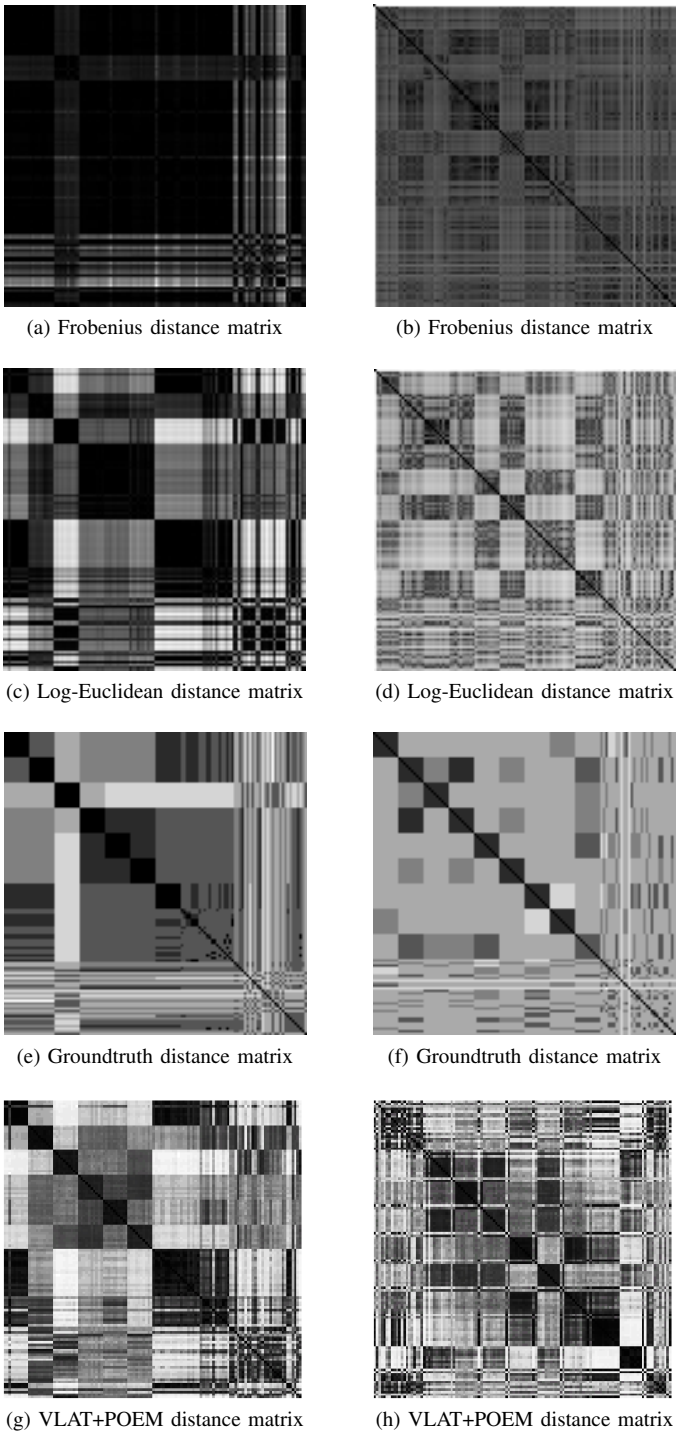
The Log-Euclidean distance has nice properties: it is a metric and corresponds to a geodesic distance on $\mathbb{S}_{++}$ [18].

The incentives to use the Log-Euclidean distance are the following: We consider the eigen-decomposition of a matrix $\Gamma_{I,c}$. Each of its eigenvectors corresponds to a specific combination of the filters in the bank, while the related eigenvalue encodes the contribution (or energy) of this combination. When using the Euclidean distance, the direction of the eigenvectors (*i.e.*, the weights of the filters for each combination) and the contribution are considered equivalently. When using the Log-Euclidean distance, the eigenvalues relative differences are shrunk, whereas the directions remain. We claim that the direction is more important than the contribution, since slight variations in direction lead to textures based on different pattern, while variations on the contributions lead textures based on the same pattern, albeit in different proportions. Using the Log-Euclidean distance, the eigenvalues are processed through a logarithm, which lower their effect on the total distance. Using this distance, the changes in the direction of the eigenvectors are better taken into account.

IV. Experiments

We carried out experiments on two datasets of photographic paper textures named *inkjet* and *b&w* [1]. Each of these datasets contains 120 images grouped in 9 classes based on their physical properties (same sheet, same package, same manufacturer, *etc*). Each class contains 10 images, and the last 30 images are distractors and do not belong to any class. Images of representatives textures from these datasets are shown in Figure 1. As we can see, the textures look already similar, even though they are all drawn from different classes.

We compute local descriptors using a bank of 8 Gabor filters using 4 orientations and 2 scales. These filters are shown in Figure 2. The signatures are computed using a Gaussian mixture model with 8 components.



(a) Frobenius distance matrix (b) Frobenius distance matrix

(c) Log-Euclidean distance matrix (d) Log-Euclidean distance matrix

(e) Groundtruth distance matrix (f) Groundtruth distance matrix

(g) VLAT+POEM distance matrix (h) VLAT+POEM distance matrix

Fig. 3. Distance matrices for the *inkjet* dataset (left) and the BW dataset (right). Dark colors indicate small distances. The last row is taken from [19].

To evaluate our method, we compute the pairwise distance matrices of each dataset. We perform the computation with the Euclidean distance and the Log-Euclidean distance. On Figure 3, we show the corresponding distance matrices, as well as a reference distance matrix provided by the expert

P. Messier. As we can see, Using the Log-Euclidean distance allows to obtain results more consistent with the groundtruth.

We also compared our results with the best distance matrices from [19]. As we can see, the results from [19] are very close to the groundtruth matrices. However, they tend to be too discriminative (very diagonal structure), especially with the textures from the distractors class (last 30 row/columns). The distance obtained by our proposed method seems to better take into account the true distance to the distractor textures.

V. CONCLUSION

In this paper, we presented a new method to perform texture identification, primarily designed to handle the recognition of photographic paper textures. Our approach is divided in three steps. First, we compute and sum the response to a bank of Gabor filter, leading to a set of local descriptors. Then, we model the space of local descriptors with a Gaussian mixture model, and we compute the covariance matrix of the descriptors of each image, weighted by their likely hood with respect to each component of the model. Finally, we compare the matrices of different textures using the sum of component-wise Log-Euclidean distances to better take into account the topology of the space of covariance matrices. We show in the experiments that this approach has promising results on this challenging task.

ACKNOWLEDGMENT

The authors would like to thank P. Messier and C.R. Johnson Jr for the access to the data basis and related discussions, and A. Klein and P. Abry for their interactions on this work.

REFERENCES

[1] P. Messier, R. Johnson, H. Wilhelm, W. A. Sethares, A. G. Klein, P. Abry, S. Jaffard, H. Wendt, S. Roux, N. Pustelnik *et al.*, "Automated surface texture classification of inkjet and photographic media," in *NIP & Digital Fabrication Conference*, vol. 2013, no. 1. Society for Imaging Science and Technology, 2013, pp. 85–91.

[2] M. R. Turner, "Texture discrimination by gabor functions," *Biological Cybernetics*, vol. 55, no. 2-3, pp. 71–82, 1986.

[3] J. Duvernoy, "Optical digital processing of directional terrain textures invar- iant under translation, rotation, and change of scale," *Appl. Opt. 23*, vol. 6, pp. 828—-837, 1984.

[4] L. S. Davis, "Polarograms - a new tool for image texture analysis," *Pattern Recognition*, no. 3, pp. 219–223.

[5] G. Eichmann and T. Kasparis, "Topologically invariant texture descriptors," *Comput. Vision Graphics Image Processing 41*, vol. 3, pp. 267—-281, 1988.

[6] M. Varma and A. Zisserman, "A statistical approach to material classification using image patch exemplars," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 11, pp. 2032–2047, 2009.

[7] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 7, pp. 971–987, 2002.

[8] Z. Guo, L. Zhang, and D. Zhang, "A completed modeling of local binary pattern operator for texture classification," *IEEE Trans. Image Processing*, vol. 19, no. 6, pp. 1657–1663, Jun. 2010.

[9] N.-S. Vu and A. Caplier, "Enhanced patterns of oriented edge magnitudes for face recognition and image matching," *Image Processing, IEEE Transactions on*, vol. 21, no. 3, pp. 1352–1365, 2012.

[10] S. Lazebnik, C. Schmid, and J. Ponce, "A sparse texture representation using local affine regions," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 8, pp. 1265–1278, 2005.

[11] J. Sivic and A. Zisserman, "Video google: A text retrieval approach to object matching in videos," in *International Conference on Computer Vision (ICCV)*, vol. 2, 2003, pp. 1470–1477.

[12] H. Jégou, M. Douze, C. Schmid, and P. Pérez, "Aggregating local descriptors into a compact image representation," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, June 2010, pp. 3304–3311.

[13] F. Perronnin, J. Sánchez, and T. Mensink, "Improving the fisher kernel for large-scale image classification," in *ECCV*, 2010, pp. 143–156.

[14] D. Picard and P. Gosselin, "Efficient image signatures and similarities using tensor products of local descriptors," *Computer Vision and Image Understanding*, vol. 117, p. 680–687, 2013.

[15] R. Negrel, D. Picard, and P. Gosselin, "Web scale image retrieval using compact tensor aggregation of visual descriptors," *IEEE Multimedia*, vol. 20, no. 3, pp. 24–33, 2013.

[16] ——, "Using spatial pyramids with compacted vlat for image categorization," in *International Conference on Pattern Recognition*, Tsukuba Science City, Japan, November 2012.

[17] D. Gabor, "Theory of communication. part 1: The analysis of information," *Electrical Engineers-Part III: Radio and Communication Engineering, Journal of the Institution of*, vol. 93, no. 26, pp. 429–441, 1946.

[18] S. Jayasumana, R. Hartley, M. Salzmann, H. Li, and M. Harandi, "Kernel methods on the riemannian manifold of symmetric positive definite matrices," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*. IEEE, 2013, pp. 73–80.

[19] D. Picard, N.-S. Vu, and I. Fijalkow, "Photographic paper texture classification using model deviation of local visual descriptors," in *IEEE Int. Conf. on Image Processing*, Paris, France, France, Oct. 2014, pp. 5701–5705.