Locally Invariant Texture Analysis from the Topographic Map

Gui-Song Xia, Julie Delon and Yann Gousseau
LTCI, CNRS, TELECOM ParisTech
46 rue Barrault, 75013 Paris, France
$\{$*xia,delon,gousseau*$\}$*@telecom-paristech.fr*

Abstract

In this paper, we present a set of texture features that are locally invariant to similarity or affinity. The proposed indexing scheme relies on the topographic map, a shape-based representation of images. Thanks to the hierarchical organization of the topographic map, the approach gives a grip on the multi-scale structure of textures. Using simple one dimensional histograms, the method is shown to achieve state-of-the-art performances among locally invariant methods, both on the whole Brodatz and UIUC databases.

1. Introduction

Texture analysis has been an active research field for about four decades. Textured surfaces can be viewed under various viewpoints and illumination conditions. Therefore, many works have addressed the problem of defining invariant features to characterize textures. Classically, one seeks for features that are globally invariant to similarity (rotation, translation and scaling) or affinity, as well as to some contrast changes, usually affine contrast changes. However, viewpoint changes can be drastic, or the texture can lie on a non-flat surface, so that one can be tempted to define features that are, in a sense, locally invariant to image transformations. Surprisingly enough, recent studies have shown that it is possible to achieve some local invariance while still being very discriminative between textures, see [5, 13, 12, 6]. In [5] and [13], a set of interest local affine regions are used to form a sparse texture representation for achieving locally affine invariant texture descriptors. A simple and efficient method relying on filter banks has been reported in [6], achieving local invariance to similarity. New developments of locally invariant texture analysis based on fractal approaches are reported in [12].

In a different direction, the mathematical morphology school has proposed the use of granulometry has a tool for texture analysis [10]. The basic idea of these approaches is to analyze images thanks to the global size distribution of geometric entities. An interesting related analysis is proposed in [11], relying on the shape-size pattern spectrum, obtained from connected components of level sets. In this paper, we propose to use a geometric and complete representation of images, the topographic map [2, 7], to derive locally invariant texture features. The approach takes advantage of well known invariant shape descriptors [4, 3] and of the hierarchical structure of the topographic map. It is demonstrated on two texture databases, Brodatz and UIUC, that the resulting features permit to efficiently retrieve or classify texture images, achieving state-of-the-art performances among locally invariant texture analysis methods.

2. The topographic map

This section presents the basic tool of this work, namely the topographic map [2, 7]. The upper and lower level sets of an image u are defined, for $\lambda \in \mathbf{R}$, respectively as $\chi_\lambda(u) = \{u \geq \lambda\}$ and $\chi^\lambda(u) = \{u \leq \lambda\}$. Level lines are defined as the connected components of the boundaries of upper level sets (they could be similarly defined with the help of lower level sets). The topographic map is then defined as the collection of all level lines and can be given a tree structure. Each node of the tree corresponds to a level line. The interior of a level line is called a *shape*. The topographic map (also called tree of shapes) of a digital image can be efficiently computed using a fast algorithm as detailed in [7].

This representation of images is closely linked to other morphological image representations. In particular, the Max-tree and Min-tree introduced in [9] consist in the collection of connected components of upper and lower level sets, respectively. In contrast, the topographic map merges both trees into a single one by considering level lines. Such image representations have

978-1-4244-2175-6/08/$25.00 ©2008 IEEE

three key advantages in our context. First, they represent images as a collection of shapes. This enables one to deal in a simple and direct way with geometric invariances. Second, their hierarchical structure permits to extract information at several scales. This is actually the basic idea behind the use of granulometry to index texture [10]. Last, these representations are invariant to local contrast changes, see [2]. An example of topographic map on a simple synthetic image is shown in Figure 1.

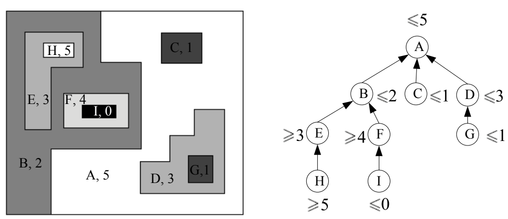


Figure 1: A synthetic image and its topographic map

3. Invariant texture descriptors

The basic idea of this paper is to use the topographic map of images to compute geometric invariant features for textures. We focus on two type of invariances: local invariance to similarity or affinity. In order to do so, we first rely on classical shape moments, then make use of the hierarchical structure of the topographic map to define a new invariant.

3.1. Marginals of shape moments

A simple way to compute invariant features from the topographic map is to compute histograms of shape moments globally for all shapes. Recall that the two-dimensional $(p+q)$th order central moments $\mu_{pq}(s)$ of a shape s are defined as

$$\mu_{pq}(s) = \int\int_s (x-\overline{x})^p (y-\overline{y})^q \, dxdy, \qquad (1)$$

where $(\overline{x}, \overline{y})$ is the center of mass of the shape. Obviously, these moments are invariant to translations. There are many moment invariants, such as Hu's moments [4] and Flusser's moments [3]. For robustness, we only consider second order moments and therefore focus on the inertia matrix, defined as

$$C = \frac{1}{\mu_{00}} \begin{pmatrix} \mu_{20} & \mu_{11} \\ \mu_{11} & \mu_{02} \end{pmatrix}. \qquad (2)$$

Let the eigenvalues of C be $\lambda_1 \geq \lambda_2$. Define the two following shape attributes:

$$\textit{Elongation:} \quad \epsilon = \lambda_2/\lambda_1, \qquad (3)$$

$$\textit{Compactness:} \quad \kappa = \frac{\mu_{00}}{4\pi\sqrt{\lambda_1\lambda_2}}. \qquad (4)$$

Both attributes are invariant to translation, scaling and rotation, and take values in $(0,1]$. In addition, the second one is invariant to affine transformations. Observe that κ^{-2} is the first affine invariant of Flusser and Suk [3]. We prefer to use κ in order to have a range of values between 0 and 1. Another classical attribute that could be used is given by the compactness $p^2/4\pi\mu_{00}$, where p stands for the perimeter. This attribute has proven quite correlated with attribute κ in our experiments on texture databases (even if not affine invariant) and is therefore not considered in this work.

The distributions of the two features (3) and (4) over all shapes in the topographic map are represented as 1D-histograms and are denoted respectively as **elongation histogram (EH)** and **compactness histogram (CpH)**.

3.2. Higher-order statistics on shapes

In order to further extract relevant information from the topographic map, we consider parent-children relationships in the tree of shapes. These can be seen as high order information, involving shape dependency within the image. For dealing with these, a partial neighborhood on the tree of shapes is defined as follows. Let s_i be a shape of the image.

Definition (*Partial Neighborhood* $\mathcal{N}_i^M$) Let s_i^f denote the parent shape of s_i, and $s_i^{f^m}$ be the *m-th* cascaded parent of s_i. For $M \geq 1$, the partial neighborhood of order M of s_i is $\mathcal{N}_i^M = \{s_i^{f^m}, 1 \leq m \leq M\}$.

Recall that $\mu_{00}(s_i)$ is the area of the shape s_i. We define

$$\alpha_i = \frac{\mu_{00}(s_i)}{\langle \mu_{00} \rangle_{\mathcal{N}_i^M}}, \qquad (5)$$

where $\langle \cdot \rangle_{\mathcal{N}_i^M}$ is the mean operator on $\mathcal{N}_i^M$. Observe that $0 < \alpha_i < 1$. Since under an affine transformation $T = AX_i + b$, a plane shape s_i with area $\mu_{00}(s_i)$ is transformed into a shape with area $det(A)\mu_{00}(s_i)$, the resulting feature is affine invariant. In fact, its is locally affine invariant, in the sense that for each shape, it is only sensitive to transformations applied to its M direct ancestors. The distribution of $(\alpha_i)_{i=1}^N$ is again represented by a 1D-histogram, named **scale histogram (SH)**.

It is worth noticing that the proposed affine scale histogram is a locally affine invariant extension of the granulometric approach to texture analysis [10]. In particular, it is related to a recent image classification method [11] relying on a *shape pattern spectra*, a two-dimensional size-shape descriptor build from both Max-tree and Min-tree [9]. This feature does not share the same invariants as ours (the shape part is globally similarity invariant, while the scale part is globally rotation invariant). However, results from [11] indicate that the Max-Min-tree approach could also serve as a basis for locally invariant texture analysis.

3.3. Combining invariant descriptors

So far, we have introduced the following descriptors for indexing textures : the elongation histogram (EH), the compactness histogram (CpH) and the scale histogram (SH) of order M. Observe that these histograms are invariant to any increasing contrast change, as defined in [2]. In order to add some contrast information to our texture indexing scheme, we split each feature histogram into two parts: one for shapes originating from upper level sets (bright shapes) and one for shapes originating from lower level sets (dark shapes).

Additional contrast information is computed in the following way: at each pixel, the gray value is normalized by substracting the mean and dividing by the standard deviation of the smallest shape containing the pixel. The resulting **contrast histogram (CtH)**, computed by scanning all pixels of the image, is invariant to local affine contrast changes, as the features in [5, 6]. We then combine these descriptors as follows:

- *affine invariant (AI)*: CtH+SH+CpH;
- *similarity invariant (SI)*: CtH+SH+CpH+EH.

In order to compare texture samples, we use the *Kullback-Leibler* divergence to compute distances between single descriptors. These distances are then added, after normalizing each of them by its arithmetic mean and standard deviation over the database.

4. Retrieval and classification results

The proposed indexing scheme is evaluated on two different databases both for retrieval and classification tasks. In order to be able to compare with the results of [5, 6], we use the same databases and follow exactly the same protocol as in these two papers. The first database is the complete Brodatz's album [1]. It consists of 111 texture classes, one image for each class. Each texture is cut into nine 215×215 samples. The second one is the UIUC Database [5]. This database is made of 25 texture classes, each containing 40 samples of size 640×480. The particularity of this database is that each class consists of samples that are acquired with strongly varying poses or lay on different surfaces, therefore illustrating the need for local invariance (see Figure 2). It is interesting to check on this database, which is much more demanding regarding geometric invariances, that the methodology is able to discriminate between textures despite strong invariances.

For all experiments, the order of the affine scale histogram is set to $M = 3$. All histograms are uniformly quantized using 50 bins (25 for bright shapes and 25 for dark shapes).

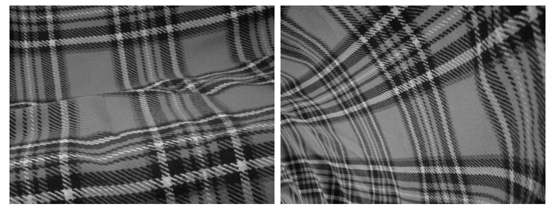


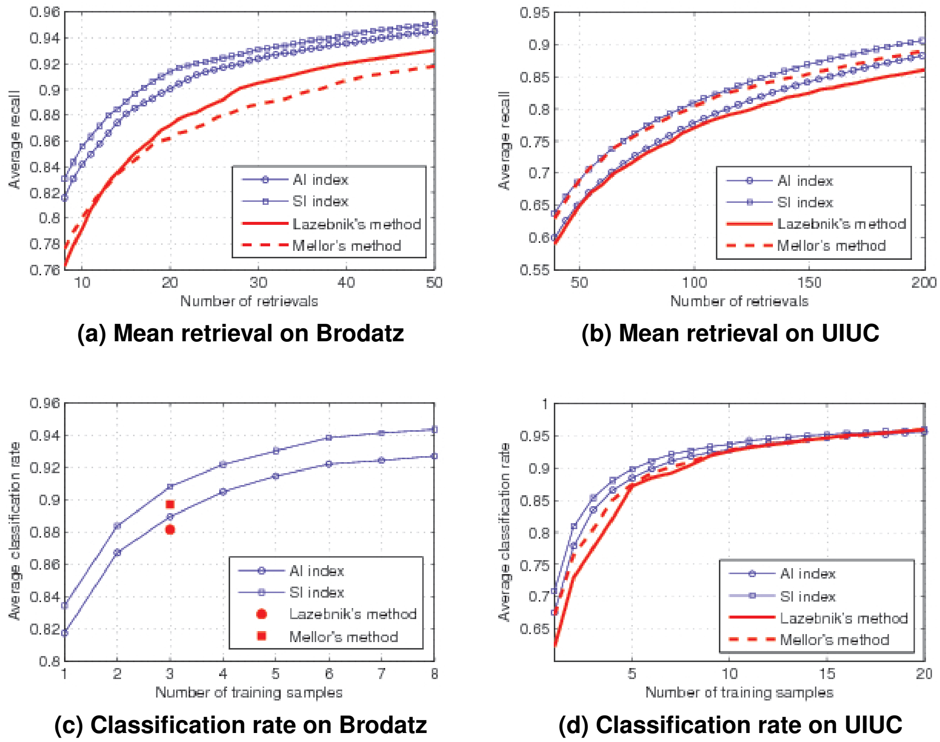
Figure 2: Two samples of the same texture in the UIUC database [5]

The retrieval protocol is the following: a sample is used as a query image (thus removed from the database) and the N_r most similar samples are retrieved. One after the other, all samples are used as query images, and the average recall is computed in function of the number N_r of retrievals.

In the classification experiment, N_c samples from each class are randomly chosen as a training set and the remaining samples are classified thanks to a nearest-neighbor criterion. The rate of correct classifications is then computed in function of N_c. In order to compare directly our results to those of [5] and [6], classification rates are averaged on a sequence of 84 random training sets for Brodatz and 200 random training sets for UIUC.

In Figure 3, we compare the performances of the two descriptor combinations AI and SI (locally invariant to respectively affinities and similarities) introduced in Section 3.3 with the best performances reported in [5] and [6]. Recall that the feature combination introduced in [5] is locally invariant to affinities and that the one from [6] is locally invariant to similarities. In the case of classification on Brodatz, the only results available in [5] and [6], reported in Figure 3 (c), are for $N_c = 3$.

Our descriptors outperform [5] and [6] on Brodatz and UIUC database, both for classification and retrieval. Observe that on UIUC, the mean retrieval performances seem to be correlated with the amount of invariance of descriptors: our AI descriptor achieves performances slightly better than those of Lazebnik *et al.*, while our SI descriptor achieves performances slightly better than those of Mellor *et al.* This confirms the results presented in [6]: despite the strong geometrical variations present in the UIUC database, features that are only in-



(a) Mean retrieval on Brodatz **(b) Mean retrieval on UIUC**

(c) Classification rate on Brodatz **(d) Classification rate on UIUC**

Figure 3: Mean retrieval and classification performances for the AI and SI features (blue lines) and for features from [5] and [6] (red lines).

variant to local similarities show best results for this database. This suggests that achieving invariance to local similarities may be enough to account for viewpoint variations or non-rigid deformations in the framework of texture analysis. Note also that, in this paper as well as in [5, 6], the notion of local invariance is somehow ambiguous as it strongly depends on the chosen texture descriptors. This somehow questions the real usefulness and discriminative power of local invariant indexing schemes, which probably deserves further experiments. We also plan to investigate further how the scale-space property of the topographic map [8] permits to analyze the multi-scale structure of textures.

References

[1] P. Brodatz. *Textures: A Photographic Album for Artists and Designers*. Dover, New York, 1966.

[2] V. Caselles, B. Coll, and J.-M. Morel. Topographic maps and local contrast changes in natural images. *IJCV*, 33(1):5–27, 1999.

[3] J. Flusser and T. Suk. Pattern recognition by affine moment invariants. *PR*, 26:167–174, 1993.

[4] M. K. Hu. Visual pattern recognition by moment invariants. *IRE Trans. Information Theory*, 8:179–187, 1962.

[5] S. Lazebnik, C. Schmid, and J. Ponce. A sparse texture recognition using local affine regions. *IEEE Trans. PAMI*, 27(8):1265–1278, 2005.

[6] M. Mellor, B.-W. Hong, and M. Brady. Locally rotation, contrast, and scale invariant descriptors for texture analysis. *IEEE Trans. PAMI*, 30(1):52–61, 2008.

[7] P. Monasse and F. Guichard. Fast computation of a contrast invariant image representation. *IEEE Trans. Image Procesing*, 9(5):860–872, 2000.

[8] P. Monasse and F. Guichard. Scale-space from a level lines tree. *Journal of Visual Communication and Image Representation*, 11(2):224–236, 2000.

[9] P. Salembier and L. Garrido. Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval. *IEEE Trans. Image Processing*, 9(4):561–576, 2000.

[10] J. Serra. *Image Analysis and Mathematical Morphology*, volume 2. Academic Press, 1982.

[11] E. R. Urbach, J. B. T. M. Roerdink, and M. H. F. Wilkinson. Connected shape-size pattern spectra for rotation and scale-invariant classification of gray-scale images. *IEEE Trans. PAMI*, 29(2):272–285, 2007.

[12] M. Varma and R. Garg. Locally invariant fractal features for statistical texture classification. In *Proc. ICCV*, 2007.

[13] J. Zhang, M. Marszalek, S. Lazebnik, and C. Schmid. Local features and kernels for classification of texture and object categories: A comprehensive study. *IJCV*, 73(2):213–238, 2007.