
Fusion of textural statistics using a similarity measure: application to texture recognition and segmentation

I. Karoui^{1,*}, R. Fablet², J.-M. Boucher¹, W. Pieczynski³ and J.-M. Augustin²

¹ Département SC, Institut Telecom, Telecom Bretagne, FRE CNRS laboratory, STICC 3167, Technopôle Brest-Iroise, CS 83818-29238 Brest Cedex, France

² Ifremer Centre de Brest, B.P. 70 29280, Plouzane, France

³ Institut National des Télécommunications Département CITI, CNRS UMR 5157, 91011 Evry Cedex, France

*: Corresponding author : I. Karoui, email address : imen.karoui@telecom-bretagne.eu

Abstract:

Features computed as statistics (e.g. histograms) of local filter responses have been reported as the most powerful descriptors for texture classification and segmentation. The selection of the filter banks remains however a crucial issue, as well as determining a relevant combination of these descriptors. To cope with selection and fusion issues, we propose a novel approach relying on the definition of the texture-based similarity measure as a weighted sum of the Kullback–Leibler measures between empirical feature statistics. Within a supervised framework, the weighting factors are estimated according to the maximization of a margin-based criterion. This weighting scheme can also be considered as a filter selection method: texture filter response distributions are ranked according to the associated weighting factors so that the problem of selecting a subset of filters reduces to picking the first features only. An application of this similarity measure to texture recognition is reported. We also investigate its use for texture segmentation within a Bayesian Markov Random Field (MRF)-based framework. Experiments carried out on Brodatz textures and sonar images show that the proposed weighting method improves the classification and the segmentation rates while relying on a parsimonious texture representation.

Keywords: Non-parametric feature statistics - Feature fusion and selection - Texture recognition - MRF-based texture segmentation

Introduction

Texture describes a visual information which is related to local spatial variations of color, orientation and intensity in an image. It is usually described by qualitative adjectives such as smooth or rough, coarse or fine, homogeneous or random, and etc. This information is fundamental in image analysis and interpretation and it helps at solving a wide range of applied recognition, segmentation and synthesis problems related to several areas of application including biomedical image analysis, industrial inspection, analysis of satellite, sonar or aerial imagery, content-based retrieval from image databases, scene analysis for robot navigation, texture synthesis for computer graphics and animation etc. A basic goal of texture research in computer vision is to develop automated computational methods for retrieving textural properties in images, and deriving efficient quantitative texture descriptions. Due to the complexity and the diversity of natural textures, there is a wide range of texture features used in the literature. Texture features can however be classified into four main groups: – Statistical approaches that describe a texture via image signal statistics which describe the spatial distribution of image; – Filtering methods that extract texture features from the filter texture response domain; – Structural approaches that consider a texture as a hierarchy of spatial arrangements of well-defined texture primitives; – Probability models which describe the underlying stochastic process that generates textures. The main comparative studies related to texture features reported that none of these feature classes outperforms the others for all textures but have however noted the effectiveness of the co-occurrence matrices, the wavelet frames, the quadrature mirror filter-banks and the Gabor filters [4–10]. Some of these studies [11, 12] stressed that the fusion of different feature types and parame-

ters may improve texture characterization. In fact, for example, Gabor filters, known as good models of the frequency and orientation sensitivity of the human visual system, are parametrized by an angle, corresponding to the expected main orientation of edges and structural elements, and a variance, associated with the scale at which the structural elements of the texture are analyzed. Hence, the parameters of the Gabor function must be varied to describe different types of textures. Changes in the orientation angle lead to different sensitivities to edge and texture orientations. Similarly, co-occurrences [13] are computed as the frequencies of the image values for pixel pairs at a given distance in a given direction. The selection of co-occurrence parameters permits to capture the characteristics of different textures in terms of orientation, smoothness, coarseness and regularity. Recent studies have stressed the effectiveness of the marginals of a large set of filters for characterizing and discriminating texture samples compared to texture features computed as moments of such statistical distribution (for example Gabor features, Haralick parameters, etc) [14–18]. For instance, Zhu et al. [19] proposed a maximum entropy principle for learning probabilistic texture models from a set of empirical distributions of filter responses. Gimel’farb used the “difference co-occurrence” statistics to model texture [20] and later, Xiuwen et al. [14] proposed a local spectral histogram, defined as the marginal distributions of feature statistics for texture classification. Motivated by these studies, we describe texture by a wide set of statistics of texture filter responses and a set of co-occurrence distributions.

Given the variety of the filter types and the associated parameterizations, the selection and the combination of the most relevant filters arise as major issues with a view to improve the classification accuracy and to shorten the learning time. Existing methods however, generally employ filter banks consisting of a large set of filters selected in an ad-hoc manner with pre-determined parameters [21, 22], and only few studies cope with the filter selection for texture synthesis and discrimination. In [19, 20] the selection is based on a synthesis criterion. It is worth noting that this selection may not be appropriate for texture classification. More broadly, the comparative study on texture classification conducted by Randen et al. [5] suggests that it should be preferred to extract texture features based on a discrimination criterion rather than on a characterization criterion.

In the subsequent, we treat the filter selection issue using a texture discrimination based criterion. We describe textures by a set of distributions of local filter responses with respect to different predefined filter banks, and we use the Kullback-Leibler divergence [23] to evaluate the similarity for each type of statistics. The texture similarity measure is then computed as a weighted sum of these Kullback-Leibler divergences. According to a supervised strategy, the weights are estimated such that they lead to the maximization of a margin-based criterion.

The proposed metric serves at texture segregation and can then be used at many field of image processing like pattern classification, object identification, texture synthesis etc. In this paper we report a texture based classification and segmentation of synthetic and natural images based on the proposed similarity measure.

The paper is organized as follows. The proposed texture based similarity measure is introduced in Section 2. A minimum distance classification and a Bayesian Markov Random Field (MRF) based segmentation method using this similarity measure are detailed in Section 3 and some experiments are reported in Section 4.

2 Texture based similarity measure

We characterize each texture sample T by a set of F first-order statistics of filter texture responses $\{Q_f(T)\}_{f=1:F}$. Note that F accounts both for given filter and associated parameterizations. In the subsequent, filter type will refer to a given filter category with some parameterization. The computation of these first-order statistics is issued from a Parzen non parametric estimation [24]. We focus on supervised texture characterization and discrimination such that each texture type k is associated with a texture sample T^k characterized by $Q^k = \{Q_f^k\}_{f=1:F}$.

Here, we propose to use the Kullback-Leibler divergence to compare texture features. The Kullback-Leibler divergence (KL) between two distributions α and β is defined as [23]:

$$KL(\alpha, \beta) = \int \alpha(x) \log \left(\frac{\alpha(x)}{\beta(x)} \right) dx \quad (1)$$

For an unlabeled texture sample T , we compute the set of similarity measures $\{KL(Q_f^k, Q_f(T))\}_{f=1:F}$ for each texture type T^k . Our goal is to determine a relevant combination of these measures to ensure a better discrimination from the resulting similarity measure.

Algorithms for feature selection can be categorized into two classes according to their tie to the induction algorithm: “Filter methods” and “Wrapper methods” [25]. “Filter methods for feature selection” [25] are independent on any induction algorithm, since irrelevant features are filtered out of the data before induction takes place. “Wrapper methods for feature selection” rely on the induction algorithm along with a statistical re-sampling techniques, such as cross-validation, to evaluate feature subsets. The latter methods, generally outperform Filter methods in terms of prediction accuracy, but are generally more complex and more time consuming. We propose a “Filter method” [25] based on a weighting scheme. In opposition to classical “Filter methods for feature selection”, our method fits the characteristics of our similarity measure based on the use of the Kullback-Leibler divergence [23]. The method relies on weighting each filter type according to its relative relevance for the discrimination task. For a given weighting

vector w , the similarity measure between two distribution sets: $Q = \{Q_f\}_{f=1:F}$ and $R = \{R_f\}_{f=1:F}$ is defined as follows:

$$KL_w(Q, R) = \sum_{f=1}^F w_f^2 KL(Q_f, R_f) \quad (2)$$

We aim at determining weights $w_f^2, f = 1, \dots, F$ that maximize the feature space separation. To this end, the proposed procedure comes to the maximization of the global margin expressed by (Eq.4). Given a texture sample T characterized by $Q = \{Q_f\}_{f=1:F}$, a set of prototypes $\{Q^k\}_{k=1:K}$ and a weight vector w , the margin of T is defined as:

$$M_w(T) = KL_w(Q^{D_T}, Q(T)) - KL_w(Q^{S_T}, Q(T)) \quad (3)$$

where S_T is the texture class of T and D_T is the texture type different from S_T the closest to T with respect to KL_w : $D_T = \arg \min_{k \neq S_T} KL_w(Q^k, Q(T))$. Unlike existing margin based ‘‘Filter methods for feature selection’’ [1–3] that estimate the quality of attributes according to how well their values distinguish between samples that are near to each other with respect to the euclidean distance, we compute the margin according to Kullback-Leibler and we evaluate the margin with respect to pre-defined prototypes. This scheme is more appropriate to our application because texture prototypes (empirical non parametric marginal distributions) are estimated on larger support than samples and this parsimonious representation is faster than classical methods. For a training set \mathcal{T} , the total margin is given by:

$$M_w^{\mathcal{T}} = \sum_{T \in \mathcal{T}} M_w(T) \quad (4)$$

The maximization of criterion $M_w^{\mathcal{T}}$ is carried out using a gradient-based approach. For filter type f , the first-order derivative of $M_w^{\mathcal{T}}$ with respect to w_f is given by:

$$\begin{aligned} (\nabla M_w^{\mathcal{T}})_f &= \frac{\partial M_w^{\mathcal{T}}}{\partial w_f} = \sum_{T \in \mathcal{T}} \frac{\partial M_w(T)}{\partial w_f} \\ &= 2w_f \sum_{T \in \mathcal{T}} \left(KL(Q_f^{D_T}, Q_f(T)) - KL(Q_f^{S_T}, Q_f(T)) \right) \end{aligned} \quad (5)$$

Rather than exploiting a purely deterministic gradient ascent, we adopt a stochastic framework with a view to improving convergence performance. It relies on iteratively updating weight vectors with respect to the gradient direction associated with one randomly selected texture sample. More precisely, we proceed as follows:

1. Initialization $w = [1, 1, \dots, 1]$;
2. For $t = 1, \dots, \text{maxITer}$
 - (a) Pick randomly an instance T from \mathcal{T} ;
 - (b) Determine D_T with respect to the weighting vector w ;
 - (c) For $f = 1, \dots, F$, calculate

$$\begin{aligned} \Delta_f &= \\ &2w_f^{t-1} \left(KL(Q_f^{D_T}, Q_f(T)) - KL(Q_f^{S_T}, Q_f(T)) \right) \\ \text{(d) } w^t &= w^{t-1} + \gamma \Delta; \\ \text{3. } w^t &= (w^t)^2 / \|(w^t)^2\|, \text{ where } ((w^t)^2)_f = ((w^t)_f)^2. \end{aligned}$$

The resulting weighting factors are exploited on the one hand for filter selection keeping only the distributions corresponding to the greatest weights, and, on the other hand, for the definition of an optimized texture-based similarity measure KL_w given the selected distributions.

3 Application to distance-based texture recognition and Bayesian based MRF texture segmentation

3.1 Distance-based texture recognition

We first consider an application of the proposed texture based metric to texture recognition. It aims at determining the class of unlabeled texture samples. Considering a supervised issue, texture types $\{T^k\}_{k=1:K}$ are characterized by a set of labeled prototypes $\{Q^k\}_{k=1:K}$. Texture classification can be applied to further image processing and analysis, for instance, to object recognition and image retrieval.

The classification of a texture sample T with respect to the set of reference classes is based on the comparison of the feature statistics through our similarity measure KL_w between texture prototypes $\{Q^k\}_{k=1:K}$, and the statistics estimated on the texture sample and denoted by: $Q(T) = \{Q_f(T)\}_{f=1:F}$. We assign a sample T to class label_T that minimizes $KL_w(Q^k, Q(T))$:

$$\text{label}_T = \arg \min_k KL_w(Q^k, Q(T))$$

3.2 Bayesian MRF based texture segmentation

The second proposed application of the use of the similarity measure KL_w is a Bayesian textured image segmentation. Similar to classification, segmentation of texture also involves extracting features and deriving metrics to segregate textures. However, segmentation is generally more difficult than classification, since boundaries that separate different texture regions have to be detected in addition to recognizing texture in each region. Results of segmentation can be applied to further image processing and analysis, for instance, to object recognition, lesion detection in biomedical imaging, acoustic image segmentation, etc.

We propose a statistical segmentation based on KL_w . Let S be a 2D lattice of pixels, Y the set of observations and X the hidden (unobserved) label field, defined on the same lattice S . Realizations of fields Y and X will be denoted y and x , respectively. The segmentation issue then

comes to estimating x from the observation of the field y . Segmentation we consider in this paper is based on the Markovianity of the posterior distribution $P_{(X/Y)}$. The latter Markovianity is obtained from a Markov field P_X , a prior on the region label map, and $P_{(Y/X)}$ the distribution of Y conditional to $X = x$. Under some assumptions on $P_{(Y/X)}$, the posterior distribution is a Markov distribution and different Bayesian segmentation techniques like the Maximum Posteriori Mode (MPM) [26], Maximum A Posteriori (MAP) [27] or Iterated Conditional Mode (ICM) [28] can be applied.

Here each texture type T^k is characterized by a set of features Q^k and each pixel s is associated with a set of features $Q(W_s) = \{Q_f(W_s)\}_{f=1:F}$ estimated according to a Parzen estimation method [24], within a square window W_s centered at s . Then we consider $Y = \{Q(W_s)\}_{s \in S}$ and we propose to consider that the likelihood of each pixel x_s to each class k is defined from the similarity measure KL_w by:

$$p(y_s/x_s = k) = \frac{\exp^{-KL_w(Q^k, Q(W_s))}}{\sum_{k=1}^K \exp^{-KL_w(Q^k, Q(W_s))}} \quad (6)$$

We use as a prior P_X , a Markov model associated to an 8-neighborhood system with potential functions given by:

$$U_2(x) = \sum_{s \in S} \sum_{t \in c_s} \alpha_c (1 - \delta(x_s, x_t)) \quad (7)$$

where δ is the delta function, $\alpha_c \in \{\alpha_H, \alpha_V, \alpha_D\}$ are real parameters assigned respectively to horizontal, vertical and diagonal cliques.

Using Bayes rule the posterior distribution is expressed as follows:

$$P_{X/Y}(x/y) = \frac{1}{Z} \exp^{\sum_{s \in S} - \sum_{t \in c_s} \alpha_c (1 - \delta(x_s, x_t)) + \log(p(y_s/x_s))} \quad (8)$$

where Z is a normalization function. According to the general Bayesian theory, the MPM segmentation method is optimal with respect to the classification error rate (the number of mis-classified pixels). The decision rule associated to the MPM segmentation strategy is defined for each pixel $s \in S$ as follows [26]:

$$\hat{x}_s^{MPM} = \arg \min_{k \in \Lambda} p_{X/Y}(x_s = k/y_s) \quad (9)$$

where $\Lambda = \{1, \dots, K\}$.

Therefore we have:

$$\hat{x}_s^{MPM} = \arg \max_{k \in \Lambda} - \left(\sum_{t \in c_s} \alpha_c (1 - \delta(x_s = k, x_t)) + KL_w(Q^k, Q(W_s)) - \log \sum_{k=1}^K \exp^{-KL_w(Q^k, Q(W_s))} \right) \quad (10)$$

Model parameters are $\{Q^k\}_{k=1:K}$ and $\alpha_c \in \{\alpha_H, \alpha_V, \alpha_D\}$ are estimated according to Iterative Conditional Estimation (ICE) method [29].

4 Experimental results

Reported experiments have been carried out using 219 texture-related features as follows: a set of 121 distributions issued from co-occurrence statistics [13] computed for the eight main directions $\theta \in \{k\pi/8\}_{k=0:7}$ and for displacements varying between 0 and 10 pixels; a set of 50 distributions of the magnitude of Gabor filter responses, computed for several parameterizations (six normalized radial frequencies $\{\sqrt{2}/2^k\}_{k=1:6}$, and five orientations: $\theta \in \{k\pi/4\}_{k=0:4}$) and a set of 48 distributions of the energy of the image wavelet packet coefficient computed for different bands for three wavelet types: Haar, Debauchies and Coiflet. The computation of the co-occurrence distributions is issued from a k-means [30] based quantization of the images into 10 gray levels, where 100 bins are exploited for Gabor and wavelet-based distributions.

4.1 Distance-based texture recognition

Texture recognition experiments for test images of various complexities have been carried out to evaluate the performance of the proposed texture-based similarity measure. Here, we show the results of the classification of 32×32 texture samples belonging to ten different classes issued from the Brodatz album [31] (D3, D4, D6, D21, D24, D49, D68, D71, D82 and D87) (figure 1). The chosen textures are of different types: D3, D6, D21 and D82 are regular textures formed by regular tiling organized into periodic patterns, but with different scales. D4 and D24 are stochastic textures without noticeable structures, displaying rather random patterns. D49 is an homogeneous texture with accurate orientation. D6 and D21 are fine textures and D68 and D71 are coarse textures.

In order to test for texture discrimination regardless local gray level means or variances, we used two test sets: original texture images and globally histogram-equalized images. Within the training stage, the distributions of the filter responses for each filter type are first estimated for each texture class, as well as the estimation of optimal weighting factors w using the proposed separability-based criterion. We then test for classification performances on sub-images which are not part of the training samples.

For the experiments reported below, we first report the plot of the estimated weighting factors ranked according to the filter type (from 1 to 121 for co-occurrence distributions, from 122 to 171 for Gabor distributions and from 172 to 219 for wavelet distributions). Besides, the classification error rate is used as the evaluation criterion of the relevance of the proposed approach. The proposed margin-based weighting criterion is compared to two random selection procedures (2 random feature permutations) and reverse order ranking with respect to

w . Figure 3 presents these results for the experiments with the ten Brodatz samples of figure 1.

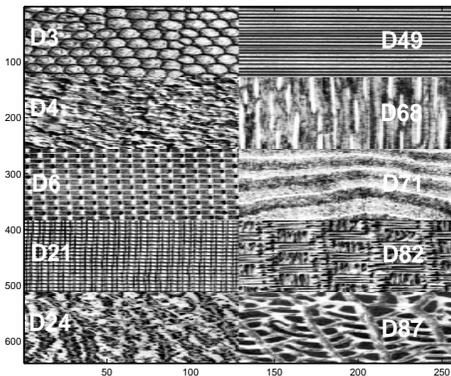


Fig. 1 Test image: 10-Brodatz texture mosaic (D3, D4, D6, D21, D24, D49, D68, D71, D82 and D87)

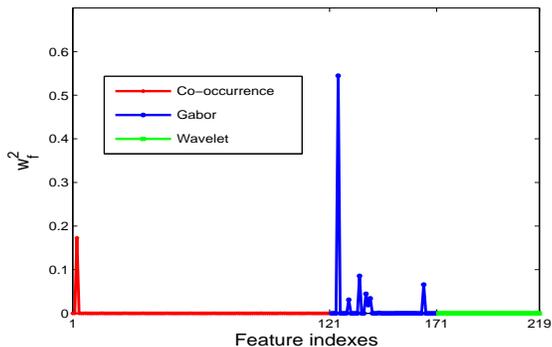


Fig. 2 Estimation of $\{w_f^2\}$ for the definition of KL_w (Eq.2) for the histogram equalized mosaic of figure 1. Feature indexes are as follows. Indexes from 1 \rightarrow 121 we refer to co-occurrence distributions, from 122 \rightarrow 171 to Gabor energy distributions and from 172 \rightarrow 219 to wavelet energy distributions. Co-occurrence distribution computed for parameters $(d, \theta) = (3, 0)$ and 2 Gabor magnitude distributions computed for parameters $(f_0, \theta) \in \left\{ \left(2^{-6} \sqrt{2}, 0 \right), \left(2^{-5} \sqrt{2}, \pi/2 \right) \right\}$ have a weight sum more than 80% of the total weight sum.

The reported classification error rates (τ) are very good compared to the results given in the literature ($\tau = 15\%$ for the ten histogram equalized Brodatz texture). As expected, the best results are obtained by the proposed discriminant ranking and weighting of the computed feature statistics. Only less than 15% of the features have a total weight greater than 80% of the sum of weights of all features (see figure 2 and figure 4). In addition, the optimized similarity measure leads to better discrimination performance compared to a simple filter selection step: we notice a decrease of the error rate classification if we use only the selected set of features compared to the case when we use all distributions with constant weights (from $\tau = 22.5\%$ to $\tau = 15\%$ for

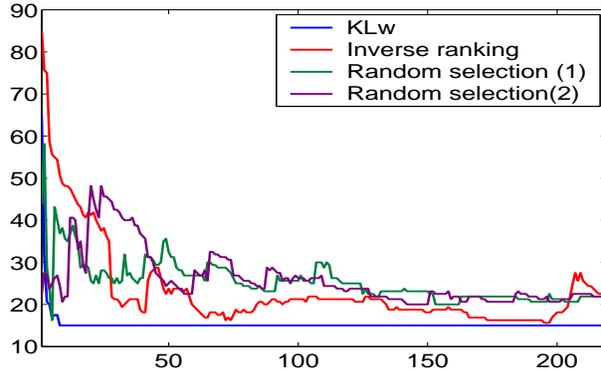


Fig. 3 Comparison of the evolution of classification error rates of the histogram equalized mosaic of figure 1 w.r.t to the number of exploited features respectively for the proposed similarity measure KL_w , 2 random feature selections and a selection according to the inverse feature ranking w.r.t w .

the normalized Brodatz samples and from $\tau = 6.25\%$ to $\tau = 3.75\%$ for non histogram equalized mosaic). It is worth noting that if we do not equalize the texture histograms, we get different weights in favor of the co-occurrence matrices and an improvement of the classification results (a gain of 11.25%) (see figure 4 and figure 5).

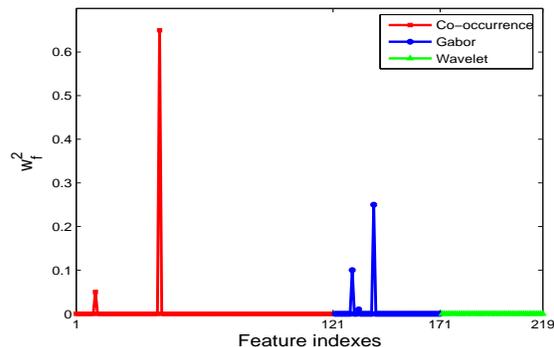


Fig. 4 Estimation of $\{w_f^2\}$ for the definition of KL_w (Eq.2) for the non histogram equalized mosaic of figure 1. Feature indexes are as follows. Indexes from 1 \rightarrow 121 we refer to co-occurrence distributions, from 122 \rightarrow 171 to Gabor energy distributions and from 172 \rightarrow 219 to wavelet energy distributions. Co-occurrence distribution computed for parameters $(d, \theta) = (3, \pi/4)$ and 1 Gabor magnitude distributions computed for the parameters $(f_0, \theta) \in \left\{ \left(2^{-5} \sqrt{2}, \pi/4 \right) \right\}$ have a weight sum more than 80% of the overall weight sum.

Let us further stress that the larger misclassification rates do not always refer to the reserve ranking strategy. This illustrates that the proposed approach does not assign a weight to each feature type independently. Contrary to classical methods based on independent and individual ranking criteria [32, 33]. The computation of the weights of the Kullback-Leibler divergences is issued

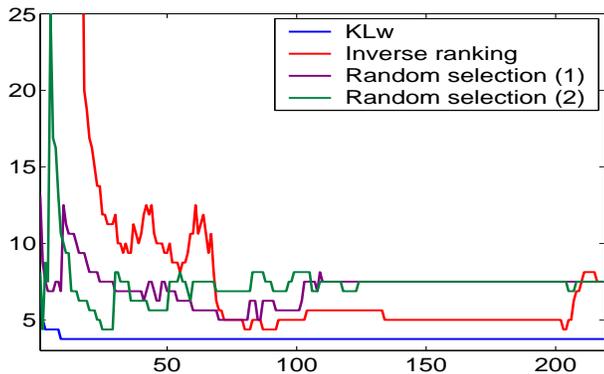


Fig. 5 Comparison of the evolution of classification error rates of the non histogram equalized mosaic of figure 1, w.r.t to the number of exploited features respectively for the proposed similarity measure KL_w , 2 random feature selections and a selection according to the inverse feature ranking w.r.t w .

from a global margin-based criterion, which leads to an efficient feature combination evaluating the redundancy or the complementary properties of the considered statistics.

4.2 Bayesian texture-based segmentation

The segmentation algorithm described in section 3.2 is applied to a set of Brodatz mosaics and real sonar images. Figure 7 shows several MPM based segmentations respectively of the image of figure 6 using a window of size $T_W = 11 \times 11$ and the selected features (we use only features that the weighting sum exceeds 80% of the total weight sum). The image is composed of five textures with different types (D6, D24, D16, D92 and D29) but with some visually close textures. For this image, the selected features are the co-occurrence distribution computed for parameters $(d, \theta) = (7, 0)$ and two Gabor magnitude distributions computed for the parameters $(f_0, \theta) \in \left\{ \left(2^{-6}\sqrt{2}, \pi/4 \right), \left(2^{-5}\sqrt{2}, \pi/2 \right) \right\}$. Figure

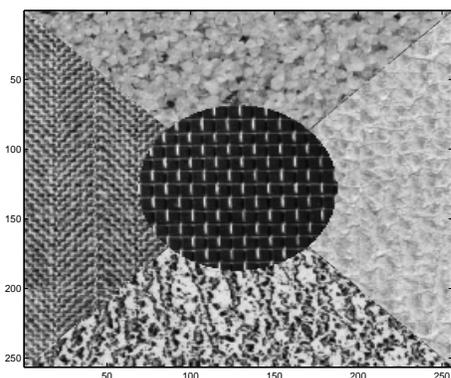


Fig. 6 5-Brodatz texture mosaic (D6, D24, D16, D92 and D29).

	selected features	1 co-occurrence	1 Gabor
MPM	3.1%	4.45%	22.87%

Table 1 Classification error rates

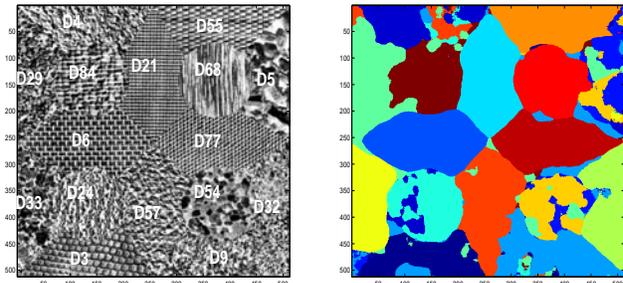


Fig. 8 MPM MRF based segmentation of the 16-texture mosaic, $T_W = 33 \times 33$, $(\hat{\alpha}_H, \hat{\alpha}_V, \hat{\alpha}_D) = (3, 2, 0.5)$.

7.(a) shows the segmentation of the image with the selected features, in figure 7.(b), we report the segmentation obtained when the co-occurrence distribution with the highest weight $(d, \theta) = (7, 0)$ is only used and figure 7.(c) represents the segmentation using one magnitude Gabor distribution computed for Gabor filter having the highest weight among the used Gabor parameters $\left((f_0, \theta) = \left(2^{-6}\sqrt{2}, \pi/4 \right) \right)$. The best classification rate are obtained using all the selected features (see Table 1).

In figure 8, we show the results for the segmentation of a complex mosaic composed of sixteen Brodatz texture [31] with wide variability (D3, D4, D5, D6, D9, D21, D24, D29, D32, D33, D54, D55, D57, D68, D77 et D84): regular, weakly-homogeneous, oriented, coarse, fine and stochastic textures. The segmentation is done using a window of size $T_W = 33 \times 33$. For this image, only three co-occurrence distributions computed for parameters $(d, \theta) \in \{(3, \pi/2), (6, \pi/2), (2, \pi/4)\}$ are selected. The segmentation error rate is about 16.3% which is a very good rate compared to results reported in the comparative study carried out by Randen et al. [4, 5] (the best classification error rate for the compared methods is about $\tau = 37\%$)

In figure 9, we show MPM based segmentation of a real sidescan image further used in seafloor cartography. The image is composed of three seafloor types [34]: a coarse texture of rock, an homogeneous class of mud and oriented texture associated to sand ripples. The co-occurrence distribution computed for $(d, \theta) = (1, \pi/4)$ is the only selected and exploited texture feature. For this image, we use the geologist segmentation as a reference and we compute the classification error rate. It is about 12%.

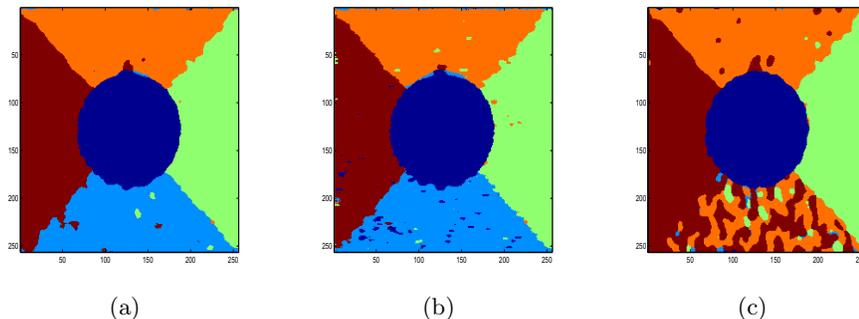


Fig. 7 MPM MRF based segmentation of the 5-texture mosaic of figure 6. (a) Segmentation with the selected features, (b) Segmentation with a co-occurrence distribution, (c) Segmentation with a magnitude Gabor distribution.

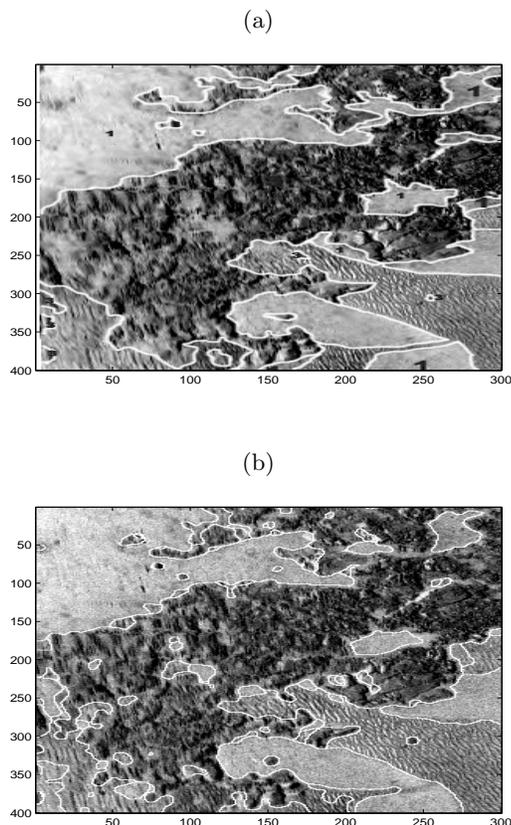


Fig. 9 Sidescan sonar image segmentation (Project Rebent, Ifremer), the boundaries of the different regions given by the segmentation map are superimposed on the original image. (a) Geologist segmentation, (b) MPM based segmentation: $T_W = 9 \times 9$, $T_{max} = 100$ and $(\alpha_H, \alpha_V, \alpha_D) = (0.99, 1.18, 0.14)$ ($\tau = 12\%$).

5 Conclusion

In this paper, we proposed a simple effective histogram selection algorithm for supervised texture discrimination problems. The proposed method is based on a weighted sum of Kullback-Leibler measure of similarity between different texture filter response distributions, where the weights are computed in order to maximize the feature

space separation. We used the proposed similarity measure with a minimum distance classifier on the one hand, and with a hidden Markov field based segmentation, on the other hand. Different experiments, whose results show the effectiveness and the performance of the new models and related processing methods, have been provided.

As perspectives for further works, we can mention that our study opens different ways for other applications based on segmentation or analysis techniques involving filter response statistics. Making the proposed methods unsupervised, is undoubtedly among the most important issues to be addressed in further investigations.

Originality and Contribution

The originality of this work lies in the use of a simple and effective similarity measure between textures in the space of probability distributions within a supervised framework. This similarity measure leads to the fusion and the selection of the most discriminant features. This method is different from the existing feature selection methods in the sense that:

- It does not exploit classical texture features: it relies on empirical marginal distributions of local texture features (co-occurrence distributions, Gabor magnitude distributions, etc) which have been reported to outperform other texture features (Haralick parameters, mean Gabor magnitude, etc).
- In contrary to many feature selection approaches, the method is not restricted to one type of features but it is able to deal at the same time with features of different types and parameterizations (e.g. Gabor, wavelet, co-occurrence, etc).
- The proposed method is an extension of margin based selection methods (Relief, ReliefF, Simba, etc [1–3]) to probability distribution features and to non-euclidean distances. Unlike the existing methods, we do not compute the margin according to the nearest neighbor classifier within a learning sample set but according to the proposed similarity measure between predefined prototypes and samples. Our scheme is more parsimonious and in subsequent faster than other

methods. In fact, for classical methods, in order to compute the margin for a sample according to a training set of N samples, we have to compute the similarities between the sample and all other samples of the set, whereas in our case we simply compute the similarities between the sample and only the K prototypes. Besides, our method is more robust with respect to noise because the similarity between efficient prototypes and samples is more reliable than similarity between samples. Finally, thanks to the non parametric feature aspect, the method allows to deal with multi-modal features.

- The method is a “Filter method” because feature subsets selection is based on a general criterion (margin maximization) and the features are filtered out of the data before induction takes place. However, unlike classical “Filter methods”, the selection is related to the induction algorithm in the sense that we use the same similarity measure to evaluate the margin criterion and to proceed to the classification or to the segmentation.

In this work, only textural features are considered. The extension to other type of features e.g. color, etc would be straightforward. Similarly, the Kullback-Leibler divergence has been chosen as the measure between statistical distributions, but other statistical distances can also be considered, such as Bhattacharya, χ^2 , etc.

Many segmentation, clustering, classification, image retrieval and analysis techniques involving filter response statistics will benefit from this approach. This similarity measure is exploited for texture recognition and Bayesian Markov Random Field (MRF) based segmentation. Reported results shown the relevance and the effectiveness of the method.

References

1. K. Kira and L. Rendell, (1992) A practical approach to feature selection. Proceedings of the International Conference on Machine Learning. Vol.1, pp: 249–256.
2. M. Robnik-Sikonja and I. Kononenko, (2003) Theoretical and Empirical Analysis of ReliefF and RReliefF. Journal of Machine Learning Research. Vol.53, no.1-2, pp: 23–69.
3. R. Gilad-Bachrach, A. Navot and N. Tishby, (2004) Margin based feature selection - theory and algorithms. ACM International Conference Proceeding Series. Vol. 69, pp: 43.
4. T. Randen, (1997) Filter and Filter Bank Design for Image Texture Recognition. Thesis report, Norwegian University of Science and Technology, Stavanger College.
5. T. Randen and J. Husoy, (1999) Filtering for texture classification: A comparative study. IEEE Transaction on Pattern Analysis and Machine Intelligence. Vol.21, no.4, pp: 291–310.
6. R.L. Kashyap, (1984) Characterization and estimation of two-dimensional ARMA models. IEEE Transactions on Information Theory. Vol.30, pp: 736–745.
7. P.P. Ohanian and R.C. Dubes, (1992) Performance evaluation for four class of texture features. Pattern Recognition. Vol.25, no.8, pp: 819–833.
8. R.W. Connors and C.A. Harlow, (1980) A theoretical comparison of texture algorithms. IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol.2, no.3, pp: 204–222.
9. J. Strand and T. Taxt, (1994) Local frequency features for texture classification. International Conference on Pattern Recognition. Vol.27, no.10, pp: 1397–1406.
10. M. Sharma and S. Singh, (2001) Evaluation of Texture Methods for Image Analysis. In 7th Australian and New Zealand Intelligent Information Systems Conference. pp: 117–121.
11. M. Singh and S. Singh, (2002) Spatial texture analysis: A comparative study. International Conference on Pattern Recognition. Vol.1, pp: 676–679.
12. A. Clausi and H. Deng, (2004) Feature fusion for image texture segmentation. International Conference on Pattern Recognition. Vol.1, pp: 580–583.
13. R. Haralick, (1979) Statistical and structural approaches to texture. Proc. of the IEEE. Vol.67, no.5, pp: 786–804.
14. L. Xiuwen and W. DeLiang, (2003) Texture classification using spectral histograms. IEEE Transaction on Image Processing. Vol.12, no.6, pp: 661–670.
15. O.G. Cula and K. Dana, (2003) 3D Texture Recognition Using Bidirectional Feature Histograms. IEEE Transaction on Pattern Analysis and Machine Intelligence. Vol.25, no.12, pp: 1619–1624.
16. P. Nammalwar, O. Ghita and P.F. Whelan, (2005) Integration of feature distributions for color texture segmentation. Conference on Pattern Recognition. Vol.1, pp: 716–719.
17. Q. Xu, J. Yang and S. Ding, (2004) Texture Segmentation using LBP embedded Region Competition. Electronic Letters on Computer Vision and Image Analysis. Vol.5, no.1, pp: 41–47.
18. R. Fablet and P. Bouthemy, (2003) Motion recognition using non parametric image motion models estimated from temporal and multiscale cooccurrence statistics. IEEE Transaction on Pattern Analysis and Machine Intelligence. Vol. 25, no.12, pp: 1619–1624.
19. X. L. S.C. Zhu and Y.Wu, (1998) Filters, random fields and maximum entropy (frame): Toward a unified theory for texture modeling. Inter. Journal of Computer Vision. Vol.27, no.2, pp: 107–126.
20. G. Gimel’farb, (1996) Texture modeling by multiple pairwise pixel interactions. IEEE Transaction on Pattern Analysis and Machine Intelligence. Vol.18, no.11, pp: 1110–1114.
21. A. Jain and F. Farrokhanian, (1991) Unsupervised textures segmentation using gabor filters. Pattern Recognition. Vol.24, no.12, pp: 1167–1186.
22. T. Chang and C. kuo, (1993) Texture analysis and classification with tree-structured wavelet transform. IEEE Transaction on Image Processing. Vol.2, no.4, pp: 429–441.
23. S. Kullback, (1959) Information theory and statistics. Wiley, New York.
24. E. Parzen, (1962) On the Estimation of a Probability Density Function and the Mode. Annals of Mathematical Statistics. Vol.33, pp: 1065–1076.
25. R. Kohavi and G. John, (1997) Wrappers for feature subset selection. Artificial Intelligence. Vol.97, no.12, pp: 273–324.
26. J. Marroquin, S. Mitter and T. Poggio, (1987) Probabilistic solution of ill-posed problems in computational vision.

- Journal of the American Statistical Association. Vol.82, pp: 76–89.
27. S. Geman and G. Geman, (1984) Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transaction on Pattern Analysis and Machine Intelligence*. Vol.6, no.6, pp: 721–741.
 28. J. Besag, (1986) On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society, Series B*. Vol.48, pp: 259–302.
 29. W. Pieczynski, (1992) Statistical image segmentation. *Machine Graphics and Vision*. Vol.1, no.2, pp: 261–268.
 30. R.O. Duda and P. E. Hart, (1973) *Pattern classification and scene analysis*. John Wiley et Sons Inc..
 31. P. Brodatz, (1966) *Textures: A photographic Album for Artists and Designers*. New York: Dover.
 32. C.A. Jensen and M.A. El Sharkawi and R.J Marks, (2001) Power system security assessment using neural networks: feature selection using Fisher discrimination. *IEEE Transactions on Power system*. Vol.16, no.4, pp: 757–763.
 33. M. Lianantonakis and Y.R. Petillot, (2005) Sidescan sonar segmentation using active contours and level set methods. *Conference Proceedings OCEANS '05*. Vol.1, pp: 719–724.
 34. A. Ehrhold and D. Hamon and B. Guillaumont, (2006) The Rehent monitoring network, a spatial integrated acoustic approach to survey nearshore macrobenthic habitats: application to the bay of Concarneau (South Brittany, France), *ICES Journal of Marine Science*, Vol.63, pp.1604–1615.