Image Segmentation by Multidimensional Anisotropic Diffusion

Maria João Rendas[*]
Laboratoire I3S, CNRS-UNSA, 250 Av. A. Einstein, Bât. 4 SPI, 06560 Valbonne (France)

Jean-Michel Bruneau
ISPACE, 20bis Avenue de Montréal, 06200 Nice (France)

La texture est un des éléments clefs de la segmentation d'images pour l'interprétation de scène. Dans cet article nous proposons un algorithme non-supervisé de segmentation d'images en zones de texture "homogène". Cette opération est une tâche fondamentale pour l'analyse d'images en cartographie, où elle permet d'identifier aisément d'importantes régions à la texture similaire (zone urbaine, cultures, forêts, ...). L'approche adoptée consiste à effectuer une segmentation basée sur les caractéristiques des textures. Ces caractéristiques sont diffusées de manière anisotropique sur l'ensemble de l'image à analyser. Les zones de texture "similaire" sont alors itérativement approximées par une texture unique, identifiant ainsi une zone de texture "homogène". Des examples d'application de l'algorithme proposé à des images de synthèse et réelle sont présentés.

Texture is a basic element in image segmentation for scene interpretation. This paper describes an algorithm for unsupervised image segmentation in regions of homogeneous texture. This is a fundamental task for the analysis of images for cartography purposes, where it enables the identification of the important homogeneous texture regions (urban areas, agriculture fields, forests, ...). In our algorithm, multidimensional texture characterisics are diffused in an anisotropic manner over the whole analyzed image. The regions of "similar" texture are iteratively associated to a single texture, identifying in this way a zone of "homogeneous" texture. Examples of application of the proposed algorithm to synthetic and real images are presented.

1. Introduction

Segmentation of complex images in regions of homogeneous texture is often the initial step in automatic image analysis, enabling the a posterior choice of tools adequate for further interpretation of each region. It is, for instance, a fundamental task in image analysis for cartography purposes, separating the whole image its major distinct regions such as urban areas, agriculture fields, forests, etc. Many existing segmentation algorithms are based on the grey level value at each individual pixel, and not on texture characteristics, and are thus only appropriate for segmentation of smooth surfaces, typical of man-made objects, and not for processing views of natural scenes. Amongst other approches , such as the fitting of parametric models, the co-occurence matrix has demonstrated its value for texture discrimination. In particular, the subspace method **[Oja82]** has shown its efficiency for natural texture classification. However, in its present form, the algorithm requires an initial learning phase, requiring the definition of a representative set of distinct visual appearances of each texture, which may not be possible in some cases or else require a time consuming inventory task.

In this paper, we present an unsupervised version of the subspace method, that combines the advantages of the subspace approach with the controlled smoothing of anisotropic diffusion methods. This algorithm, as it is shown in the paper, is able to yield a good identification of the principal regions of images of natural scenes without needing the *a priori* creation of a *texture data base.*

The subspace algorithm is briefly described in Section 2. In Section 3 we present our algorithm, and in Section 4 we give examples of its application to a digitized photography including an agricultural region and several zones of distinct ground characteristics (rocky regions, bare fields, ...), demonstrating its ability in partitioning the global image into its principal components. Finally, we present our conclusions and some ideas for future work.

2. The Subspace Algorithm

The subspace algorithm for texture classification is based on the association of a vector subspace to each distinct texture type. It is ultimately based on the discriminating properties of co-occurrence matrices, which are local estimates of the second order statistics of the grey level distribution. Let $=(i,j)$ be a generic pixel location, and let i_s be the grey level at s, $i_s \in \{1,2,\cdots,N\} = I$, where N is the number of grey levels. Let $\vec{d} = (d_i, d_j)$, $d_i, d_j \in \mathbf{Z}$, be a fixed vector. The co-occurrence matrix of a closed region A, $\boldsymbol{C}_{\vec{d}}$ is a $(N \times N)$ non-symmetric square matrix, whose generic (n,m) element is an estimate of the second order joint distribution of pixel intensity

$$\boldsymbol{C}_{\vec{d}}(n,m) \approx \Pr\left[i_s = n, i_{s+\vec{d}} = m\right], \quad s, s+\vec{d} \in A.$$

To each possible *pure* texture it corresponds a distinct co-occurrence matrix. The feature of the subspace method which makes it adequate for natural texture segmentation is

[*] On leave from ISR, Instituto Superior Técnico, Universidade Técnica de Lisboa, Portugal.

the recognition that natural, *non-pure* textures should be locally characterized by "similar" co-occurrence matrices. In fact, during its learning phase, this algorithm associates to a given texture a proper subspace of the space of all possible co-occurrence matrices, defined in the following way. Let $S_A = \{s_1, s_2, \ldots, s_M\}$ be a collection of sites belonging all to the same texture A, and let $\boldsymbol{C}_{s_i}$, $i = 1, 2, \ldots, M$ be the co-occurrence matrices computed over a small neighborhood of each site in S_A. Define the N^2 vectors obtained by stacking the columns of the matrices $\boldsymbol{C}_{s_i}$

$$v_{s_i} = \mathrm{vect}(\boldsymbol{C}_{s_i}).$$

Finally, let $\boldsymbol{X}_A = \left[v_{s_1} v_{s_2} \ldots v_{s_M}\right]$ $(N^2 \times M)$ be the matrix collecting all vectors corresponding to texture A, and let $\boldsymbol{U}_A, \boldsymbol{V}_A, \boldsymbol{D}_A$ be obtained by singular value decomposition of $\boldsymbol{X}_A$:

$$\boldsymbol{X}_A = \boldsymbol{U}_A \boldsymbol{D}_A \boldsymbol{V}_A^T = \sum_{i=1}^{N^2} \lambda_i \boldsymbol{u}_i \boldsymbol{v}_i^T.$$

According to the basic assumption of the subspace method, the number of non-zero singular values of $\boldsymbol{X}_A$ must be much less than N^2, i.e., $\boldsymbol{X}_A$ can be well approximated by

$$\boldsymbol{X}_A = \sum_{i=1}^{r_A} \lambda_i \boldsymbol{u}_i \boldsymbol{v}_i^T,$$

meaning that each v_{s_i} belongs - approximately - to the subspace Π_A spanned by singular vectors corresponding to the r_A largest singular values. Once a subspace is associated to each texture, classification of each pixel is done by finding the subspace on which the corresponding co-ocurrence matrix has the largest projection.

3. Anisotropic Diffusion of Texture Attributes

The algorithm proposed in this paper adds one hypothesis to the fundamental assumption of the subspace method. Namely, it assumes that whithin one region of homogeneous texture the *local co-ocurrence matrix changes continuously*, while accross boundaries of distinct textures there should be an abrupt change. This same assumption, but on the grey level of each pixel, is the basis of the anisotropic diffusion methods for image segmentation. Our algorithm extends these methods to work on multidimensional features, replacing the grey-level at each image site by a vector subspace to which belongs the co-occurrence matrix computed in its immediate neighborhood. The key idea of the method is to gradually diffuse texture information inside homogeneous regions, such that at convergence, the subspaces Π_A learned in the initial step of the subspace method are identified. In this way, at the end of the diffusion process, the global characteristics of each texture are learned in an unsupervised manner.

This extension of the anisotropic diffusion methods requires the definition of a metric in the space of all subspaces of a given finite dimensional vector space, to be able to compare different textures amongst them. Here, we give an heuristic description of the algorithm, reserving a more rigorous presentaiton to a forthcoming paper.

Let Π_A and Π_B be two proper subspaces of $\mathbf{R}^{N^2}$, of dimension r_A and r_B respectively and let $\boldsymbol{U}_A$ and $\boldsymbol{U}_B$ be matrices whose columns are orthonormal basis for Π_A and Π_B. Let γ_i be the singular values of the matrix $\boldsymbol{U}_A^T \boldsymbol{U}_B$. We define the following normalized measure of similarity between the two subspaces

$$\rho(\Pi_A, \Pi_B) = 1 - \frac{1}{\max(r_A, r_B)} \sum_{i=1}^{\min(r_A, r_B)} \gamma_i^2 \equiv \sin^2 \theta.$$

This similarity measure is related to the notion of principal angles between two subspaces [**Golub**]. In particular, $\rho(\Pi_A, \Pi_B) = 1$ if $\Pi_A \perp \Pi_B$ and $\rho(\Pi_A, \Pi_B) = 0$ if $\Pi_A \equiv \Pi_B$.

The algorithm is initialized by defining a regular lattice $\mathcal{L}$ over the image. This one is partitioned in a set of fixed size windows $\mathcal{W}_s$, a site s of the lattice corresponding to a window. For each site $s = (i, j)$ of this lattice a unidimensional subspace $\Pi_s^0 = \mathrm{span}\left(\mathrm{vect}\left(\boldsymbol{C}_s\right)\right)$, where the co-occurrence matrix is computed over $\mathcal{W}_s$. From this initial assignment, the subspaces associated to each site are slowly deformed and/or enlarged making them more similar to their close - in texture sense - neighbors.

Let $\mathcal{V}_s$ denote a neighborhood of site s in the regular lattice $\mathcal{L}$. At iteration k, the set of sites with which site s is going to diffuse is obtained by comparing Π_s^{k-1} to all Π_t^{k-1}, $t \in \mathcal{V}_s$

$$\boldsymbol{X}_s^k = [\ ]$$

$$\forall t \in \mathcal{V}_s, \text{ if } \rho\left(\Pi_s^{k-1}, \Pi_t^{k-1}\right) < \beta \Rightarrow \boldsymbol{X}_s^k = \left[\boldsymbol{X}_s^k \middle| \boldsymbol{U}_t^{k-1}\right]$$

where $0 \le \beta \le 1$ is a fixed threshold that controls the amount of smoothing performed by the algorithm. Π_s^k is obtained by principal component analysis of $\boldsymbol{X}_s^k$, as the subspace spanned by the vectors associated to its most important singular values - which satisfy a given threshold condition.

This iteration continues until the following stopping condition is achieved:

$$\left(\chi^k - \chi^{k-1}\right) / \chi^k < \varepsilon \quad \text{with} \quad \chi^k = \sum_{s \in S} \rho\left(\Pi_s^k, \Pi_s^{k-1}\right).$$

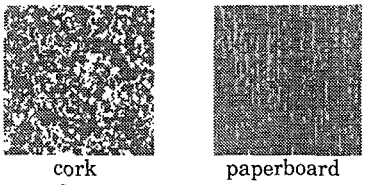
4. Examples

To study the effect of boundary shapes, we applied this algorithm first to a series of images synthetized by combining two natural textures, cork and paperboard, with sharp boundaries of three different shapes between them, see **Figures-1** and **2**.

In these trials windows $\mathcal{W}_s$ are square of size 10, the neighborhoods $\mathcal{V}_s$ correspond to the 4 neighbors system and the cliques are the 2nd order ones. To reduce the computational load, the number of grey levels has been decreased to 4 by linear quantization. For case **(a)** and **(b)**, the co-occurrence matrices were computed using a circulaly symmetric definition, where element (n, m) is an estimate of the second order joint distribution of pixel intensity

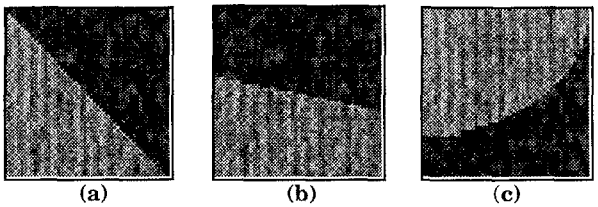
$$\boldsymbol{C}_d(n, m) \approx \Pr\left[i_s = n, i_t = m / \|s - t\| = d\right], \quad s, t \in A,$$

with $d = 1$. For case **(c)** we used the previous definition with a displacement vector $\vec{d} = (1,1)$. The smoothing threshold β was set to 0.2.



cork paperboard

Figure-1: test textures.



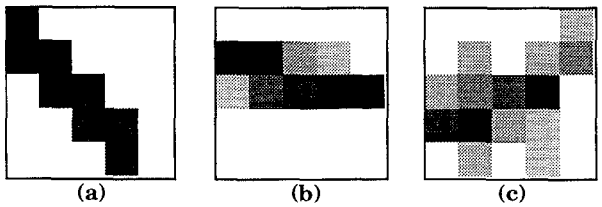
(a) (b) (c)

Figure-2: test images 50×50.

To visualise the result of the algorithm we plot in **Figure-3** the value of

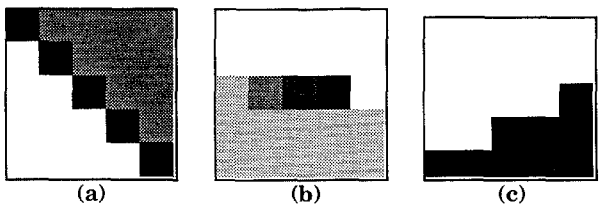
$$\eta(i,j) = \rho\left(\Pi_{i,j}, \Pi_{i-1,j}\right)^2 + \rho\left(\Pi_{i,j}, \Pi_{i,j-1}\right)^2,$$

at the end of the algorithm, which gives an idea of the variation of the solution found across the whole image. Zero is coded as white and large values correspond to the dark regions. Finally, in **Figure-4** we show the result of applying a simple clustering procedure to the subspaces found by the segmentation algorithm, with distinct classes being coded by different grey levels. **Figure-3** shows that the boundary has been well indentified in all cases and that inside each homogenous region an almost constant solution has been found - corresponding to values of $\eta(i,j)$ close to zero. The number of distinct classes in **Figure-4** varies from two to five, but in all cases the image was successfully split into two large homogenous regions. We conclude that boundaries can lead to the definition of additional classes, corresponding to different mixing percentages of the individual textures inside the windows W_s that fall on the border of distinct classes, whenever there is not a clear predominance of one of the classes, as in case **(c)**.



(a) (b) (c)

Figure-3: result of segmentation 5×5.



(a) (b) (c)

Figure-4: result of clustering 5×5.

The following figures show the result of the anisotropic texture diffusion in a real image obtained by scanning an original 22×22 cm black and white aerial photography with a resolution of 300 dpi. We preprocess the digitalized image by reducing its size to 250×270 (**Figure-5**), performing histogram flatenning (**Figure-6**) and reducing the number of grey levels to 8 by linear quantization.

The algorithm was applied with the same neighborhood system as the one use for the test images of **Figure-2**. The co-occurrence matrices were computed using a displacement vector $\vec{d} = (1,1)$. The smoothing threshold β was set to 0.4. We show in **Figures-7** and **8** the values of $\eta_{i,j}$ after 8 and 32 iterations, respectively. **Figures-9** and **10** display the classes found in each case by the clustering procedure.

Figure-5: original image

Figure-6: image after histogram flatenning.

As it should be expected, the homogenous regions grew during the diffusion between iterations 8 and 32. The agriculture fields at the right top of the image were rapidly grouped together into one homogenous region, as well as the dense forest area at the left top. The bare fields at the left

bottom are successfully recognized after 32 iterations, while the iregular rocky portion on the right bottom begins to appear only at iteration 32. The small distinct areas at the center of the image (above the road) are not included in the other larger zones. The largest of these two - upper triangular region - is large enough to be considered as a separate region, while the other one appears under the form of boundary artefacts.

The sparse forest region at the center of the bottom of the image has disappeared during the diffusion process

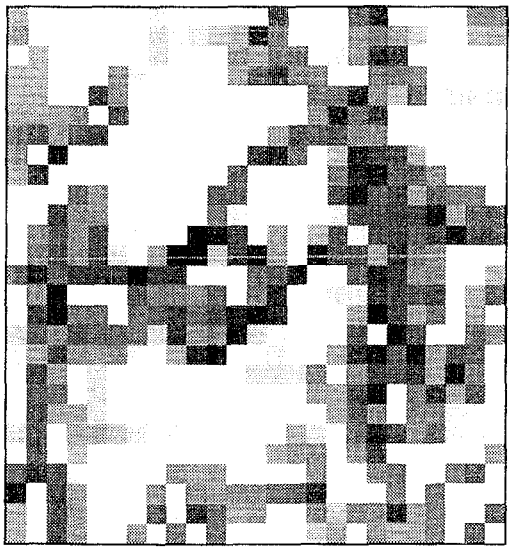


Figure-7: segmented image (8 iterations).

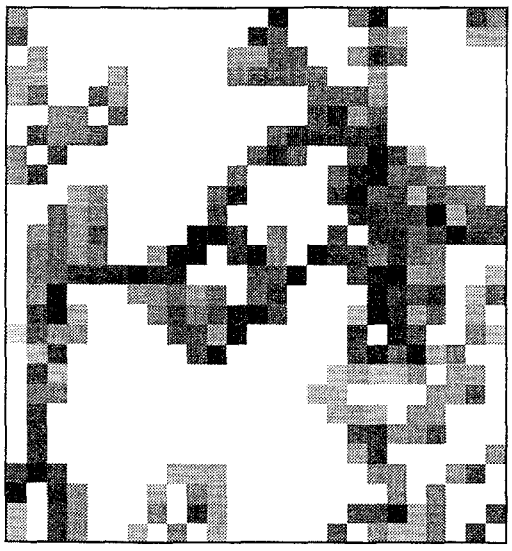


Figure-8: segmented image (32 iterations).

between iterations 8 and 32. This is explained by the fact that the original data $\{\Pi_s^0\}$ was only used in the initialisation step and can possibly be corrected by considering and additional term that measures the fit to the data - i.e. distance between Π_s^0 and Π_s^k - in the present diffusion algorithm.

5. Conclusion

We presented an algorithm for unsupervised segmentation of real images according to their texture characteristics. It is inspired on the subspace method for texture classification, replacing its learning phase by a multivariable anisotropic diffusion procedure. The performance of the algorithm was assessed on a real aerial photography, demonstrating its ability to decompose the image in its important regions. In its actual form, the algorithm is based on a non-optimal definition of the diffusion operator. Future work will address the definition of a framework to formalize the diffusion process.

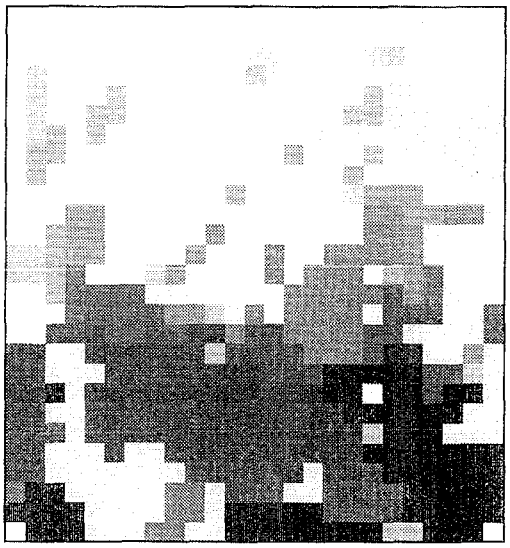


Figure-9: clustered image (8 iterations).

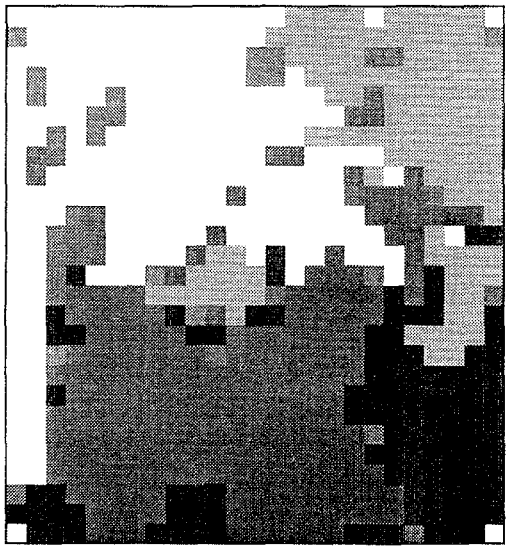


Figure-10: clustered image (32 iterations).

6. References

[Catté-92] F. Catté, P. Lions, J.M. Morel, "Image Selective Smoothing and Edge Dectection by Non-Linear Diffusion", SIAM J. NUMER ANAL., February 1992, Vol. 29, N°1, pp. 182-193.

[Golub] Gene H. Golub and Charles F Van Loan "Matrix Computations", John Hopkins University Press, Baltimore, 1983.

[Oja-87] Oja, Erkki, Parkkinen, Jussi, "Texture Subspaces", NATO ASI Series, Vol. F30, Pattern Recognition Theory and Applications, Springer-Verlag, 1987.