??, ??, 1–33 (??)

Affine Invariant Texture Segmentation and Shape From Texture by Variational Methods

COLOMA BALLESTER & MANUEL GONZÁLEZ

dmicbn0@ps.uib.es, dmimgh0@ps.uib.es

Dep. de Matemàtiques i Informàtica, Universitat de les Illes Balears, 07071 Palma de Mallorca, Spain.

Received ??; Revised ??

Editors: ??

Abstract. We address the problem of texture segmentation by using a novel affine invariant model. The introduction of affine invariance as a requirement for texture analysis goes beyond what is known of the human performance and also beyond the psychophysical theories. We propose to compute texture features using affine invariant intrinsic neighborhoods and affine invariant intrinsic orientation matrices. We discuss several possibilities for the definition of the channels and give comparative experimental results where an affine invariant Mumford-Shah type energy functional is used to compute the multichannel affine invariant segmentation. We prove that the method is able to retrieve faithfully the texture regions and to recover the shape from texture information in images where several textures are present. The numerical algorithm is multiscale.

1. Introduction

Texture is an important characteristic in order to analyse images. Texture analysis has been used in classification tasks concerning 2D images in general and segmentation of aerial or medical images in particular. Texture segmentation has been attempted in many different ways. Most of them follow the same general strategy: a process of feature extraction followed by a segmentation. The main purpose of feature extraction is to map differences in spatial structures, either stochastic or geometric, into difference values in feature space. Then segmentation methods analyse the feature space in order to extract homogeneous regions.

Feature extraction has been effectuated by computing characteristics of the autocorrelation function, partitioning of the Fourier transform energy into radial and angular bins [12], using texton statistics [65], DOG filters [50], using the dominant local orientation estimation at different scales applying the Laplacian pyramid [11], computing the energy of special texture masks [46], [64], interpretation of coocurrence matrices through their moments [36], [22], identification of Markov random process parameters [31], using a Coupled-membrane model in which the texture features are the power responses of quadrature Gabor filters [48], using decompositions in wavelets [43], and many other methods which can be grouped loosely into those based on statistical methods and those using spatial-frequency or spatial-spatial frequency methods. For a detailed review we refer, for instance, to [55].

Our goal in this paper is to propose affine invariant methods for texture segmentation and to prove that they work (at least) as well as the euclidean ones. Since images are the projections of physical objects onto a two dimensional planar surface, affine invariance appears as a simplified form of projective invariance (when the focal distance tends to infinity). All the methods we

propose are unified by a simple affine invariant energy model, with only two terms. One term is allways the same, the Affine Total Variation, and the other one uses affine invariant texture features related to, either the dominant local orientation notion −oriented patterns are common in nature and orientation selective mechanisms are of interest since Hubel and Wiesel's [38] discovery of orientation selective cells in mammalian visual cortex−, or moments. Moreover, using the affine invariant texture information comming from dominant local orientation, the method is able to face the Shape from Texture problem in a similar way as the one proposed by Garding-Lindeberg in [33].

Let us briefly introduce our approach. Coming back to the literature, methods based on coocurrence matrices or the autocorrelation function exploit the fact that the texture is repetitive. As a consequence, peaks occur in the autocorrelation function and concentration of energy is observed in the frequency domain. These non-local methods are able to take into account textures with repetition frequencies ranging from low to high frequencies while other methods are restricted to the size of a prescribed neighborhood. The textural structures which can be described within a neighborhood are naturally restricted to those which can be observed in the neighborhood. Hence, features based on measurements on a neighborhood of fixed size have poor discrimination power when applied to textural structures not observable within the size of the neighborhood, because of the wrong scale selected. In general, this size information is not available. The psychophysical experiments indicate that there exists frequency and orientation selectivity in the human visual system, giving a hint of how this scale problem can be solved [38], [21]. The scaling problem has been treated, in image processing, by filtering with Laplacian filters of different sizes ([16] or using different decompositions in Gabor ([27], [23]), wavelet ([34], [43]) and other basis functions (see [8], [55]). The orientation selectivity within a frequency band can be obtained by filtering the Laplacian image with a directional cosine filter ([26]). [8] observed that this is not desirable in texture analysis because one would like to have a uniform response at the regions of the image which have a dominant orientation. [8] proposed to compute at each level of the Laplacian pyramid a linear symmetry measure or estimation of the dominant local orientation to produce the image features. The number of features is then reduced by using the Karhunen-Loeve transform as in [10], and the clustering algorithm proposed in [61] is applied to these features to obtain a unsupervised segmentation.

As noticed in [55] statistical methods have in the past proved to be superior to frequency domain techniques [18], [66] due to the lack of locality in these earlier frequency analysis methods. But the new joint spatial-spatial frequency techniques are inherently local in nature and are based on image representations that indicate the frequency content in localized regions of the spatial domain. They are able to achieve resolution in both the spatial and spatial-frequency domain, and are consistent with recent theories of human vision. By now, they have characteristics that can be compared favorably with those of the statistical methods. In this direction, Lopez and Morel [47], [49], reformulated the texton theory of Julesz leading to a single multiscale analysis, governed by a parabolic partial differential equation. Indeed, they observed that the textons characteristics indicated by Julesz are shape elements based on curvature and orientation. Considering these elements multiscale because of the *a priori* unknown size of the textons, they concluded that texton densities are nothing but multiscale curvature and orientation densities, which have to be computed using the unique the local, causal, morphological, homogeneous, isotropic and affine invariant scale space (instead of the usual heat equation associated with the Laplacian pyramid), introduced and axiomatically justified in [3] -also proposed in [60]- . In their experiments, Lopez and Morel [47] used multiscale curvature densities at each position and did not take orientation into account.

The main approach we are going to follow in this paper uses orientation densities but it will be slightly different. Partly, because of the desire of constructing affine invariant energy functionals based on the dominant local orientation. Partly inspired by the discusion in [33] concerning scale selection in the context of shape from texture. We are interested in making evident a relation between dominant orientation and an appropriate scale. Collecting these ideas, we shall generate an

oriented texture field, also called a representation of dominant local orientation, computed on adaptative intrinsic elliptic neighborhoods which, as we shall see, give an intrinsic scale addressing the scale selection problem -therefore, the orientation densities do not need to be multiscale (a multiscale variant will be given at the end of Subsection 3.1)-. Indeed, at each point of the image domain, we will try to find an elliptic neighborhood determined by the texture pattern which represents in some sense, a particular length scale related to the texture pattern and to the dominant local orientation (see Figure 2). From it we construct the feature channels for a segmentation algorithm. Several affine invariant Mumford-Shah type energy functionals will be constructed: we propose to use an affine invariant analogue of the second moment matrix, dominant local orientations or statistics of them as channels for a Mumford-Shah type functional. Moreover, the adaptative intrinsic elliptic neighborhood we propose will be used to compute a set of affine invariant features based on Zernike moments. Although the texton doctrine of Julesz is basically euclidean invariant we retrieve the affine invariance in all cases by using the notion of affine total variation as introduced in [6] and a suitable normalization of the image usual in shape analysis ([25], [37]). For a detailed account of affine invariant multiscale methods, we refer to [7], where several possibilities for the definition of the channels are discussed and comparative experimental results are given. Euclidean invariant texture segmentation algorithms based on the simple and elegant Mumford-Shah type energy, were proposed in [43], [48], [49].

On the other hand, texture can be used, in particular, to infer the 3D structure of surfaces in the scene. If we wish to recover information about the 3D shape of visible surfaces, we must distinguish the distorting effects due to projection from the properties of the texture on which the distortion acts. In fact, we can imagine the following two situations: We have an unique texture pattern on a surface with different orientation planes, or we have several different texture patterns. In the first situation, it seems reasonable to accept the strategy of attributing as much as possible of the variations observed in the image to projection. In the second case, we face an ambiguous situation. Either we interpret them as textures with different properties or we suppose as above that the maximum of variation observed in the image is due to projection effects. We shall assume that indeed the first case occurs. In a few words, texture does not mimic projective effects. This permits us to settle a strategy for the shape from texture problem which will be described in Section 5. Our approach to understanding the surface information contained on a textured image will use the affine invariant intrinsic elliptic neighborhoods and the affine invariant analogue of the second moment matrix combined with our affine invariant segmentation model.

This paper is organized as follows. We introduce in Section 2 a multichannel affine invariant energy model which generalizes the one introduced in [6] to segment grey level images and which is our general model to segment textured images. The following two sections, Section 3 and 4, are devoted to introduce several variants of this general model. In Section 3, we start by reviewing the dominant local orientation notion. Then, using the representation of dominant local orientation as feature space, we construct several energy functionals for texture segmentation and discuss its invariance properties. The possibility of using affine invariant moments computed on our intrinsic elliptic neighborhood as channels, will be explored in Section 4. We address in Section 5 the shape from texture problem. The object of Section 6 will be the description of the algorithm used to minimize the energy functionals and the displaying of experimental results. The segmentation algorithm is multiscale but we shall show an important experimental evidence: studying the number of regions obtained from the segmentation algorithm with respect to the scale parameter, we obtain that with any value of the scale parameter in a large interval, the algorithm achieves a correct number of regions in the segmentation.

2. Texture segmentation by a multichannel affine invariant method

In this section we are going to propose an affine invariant variational model we shall use in this work to achieve texture segmentation and to retrieve information about the shape of a textured sur-

face. The affine invariant analogue of Mumford-Shah functional for grey level image segmentation was introduced in [6]. Let us adapt it to the present multichannel segmentation. Since it will also be useful below, let us review in some detail the framework introduced in [6].

As usual, let Ω be an open connected set in $\mathbb{R}^2$ whose boundary is a smooth simple Jordan curve (usually, Ω is an open rectangle). Let $g : \Omega \to \mathbb{R}_+$ be the given image to which we associate a function G with values in feature space, i.e., let $G : \Omega \to E$ be a bounded measurable function with values in a Hilbert space E representing the feature space. Recall that a Jordan curve is a continuous curve $c : [a, b] \to \mathbb{R}^2$ such that for all $t, t' \in]a, b[$, $c(t) \neq c(t')$ if $t \neq t'$ $(a < b)$. If $c(a) = c(b)$ the Jordan curve is said to be closed. The points $c(a)$ and $c(b)$ are called tips of the curve, all other points in the range of c are interior points. Let $\Im$ be the following family of sets

$$\Im = \Big\{ B \subseteq \bar{\Omega} \ : \ B \text{ is a finite union of rectifiable Jordan curves whose interiors are disjoint and contained in } \Omega \Big\}.$$

To introduce the admissible class of segmentations we need the following definition.

Definition 1. . Let $u \in L^2(\Omega, E)$. We say that u is *a cylindrical function in the direction* $v \in \mathbb{R}^2, v \neq 0$, if $\nabla < u, \varphi > \cdot v = 0$ in the sense of distributions, for any $\varphi \in E$. We say that u is *cylindrical* if u is cylindrical in some direction $v \in \mathbb{R}^2, v \neq 0$.

A simple argument shows that u is cylindrical in the direction $v \neq 0$ if and only if, after a possible modification of u in a set with zero measure, $u(x + \lambda v) = u(x)$ for almost every x and all $\lambda \in [0, 1]$, i.e., u is constant on lines parallel to the direction v. Since $u \in L^2(\Omega, E)$, almost all points $x \in \Omega$ are Lebesgue points of u. To choose a particular representative of u we use the following rule: if for $x \in \Omega$ there exists some $\lambda \in E$ such that

$$\lim_{r \to 0} \frac{1}{\pi r^2} \int_{D(x,r)} \|u(y) - \lambda\|_E dy = 0,$$

where $D(x, r) = \{y \in \Omega \ : \ ||y - x|| \leq r\}$, then we define $u(x) = \lambda$. Hence when, for a cylindrical function, we speak of the discontinuity set of u we mean the discontinuity set of its chosen representative.

Let

$$\zeta_0 := \Big\{ u \ : \ \text{there exists } B \in \Im \text{ such that } u : \Omega \to E \text{ is constant on each connected component of } \Omega \setminus B \text{ and } u \text{ is discontinuous on } B \Big\}$$

$$\zeta_1 := \Big\{ u \ : \ u : \Omega \to E \text{ is a cylindrical function } \Big\}.$$

Let

$$\zeta = \zeta_0 \cup \zeta_1.$$

We shall call elements (u, B) of ζ *segmentations*. Function u will be referred to as the *segmented image* and its discontinuity set B as the *segmentation boundaries* or, simply, *segmentation*. Let us observe that segmentations in ζ_0 are Mumford-Shah type segmentations while segmentations in ζ_1 are affine degenerate segmentations. Such affine degenerate segmentations correspond to a underlying transformation of the image by a linear map A with one of its eigenvalues equal to zero.

To introduce the $ATV(\cdot)$, let us define:

Definition 2. Let , , $\tilde{}$ be two rectifiable Jordan curves. We define the *interaction of* , *and* $\tilde{}$ by

$$Inter(, , \tilde{}) = \int_{\Gamma} \int_{\tilde{\Gamma}} |\tau(x) \wedge \tilde{\tau}(y)| \, d\sigma(x) \, d\tilde{\sigma}(y) \quad (1)$$

where $\sigma, \tilde{\sigma}$ denote, respectively, the arclength parameters on each curve , , $\tilde{}$ and $\tau(x), \tilde{\tau}(y)$ denote the tangent vectors at $x \in$, , $y \in \tilde{}$, respectively.

For convenience in notation, given $u \in \zeta_0$, let us consider B as the set of discontinuity of u and write $(u, B) \in \zeta_0$ instead of $u \in \zeta_0$. If u is in ζ_1, the discontinuity set of u may be very wild. On the other hand, it will not play any role in what follows. But, for a uniform notation below, it will be convenient to write also B as the discontinuity set of u and write $(u, B) \in \zeta_1$ instead of $u \in \zeta_1$. We also refer to pairs $(u, B) \in \zeta$ as segmentations.

We now define the ATV functional. Let $(u, B) \in \zeta$. If $(u, B) \in \zeta_0$, then $B = \bigcup_{i=1}^{N}$, $_i$ where , $_i$ are rectifiable Jordan curves whose interiors are disjoint. We set

$$ATV(B) = \sum_{i,j=1}^{N} Inter(, _i, , \tilde{} _j).$$

If $(u, B) \in \zeta_1$, then we set $ATV(B) = 0$. In any case, given $(u, B) \in \zeta$, we define

$$E(u, B) = \int_\Omega ||u - G||_E^2 dx + \lambda ATV(B) \qquad (2)$$

where $\lambda > 0$ and we want to minimize it on the class of segmentations ζ. As observed in [6], the ATV functional is affine invariant. Indeed, for any linear map A one has

$$ATV(A(B)) = |\det A|\, ATV(B). \qquad (3)$$

where $|\det A|$ denotes the absolute value of the determinant of the matrix A. Moreover, let us suppose that G is an affine invariant quantity in the following sense:
If $g, h : \Omega \to I\!R_+$ are images such that $h(x) = g(Ax)$, A being any linear map with $\det A \neq 0$, and $G, H : \Omega \to E$ are the vector features with values in the Hilbert space E associated to g, h, respectively, then there exists an isometry U of E depending on g, A, h such that

$$H(x) = UG(Ax).$$

Then, letting $u'(x) = Uu(Ax)$,

$$\begin{aligned} &\int_\Omega ||u'(x) - H(x)||_E^2 dx \\ &= \int_\Omega ||Uu(Ax) - UG(Ax)||_E^2 dx \\ &= |\det A|^{-1} \int_{A\Omega} ||u(x) - G(x)||_E^2 dx \end{aligned}$$

which, together with (3), proves the affine invariance of Functional (2).

Let us recall that the ATV was not an arbitrary functional. It is shown in [6] to be the only positive functional, up to a scaling factor, associating to each pair of Jordan curves a quantity which is geometric, affine invariant, biadditive and continuous (in the $W^{1,1}$ topology of the space of parametric curves).

As in [6], we have the following theorem concerning the existence of minimizers for Functional (2) which can be proved along the lines of [6].

Theorem 1. *The functional E attains its minimum at some $(u, B) \in \zeta$.*

On the other hand, recall that, as in [6], the problem

$$\min\{E(u, B) \ : \ (u, B) \in \zeta\}, \qquad (4)$$

can be written as

$$\min_{B \in \Im} \min_{u:(u,B)\in\zeta_0} E(u, B). \qquad (5)$$

Since for each $B \in \Im$,

$$\min\{E(u, B) : u \text{ such that } (u, B) \in \zeta\}, \qquad (6)$$

is attained for

$$\bar{u}_B = \sum_O u_O \chi_O, \qquad (7)$$

where

$$u_O = \frac{1}{|O|} \int_O G(x) dx \in E, \qquad (8)$$

$|O|$ denotes the area of the connected component O and the sum extends to all connected components of $\Omega \setminus B$, then Problem (4) is reduced to

$\min\{E(\bar{u}_B, B) \ : \ B$ is either in $\Im$ or is an affine degenerate segmentation and $\bar{u}_B$ is given by (7), (8)$\}$.

3. Texture segmentation using the dominant local orientation

As argued in the Introduction, the representation of dominant local orientations seems to fit to the physiological discoveries of Hubel and Wiesel [38], [4]. Dominant local orientations have been used for texture description [57], texture discrimination [8],[57], and to retrieve information about surface orientation and depth ([67], [57], [33]). In all these cases, they used the same estimate on the dominant local orientation for a textured image.

Consider an image $g : I\!R^2 \to I\!R_+$. Let us denote by $g(x)$ the value of g at $x \in I\!R^2$. We compute the gradient of $g(x)$ by using finite differences. We are interested in giving an estimate of the orientation of $\nabla g(\cdot)$ (modulo π) at some point x_0. In principle this estimate should be given by some average of $\nabla g(\cdot)$ in a neighborhood of x_0. As pointed out

in [45], one cannot smooth the gradient vectors as they tend to cancel each other out at intensity ridges. To avoid such cancellations, estimates of the dominant local orientation have been developped in [45], [57], [8], with a common point of view. These estimates coincide with the one obtained by the method of moments which, moreover, works in any dimension. It tries to give an answer to the following problem. Given the vectors $\{v_1, v_2, \ldots, v_N\}$ how does one determine the dominant orientation of this system? The first answer could be to sum up these vectors and take the resulting direction. But there is the danger that vectors pointing in opposite directions cancel each other instead of influencing each other to produce a dominant orientation. Therefore, to answer the previous question we compute the second moment tensor (or inertia matrix)

$$M = \sum_{i=1}^{N} v_i v_i^t \tag{9}$$

and propose as average orientation the vector v minimizing

$$\min_v \operatorname{tr}\left((M - vv^t)^2\right). \tag{10}$$

Then, we have

Lemma 1. *The vector v minimizing (10) is given by the principal eigenvector of M with norm $\sqrt{\lambda}$, λ being the principal eigenvalue of M.*

The proof of this lemma is given in Appendix A.

Coming back to our purposes, by considering a continuous image g and interpreting v_i above as ∇g at some point x, the analogous of the matrix M could be

$$\frac{1}{|\varepsilon(x)|}\int_{\varepsilon(x)} \nabla g(y)\nabla g(y)^t dy,$$

where the integral is taken on a neighborhood $\varepsilon(x)$ of x and, for convenience, we divided the integral above by the area of $\varepsilon(x)$, $|\varepsilon(x)|$. Some authors use a weighting function $\omega(x)$ and the matrix

$$\mu(g;\omega)(x) = \int_{I\!R^2} \omega(x-y)\nabla g(y)\nabla g(y)^t dy,$$

which is called the second moment matrix, taking its principal eigenvector as the dominant orientation. In practice w can be a Gauss function. We refer to Section 5 for connections of the second moment matrix with Shape from Texture.

The smoothing function ω, usually G_σ, introduces a local scale which is, *a priori*, unknown. This could be, in principle, solved by a multiscale approach. The approach we are going to follow here will be slightly different. Partly, because of the desire of constructing affine invariant energy functionals. Partly inspired by the discusion in [33] concerning scale selection in the context of shape from texture. At each point x, we try to find an elliptic neighborhood $\varepsilon(x)$ determined by the texture pattern in which the histogram is stable. This neighborhood $\varepsilon(x)$ represents in some sense, a particular length scale related to the texture pattern. With this neighborhood, we compute the matrix

$$Q_g(x) = \frac{1}{|\varepsilon(x)|}\int_{\varepsilon(x)} \nabla g(y)\nabla g(y)^t dy$$

to obtain an estimate of the dominant local orientation at x along the lines above. The actual construction of $\varepsilon(x)$ will be explicited in subsection 3.2. This matrix is the basic tool to construct a series of energy functionals for texture discrimination.

3.1. Some energy functionals based on the dominant orientation

In order to explote the previous idea and construct some features vectors, G, we start by introducing the definition of intrinsic elliptic Δ-neighborhood which is at the basis of our affine invariant estimation of the dominant local orientation

Definition 3. (Intrinsic elliptic Δ-neighborhood). Let $g : \Omega \to I\!R_+$ be a given image. Suppose that we have for each point a neighborhood $\varepsilon_g(x)$, a matrix $Q_g(x)$ and a real number $\rho_g(x)$ such that

$$Q_g(x) = \frac{1}{|\varepsilon_g(x)|}\int_{\varepsilon_g(x)} \nabla g(y)\nabla g(y)^t dy \tag{11}$$

$$\varepsilon_g(x) = \left\{y \in I\!R^2 : \langle Q_g(x)(y-x), y-x\rangle \leq \rho_g(x)\right\}, \tag{12}$$

where $\rho_g(x)$ has been selected large enough so that the histogram of g in $\varepsilon_g(x)$ is stable in some sense to be precised below (see Subsection 3.2). Then $\varepsilon_g(x)$ is called an intrinsic elliptic Δ-neighborhood of the image g at point x.

This notion can be seen like a generalization in an affine framework of the Julesz notion of Δ-neighborhood ([39]), which is circular in the original definition. In the sequel $Q_g(x)$ will be called *affine invariant intrinsic orientation matrix* of the image g at point x.

Let us, for convenience, use also the notation $\varepsilon_g\big(x;\rho_g(x)\big)$ to refer to $\varepsilon_g(x)$. We notice that, if A is any linear map in $\mathbb{R}^2$ with $\det A \neq 0$ and $h : A^{-1}\Omega \to \mathbb{R}_+$, $h(x) = g(Ax)$, then we can define an Δ-neighborhood for h, $\varepsilon_h\big(x;\rho_h(x)\big)$, $\rho_h(x)$ and $Q_h(x)$, by

$$Q_h(x) = A^t Q_g(Ax) A \tag{13}$$

$$\varepsilon_h\big(x,\rho_h(x)\big) = A^{-1}\varepsilon_g\big(Ax,\rho_g(Ax)\big) \tag{14}$$

$$\rho_h(x) = \rho_g(Ax). \tag{15}$$

Following Julesz, we assume along this section the existence of an intrinsic elliptic Δ-neighborhood for g at any x. With this on hand, one can construct several affine invariant energy functionals based on the matrix Q_g. Thus, given the image $g : \Omega \to \mathbb{R}_+$ let $Q_g(x), \varepsilon_g(x), \rho_g(x)$ be as above. Let

$$Q_g = \frac{1}{|\Omega|}\int_\Omega Q_g(x)dx. \tag{16}$$

Let $MS_2 = \big\{A :\, 2\times 2\, matrix,\, A = A^t > 0\big\}$ and let $S = \Big\{M : \Omega \to MS_2 \,:\, M = \sum_O \chi_O(x) M_O,$ $M_O \in MS_2$ and the sum extends to all connected components of $\Omega \setminus B$ for some $B \in \Im\Big\}$. S is the set of all piecewise constant images whose range are in MS_2. Observe that $Q_g \in MS_2$. Without loss of generality, we may assume that Q_g is invertible. Otherwise, $g(x)$ would be constant on lines parallel to some direction, that is, g would be cylindrical.

We propose to minimize the functional

$$E_1(Q,B) = \int_\Omega \mathrm{tr}\big(Q_g^{-1/2} Q_g(x) Q_g^{-1/2} - Q\big)^2 dx + \lambda ATV(B), \tag{17}$$

in the set of pairs (Q,B) with $Q \in S$, $B \in \Im$. The dominant orientation corresponds to the principal eigenvalue of $Q_g(x)$. The matrix $Q_g(x)$ also contains information about surface shape, as we shall describe in Section 5.

The matrix Q_g is used in (17) as a normalization factor. It plays an essential role to recover affine invariance (see [25], [20]). Indeed, for any linear map A such that $\det A \neq 0$, and setting $h(x) = g(Ax)$, $x \in A^{-1}\Omega$, then

$$Q_h(A^{-1}\Omega) = A^t Q_g(\Omega) A. \tag{18}$$

The identity (18) is the basis of the normalization theory for images, which permits the introduction of affine invariant quantities associated with usual euclidean invariant ones. For the sake of completeness we shall review a little bit more than is needed. We believe this to be helpful for a better understanding of the developpements below.

Thus, making abstraction of the actual definition of Q_g, let us suppose that with any image $g : A\Omega \to \mathbb{R}_+$ and $h(x) = g(Ax)$, $x \in \Omega$, with A as above, we associate symmetric matrices $M_g(A\Omega)$, $M_h(\Omega)$ satisfying

$$M_h(\Omega) = A^t M_g(A\Omega) A. \tag{19}$$

Let us observe the first simple but useful implication of (19).

Fact 0: $M_h^{-1/2}(\Omega) A^t M_g^{1/2}(A\Omega)$ is a rotation matrix. Indeed, from (19)

$$I = \Big(M_h^{-1/2}(\Omega) A^t M_g^{1/2}(A\Omega)\Big)\Big(M_g^{1/2}(A\Omega) A M_h^{-1/2}(\Omega)\Big),$$

which is a translation of Fact 0.

Let us introduce the following definition. For simplicity we suppose all images defined in $\mathbb{R}^2$.

Definition 4. ([25], [20]) We say that the matrix A with $\det A \neq 0$ normalizes an image $g : \mathbb{R}^2 \to \mathbb{R}$ if $h(x) = g(Ax)$ is such that $M_h = I$.

Since $M_h = A^t M_g A$, A normalizes g if and only if $A^t M_g A = I$. In that case $M_g = (A^{-1})^t A^{-1}$. Writing $C = A^{-1}$ we see that $M_g = C^t C$. Moreover, any invertible matrix C such that $M_g = C^t C$ permits to construct a matrix A normalizing g. It suffices to take $A = C^{-1}$. In particular:

Fact 1: $M_g^{-1/2}$ normalizes g.

The process of normalizing the images was introduced in the context of shape analysis to reinforce affine invariance. Thus, it is not surprising to find it here. Two important and related facts in the context of normalization are the following:

Theorem 2.

(a) Uniqueness of the normalizing matrix. *Given an image $g : \mathbb{R}^2 \to \mathbb{R}$ the matrix A normalizing g is unique, up to orthogonal matrices.*

(b) Uniqueness of the normalized image. *If two images g, h are related by a linear matrix A, $h(x) = g(Ax)$, then they belong to the same normalization class, i.e., if $\bar{h}$, $\bar{g}$ are the normalizations of h, g respectively, then $\bar{h}(x) = \{U\}\bar{g}(x) = \bar{g}(Ux)$ for some orthogonal matrix U.*

The proof of these facts is elementary. Let us give them in Appendix B for the sake of completeness. Theorem 2 says that the normalization is a well defined and intrinsic process.

Let us prove the affine invariance of functional E_1.

Proposition 1. *The functional E_1 is affine invariant.*

Proof: Let A be a linear map in $\mathbb{R}^2$, $g : \Omega \to \mathbb{R}_+$, $h : A^{-1}\Omega \to \mathbb{R}_+$ images related by $h(x) = g(Ax)$. Let $Q_h(x), Q_g(\bar{x})$, $x \in A^{-1}\Omega, \bar{x} \in \Omega$, Q_h, Q_g be defined as in (11), (16). Using (13) one easily sees that

$$Q_h(A^{-1}\Omega) = A^t Q_g(\Omega) A$$

which implies, by Fact 0, that $U_A = Q_h(A^{-1}\Omega)^{-\frac{1}{2}} A^t Q_g(\Omega)^{\frac{1}{2}}$ is a rotation. Hence,

$$\int_{A^{-1}\Omega} \mathrm{tr}\big(Q_h^{-\frac{1}{2}}(A^{-1}\Omega) Q_h(x) Q_h^{-\frac{1}{2}}(A^{-1}\Omega) - Q\big)^2 dx = \int_{A^{-1}\Omega} \mathrm{tr}\big(Q_h(A^{-1}\Omega)^{-\frac{1}{2}} A^t Q_g(\Omega)^{\frac{1}{2}} Q_g(\Omega)^{-\frac{1}{2}} Q_g(Ax) Q_g(\Omega)^{-\frac{1}{2}} Q_g(\Omega)^{\frac{1}{2}} A Q_h^{-\frac{1}{2}}(A^{-1}\Omega) - Q\big)^2 dx = \int_{A^{-1}\Omega} \mathrm{tr}\big(U_A Q_g(\Omega)^{-\frac{1}{2}} Q_g(Ax) Q_g(\Omega)^{-\frac{1}{2}} - U_A^t Q U_A\big)^2 dx =$$

writting $Q' = U_A^t Q U_A$, $y = Ax$, $dy = |\det A| dx$,

$$= |\det A^{-1}| \int_\Omega \mathrm{tr}\big(Q_g(\Omega)^{-\frac{1}{2}} Q_g(y) Q_g(\Omega)^{-\frac{1}{2}} - Q'\big)^2 dy.$$

On the other hand, since $ATV\big(A^{-1}(B)\big) = |\det A^{-1}| ATV(B)$, we see that

$$E_1\big(Q, A^{-1}(B)\big) = |\det A|^{-1} E_1\big(Q', B\big).$$

Since the map $Q \to Q' = U_A^t Q U_A$ is a rotation in the space of symmetric matrices, Q' describes the set of all positive definite matrices and the infimum is the same in both cases. □

As we have seen, the main role of Q_g was to reinforce affine invariance. But it also plays a role as a normalization factor with respect to the magnitude of the gradient.

Remark 3.1. To construct an affine invariant energy functional on the basis of the dominant orientation we also need a previous normalization of $Q_g(x)$. Thus, let for each $x \in \Omega$, $w_g(x)$ be the eigenvector of $Q_g^{-1/2} Q_g(x) Q_g^{-1/2}$ associated to its largest eigenvalue $\lambda_g(x)$. To avoid the sign indetermination implicit in the definition of the principal eigenvector, we use the map $Tw_g(x) = Re^{2i\theta}$ if $w_g(x) = Re^{i\theta}$. Then, to each image g we associate the feature vector $G \in L^2[0, 2\pi]$, where

$$G(x,\theta) = \begin{cases} 1 & \text{if } Tw_g(x) \in S_\theta \\ 0 & \text{otherwise} \end{cases} \qquad (20)$$

where S_θ is the ray of $\mathbb{R}^2$ of angle θ, i.e., $S_\theta = \{\rho e^{i\theta} : \rho \geq 0\}$. Then, given $B \in \Im$, we associate the function $u_B \in L^2[0, 2\pi]$ where

$$u_B(x,\theta) = \sum_O n(O,\theta) \chi_O(x) \qquad (21)$$

and

$$n(O,\theta) = \frac{1}{|O|} \int_O G(y,\theta) dy, \qquad \theta \in [0, 2\pi[,$$

the sum in (21) being extended to all connected components of $\Omega \setminus B$.

Using these statistics of the dominant local orientation, we define the following affine invariant energy functional

$$E_2(B) = \frac{1}{2\pi}\int_\Omega \int_0^{2\pi} \big(G(x,\theta) - u(x,\theta)\big)^2 d\theta dx + \lambda ATV(B). \tag{22}$$

Remark 3.2. If we take

$$\|w_g(x)\| = \big(\lambda_g(x)\big)^{-1/2}, \tag{23}$$

we can construct the following affine invariant functional based on the dominant local orientation

$$E_3(w,B) = \int_\Omega \|Tw_g(x) - w\|^2 dx + \lambda ATV(B), \tag{24}$$

where $B \in \Im$ and $w(x) = \sum_O w_O \chi_O(x)$, the sum being extended to all connected components of $\Omega \setminus B$ and $w_O \in I\!R^2$.

Remark 3.3. The previous functionals are related to the dominant local orientation and address the scale selection problem by using a neighborhood giving an intrinsic scale. Another possibility could be to use a multiscale representation of local orientation or statistics of them. As observed by C. Lopez and J.M. Morel ([47]) the textons characteristics indicated by Julesz are shape elements based on curvature and orientation. Considering these elements multiscale because of the *a priori* unknown size of the textons, they concluded that texton densities are nothing but multiscale curvature and orientation densities. And they proved that there is only one way to compute multiscale curvature and orientation: It is using the Affine Invariant Morphological Scale-Space, introduced and axiomatically justified in [3] (also proposed in [60]). In their experiments [47], they used multiscale curvature densities at each position and they do not take into account orientation. Now, we shall use this multiscale orientation densities and, in order to obtain the affine invariance property, we also use some normalization matrices.

From [3], let $g(t,\cdot)$ be the local, causal, morphological, affine invariant multiscale representation of $g(x)$, i.e., $g(t,x)$ is given by the unique viscosity solution of

$$g_t = |\nabla g| \left(t div \left(\frac{\nabla g}{|\nabla g|} \right) \right)^{1/3} \tag{25}$$

with initial datum $g(0,x) = g(x)$, the given image. Let us choose a discrete set of scales $t_j : j = 0,\dots,N$, with $t_0 = 0$. Moreover, we suppose that $g(t_j,\cdot) \in W^{1,2}(\Omega \setminus B)$ for some B with $\mathcal{H}^2(B) = 0$, where $\mathcal{H}^2$ denote the 2-dimensional Haussdorff measure (for that, it suffices take $g(x) \in W^{1,\infty}(\Omega)$). Hence $\nabla g(t_j,x)$ is defined almost everywhere in Ω. The vectors $\{\nabla g(t_j,x)\}_{j=0}^N$ are called the multiscale representation of local orientation of the image $g(x)$. We associate to it a vector $H \in L^2[0,2\pi]^{N+1}$. Let M_{g_j} be the matrix

$$M_{g_j}(\Omega) = \frac{1}{|\Omega|}\int_\Omega \nabla g(t_j,x) \nabla g(t_j,x)^t dx. \tag{26}$$

Observe that $M_{g_j}(\Omega)$ is a symmetric positive definite matrix, $j = 0,\dots,N$. If no confusion arises, we shall denote $M_{g_j}(\Omega)$ by M_{g_j}. We shall always assume that M_{g_j} is invertible. Otherwise, there would be no problem since $g(t_j,x)$ would be cylindrical.

Let H_j be the function

$$H_j(x,\theta) = \begin{cases} 1 & \text{if } M_{g_j}^{-1/2}\nabla g(t_j,x) \in S_\theta \\ 0 & \text{otherwise,} \end{cases}$$

and we take

$$u_{Bj}(x,\theta) = \sum_O u_{Bj}(O,\theta)\chi_O(x), \tag{27}$$

$$u_{Bj}(O,\theta) = \frac{1}{|O|}\int_O H_j(y,\theta)dy \in L^2[0,2\pi] \tag{28}$$

where S_θ is as above. Then one can construct the affine invariant functional

$$E_4(B) = \sum_{j=0}^N \frac{1}{2\pi}\int_\Omega \int_0^{2\pi} (\bar{u}_j(x,\theta) - H_j(x,\theta))^2 d\theta dx + \lambda ATV(B). \tag{29}$$

3.2. Determination of the intrinsic elliptic neighborhood

We are going to give an algorithm to construct for each x a matrix $M_g(x)$, a radius $\rho_g(x)$ and an elliptic neighborhood $\varepsilon_g\big(x,\rho_g(x)\big)$ where the histogram of g is stable -in some sense to be precised below- and satisfying (13) – (15). The algorithm is an iterative one trying to reinforce both conditions (11), (12). This will be the case whenever the iterative procedure converges. In practice, if the procedure does not converge, we stay at some stage of the iterative process and guarantee, at least, (13), (14), (15). In Figure 2 we display some intrinsic elliptic Δ-neighborhoods for some textured images used in the numerical experiments displayed in Section 6.

To proceed to the construction, let us precise our notion of histogram stability. Let $g:\Omega\to I\!R_+$ be a given image.

Definition 5. Let M be a positive definite matrix. Let $x\in\Omega$, $\rho>0$, $B(M,x,\rho)=\big\{y\in I\!R^2:\langle M(y-x),y-x\rangle\le\rho\big\}$. Let $H(g,M,\rho,\lambda)$ be the distribution function of g in $B(M,x,\rho)$; i.e.,

$$H(g,M,\rho,\lambda)=\frac{\big|\big\{y\in B(M,x,\rho):g(y)\le\lambda\big\}\big|}{|B(M,x,\rho)|},\qquad \lambda\in I\!R_+.$$

We say that $H(g,M,\rho,\cdot)$ is stable at the point x with respect to (M,ρ) if for any $\epsilon>0$, there exists $\delta=\delta(\epsilon,g,x)$ (not depending on (M,ρ)) such that for any (M',ρ') with $\|M-M'\|+|\rho-\rho'|<\delta$ we have

$$\|H(g,M',\rho',\cdot)-H(g,M,\rho,\cdot)\|_2\le\epsilon\|H(g,M,\rho,\cdot)\|_2.$$

As usual, $\|\cdot\|_2$ denotes the L_2 norm. For simplicity, we shall say that $H(g,M,\rho,\cdot)$ is stable at x with the implicit understanding of stability at x with respect to (M,ρ).

We shall need the following observation (proved in Appendix C).

Lemma 2. *Let M_n, $M\in MS_2$ be such that $M_n\to M$. Let $\rho_n,\rho>0$, $\rho_n\to\rho$. Suppose that $H(g,M_n,\rho_n,\cdot)$ is stable at the point $x\in\Omega$. Then $H(g,M,\rho,\cdot)$ is also stable at x.*

Now let us introduce the elliptic neighborhood at the point $x\in\Omega$. We suppose that $\nabla g(x)$ is defined a.e. in Ω, say in $\Omega\setminus B$ for some B with $\mathcal{H}^1(B)<+\infty$ and that $\nabla g(x)\in L^2(\Omega)$, so that the following matrix is well defined

$$Q_g^*=\int_\Omega\nabla g(x)\nabla g(x)^t dx\in MS_2.$$

For our purposes it will be harmless to assume that

$$Q_g^*\not\equiv 0.$$

Otherwise $\nabla g(x)=0$ a.e., i.e. $g(x)$ would be a piecewise constant function in $\Omega\setminus B$, case which we exclude since it could be handled with an energy functional for grey level segmentation.
For convenience, we extend $\nabla g(x)=0$, $\forall x\in I\!R^2\setminus\Omega$. Set $\rho_{0g}(x)=+\infty$, $\varepsilon_0(x;\rho_{0g}(x))=I\!R^2$. Set

$$Q_1(g,x)=\frac{1}{|\Omega|}\int_{I\!R^2}\nabla g(y)\nabla g(y)^t dy\ \Big(=\frac{1}{|\Omega|}Q_g^*\Big).$$

Set $\varepsilon_1(g,x,\rho):=\big\{y\in\Omega:\ \langle Q_1(g,x)(y-x),y-x\rangle<\rho\big\}$.
Now, we choose $\rho>0$ such that $H(g,Q_1,\rho,\cdot)$ is stable at x. Two situations are possible at a point x. Either, there exists a neighborhood N of x such that

$$\int_N\nabla g(y)\nabla g(y)^t dy=0,\qquad(30)$$

in which case we define as $N_g(x)$ as the largest such neighborhood, or for any neighborhood N of x

$$\int_N\nabla g(y)\nabla g(y)^t dy>0.\qquad(31)$$

In the first case we choose

$$\rho_g(x)=\sup\{\rho\ge0:\ \varepsilon_1(g,x,\rho)\subseteq N_g(x)\}\qquad(32)$$

and we set $\varepsilon_g\big(x,\rho_g(x)\big)=\varepsilon_1\big(g,x,\rho_g(x)\big)$. We can take $Q_1(g,x)$ as $Q_g(x)$ so that (12) holds. We shall see below that (13)-(15) also hold. It is also clear that $H(g,Q_g(x),\rho_g(x),\cdot)$ is stable at x. If we use (11) to define $Q_g(x)$, we would have $Q_g(x)\equiv 0$. In this case, (11) does not hold, but formula (11) is not essential for the developments of this case.

In the second case, to avoid intrinsic neighborhoods reduced to a point, we choose a lower scale ρ_m at the point x (this quantity $\rho_m > 0$ can be fixed constant for all x). Then, we choose

$$\rho_{1g}(x) = \inf\{\rho :\ \rho \geq \rho_m \text{ such that } H(g, Q_1(g,x), \rho, \cdot) \text{ is stable at } x\}, \tag{33}$$

with the implicit convention that the infimum is $+\infty$ when the above set is empty, i.e., when the histogram does not stabilize in the neighborhoods above. Then $\varepsilon_{1g}\big(x, \rho_{1g}(x)\big) = \varepsilon_1\big(g, x, \rho_{1g}(x)\big)$. Observe that $|\varepsilon_{1g}\big(x, \rho_{1g}(x)\big)| > 0$. Indeed, for any $\rho > 0$,

$$B\Big(x, \sqrt{\frac{\rho}{\|Q_1(g,x)\|}}\Big) \cap \Omega \subseteq \varepsilon_1(g, x, \rho). \tag{34}$$

For that, let $y \in \Omega, \|y - x\| < \sqrt{\frac{\rho}{\|Q_1(g,x)\|}}$. Then

$$\langle Q_1(g,x)(y-x), y-x\rangle \leq \|Q_1(g,x)\|\|y-x\|^2 < \rho \tag{35}$$

i.e., $y \in \varepsilon_1(g, x, \rho)$. This implies that

$$B\Big(x, \sqrt{\frac{\rho_m}{\|Q_1(g,x)\|}}\Big) \cap \Omega \subseteq \varepsilon_1\big(g, x, \rho_{1g}(x)\big). \tag{36}$$

To complete the construction in the second case, suppose that $\rho_{1g}(x)$ has been defined by (34). We set

$$Q_2(g,x) = \frac{1}{|\varepsilon_{1g}(x, \rho_{1g}(x))|} \int_{\varepsilon_{1g}(x, \rho_{1g}(x))} \nabla g(y) \nabla g(y)^t dy.$$

Let $\varepsilon_2(g, x, \rho) = \big\{y \in \Omega :\ \langle Q_2(g,x)(y-x), y - x\rangle \leq \rho\big\}$ and define

$$\rho_{2g}(x) = \inf\big\{\rho : \rho \geq \rho_m \text{ such that } H(g, Q_2(g,x), \rho, \cdot) \text{ is stable at } x\big\}.$$

As above one checks that

$$B\Big(x, \sqrt{\frac{\pi}{|\Omega|}\big(\frac{\rho_m}{\|Q_1(g,x)\|}\big)^2}\Big) \cap \Omega \subseteq \varepsilon_2(g; x, \rho_{2g}(x)).$$

We set $\varepsilon_{2g}\big(x; \rho_{2g}(x)\big) = \varepsilon_2\big(g; x, \rho_{2g}(x)\big)$.
Suppose that we already have $\rho_{n-1g}(x)$, $\varepsilon_{n-1g}\big(x; \rho_{n-1g}(x)\big)$, $Q_{n-1}(g,x)$ such that $H\big(g, Q_{n-1}(g,x), \rho_{n-1g}(x), \cdot\big)$ is stable at x and we have

$$B\Big(x, \sqrt{\big(\frac{\pi}{|\Omega|}\big)^{n-2}\big(\frac{\rho_m}{\|Q_1(g,x)\|}\big)^{n-1}}\Big) \cap \Omega \subseteq \varepsilon_{n-1g}(x, \rho_{n-1g}(x)).$$

Then, we define

$$Q_n(g,x) = \frac{1}{|\varepsilon_{n-1g}(x, \rho_{n-1g}(x))|} \int_{\varepsilon_{n-1g}(x, \rho_{n-1g}(x))} \nabla g(y) \nabla g(y)^t dy \tag{37}$$

$$\varepsilon_n(g, x, \rho) = \big\{y \in \Omega :\ \langle Q_n(g,x)(y-x), y-x\rangle \leq \rho\big\} \tag{38}$$

$$\rho_{ng}(x) = \inf\big\{\rho : \rho \geq \rho_m \text{ such that } H(g, Q_n(g,x), \rho, \cdot) \text{ is stable at } x\big\}. \tag{39}$$

Then, set $\varepsilon_{ng}\big(x, \rho_{ng}(x)\big) = \varepsilon_n\big(g, x, \rho_{ng}(x)\big)$ and check that

$$B\Big(x, \sqrt{\big(\frac{\pi}{|\Omega|}\big)^{n-1}\big(\frac{\rho_m}{\|Q_1(g,x)\|}\big)^{n}}\Big) \cap \Omega \subseteq \varepsilon_{ng}\big(x, \rho_{ng}(x)\big). \tag{40}$$

By definition, (12) holds. Let us check (13), (14), (15). Let us start by the degenerate case where (30) holds. For that, let A be a linear map such that $\det A \neq 0$ and $h : A^{-1}\Omega \to I\!R$, $h(x) = g(Ax)$. Let $x \in A^{-1}\Omega, \bar{x} = Ax \in \Omega$. Let $N_g(\bar{x}), N_h(x)$ be as above. Then, for any neighborhood N of $\bar{x}$

$$\begin{aligned}
&\int_{A^{-1}N} \nabla h(y) \nabla h(y)^t dy \\
&= A^t \int_{A^{-1}N} \nabla g(Ay) \nabla g(Ay)^t dy\, A \\
&= \frac{A^t}{|\det A|} \int_N \nabla g(z) \nabla g(z)^t dz\, A,
\end{aligned}$$

which implies that $N_h(x) = A^{-1} N_g(\bar{x})$. On the other hand

$$Q_h(x) = \frac{1}{|\det A^{-1}\Omega|} \int_{A^{-1}\Omega} \nabla h(y) \nabla h(y)^t dy$$

$$= \frac{A^t}{|\det A^{-1}\Omega|}\int_{A^{-1}\Omega}\nabla g(Ay)\nabla g(Ay)^t dy\, A$$
$$= \frac{A^t}{|\Omega|}\int_{\Omega}\nabla g(z)\nabla g(z)^t dz\, A = A^t Q_g(\bar{x})A,$$
$$\varepsilon_1(h,x,\rho) = A^{-1}\varepsilon_1(g,\bar{x},\rho) \quad \text{and}$$
$$\rho_h(x) = \sup\{\rho \geq 0 :\ \varepsilon(h,x,\rho) \subseteq N_h(x)\} =$$
$$\sup\{\rho \geq 0 :\ A^{-1}\varepsilon_1(h,\bar{x},\rho) \subseteq A^{-1}N_g(\bar{x})\} = \rho_g(\bar{x}).$$

To consider the nondegenerate case, we observe the following simple lemma (proved in Appendix D).

Lemma 3. *Let $M, M' \in MS_2$ be such that $M' = A^t M A$ for some linear map A with $\det A \neq 0$. Let $g : \Omega \to \mathbb{R}_+$, $h : A^{-1}\Omega \to \mathbb{R}_+$ be such that $h(x) = g(Ax)$. Let $x \in A^{-1}\Omega, \bar{x} = Ax$. Let $\rho > 0$ and $\varepsilon(\bar{x}) \equiv \{y \in \Omega :\ \langle M(y-\bar{x}), y-\bar{x}\rangle \leq \rho\}$, $\varepsilon'(x) \equiv \{z \in A^{-1}\Omega :\ \langle M'(z-x), z-x\rangle \leq \rho\}$ $\big(= A^{-1}\varepsilon(\bar{x})\big)$. Then, $H(g,M,\rho,\cdot)$ is stable at $\bar{x}$ if and only if $H(h,M',\rho,\cdot)$ is stable at x.*

With this lemma it is immediate to see that for each n

$$Q_{nh}(x) = A^t Q_{ng}(Ax)A \tag{41}$$
$$\varepsilon_{nh}\big(x,\rho_{nh}(x)\big) = A^{-1}\varepsilon_{ng}\big(Ax,\rho_{ng}(Ax)\big) \tag{42}$$
$$\rho_{nh}(x) = \rho_{ng}(Ax) \tag{43}$$
$$H\big(g,Q_{ng}(Ax),\rho_{ng}(Ax),\cdot\big) \text{ is stable at } Ax \text{ and consequently} \tag{44}$$
$$H\big(h,Q_{nh}(x),\rho_{nh}(x),\cdot\big) \text{ is stable at } x \tag{45}$$

where $Q_{ng}(Ax), Q_{nh}(x), \ldots$ are defined as in the previous construction.

Let us make some comments on the question of convergence of the sequence above. Under the assumption that $\|\nabla g\|_\infty = \sup_{x\in\Omega}|\nabla g(x)| < +\infty$, we have that

$$0 \leq Q_{ng}(x) \leq \|\nabla g\|_\infty^2 I,$$

which implies the precompactness of the set of matrices $\big\{Q_{ng}(x)\big\}_{n=1}^{\infty}$. But the question of convergence is still there. Convergence is an assumption which must be experimentally checked. If $Q_{ng}(x), \rho_{ng}(x), \varepsilon_{ng}(x,\rho_{ng}(x))$ converge to $Q_g(x), \rho_g(x), \varepsilon_g(x,\rho_g(x))$ then (11)-(15) are a consequence of (37), (38), (41), (42), (43), (44) and Lemma 3.

4. Texture segmentation using moments

Moments and functions of moments have since long been used as features in a number of applications to pattern recognition. Our ability to recognize objects independently of their euclidean motion in $3D$ space forces us to design methods for pattern recognition which are invariant under affine transformations. This has been attempted in several ways. Hu [37] used nonlinear combinations of usual geometric moments to construct a set of invariants. Later, moments based on bases of orthogonal polynomials were introduced with the same purpose [63], [9], [42]. Abu-Mustapha and Psaltis [1] introduced the notion of complex moments as a simple way to derive moment invariants. In a few words, equivalent sets of moment invariants can be defined in terms of Zernike polynomials, pseudo-Zernike polynomials, rotational or complex moments. On one hand, since they are bases of orthogonal polynomials, most of the image information can be recaptured by using a sufficiently large number of image moments. But, on the other hand, the higher order moments are more sensitive to noise. A detailed comparative study of several generating sets of polynomials to compute moments and their sensibility to noise can be seen in [62]. The sensivity to noise increases with the order but high orders are necessary for a better representation of the image. In this section, our purpose is to use the method of moments to compute locally at the intrinsic elliptic Δ-neighborhood of each point, a set of affine invariant features which will be used as channels in an affine invariant Mumford-Shah type segmentation functional for texture segmentation. In order to do that, we give a description of Zernike moments and their properties briefly summarized. The reason for selecting them from among the other othogonal moments is that they possess a useful rotation invariance property. On the other hand, as they are defined into the unit circle, it is easy, as we can see below, to define them on elliptics neighborhoods.

4.1. Zernike moments

Zernike polynomials were first introduced by Zernike in 1934. In [17] we can see the deriva-

tion of them from the requirement of orthogonality and invariance properties. The different bases of moments are in principle equally suited to represent functions but the orthogonal moments are better than others in terms of information redundacy ([62]).

The complex Zernike moment of order n with repetition l for a continuous image function $f(x,y)$ that vanishes outside the unit circle is defined by:

$$A_{nl} = \frac{n+1}{\pi} \int_0^{2\pi} \int_0^{\infty} \overline{V_{nl}(r,\theta)} f(r\cos\theta, r\sin\theta) r dr d\theta, \tag{46}$$

where $n \geq 0$, $l \in \mathbb{Z}$, $|l| \leq n$ and $n - |l|$ even. By $\overline{V_{nl}(r,\theta)}$ we denote the complex conjugate.

The complex Zernike polynomials are defined by

$$V_{nl}(x,y) = V_{nl}(r,\theta) = R_{nl}(r) e^{il\theta}, \qquad x^2 + y^2 \leq 1$$

with

$$R_{nl}(r) = \sum_{s=0}^{\frac{n-|l|}{2}} \frac{(-1)^s (n-s)!}{s! \left(\frac{n+|l|}{2} - s\right)! \left(\frac{n-|l|}{2} - s\right)!} r^{n-2s} = \sum_{\substack{k=|l| \\ n-k \ even}}^{n} B_{n|l|k} r^k.$$

The $R_{nl}(r)$ functions, called radial polynomials, satisfy the relations

$$< R_{nl}, R_{ml} >= \frac{1}{2(n+1)} \delta_{mn}, \tag{47}$$

and $\{V_{mn}\}$ is an orthogonal set. Indeed,

$$\int_0^{2\pi} \int_0^1 \overline{V_{nl}(r,\theta)} V_{mk}(r,\theta) r dr d\theta = \frac{\pi}{n+1} \delta_{mn} \delta_{kl}, \tag{48}$$

with $\delta_{mn} = 1$ if $m = n$ and $\delta_{mn} = 0$ otherwise. Then Zernike moments are the projection of the image function f onto these basis functions and they constitute an orthogonal basis of $L^2(x^2+y^2 \leq 1)$. Moreover, observe that $\overline{A_{nl}} = A_{n,-l}$ and that they are invariant under rotations.

4.2. An energy functional based on local moment features

We are going now to construct a set of affine invariant features based on Zernike moments.

As above, let $g : \Omega \to \mathbb{R}_+$ be a given image. At each point $x \in \Omega$, we consider the elliptic neighborhood $\varepsilon_g(x, \rho_g(x))$ defined in 3.1 and redefined in terms of the matrix Q_g by

$$\varepsilon_g(x, \rho_g(x)) := \{y \in \Omega :< Q_g(y-x), y-x > \leq \rho_g(x)\}.$$

For simplicity we write $N_g(x) = \rho_g(x)^{-1} Q_g$. Then

$$\varepsilon_g(x, \rho_g(x)) = \{y \in \Omega :< N_g(x)^{1/2}(y-x), N_g(x)^{1/2}(y-x) >\leq 1\}.$$

Recall that if A is a linear map with $\det A \neq 0$ and $h : A^{-1}\Omega \to \mathbb{R}_+$ is given by $h(x) = g(Ax)$, we have

$$N_h(x) = A^t N_g(Ax) A$$
$$\rho_h(x) = \rho_g(Ax) \tag{49}$$
$$\varepsilon_h(x, \rho_h(x)) = A^{-1} \varepsilon_g(Ax, \rho_g(Ax)). \tag{50}$$

Since

$$\varepsilon_g(x, \rho_g(x)) = x + N_g(x)^{-1/2} D(0,1),$$

where $D(0,1) = \{z \in \mathbb{R}^2 : ||z|| \leq 1\}$, we can exploit this identity to define the following moments:

$$m_{nl}(g,x) = \int_{||z|| \leq 1} V_{nl}(z) g(x + N_g^{-1/2}(x) z) dz, \tag{51}$$

where $z = (z_1, z_2)$, $x = (x_1, x_2) \in \Omega$ and $V_{nl}(z)$ are the Zernike polynomials. Let us check the formula

$$m_{nl}(h,x) = m_{nl}(g, Ax) e^{il\alpha}, \qquad x \in A^{-1}\Omega \tag{52}$$

for some α related to g, A (hence to h). Indeed,

$$\begin{aligned} m_{nl}(h,x) &= \int_{||z|| \leq 1} V_{nl}(z) h(x + N_h^{-1/2}(x) z) dz \\ &= \int_{||z|| \leq 1} V_{nl}(z) g(Ax + A N_h^{-1/2}(x) z) dz \\ &= \int_{||z|| \leq 1} V_{nl}(z) g(Ax + N_g^{-1/2}(Ax) \\ &\quad N_g^{1/2}(Ax) A N_h^{-1/2}(x) z) dz. \end{aligned}$$

Since, by Fact 0 (see Subsection 3.1), $U_A = N_g^{1/2}(Ax)AN_h^{-1/2}(x) \in O(2)$, letting $z' = U_A z$, $z = U_A^t z'$, $dz = dz'$ this last quantity is equal to

$$\int_{||z'||\leq 1} V_{nl}(U_A^t z')g(Ax + N_g^{-1/2}(Ax)z')dz'.$$

By the properties of the Zernike polynomials,

$$V_{nl}(U_A^t z') = R_{nl}(|U_A^t z'|)e^{il(\theta+\alpha)} = R_{nl}(|z'|)e^{il\theta}e^{il\alpha} = V_{nl}(z')e^{il\alpha},$$

where α denotes the angle associated with the rotation U_A^t.
Hence

$$m_{nl}(h,x) = \int_{\|z'\|\leq 1} V_{nl}(z')g(Ax + N_g^{-\frac{1}{2}}(Ax)z')dz' \cdot e^{il\alpha} = m_{nl}(g, Ax)e^{il\alpha}.$$

We use the moments m_{nl} to define the following functional

$$E_Z(\{m_{nl}\}, B) = \sum_{n,l}\int_\Omega \|m_{nl}(g,x) - m_{nl}\|^2 dx + \lambda ATV(B), \tag{53}$$

where $m_{nl}(g,x)$ are defined as in (51), $B \in \zeta$ and

$$m_{nl} = \sum_O m_{nl}(O)\chi_O,$$

$m_{nl}(O) \in \mathbb{C}$, $n \geq 0$, $l \in \mathbb{Z}$, $|l| \leq n$, $n - |l|$ even, the sum being extended to all connected components O of $\Omega \setminus B$.
Using (52) it is inmediate to see that the Functional (53) is affine invariant. The normalization with the matrix $N_g(x)$ is needed to use the Zernike moments on the elliptic neighborhood of each point and it gives for free the affine invariance of the energy functional.

5. Applications to shape from texture

The image of a slanted textured surface provides important information about the shape and orientation of the surface. This fact has attracted considerable attention in the last 40 years. Perhaps the main impetus for this interest was the hypothesis formulated by Gibson [32], which states that texture is a mathematically and psychologically sufficient stimulus for surface perception. Based on this observation, many quantitative methods for recovery of surface orientation from projective distorsion of texture have been proposed ([2], [5], [41], [14], etc).

Witkin [67] pointed out that the foreshortening effect, i.e. the compression of a slanted pattern in the direction of slant can also be a cue to surface orientation. Indeed, the image of an slanted circle is an ellipse and the orientation and eccentricity of the ellipse indicate the magnitude and direction of the slant. This effect is caused by the angle between the line of sight and the surface in the scene. Hence, the foreshortening effect can also be observed in orthographical projection of a planar surface pattern. Various developpements and extensions of Witkin's method have been studied ([24], [40], [14], [28], [33], etc.).

To understand how a local texture description can be interpreted in terms of three-dimensional surface shape, let us review the surface and viewing geometry. For a better description of the assumptions involved, we follow the formulation of Garding [30], Garding-Lindeberg [33].

Consider a smooth surface S viewed in perspective projection. The smooth surface S is mapped by central projection onto the unit sphere Σ centered at the focal point which is taken as the origin. This spherical projection model is equivalent to the ordinary perspective projection onto a flat image plane in the sense that if one of these projections is known the other can be computed from it. Recall that the local perspective distorsion of the projected surface pattern results from two factors. First the distance and orientation of the surface with respect to the line of sight. Second, the angle between the line of sight and the image surface. This second effect is known as the "position effect". This effect depends only on the camera geometry and it can be eliminated. The main advantage of the spherical projection model is that it treats all parts of the visual field equally and it avoids the artificial position effect.

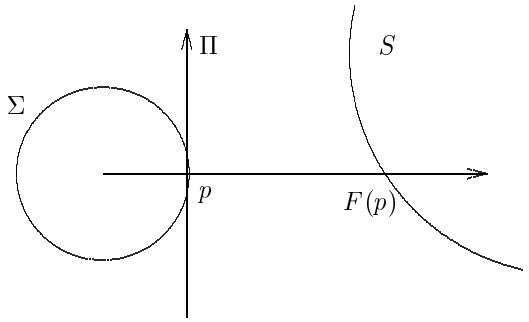

Fig. 1. Local surface geometry and imaging model. Let S be a smooth surface, Σ the unit sphere, Π the tangent plane to S at some point p in S and $F : \Sigma \to$ the perspective backprojection.

Fortunately, in order to estimate the local surface shape from texture, it suffices to consider the first order terms of the perspective projection at each image point. Following Garding [30], [33] we consider a spherical camera mapping a smooth surface S onto a unit sphere Σ and let $F : \Sigma \to S$ be the perspective backprojection from Σ to S (Figure 1 illustrates the basic geometry). Then $F(p) = d(p)p$, $p \in \Sigma$, where $d(p)$ is the distance of p to the surface S along the direction p. At each point $p \in \Sigma$, we consider a local coordinate system $\{p, t, b\}$ defined by: p is the unit vector in the view direction, t is parallel to $\nabla d(p)$ (assuming $\nabla d(p) \neq 0$) and $b = p \times t$. Let $T_p\Sigma, T_{F(p)}S$ be the tangent planes at Σ at the point p and S at the point $F(p)$ respectively. Then the tangent map F_{*p}

$$F_{*p} : T_p\Sigma \to T_{F(p)}S$$

Let $\vec{T} = \frac{F_{*p}(t)}{\|F_{*p}(t)\|}$, $\vec{B} = \frac{F_{*p}(b)}{\|F_{*p}(b)\|}$.
The matrix of the map F_{*p} written on the basis $\{t, b\}, \{\vec{T}, \vec{B}\}$ is

$$F_{*p} = \begin{pmatrix} r/\cos\sigma & 0 \\ 0 & r \end{pmatrix} = \begin{pmatrix} 1/m & 0 \\ 0 & 1/M \end{pmatrix} \qquad (54)$$

where $r = d(p)$ and σ is the slant of the surface, i.e., the angle between the viewing direction and the surface normal. The characteristic values m, M are the inverse eigenvalues of F_{*p}. They describe how a unit circle in $T_{F(p)}S$ is mapped to $T_p\Sigma$ by F_{*p}^{-1}. It becomes an ellipse with m as minor axis in the direction of t and M as major axis in the direction of b. We see how several useful relations between local perspective distorsion and surface shape can be derived: surface orientation is directly related to (m, M) and its eigenvectors (t, b). The tilt direction, defined as the gradient of the distance from Σ to the surface along the wiewing ray is parallel to t, the eigenvector associated to the smallest eigenvalue of F_{*p}^{-1}, m. The foreshortening is defined by the ratio $\frac{m}{M}$ and is related to surface slant σ by the relation $\cos\sigma = \frac{m}{M}$. Together, slant and tilt determine surface orientation (up to the sign of tilt, since both t and $-t$ correspond to the eigenvalue $1/m$). Other texture gradients can be computed from the spatial rate of change of certain measures derived from F_{*p} (see [30], [33]).

If the brightness data are available in a planar image Π and the internal camera geometry is known, then we can compute the mapping $G : \Pi \to \Sigma$ from the image plane Π to the view sphere Σ. Then $A = F \circ G : \Pi \to S$ maps the image plane onto S and if $x \in \Pi$, $p = G(x) \in \Sigma$, then

$$A_{*x} = F_{*p} \circ G_{*x}$$

If G_{*x} is known, $F_{*p} = A_{*x} \circ G_{*x}^{-1}$.

The main tool used in the analysis will be the affine intrinsic orientation matrix introduced in Section 3, in the same sense that in [33], where they use the second moment matrix. The inaccuracies inherent in the measuring process and in the scale selection are corrected in some sense by using the intrinsic elliptic Δ-neighborhood. Let us mention that different second order descriptors have been used for similar purposes ([15], [29], [58], [40], [13], [30]). If $h : I\!R^2 \to I\!R_+$, the second order matrix is defined by

$$\mu_h(x) = \int_{I\!R^2} w(x-y)\nabla h(y)\nabla h(y)^t dy, \qquad x \in I\!R^2, \qquad (55)$$

for some regularizing window $w \geq 0$, $\int_{I\!R^2} w = 1$. Let g be an image on the view sphere Σ, i.e., $g : \Sigma \to I\!R_+$. Then

$$\begin{array}{ccc} & h = g \circ G & \\ G & & g \\ \Pi & \to \Sigma \to & I\!R_+ \end{array}$$

is the corresponding image defined on the image plane Π. Let $x \in \Pi$ and $p = G(x)$. Then $\mu_h(x)$ is

given by (55). Observe that for $y \in \Pi$, $\nabla h(y) = g_{*G(y)} \circ G_{*y} = G_{*y}^t \nabla g\big(G(y)\big)$. Hence

$$\mu_h(x) = \int_{I\!R^2} w(x-y) G_{*y}^t \nabla g\big(G(y)\big) \nabla g\big(G(y)\big)^t G_{*y} dy.$$

In a first approximation, $G_{*y} = G_{*x}$, $z = G(y) = p + G_{*x}(y-x)$ for y near x (which is reasonable if $w = G_\sigma$ with $\sigma \to 0+$) and we may write

$$\mu_h(x) = \frac{G_{*x}^t}{|G_{*x}|} \int_{T_p\Sigma} w\big(G_{*x}^{-1}(p-z)\big) \nabla g(z) \nabla g(z)^t dz G_{*x}.$$

Writting $w'(u) = \frac{1}{|G_{*x}|} w\big(G_{*x}^{-1}(u)\big)$ and

$$\mu_g(p) = \int_{T_p\Sigma} w'(p-z) \nabla g(z) \nabla g(z)^t dz, \qquad (56)$$

we have

$$\mu_h(x) = G_{*x}^t \, \mu_g(p) \, G_{*x} \qquad (57)$$

which holds as a first order approximation. This could be made rigorous by using the exponential map on a neighborhood of $T_p\Sigma$ and some ϵ argument. A similar formula can be derived from

$$\begin{array}{c} g = f \circ F \\ F \quad\quad f \\ \Sigma \to S \to I\!R \\ p \to F(p) \end{array}$$

(we assume that the image brightness on Σ is proportional to the surface reflectance). In this case

$$\mu_g(p) = F_{*p}^t \mu_f\big(F(p)\big) F_{*p} \qquad (58)$$

where

$$\begin{aligned} \mu_f\big(F(p)\big) &= \int_{T_{F(p)}S} w''\big(F(p) - z'\big) \\ &\quad \nabla f(z') \nabla f(z')^t dz' \qquad (59) \\ w''(u) &= \frac{1}{|F_{*p}|} w'\big(F_{*p}^{-1}(u)\big). \end{aligned}$$

Combining (57) and (58) we see that to estimate the slant and tilt parameters contained in F_{*p} we need to estimate $\mu_h(x)$ in the image plane and some assumptions about the surface reflectance pattern reflected in $\mu_f\big(F(p)\big)$ in order to infer the structure of F_{*p}.

A simple but useful assumption is to suppose that μ_f is proportional to the identity matrix

$$\mu_f = cI \qquad \text{for some} \, c > 0.$$

This is like saying that the surface texture is isotropic, i.e., there is no single direction on the surface texture. With this assumption (58) writes

$$\mu_g(p) = c F_{*p}^t F_{*p} \, .$$

This means that the eigenvalues of F_{*p} are proportional to the square roots of the eigenvalues λ_1, λ_2 of $\mu_g(p)$ and the eigenvectors are the same, i.e.,

$$m \sim \frac{1}{\sqrt{\lambda_1}} \qquad M \sim \frac{1}{\sqrt{\lambda_2}} \, .$$

As discussed above, the tilt direction coincides with the eigenvector associated to the maximum eigenvalue λ_1 and the slant is given by

$$\cos\sigma = \frac{m}{M} \, .$$

A further assumption is required to exploit formula (57). Thus, we assume that the focal distance tends to $+\infty$ and we may assume that $G_{*x} = I$. Hence

$$\mu_h(x) = \mu_g(p)$$

and, under the previous assumptions, we are able to extract from the second order matrix useful information concerning the shape of the surface S. Of course, in many cases these assumptions are violated and other methods are required (see [33]).

With this excursion in the shape from texture analysis and the developpements in the previous sections, it is not surprising if we use the functional

$$E(\mu, B) = \int_\Omega \mathrm{tr}\big(\mu_g^{-1/2} \mu_g(x) \mu_g^{-1/2} - \mu\big)^2 dx + \lambda ATV(B), \qquad (60)$$

with purposes of computing the slant and tilt parameters from an image $g(x)$ where $\mu_g(x)$ is the second order matrix given by (55), μ_g being a normalization matrix given, for instance, by

$$\mu_g = \frac{1}{|\Omega|} \int_\Omega \nabla g(x) \nabla g(x)^t dx.$$

In practice, we use as matrix $\mu_g(x)$ the particular matrix

$$\mu_g(x) = Q_g(x) = \frac{1}{|\varepsilon_g(x)|}\int_{\varepsilon_g(x)} \nabla g(y)\nabla g(y)^t dy,$$

which is related to the particular length scale given by the intrinsic elliptic Δ-neighborhood. Summarizing our strategy:

1. We compute the matrix $\mu_g = \frac{1}{|\Omega|}\int_\Omega \mu_g(x)$.
2. We minimize $E(\mu, B)$ on the set of matrix functions $\mu(x) = \sum_O \mu_O \chi_O(x)$ with $\mu_O \in MS_2$, $B \in \Im$, and the sum in $\mu(x)$ is extended to all connected components of $\Omega \setminus B$.
4. We use the matrix functions $\mu_g^{1/2}\mu\mu_g^{1/2}$ to compute the slant and tilt at each region O of the segmentation as described above. The experiments displayed in Section 6 show that this strategy permits to recover in an efficient way surface information.

6. Numerical analysis and experiments

This section is devoted to give the numerical analysis and to display several experiments. The textured images used in the experiments are, most of them, compositions of natural textures: Figure 3 and Figure 8 show a composition of three differents wallpaper textures. Figure 13 displays a composition of three different types of cork. Figure 12 has been obtained by photographying a pullover (with just one texture) on a pyramid with four sides. Finally, Figure 5 is a composition of a piece of cheetah leather on a field of grass.

All input images are gray level images of 128×128 or 256×256 pixels size with eight bits by pixel. The results have been obtained using software written in C programming lenguage on the UNIX operating system running on a IPC SUN workstation and we have used functions of the software MEGAWAVE (author: J. Froment). For a given image, the output of the algorithm consists of the segmentation boundaries and the superposition of the original image together with the segmentation boundaries. The segmentation plots shows how the underlying segmentation fits with the real regions of the texture image, both, alone and superposed to the original image.

The original image and the number of desired regions (or, equivalently, the final scale parameter) are the only variable inputs for the algorithm and each functional. The experiments displayed correspond, in general, to an initial partition of the image in regions of size 1×1 pixels. One can also take a grid of 2×2 or 4×4 pixels size, which makes the algorithm faster.

In the numerical approach, a segmentation, B, consists of a finite union of piecewise affine Jordan curves with disjoint interiors together with a piecewise constant vector function, noted $u^B : \Omega \to I\!R^p$, $p \in I\!N$, such that $u^B(x) = \sum_{O \subset \Omega\setminus B} u^B(O)\chi_O(x)$ with $u^B(O) = (u_1^B(O), \ldots, u_p^B(O))$, $u_i^B(O) \in I\!R$ obtained from G_i where $G = (G_1, \ldots, G_p)$ is the feature vector. We want to find numerically the pair $(\bar{u}^{\bar{B}}, \bar{B})$ where the minimum (in a certain sense) of each energy functional is attained.

We can express the energy functional (2) by

$$E(u^B, B) = \sum_{i=1}^{p}\sum_{k=1}^{R}\int_{O_k}(u_i^B(O_k) - G_i(x))^2 dx + \lambda ATV(B) \tag{61}$$

where:

- B is the set of piecewise affine Jordan curves with disjoint interiors that defines the discrete segmentation.
- $\{O_k\}_k$ denotes the set of connected components or regions of $\Omega \setminus B$.
- R is the number of regions.
- G is the feature vector.
- $u_i^B(O_k)$ is the mean value of G_i in the region O_k.
- $ATV(B)$ denotes the Affine Total Variation of B. Since B is made of a finite union of piecewise affine curves $\{, _l\}$ with disjoint interiors, we may write

$$B = \bigcup_{finite} , _l.$$

Then, we may express the second term in the functional as

$$ATV(B) = \sum_{\Gamma_l, \Gamma_p \in B} Inter(, _l, , _p).$$

Our plan, in this section, is the following. The discretization of the ATV term of the functional (61) is explained in subsection 6.1 and, in order to describe the discretization of the feature term, we detail in 6.2 the algorithm for the elliptic Δ-neighborhood construction (defined in subsection 3.2). Subsection 6.3 is devoted to the quantization of the feature vectors corresponding to the functional E_1, described in Section 2 (dominant local orientation), and the functional, E_Z, based on Zernike moments (described in Section 4). To finish subsection 6.3, we display the shape from texture experiments. Finally, in subsection 6.4, we describe the segmentation algorithm and the data structure used in the numerical experiments.

6.1. *Discrete approximation of the ATV term*

Let us recall the simple numerical scheme to compute the ATV of a set of discrete curves introduced in [6]. The ATV concept is based on the Interaction between two curves $, _1, , _2$ which is defined by

$$Inter(, _1, , _2) = \int_{\Gamma_1} \int_{\Gamma_2} |\tau_1(x) \wedge \tau_2(y)| d\sigma_1(x) d\sigma_2(y). \tag{62}$$

Discretizing this term may lead to expensive computations and, in any case, to an accumulation of errors. For that reason and taking into account that the curves, as boundaries of regions which are made of pixels, are polygonal curves made of small horizontal and vertical segments we decide to favour a finite number of fixed directions and we count the number of those segments folding over each one of these directions.

Thus, let $d_1, \cdots, d_n$ be a finite number of directions in the plane. Each direction d_i may be represented as a line segment with specified magnitude and direction. Then, we compute the number m_i of ocurrences of the direction d_i in the boundary of a region. Hence, each boundary $,$ will be determined by n integer values, $m_1, \cdots, m_n$. The numerical values of the magnitude $|d_i|$, $1 \leq i \leq n$, depend on the grid size we consider in the initial segmentation. That is, we initialize the region growing algorithm with a initial segmentation which is (see Subsection 6.4), simply, a grid on the image we want to segment, whose initial regions are the cells of the grid, small squares of size $D \times D$ (with $D \in I\!N$, for example, $D = 1, 2$ or 4). Then, in the numerical experiments displayed in the sequel we have used the next eight directions:

$$\begin{array}{ll} d_1 = D(1,0), & d_2 = D(0,1), \\ d_3 = D(1,1), & d_4 = D(-1,1), \\ d_5 = D(1,2), & d_6 = D(-1,2), \\ d_7 = D(2,1), & d_8 = D(-2,1), \end{array} \tag{63}$$

where $D(x, y)$ represents the product of D by the vector (x, y).

Now, let $, _1$ and $, _2$ be any two curves and let $d_1, \cdots, d_n$ be the given directions. According to (62), we define the discrete Interaction of $, _1$ and $, _2$ by

$$Inter_d(, _1, , _2) = \sum_{i=1}^{n} \sum_{j=1}^{n} m_i^1 m_j^2 |d_i||d_j| |\sin(d_i, d_j)|,$$

where $\{m_i^1\}_i$ and $\{m_i^2\}_i$ are the values corresponding to $, _1$ and $, _2$, respectively.

6.2. *The elliptic Δ-neighborhood construction*

In [39], Julesz defined the Δ-neighborhood as the neighborhood in which differences in texton densities (numbers) are determined and texton gradients computed, but he did not indicate how can they be computed from grey level images of natural scenes. We proposed in subsection 3.2 to construct what we called "the intrinsic elliptic Δ-neighborhood" at each point of the image domain. It is the appropriate neighborhood on which we compute the local texton features. We describe, in this subsection, an algorithm to compute the intrinsic elliptic Δ-neighborhood. The numerical experiments that we display in the next subsections show that, with texture features extracted from this neighborhood, one retrieves faithfully the homogeneous texture regions and recovers the shape from texture information in natural scenes.

Following the construction in subsection 3.2, the numerical algorithm which gives an elliptic neighborhood of each pixel where the histogram of the discrete image is stable and satisfies (13)-(15), is an iterative algorithm trying to reinforce both conditions (11), (12). Related to the notion of histogram stability given in subsection 3.2 −see Definition 3.2−, let us precise the discrete notion we shall use. Le M be a positive definite matrix. Let $x \in \Omega$. Let us denote for each $\rho > 0$

$$\varepsilon(x,\rho) = \{y \in I\!R^2 : < M(y-x), y-x > \leq \rho\}.$$

Let us note by $H(\varepsilon(x,\rho),\cdot)$ the distribution function of the given image in $\varepsilon(x,\rho)$. Let $\rho > 0$ be fixed. We say that $H(\varepsilon(x,\rho),\cdot)$ *is stable at the pixel* x *(with respect to* $\varepsilon(x,\rho)$*)* or that $\varepsilon(x,\rho)$ *verifies the Histogram Stabilization Criterion -HSC* in short- if for any $\epsilon > 0$, there exists $\delta = \delta(\epsilon, x)$ such that for any $\rho' > 0$ with $\mu(\varepsilon(x,\rho)\Delta\varepsilon(x,\rho')) < \delta$ we have

$$\sum_{\xi=0}^{255} \left(H(\varepsilon(x,\rho'),\xi) - H(\varepsilon(x,\rho),\xi)\right)^2 \leq \epsilon \sum_{\xi=0}^{255} H(\varepsilon(x,\rho),\xi)^2.$$

In this case, we say that $HSC(\varepsilon(x,\rho),\varepsilon(x,\rho'))$ is satisfied.

Let us describe the discrete iterative process giving the elliptic neighborhood at a pixel $x \in \Omega$. Let $x \in \Omega$ and $\epsilon > 0$. Let $\rho_m > 0$ be fixed representing the lower scale −see 3.2 −. We start at step 1 of the iterative process with a disk with radius equal to ρ_m. More precisely:

Step 1. Set

$$Q_1(x) = \begin{pmatrix} \rho_m^{-1} & 0 \\ 0 & \rho_m^{-1} \end{pmatrix}$$

and $\varepsilon_1(x,\rho) = \{y \in \Omega : < Q_1(x)(y-x), y-x > \leq \rho\}$, $\rho > 0$.

Now, we choose −see details below−

$$\rho_1(x) = \inf\{\rho :\ \rho \geq \rho_m \,\text{such that}\, H(\varepsilon_1(x,\rho),\cdot) \,\text{is stable at}\, x\}.$$

Finally, take $\varepsilon_{1g}(x,\rho_1(x)) = \varepsilon_1(x,\rho_1(x))$ as the elliptic neighborhood at the point x corresponding to step 1.

Step n **(**$n \geq 2$**).** First, compute

$$Q_n(x) = \frac{1}{|\varepsilon_{n-1,g}(x,\rho_{n-1}(x))|} \int_{\varepsilon_{n-1,g}(x,\rho_{n-1}(x))} \nabla g(y) \nabla g(y)^t dy.$$

In fact, we compute an approximation to the matrix $Q_n(x)$, called again $Q_n(x)$, given by

$$Q_n(x) = \frac{1}{Card(\varepsilon_{n-1,g}(x,\rho_{n-1}(x)))} \sum_{y \in \varepsilon_{n-1,g}(x,\rho_{n-1}(x))} \nabla g(y) \nabla g(y)^t,$$

where the discrete approach to the gradient of g is computed by using finite differences.

Now, set $\varepsilon_n(x,\rho) = \{y \in \Omega : < Q_n(x)(y-x), y-x > \leq \rho\}$, $\rho > 0$. Then choose

$$\rho_n(x) = \inf\{\rho :\ \rho \geq C_n \rho_m \,\text{such that}\, H(\varepsilon_n(x,\rho),\cdot) \,\text{is stable at}\, x\}, \tag{64}$$

where C_n is a positive constant selected in some way to be precised below. Finally, we set $\varepsilon_{ng}(x,\rho_n(x)) = \varepsilon_n(x,\rho_n(x))$.

Let us describe the algorithm to compute $\rho_n(x)$ as given by (64). Suppose that we have, at a given step n of the process, the positive definite matrix $Q_n(x)$ and the familly of neighborhoods $\varepsilon_n(x,\rho)$ depending on $\rho > 0$. All of these elliptic neighborhoods have center at x and the same major-axis orientation. If we denote by $\lambda_{1,n}$, $\lambda_{2,n}$ the eigenvalues of $Q_n(x)$, then $\sqrt{\frac{\rho}{\lambda_{2,n}}}$ and $\sqrt{\frac{\rho}{\lambda_{1,n}}}$ are the major-axis and minor-axis of $\varepsilon_n(x,\rho)$ and its area is equal to $\frac{\pi\rho}{\sqrt{\lambda_{1,n}\lambda_{2,n}}}$.

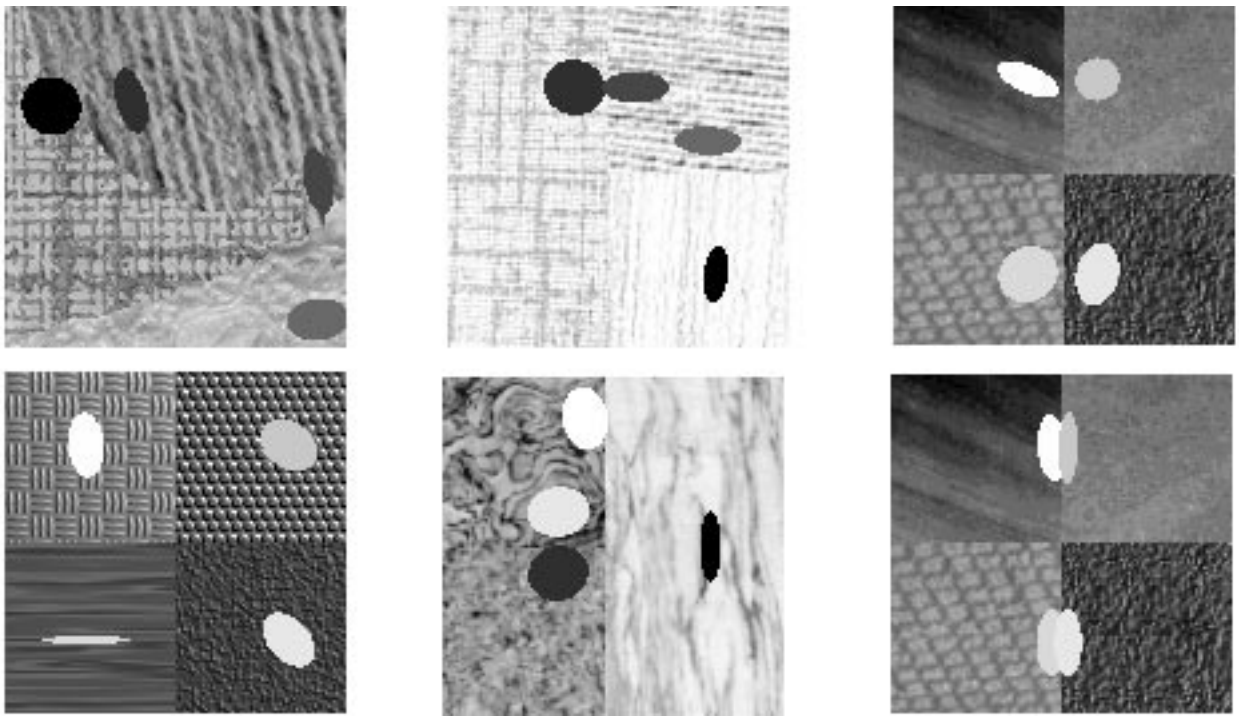

Fig. 2. Some intrinsic elliptic Δ-neighborhoods. The elliptic neighborhood for points in the same texture have approximately the same major-axis orientation and the same ratio between the major and minor-axis. But for pixels near the boundary between two textures, the associated neighborhoods are, unfortunately but obviously, slightly different from the other ones associated to interior points in the same texture.

Choose $\rho = C_n\rho_m$ where $C_n > 0$ is such that $|\varepsilon_n(x, C_n\rho_m)| = \pi\rho_m^2$. Then proceed to increase it by growing slowly the major-axis or/and the minor-axis until the associated $\varepsilon_n(x,\rho)$ satisfies the Histogram Stabilization Criterion. To be more precise, let us use the notation $\varepsilon_n(x; a_n, b_n)$ to refer to $\varepsilon_n(x,\rho)$ where $a_n = \sqrt{\frac{\rho}{\lambda_{2,n}}}$ and $b_n = \sqrt{\frac{\rho}{\lambda_{1,n}}}$ are the major and minor-axis of the elliptic neighborhood $\varepsilon_n(x,\rho)$, respectively, if $\lambda_{1,n} \geq \lambda_{2,n}$. Then, the algorithm we use to compute $\rho_n(x)$ and $\varepsilon_{ng}(x, \rho_n(x))$ can be summarized as follows:

(a) Set $C_n = \rho_m\sqrt{\lambda_{1,n}\lambda_{2,n}}$ and consider $\varepsilon_n(x, C_n\rho_m)$ -with area equal to $\pi\rho_m^2$-. Set $\varepsilon_n(x; a_n, b_n) = \varepsilon_n(x, C_n\rho_m)$.

(b) Check if $HSC(\varepsilon_n(x; a_n, b_n), \varepsilon_n(x; a_n+1, \tilde{b}_n))$ is satisfied, where $\tilde{b}_n$ is such that $\frac{\tilde{b}_n}{a_n+1} = \frac{b_n}{a_n}$.

If it is the case, set $\rho_n(x) = C_n\rho_m$, $\varepsilon_{ng}(x, \rho_n(x)) = \varepsilon_n(x; a_n, b_n)$ and stop the process.

Otherwise, check if $HSC(\varepsilon_n(x; a_n, b_n), \varepsilon_n(x; \tilde{a}_n, b_n+1))$ is satisfied, where $\tilde{a}_n$ is such that $\frac{b_n+1}{\tilde{a}_n} = \frac{b_n}{a_n}$.

If it is the case, set $\rho_n(x) = C_n\rho_m$, $\varepsilon_{ng}(x, \rho_n(x)) = \varepsilon_n(x; a_n, b_n)$ and stop the process. Otherwise,

(c) Put $a_n^* = \tilde{a}_n$, $b_n^* = b_n + 1$ and start again at step (b) with $\varepsilon_n(x; a_n^*, b_n^*)$ instead of $\varepsilon_n(x; a_n, b_n)$.

Remark 7.1. Let us note that, in the $k-$th step of the process (b)-(c), we take $\delta = \pi\frac{a_n}{b_n}(2b_n + 2k - 1)$. Obviously, if $\epsilon > 0$ is big enough, we obtain a stabilization of the histogram in a small number of steps of the process (b)-(c). For instance, in the numerical experiments showed in subsection 6.3, we take $\epsilon = 0.01$ and we get the stabilization with k small, $k = 2, 3$ or 4 (depending on the kind of texture and on the point in the texture). Of course, if ϵ is large, there is less accuracy in the later computation of texton densities in the elliptic Δ-neighborhood, but the algorithm is faster.

Figure 2 above show some neighborhoods associated with the indicated points for some textured

images used in the numerical experiments. We can observe how the elliptic neighborhood obtained for points in the same texture have approximately the same major-axis orientation and the same shape (that is to say, the same ratio between the major and minor-axis). One can observe that for pixels on the boundary of a texture the associated neighborhoods are slightly different from the other ones associated to interior points in the same texture. This last fact can be a drawback for a correct localization of the boundaries between different textures. This fact is more accentuated in textured images with a very high contrast of grey level. Finally, let us remark that the computations of the neighborhood at each point are highy paralellizable and could be used in real-time image analysis since they utilize only elementary functions.

6.3. Discretization of the feature vector

In this subsection we describe the way we compute the texture and surface shape features introduced above, that is, the way we compute from the original image $g : \Omega \to I\!R_+$ a feature vector $G : \Omega \to I\!R^p$, $G = (G_1, ..., G_p)$ for some $p \in I\!N$ to be used in the segmentation process.

Models based on dominant local orientation

Suppose that we have, for each pixel x in Ω, the intrinsic elliptic Δ-neighborhood and the symmetric positive definite matrix computed with the algorithm detailed in 6.2. For simplicity, let us forget about the subindices and denote them by $\varepsilon_g(x)$ and $Q_g(x)$, respectively. Let us consider the energy functional (17) introduced in Section 3

$$E_1(Q,B) = \int_\Omega \mathrm{tr}\big(Q_g^{-1/2} Q_g(x) Q_g^{-1/2} - Q\big)^2 dx + \lambda ATV(B),$$

where the normalization matrix

$$Q_g = \frac{1}{|\Omega|}\int_\Omega Q_g(x)dx,$$

is assumed to be invertible (otherwise, we define $Q_g^{-1/2} = Id$), and $Q = \sum_{O\subset\Omega\setminus B} M_O \chi_O(x)$, $M_O \in MS_2 = \big\{A :\ 2\times 2\, matrix,\ A = A^t > 0\big\}$. Let us recall that the matrix $Q_g^{-1/2}$ plays an important normalizing role in two senses: both to guarantee the affine invariance of the functional and to normalize the gradient effects.

If we denote by

$$\begin{pmatrix} a_{11}(x) & a_{12}(x) \\ a_{12}(x) & a_{22}(x) \end{pmatrix} \tag{65}$$

the matrix product $Q_g^{-1/2} Q_g(x) Q_g^{-1/2}$ and we set

$$Q = \sum_{O\subset\Omega\setminus B} \begin{pmatrix} b_{11}^O & b_{12}^O \\ b_{12}^O & b_{22}^O \end{pmatrix} \chi_O(x),$$

then

$$\mathrm{tr}\big(Q_g^{-1/2} Q_g(x) Q_g^{-1/2} - Q\big)^2 = \sum_{O\subset\Omega\setminus B} \big((a_{11}(x) - b_{11}^O)^2 + 2(a_{12}(x) - b_{12}^O)^2 + (a_{22}(x) - b_{22}^O)^2\big)\chi_O(x),$$

and the functional to be minimized can be written as

$$E_1(Q,B) = \sum_{i=1}^{4}\sum_{k=1}^{R}\int_{O_k} (G_i(x) - u_i(O_k))^2 dx + \lambda ATV(B),$$

where $G_1(x) = a_{11}(x)$, $G_2(x) = G_3(x) = a_{12}(x)$, $G_4(x) = a_{22}(x)$, $x \in \Omega$, $u_1(O_k) = b_{11}^{O_k}$ $u_2(O_k) = u_3(O_k) = b_{12}^{O_k}$ $u_4(O_k) = b_{22}^{O_k}$, $O_k \in \Omega \setminus B$ and R denotes the number of regions. Notice that in the present case the number of channels p (or dimension of the feature vector G) is equal to 4.
Summarizing, the algorithm we use to compute the feature channels is the following.

- First, compute the discrete approximation of the matrix Q_g, which we denote again by Q_g, given by

$$Q_g = \frac{1}{Card(\Omega)}\sum_{x\in\Omega} Q_g(x)$$

  where the sum extends to all the pixels x in the discrete domain Ω.
- Second, compute $Q_g^{-1/2}$. If $|Q_g| = 0$, take $Q_g^{-1/2} = Id$.

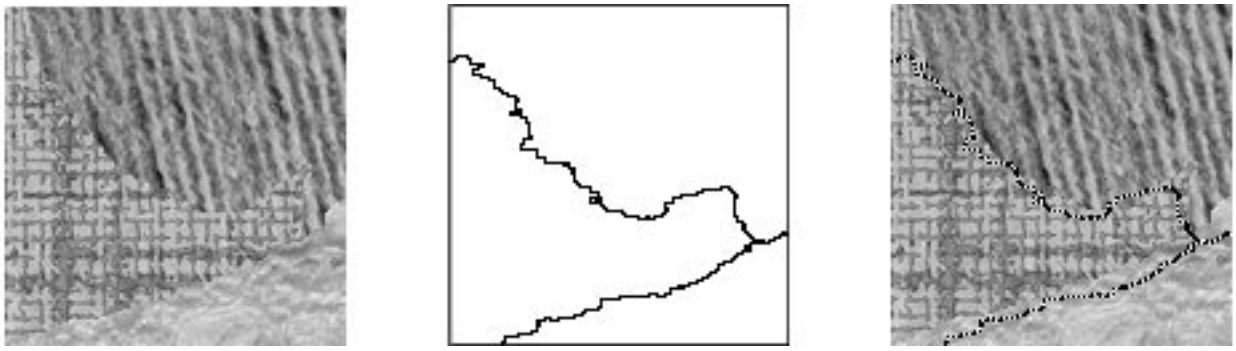

Fig. 3. Left: original image. Center: segmentation boundaries obtained from E_1. Right: segmentation superposed original image.

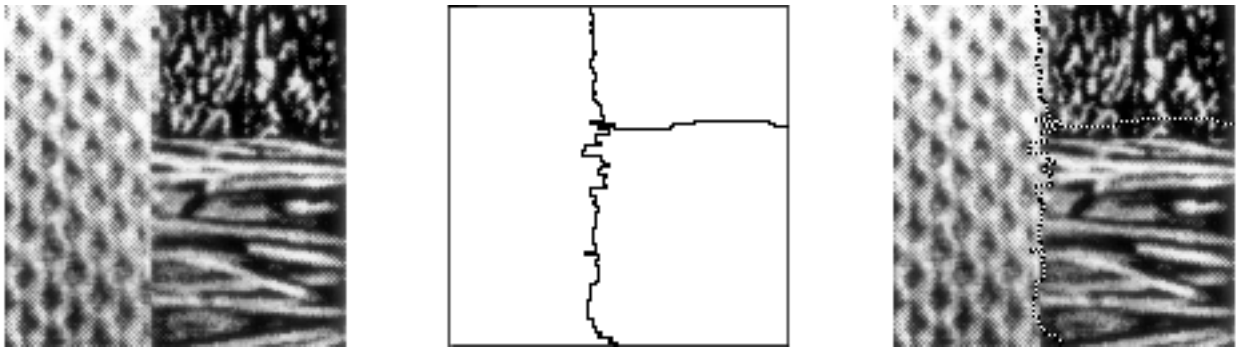

Fig. 4. Left: Original image. Center: segmentation boundaries obtained from E_1. Right: segmentation superposed original image.

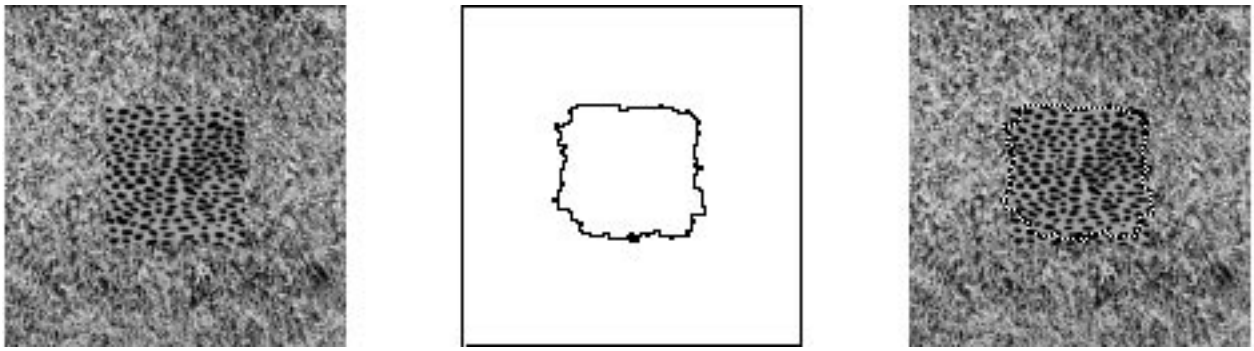

Fig. 5. Left: original image. Center: segmentation boundaries obtained from E_1. Right: segmentation superposed original image.

- Third, compute, for all $x \in \Omega$, the matrix product $Q_g^{-1/2} Q_g(x) Q_g^{-1/2}$, noted as in (65).
- Finally, set $G_1(x) = a_{11}(x)$, $G_2(x) = G_3(x) = a_{12}(x)$ and $G_4(x) = a_{22}(x)$. and use the minimization algorithm of subsection 6.4.

In Figure 3 and Figure 4 above we can observe two textured images discriminated minimizing the previous functional.

Figure 5 display an experiment using the previous functional where the input image is a piece of cheetah leather on a field of grass.

In the next figure, Figure 6, we present an experiment to test the affine invariance of the E_1 functional. First we apply to the original image displayed in Figure 3 a linear map whose matrix is $\begin{pmatrix} 2 & 0 \\ 0 & 1/2 \end{pmatrix}$ to get a new image. Figure 6 display the results obtaining with the functional E_1 over this new image.

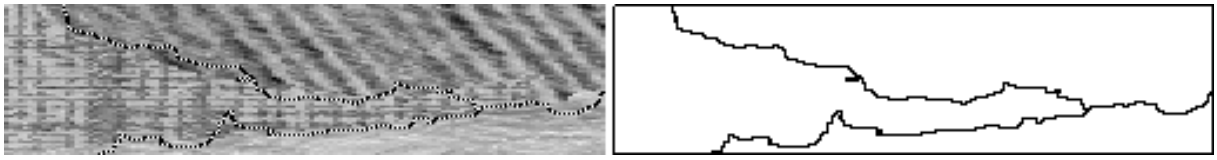

Fig. 6. Right: segmentation superposed original image. Left: segmentation boundaries obtained from E_1.

Figure 7 below shows an example of texture discrimination using the functional E_2 described in remark 3.1.

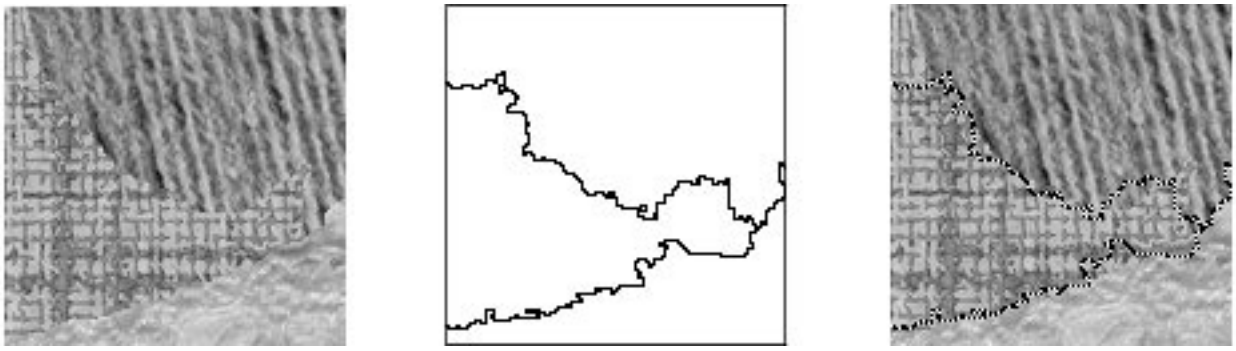

Fig. 7. Left: original image. Center: segmentation boundaries obtained from E_2. Right: segmentation superposed original image.

Functionals based on Zernike moments.

Recall that the functional based on Zernike moments is defined by

$$E_Z(\{m_{nl}\}, B) = \sum_{n,l} \int_\Omega \|m_{nl}(g,x) - m_{nl}\|^2 dx + \lambda ATV(B) \tag{66}$$

where, the moments $m_{nl}(g,x)$ are based on the intrinsic elliptic neighborhood redefined in terms of the matrix Q_g, $\varepsilon_g(x, \rho_g(x)) := \{y \in \Omega : < Q_g(y-x), y-x > \leq \rho_g(x)\}$. We take

$$m_{nl}(g,x) = \int_{\|z\| \leq 1} V_{nl}(z) g(x + N_g^{-1/2}(x) z) dz, \tag{67}$$

where $z = (z_1, z_2)$, $x = (x_1, x_2)$ is in the domain Ω and $N_g(x) = \rho_g(x)^{-1} Q_g$ is the normalization matrix needed to reinforce the affine invariance of the functional we use. For each $n \geq 0, l \in \mathbb{Z}$, $|l| \leq n$ and $n - |l|$ is even and $V_{nl}(z)$ are the complex Zernike polynomials

$$V_{nl}(z) = V_{nl}(r,\theta) = R_{nl}(r) e^{il\theta}; \quad x^2 + y^2 \leq 1. \tag{68}$$

Then, the functional (66) can be written

Fig. 8. Left: Original image. Center: segmentation boundaries obtained from E_Z. Right: segmentation superposed original image.

Fig. 9. Left: Original image. Center: segmentation boundaries obtained from E_Z. Right: segmentation superposed original image.

$$E_Z(\{m_{nl}\},B)=\sum_{n=0}^{N}\sum_{\substack{l=-n\\ n-|l|\ even}}^{n}\sum_{k=1}^{R}\int_{\Omega}\|m_{nl}(g,x)-m_{nl}(O_k)\|^2dx+\lambda ATV(B)$$

where N is a fixed integer. Replacing integrals by summations and taking the real and imaginary parts of $m_{nl}(g,x)$ in (67)

$$Re(m_{nl}(g,x))=\sum_{z_1}\sum_{z_2}R_{nl}(r(z))\cos(l\theta(z))g((x_1,x_2)+N_g^{-1/2}(x)(z_1,z_2)),$$

$$Im(m_{nl}(g,x))=\sum_{z_1}\sum_{z_2}R_{nl}(r(z))\sin(l\theta(z))g((x_1,x_2)+N_g^{-1/2}(x)(z_1,z_2)),$$

where $z_1^2+z_2^2\leq 1$, we associate with g a set of Zernike features G. By virtue of (67) and (68) we notice that $\overline{m_{nl}(g,x)}=m_{n,-l}(g,x)$, so that $Im(m_{n,-l}(g,x))=-Im(m_{nl}(g,x))$. Thus we can concentrate on $m_{nl}(g,x)$ with $l\geq 0$ as far as the defined Zernike features are concerned and because of Zernike polynomials contains $\frac{1}{2}(N+1)(N+2)$ linearly independent polynomials of degree $\leq N$, we compute $\frac{1}{2}(N+1)(N+2)$ different features. Thus, the feature vector G, for a fixed N, takes values in $\mathbb{R}^{\frac{1}{2}(N+1)(N+2)}$. The G_i components are the real or imaginary parts of the moments $m_{nl}(g,x)$, $0\leq n\leq N$, $l\geq 0$ and $n-l$ even. Figure 8 and Figure 9 show two numerical results obtained from the texture segmentation algorithm (see subsection 6.4) when the previous functional is applied. The number of moment basis used in the experiments is 4.

Remark: If we only require euclidean invariance, we can compute the Zernike moments m_{nl} in (67)

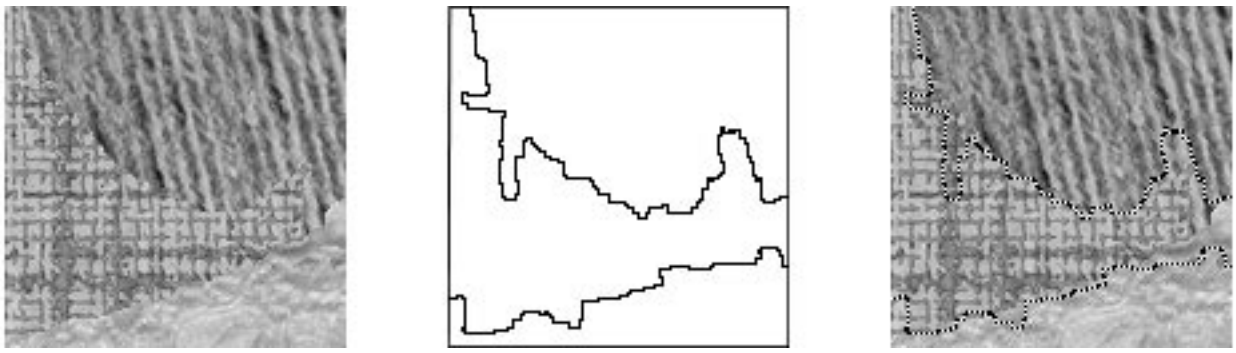

Fig. 10. Left: Original image. Center: segmentation boundaries obtained from the euclidean features. Right: segmentation superposed original image.

on a circular neighborhood of each point in the image domain

$$m_{nl}(g,x) = \int_{\|z\|\leq a} V_{nl}(z/a)g(x+z)dz,$$

and/or use in the functional the euclidean length instead of the ATV term. Figure 10 above show a numerical experiment using these euclidean features, where the number of moment basis used in the experiment is 4

Obviously, we can construct for each one of the other functionals its analogue euclideant invariant functionals (we refer to [7] for a detailed account of these).

Shape from Texture.

In order to obtain information about the shape of a 3D surface from a single image −or, more specifically (see Section 5), in order to extract the slant and tilt parameters− we use the affine intrinsic orientation matrix computed by the algorithm detailed in 6.2. We prefer to use this matrix instead of

$$\mu_g(x) = \int_{I\!R^2} w(x-y)\nabla g(y)\nabla g(y)^t dy, \qquad x\in\Omega, \tag{69}$$

with w a Gauss kernel, in order to guarantee affine invariance against the perspective affine distorsion caused by the perspective projection (assuming the focal distance $f\to+\infty$) and to avoid the problem of scale selection usual in the context of shape from texture. Thus, we use the particular length scale given by the intrinsic elliptic Δ-neighborhood $\varepsilon_g(x)$ and the matrix

$$\mu_g(x) = Q_g(x) = \frac{1}{|\varepsilon_g(x)|}\int_{\varepsilon_g(x)} \nabla g(y)\nabla g(y)^t dy,$$

and we segment the image using the affine invariant energy functional

$$E(\mu,B) = \int_\Omega \mathrm{tr}\big(\mu_g^{-1/2}\mu_g(x)\mu_g^{-1/2}-\mu\big)^2 dx + \lambda ATV(B),$$

(i.e., E_1) in order to extract the regions in the image domain Ω with a homogeneous surface shape and to estimate its slant and tilt.

Briefly, the algorithm is as follows

- First, compute the feature vector and minimize the functional E using the algorithm given above for the functional E_1. We obtain a local minimum $(\bar\mu,\bar B)$ with $\bar B$ made of a finite union of piecewise affine Jordan curves with disjoint interiors and $\bar\mu(x) = \sum_O \bar\mu_O\chi_O(x)$ with $\bar\mu_O\in MS_2$, the sum being extended to all connected components of $\Omega\setminus\bar B$.
- Second, compute $\mu_g^{1/2}$ and $\mu_g^{1/2}\bar\mu\mu_g^{1/2} = \sum_O \mu_g^{1/2}\bar\mu_O\mu_g^{1/2}\chi_O$.
  Now, for each region $O\in\Omega\setminus B$,
- Compute the maximum and minimum eigenvalues of the matrix product $\mu_g^{1/2}\bar\mu_O\mu_g^{1/2}$. Let us denote them, respectively, by $\lambda_{1,O}$, $\lambda_{2,O}$.

Compute also the eigenvector associated to the maximum eigenvalue λ_{1O}.

- Compute the slant $-\sigma_O-$ and tilt $-\tau_O-$ parameters given by

$$\cos\sigma_O = \sqrt{\frac{\lambda_{2,O}}{\lambda_{1,O}}} \qquad \text{and} \qquad t_O = (\cos\tau_O, \sin\tau_O)$$

where t_O is the unit eigenvector associated to $\lambda_{1,O}$.

- Finally, for each region $O \in \Omega \setminus \bar{B}$, estimate the local surface normal, given by (see Figure 11)

$$n_O = (\sin\sigma_O \cos\tau_O, \sin\sigma_O \sin\tau_O, -\cos\sigma_O).$$

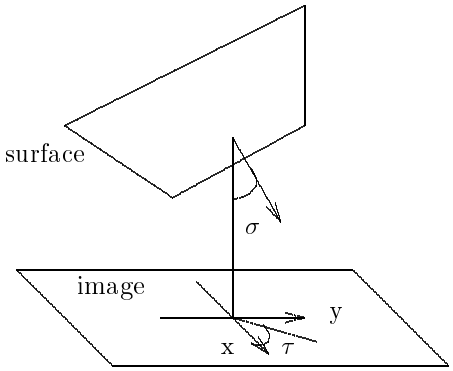

Fig. 11. Consider a coordinate system in $\mathbb{R}^3$ in such a way that the x,y-axis coincide with the coordinate system on the image plane, and let σ be the angle between the z-axis and the normal vector to the surface (i.e., the slant). The angle of tilt, τ, is the angle between the projection of the normal and the x-axis.

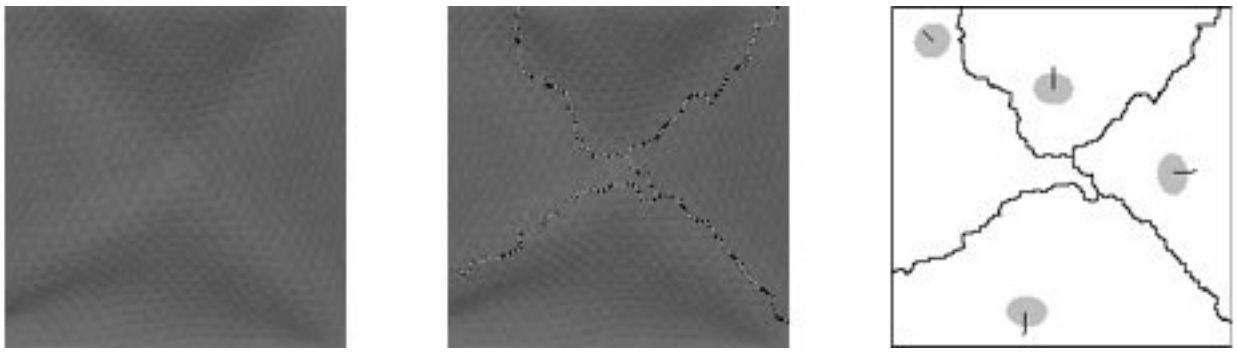

Fig. 12. Shape from texture information. Left: original image. Center: original image with the boundaries. Right: segmentation boundaries with the estimated surface orientations.

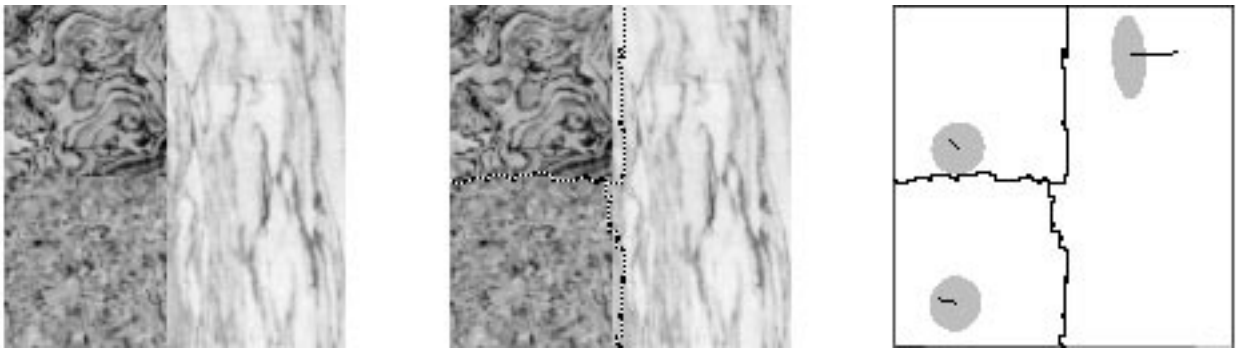

Fig. 13. Shape from texture information. Left: original image. Center: original image with the boundaries. Right: segmentation boundaries with the estimated surface orientations.

An experimental result for an image of a 3D-shape made of a same texture with different orientations planes is presented in Figure 12 and experiments recovering the shape from texture information in images where several textures are presents are displayed in Figure 13. In the figures the sur-

face orientation of each region is indicated graphically by the intrinsic elliptic Δ-neighborhood corresponding to such region (computed from *mean* affine intrinsic orientation matrix of such regions), with and attached needle parallel to the surface normal.

6.4. Optimization strategy. Segmentation multiscale algorithm

Obviously, it is not possible to find directly the global minimum of the energies defined in the previous sections by examining the whole set of possible segmentations. The principle of the computational method we use consists in generating local transformations of a given segmentation of the textured image keeping the ones which reduces the energy (lower energy means improvement of the segmentation). The tool to produce this transformation is the merging of adjacent regions according to a region growing algorithm. This simple segmentation tool applied to a Mumford-Shah type energy functional, is enough to compute a local minimum of it and to do the job of most of the classical segmentation tools.

Indeed, after the multiscale formalization of image segmentation given by [44] (see also [43]), it is clear which should be the kind of segmentation algorithm to be used: a region growing algorithm. The criterion for region "merging" would be one of the Mumford-Shah type energy functionals proposed. Then, two regions of a given segmentation will be merged if and only if the energy of the resulting segmentation decreases. This leads to a fast algorithm with a causal structure -see below-.

Let us recall that, given the set of boundaries B, the corresponding minimum u^B of $E(u^B, B)$ is completely defined by the fact the value of the u_i^B coordinate on each connected component of $\Omega \setminus B$ is equal to the mean value of G_i on this connected component. Therefore, we shall assume that with each B is associated this unique u^B and we shall write in this case $E(B)$ instead of $E(u^B, B)$ and u instead of u^B.

The concept of 2-normal segmentation synthetizes the concept of optimal segmentation we are looking for, and it lays on the basis of the computational method we use. In fact, if we follow the main idea of the region growing methods (see references [69] and [35]), we shall see that what they compute is precisely a 2-normal subsegmentation of a fine initial segmentation, obtained by recursive merging. Let us recall from [51] this concept.

Definition 6. A segmentation B is called 2-normal if, for every pair of neighbouring regions O_i and O_j, one has

$$E(B) \leq E(B \setminus \partial(O_i, O_j)) \tag{70}$$

where $\partial(O_i, O_j)$ denotes the common boundary of O_i and O_j.

Then, if B is 2-normal, (70) implies the inequality:

$$2\lambda \sum_{\Gamma \in B} Inter(\partial(O_i, O_j), ,) - \lambda ATV(\partial(O_i, O_j)) \leq \frac{|O_i||O_j|}{|O_i| + |O_j|} \sum_{k=1}^{p} (u_k(O_i) - u_k(O_j))^2 \tag{71}$$

where $u_k(O_i), u_k(O_j)$ are, respectively, the mean values of G_k in O_i, O_j. We call (71) the **Merging Criterium.** We decide to remove the common boundary $\partial(O_i, O_j)$ of O_i and O_j if (71) is not satisfied. By repeating this step, that is, by comparing the balance of energy for deciding to join any two neighbour regions, we finally get a 2-normal segmentation; a segmentation where no further elimination improves the energy.

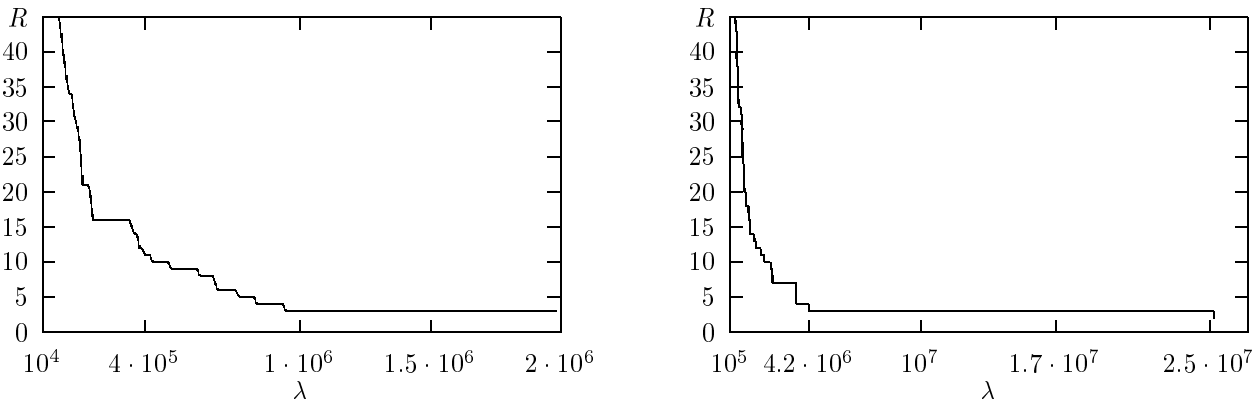

Fig. 14. The graph on the left corresponds to Figure 4. At right, graph corresponding to Figure 7. In the horizontal axis, we represent the scale parameter λ and, in the vertical one, the number of regions. The largest valley corresponds to the three main regions of the original images.

The minimizing algorithm is applied in a multiscale framework with a causal structure. In other terms, it will not only compute one segmentation, but a hierarchy of segmentations from fine to coarse scales λ. Starting at the finest level of scale, and iteratively passing from one level to the next coarser level, the minimization algorithm is performed in such a way that the output of the process at each level is used as an input for the next level. Moreover, the coarser segmentations will be deduced from the finer ones by "merging" operations, with a causal structure for the computation (see [44]). The scale parameter λ in each functional can be interpreted as a measure of the "amount of boundary" in the image: if λ is small, we allow for many boundaries in the segmentation. In this way, we will have the "causality property": if $\lambda > \lambda'$, then the boundaries provided by the algorithm for λ are contained in those obtained for λ' and the regions of the segmentation associated to λ are the unions of some of the regions obtained for λ' ([44]). In practice, for each functional, we consider an increasing sequence of parameters (e.g. $\{\lambda_k\}_{k\geq 1}$) and seek for a sequence of 2-normal segmentations (u_k, B_k) relatively with respect to the energy

$$E_{*,k}(B_k) = \int_\Omega \|u_k - G\|^2 + \lambda_k ATV(B_k).$$

We define recursively a sequence (u_k, B_k) for $k = 1, 2, \ldots$, by the following property: (u_{k+1}, B_{k+1}) is obtained from (u_k, B_k) by merging all possible pairs of regions of $\Omega \setminus B_k$. Then, the multiscale algorithm may be summarized as follows:

- **(A)** Start at the finest level, $k = 0$, with a partition of the image in regions of equal size $D \times D$, where $D = 1, 2$ or 4, usually.
- **(B)** At each level k, merge all possible regions obtained in the segmentation from level $k - 1$ if the merging produces an energy decay. Algorithmically,
  1. - Consider a region. For each one of its boundaries consider the adjacent region. If the merging produces a lower energy, proceed to merge and update the data structure.
  2. - Iterate until all the regions from the list of regions have been scanned and no other merging is possible.
     After this step a 2-normal subsegmentation B_k of the initial segmentation for the scale parameter λ_k is achieved.
- **(C)** Stop if the last level $k = p$ has been reached or if there is just one region left or if the desired number of regions is reached.

If the discrimination is successful (e. g., the regions obtained correspond approximately to the textures location) one can say that the used channels are able to discriminate the given textures. Let us observe that the stopping criterion is the number of regions. That is, in the numerical experiments, we choose a final number of regions

and we allow that the scale parameter λ grows until this number of regions is reached. Anyway, we have observed an important experimental evidence: the live expectancy of a significative region is very large with respect to the scale parameter (i.e., a significant region remains along the time). It can be seen in the Figure 14 that, for a given significant region O, the width $\lambda_{max} - \lambda_{min}$ of the interval given by $[\lambda_{min}, \lambda_{max}]$, where $\lambda_{min} = \inf\{\lambda : O \in \Omega \setminus B_\lambda\}$ and $\lambda_{max} = \sup\{\lambda : O \in \Omega \setminus B_\lambda\}$, is very large. Then, we could think at eliminating this stop criterion specially in the case that the number of texture regions of the image is unknown or there is not human supervision (it is possible to obtain a correct segmentation with a sufficiently large random scale parameter).

The main **Data Structure** needed by this algorithm is the list of regions. For each region we define

- a list of its boundaries,
- its characteristics or channels: for each considered functional, the channels are defined as we stablished above.

The boundaries are defined by

- its neighbouring regions and vertices,
- its n channels or characteristics: the channels are the values m_i $(i = 1, \cdots, n)$ corresponding to the number of times that the direction d_i appears on the boundary.

The vertices are defined by

- its coordinates,
- a description of the distance between two vertices along the geometrical curve joining them.

7. Summary and discussion

In this paper we studied the problem of texture segmentation by using affine invariant models. Indeed, we proposed an affine invariant model to, both extract affine invariant texture features from images of natural scenes and to find the boundaries between the different textures in the scene. We take an intrinsic approach to solve the usual problem in texture discrimination of choosing an appropiate window in which one computes the texture features, by incorporating what we call the intrinsic elliptic Δ-neighborhood determined by the texture pattern. From our affine intrinsic neighborhoods we compute affine intrinsic orientation matrices giving an estimate of the dominant local orientation and affine invariant texture features based on Zernike moments. These features are used in a multichannel analogue to the affine invariant energy functional introduced in [6] for grey level image segmentation. The method has been tested on a set of natural and man-made textures giving accurate discrimination.

On the other hand, we have exploited the classical idea of Gibson [32] who observed that, thanks to texture, one can extract the essential information of surface orientation in the scene, which is naturally split into two components, called slant and tilt. We showed that the affine intrinsic orientation matrix captures this essential information and that its combination with our affine invariant segmentation model is able to recover faithfully the shape from texture information in images where several textures are present.

The model we proposed combines the several steps of discriminating textures, smoothing and boundary detection into a coherent and unified framework with a simple and elegant formalism. The model requires only one parameter (two in the cases of "statistics of dominant local orientation" and of "affine invariant moments") which is the desired number of regions. On the other hand, the fact that the computation of the affine intrinsic elliptic Δ-neighborhood in each point of the image domain is an iterative process slows down the global algorithm, but it can be parallelized. On the other hand, we could think at eliminating the stop criterion in the algorithm (see subsection 6.4) specially in the case that the number of texture regions of an image is unknown or there is not human supervision.

Acknowledgements

We would like to thank V. Caselles and J.M. Morel for their help and valuable suggestions. We acknowledge partial support by DGICYT project PB94-1174.

Appendix A

Proof of Lemma 1.

Let $M = (a_{ij})$. Then

$$\mathrm{tr}\left((M - vv^t)^2\right) = \sum_{i,j}(a_{ij}^2 - 2a_{ij}v_iv_j + v_i^2v_j^2). \tag{A1}$$

Then, differentiating (A1) with respect to v_i and equating to zero, we get:

$$Mv = \|v\|^2 v.$$

Hence, if $v \neq 0$, v is an eigenvector of M associated with the eigenvalue $\lambda = \|v\|^2$, then λ must be the largest eigenvalue. Indeed, let $\lambda_1, \ldots, \lambda_d$ be the set of eigenvalues of M ordered such that $\lambda_1 \geq \ldots \geq \lambda_d$ and let e_i be the normalized eigenvector associated to λ_i. Then $M = \sum_{i=1}^d \lambda_i e_i \otimes e_i$ and

$$M - v \otimes v = \sum_{i=1}^{d} \lambda_i e_i \otimes e_i - \lambda_j e_j \otimes e_j$$

where $\lambda = \lambda_j$ for some $j = 1, \ldots, d$. Then we see that

$$\|M - v \otimes v\|_2^2 = \mathrm{tr}(M - v \otimes v)^2 = \sum_{i \neq j} \lambda_i^2$$

is a minimum when $\lambda = \lambda_1$.

Appendix B

Proof of Theorem 2.

(a) Let A_1, A_2 be two linear maps normalizing g. Let $u_1(x) = g(A_1x)$, $u_2(x) = g(A_2x)$. Observe that

$$\begin{aligned} u_2(x) = g(A_2x) &= g(A_2A_1^{-1}A_1x) \\ &= \{A_2A_1^{-1}\}g(A_1x) = \{A_2A_1^{-1}\}u_1(x). \end{aligned}$$

Hence, $M_{u_2} = (A_2A_1^{-1})^t M_{u_1}(A_2A_1^{-1})$. Since $M_{u_1} = M_{u_2} = I$, it follows that $A_2A_1^{-1}$ is an orthogonal matrix. (a) is proved.
(b) Let A_1, A_2 be two linear maps normalizing g and h, respectively, i.e., letting $\bar{g} = \{A_1\}g$, $\bar{h} = \{A_2\}h$ then $M_{\bar{g}} = A_1^t M_g A_1 = I$ and $M_{\bar{h}} = A_2^t M_h A_2 = I$. Hence $M_g = (A_1^{-1})^t A_1^{-1}$, $M_h = (A_2^{-1})^t A_2^{-1}$. On the other hand, since $h(x) = g(Ax)$, by (18) $M_h = A^t M_g A$ and we have that

$$(A_2^{-1})^t A_2^{-1} = A^t (A_1^{-1})^t A_1^{-1} A$$

and therefore

$$(A_1^{-1}AA_2)^t (A_1^{-1}AA_2) = I$$

i.e., $U = A_1^{-1}AA_2$ is a rotation matrix. Finally,

$$\begin{aligned} \bar{h}(x) = h(A_2x) = g(AA_2x) &= g(A_1A_1^{-1}AA_2x) \\ &= \bar{g}(A_1^{-1}AA_2x) = \{U\}\bar{g}(x). \end{aligned}$$

Appendix C

Proof of Lemma 2.

Let $\epsilon > 0$. Then, for some $\delta = \delta(\epsilon, g, x)$ and any (M', ρ'), $M' \in MS_2$, $\rho' > 0$ such that $\|M' - M_n\| + |\rho' - \rho_n| < \delta$ we have

$$\begin{aligned} &\|H(g, M', \rho', \cdot) - H(g, M_n, \rho_n, \cdot)\|_2 \\ &\qquad \leq \epsilon \|H(g, M_n, \rho_n, \cdot)\|_2. \end{aligned} \tag{C1}$$

Let $B_n := \left\{y \in I\!R^2 : \langle M_n(y - x), y - x\rangle \leq \rho_n\right\}$. Since $M_n \to M$, one can easily show that

$$\chi_{B_n} \to \chi_B \qquad in\ L^p(\Omega), \quad \forall p < +\infty,$$

for some convex set B. Then it is immediate to show that $B = \left\{y \in I\!R^2 : \langle M(y - x), y - x\rangle \leq \rho\right\}$. It follows that $|B_n| \to |B|$ and $\|H(g, M_n, \rho_n, \cdot)\|_2 \to \|H(g, M, \rho, \cdot)\|_2$. Thus, we are done by letting $n \to \infty$ in (C1).

Appendix D

Proof of Lemma 3.

$$\begin{aligned} H(h, M', \rho, \lambda) &= \frac{|\{y \in \varepsilon'(x) : h(y) \leq \lambda\}|}{|\varepsilon'(x)|} \\ &= \frac{|\{y \in A^{-1}\varepsilon(\bar{x}) : g(Ay) \leq \lambda\}|}{|\det A^{-1}\varepsilon(\bar{x})|} \end{aligned}$$

$$
\begin{aligned}
&= \frac{|\{A^{-1}z : z \in \varepsilon(\bar{x}),\, g(z) \le \lambda\}|}{|\det A^{-1}\varepsilon(\bar{x})|} \\
&= \frac{|\{z : z \in \varepsilon(\bar{x}),\, g(z) \le \lambda\}|}{|\varepsilon(\bar{x})|} \\
&= H(g, M, \rho, \lambda).
\end{aligned}
$$

Let $\epsilon > 0$. Let $\delta' = \delta'(\epsilon, h, x)$ be such that if

$$\|\tilde{M}' - M'\| + |\tilde{\rho} - \rho| < \delta',$$

then

$$\|H(h, \tilde{M}', \tilde{\rho}, \cdot) - H(h, M', \rho, \cdot)\|_2 \le \epsilon \|H(h, M', \rho, \cdot)\|_2.$$

Let $\tilde{M}, \tilde{\rho}$ be such that

$$\|\tilde{M} - M\| + |\tilde{\rho} - \rho| < \delta := \frac{\delta'}{max(1, \|A\|^2)} .$$

Then $\tilde{M}' = A^t \tilde{M} A$, $\tilde{\rho}$ satisfies $\|\tilde{M}' - M'\| + |\tilde{\rho} - \rho| < max(1, \|A\|^2)\delta = \delta'$.
Hence

$$
\begin{aligned}
&\|H(g, \tilde{M}, \tilde{\rho}, \cdot) - H(g, M, \rho, \cdot)\|_2 \\
=\ &\|H(h, \tilde{M}', \tilde{\rho}, \cdot) - H(h, M', \rho, \cdot)\|_2 \\
\le\ &\epsilon \|H(h, M', \rho, \cdot)\|_2 = \epsilon \|H(g, M, \rho, \cdot)\|_2 ,
\end{aligned}
$$

i.e., $H(g, M, \rho, \cdot)$ is stable at $\bar{x}$. The converse follows with the same proof.

References

[1] Y. S. Abu-Mustafa and D. Psaltis. Recognitive aspects of moments invariants. *IEEE Trans. on Pattern Analysis and Machine Intell.*, PAMI-6:698–706, Nov. 1984.

[2] J. Aloimonos. Shape from texture. *Biol. Cyber.*, 58:345–360, 1988.

[3] L. Alvarez, F. Guichard, P. L. Lions, and J. M. Morel. On the fundamental equations of image processing. *Arch. Rat. Mechanics and Dual*, 123:199–257, 1993.

[4] C. Blakemore and F. W. Campbell. On the existence of neurones in the human visual system selectively sensitive to orientation and size of retinal images. *J. Physiol.*, 203:237–260, 1969.

[5] D. Blostein and N. Ahuja. Shape from texture: integrating texture element extraction and surface estimation. *IEEE Trans. on Pattern Analysis and Machine Intell.*, 11:1233–1251, 1989.

[6] C. Ballester, V. Caselles, and M. González. Affine invariant segmentation by variational methods. In *SIAM J. Appl. Math.*, 56(1):294–325, 1996.

[7] C. Ballester and M. González. Texture segmentation by variational methods. In *Proc. Second European Workshop on Image Processing and Mean Curvature Motion*, Palma de Mallorca, Spain, Sept. 25-27, 1995.

[8] J. Bigun, G. H. Granlund, and J. Wiklund. Multidimensional orientation estimation with applications to texture analysis and optical flow. *IEEE Trans. Patt. Anal. Machine Intell.*, 13(8):775–790, 1991.

[9] J. F. Boyce and W. J. Hossack. Moment invariants for pattern recognition. *Pattern Recognition Letters*, 1(5-6):451–456, 1983.

[10] J. Bigun and C. Horne. Feature space dimensionality reduction by using local representability. In *Proc. Int. Conf. Signal Processing*, pages 555–558, Beijing, China, Oct. 22-26 1990.

[11] J. Bigun. Frequence and orientation selective texture measure using linear summetry and laplacian pyramid. In *Proc. Visual Communications and Image Procesing*, pages 1319–1331, Lausanne, Switzerland, October 1990.

[12] R. Bajcsy and L. I. Lieberman. Texture gradient as a depth cue. *Comput. Graphics Image Processing*, 5:52–67, 1976.

[13] A. Blake and C. Marinos. Shape from texture: estimation, hypothesis and moments. *Artificial Intelligence*, 45:323–380, 1990.

[14] A. Blake and C. Marinos. Shape from texture: the homogeneity hypothesis. In *Proc. IEEE 3rd Int. Conf. on Computer Vision*, pages 350–353, Osaka Japan, 1990. IEEE Computer Society Press.

[15] L. G. Brown and H. Shvaylser. Surface orientation from projective foreshortening of isotropic texture autocorrelation. *IEEE Trans. on Pattern Analysis and Machine Intell.*, 12:584–588, 1990.

[16] P. Burt. Fast filter transforms for image processing. *Comput. Vision Graphics Image Process.*, 16:20–51, 1981.

[17] A. B. Bhatia and E. Wolf. On the circle polynomials of zernike and related orthogonal sets. *Proc. Camb. Phil. Soc*, 50:40–48, 1954.

[18] R. W. Connors and C. A. Harlow. A theoretical comparison of texture algorithms. *IEEE Trans. Patt. Anal. Machine Intell.*, 2(3):204–222, 1980.

[19] M. G. Crandall, H. Ishii, and P. L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc.*, 27:1–67, 1992.

[20] T. Cohignac, C. Lopez, and J. M. Morel. Classification de formes invariante par transformation affine. In *Proc. 9ème congrès RFIA*, pages 39–50, Paris, Janvier 11-14, 1994.

[21] F. H. Crick, D. C. Marr, and T. Poggio. An information processing approach to understanding the visual cortex. A.I. Memo 557, MIT, April 1980.

[22] R. W. Conners, M. M. Trivedi, and C. A. Harlow. Segmentation of a high resolution urban scene using texture operators. *Comput. Vision Graphics Image Process.*, 25:273–310, 1984.

[23] J. G. Daugman. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *J. Opt. Soc. Am. A.*, 2(7):1160–1169, 1985.

[24] L. S. Davis, L. Janos, and S. M. Dunn. Efficient recovery of shape from texture. *IEEE Trans. on Pattern Analysis and Machine Intell.*, 5:485–492, 1983.

[25] H. Dirilten and T. G. Newman. Pattern matching under affine transformations. *IEEE Trans. on Computers*, pages 314–317, 1977.

[26] D. J. Fleet and A. D. Jepson. Hierarchical construction of orientation and velocity selective fields. *IEEE Trans. Patt. Anal. Machine Intell.*, 11(3):315–325, 1989.

[27] D. Gabor. Theory of communication. *J. Inst. Elec. Eng.*, 93:429–459, 1946.

[28] J. Garding. Estimating the orientation of multiple textured planes. In *Proc. Sixth Scandinavian Conference on Image Analysis*, pages 228–236, Oulu. Finland, 1989.

[29] J. Garding. *Shape from texture markings*. PhD thesis, Dept. of Numerical Analysis and Computing Science, Royal Institute of Technology, Stockholm, May 1991.

[30] J. Garding. Shape from texture and contour by weak isotropy. *Artificial Intelligence*, 64:243–297, 1993.

[31] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Patt. Anal. Machine Intell.*, PAMI-6(6):721–741, Nov. 1984.

[32] J. Gibson. *The perception of the visual world*. Houghton Mifflin, Boston, 1950.

[33] J. Garding and T. Lindeberg. Direct computation of shape cues by multi-scale retinotopic processing. Report from comput. vision and active perception (CVAP), Dept. of Numer. Anal. and Comp. Science., Royal Inst. of Technology. Stockholm, Sweden, 1992.

[34] A. Grossman and J. Morlet. Decomposition of hardy functions into square integrable wavelets of constant shape. *SIAM J. Math.*, 15:723–736, 1984.

[35] S. L. Horowitz and T. Pavlidis. Picture segmentation by a directed split-and-merge procedure. In *Proc. of Second IJCPR*, pages 424–433, 1974.

[36] R. M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Trans. System Man Cybernet.*, 3(1):610–621, 1973.

[37] M. K. Hu. Pattern recognition by moment invariants. *Proc. IRE*, 49:14–28, Sept. 1961.

[38] D. H. Hubel and T. N. Wiesel. Perceptive fields of single neurones in the cat's striate cortex. *J. Physiol.*, 148:574–591, 1959.

[39] B. Julesz. Textons gradients: the texton theory revisited. *Biol. Cyber.*, 54:245–251, 1986.

[40] K. Kanatani. Detection of surfaces orientation and motion from texture by stereological technique. *Artificial Intelligence*, 23:213–237, 1984.

[41] K. Kanatani and T. C. Chou. Shape from texture: general principle. *Artificial Intelligence*, 38:1–48, 1989.

[42] A. Khotanzad and Y. H. Hong. Invariant image recognition by zernike moments. *IEEE Trans. on Pattern Analysis and Machine Intell.*, 12(5):489–497, May. 1990.

[43] G. Koepfler, J. M. Morel, and C. Lopez. A multiscale algorithm for image segmentation by variational methods. To appear in SIAM J. on Numerical Analysis.

[44] G. Koepfler, J. M. Morel, and S. Solimini. Segmentation by minimizing a functional and the merging methods. To appear in SIAM J. of Num. analysis, 1993.

[45] M. Kass and A. Witkin. Analyzing oriented patterns. *Comput. Vision Graphics Image Process.*, 37:362–385, 1987.

[46] K. I. Laws. *Textured image segmentation*. PhD thesis, Dept. Elec. Eng., Univ. Southern California, Los Angeles, 1980.

[47] C. Lopez and J. M. Morel. Axiomatisation of shape analysis and application to texture hyperdiscrimination. Proceedings on the Trento Conference of Surface Tension and Movement by Mean Curvature. De Grayter Publishers. Berlin.

[48] T. S. Lee, D. Mumford, and A. Yuille. *Texture Segmentation by Minimizing Vector Valued Energy Functionals: The Coupled-Membrane Model*, volume 588 of *Lecture Notes in Computer Science*, pages 165–173. Springer-Verlag, 1992.

[49] C. Lopez. *Theorie de la discrimination de textures*. PhD thesis, CEREMADE. Université de Paris-Dauphine, 1993.

[50] J. Malik and P. Perona. Preattentive texture discrimination with early vision mechanisms. *Journal of the Optical Society of America*, 7, No 5:923–932, 1991.

[51] J. M. Morel and S. Solimini. Segmentation of images by variational methods: a constructive approach. *Rev. Matemática de la Univ. Complutense de Madrid*, 1:169–182, 1988.

[52] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and variational problems. *Communications on Pure and Applied Mathematics*, XLII(4), 1989.

[53] J. M. Morel and S. Solimini. *Variational Methods for image segmentation*. Birkhauser Verlag, 1995.

[54] M. Nitzberg, D. Mumford, and T. Shiota. *Filtering, Segmentation and Depth*, volume 662 of *Lecture Notes in Computer Science*. Springer-Verlag, 1993.

[55] T. R. Reed and J. M. Hans du Buf. A review of recent texture segmentation and feature extraction techiques. *CVGIP: Image Understanding*, 57(3):359–372, 1993.

[56] A. R. Rao and R. C. Jain. Computerized flow field analysis: Oriented texture fields. *IEEE Trans. Patt. Anal. Machine Intell.*, 14(7):693–709, 1992.

[57] A. R. Rao and B. G. Schunk. Computing oriented texture fields. *CVGIP: Graphical Models and Image Processing*, 53(2):157–185, 1991.

[58] B. J. Super and A. C. Bovik. Shape from texture by wavelet-based measurement of local spectral moments. In *Proc. IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition*, pages 296–301, Champaign, Illinois, June 1992.

[59] G. Sapiro and A. Tannenbaum. Affine invariant scale-space. *International Journal of Computer Vision*, 11(1):25–44, 1993.

[60] G. Sapiro and A. Tannenbaum. On affine plane curve evolution. *Journal of Functional Analysis*, 119(1):79–120, 1994.

[61] M. Spann and R. Wilson. A quad-tree approach to image segmentation which combines statistical and spatial information. *Patt. Recognition*, 18(3/4):257–269, 1989.

[62] C. H. Teh and R. T. Chin. On image analysis by the methods of moments. *IEEE Trans. on Pattern Analysis and Machine Intell.*, 10(4):496–513, July 1988.

[63] M. R. Teague. Image analysis via the general theory of moments. *J. Opt. Soc. Amer.*, 70:920–930, Aug. 1980.

[64] M. Unser. Local linear transforms for texture measurements. *Signal Processing*, 11:61–67, 1986.

[65] H. Voorhees and T. Poggio. Computing texture boundaries in images. *Nature*, 333:364–367, 1988.

[66] J. S. Weska, C. R. Dyer, and A. Rosenfeld. A comparison study of texture measures for terrain classification. *IEEE Trans. System Man Cybernet.*, 6(4):269–286, 1976.

[67] A. P. Witkin. Recovering surface shape and orientation from texture. *Artificial Intelligence*, 17:17–45, 1981.

[69] S. W. Zucker. Region growing: Childhood and adolescence (survey). *Comp. Graphics and Image Processing*, 5:382–399, 1976.

Coloma Ballester received the Ph.D. degree in Computer Science from the University of Illes Balears, Spain, in 1995. She is currently an Assistant Professor in the Department of Mathematics and Computer Science of this University. Her research interests include computer vision, image processing.

Manuel González received the Ph.D. degree in Computer Science from the University of Illes Balears, Spain, in 1995. He is currently an Assistant Professor in the Department of Mathematics and Computer Science of this University. His research interests include computer vision, image processing.