Available online at www.sciencedirect.com
SCIENCE DIRECT®
Pattern Recognition Letters 27 (2006) 1022–1030
Pattern Recognition Letters
www.elsevier.com/locate/patrec

Inhibition of false landmarks

Debora Gil *, Petia Radeva

Centre de Visió per Computador, Dept. Informàtica, UAB, Edifici O-Campus UAB, 08193 Bellaterra, Barcelona, Catalonia, Spain

Received 29 July 2004; received in revised form 8 November 2005
Available online 21 February 2006

Communicated by Prof. M. Lindenbaum

Abstract

Corners and junctions are landmarks characterized by the lack of differentiability in the unit tangent to the image level curve. Detectors based on differential operators are not, by their own definition, the best posed as they require a higher degree of differentiability to yield a reliable response. We argue that a corner detector should be based on the degree of continuity of the tangent vector to the image level sets, work on the image domain and need no assumptions on neither the image local structure nor the particular geometry of the corner/junction. An operator measuring the degree of differentiability of the projection matrix on the image gradient fulfills the above requirements. Because using smoothing kernels leads to corner misplacement, we suggest an alternative fake response remover based on the receptive field inhibition of spurious details. The combination of both orientation discontinuity detection and noise inhibition produce our inhibition orientation energy (IOE) landmark locator.

Keywords: Corners; T-junctions; Wavelets

1. Introduction

The ability of our receptive system to detect sudden changes in the surrounding environment has been often used in computer vision to extract the key image features. However the visual system perception inhibition principle has been scarcely used in spite of being an efficient noise remover. We suggest using inhibition of image discontinuities to design a landmark detector based on corner/junction location. Current corner detectors (see Rohr (2001) for an exhaustive review) split into those working on the image (2D) domain and those ones that analyze the (1D) geometry of some curves extracted from the image.

Most 2D operators, either explicitly or implicitly, search for corners with a particular angulation, which makes them unable to detect junctions and hinders their performance in real images. Response of operators measuring image isotropy (Harris et al., 1988; Noble, 1988) drops at acute angles and their performance substantially worsens with noise or textured backgrounds. Parametric approaches (Rohr, 1992; Rosin, 1999) fit an analytic model of a corner to the particular image they handle. Besides being time consuming, they are prone to poorly perform in real images as they assume uniform grey level within regions. Wedge filters need no assumptions on image intensity but, by their own design, they only respond to a given set of corner angulation and orientations (Robbins and Owens, 1997; Simoncelli and Farid, 1996). Only curvature-based algorithms and the simple SUSAN (Smith and Brady, 1997) comparison of grey values need no assumptions on the corner geometry. The first ones, despite being better posed for general corner detection, require of some degree of differentiability of the image level sets to yield reliable responses for the curvature operator. However corners are characterized by lack of continuity in level sets tangent direction, so that curvature based detectors may lead to ambiguous results. Although the SUSAN scheme works fine without

* Corresponding author. Tel.: +34 93 581 23 01; fax: +34 93 581 16 70.
E-mail addresses: debora@uab.es (D. Gil), petia@uab.es (P. Radeva).

0167-8655/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2005.11.018

any requirements on the image differentiability or local structure, it sometimes confuses edges and corners on real images.

On the other side, algorithms running on curves extracted from the image mainly search for discontinuities on the curve tangent vector. Wavelets (Lee et al., 1995; Qudus and Gabbouj, 2002; Yeh, 2003) are a usual tool to analyze such regularity because of their robust high response near points of discontinuity (Mallat, 1999). Although we agree with the former definition of corner, working on image contours instead of working on the image domain has several disadvantages. First, the compulsory previous extraction of image edges, which might occasionally lead to some sort of boundary tracking (Lee et al., 1995; Mokhtarian and Suomela, 1998; Mokhtarian and Bober, 2003) to fill-in edge gaps at corners. Second, computation of tangent spaces of curves in parametric form is a delicate step (Lee et al., 1995; Yeh, 2003) that would be unnecessary in their implicit level set original form.

In any case, all algorithms must reject the operator response to texture and noise by working on a smoothed version of the original image. This smoothing can be achieved by either straight convolution of the image with a gaussian kernel or, in the case of wavelets and scale-space approaches, working at coarse scales/level of detail. A major drawback of both strategies is that image prominent features are blurred in the measure noise is reduced. Although there are some criteria (Qudus and Gabbouj, 2002) to determine the proper resolution, corner misplacement is difficult to avoid unless some sort of corner tracking (Lee et al., 1995; Mokhtarian and Suomela, 1998) through all levels of detail is performed. Although corner tracking can be efficiently done (Mokhtarian and Bober, 2003), it could be skipped (or at least reduced) if the corner detector admitted a robust performance in the presence of noise and, thus, work at the finer level of detail. The point of the present article is to propose an alternative way of removing fake responses that allows working at small scales.

In this paper we characterize corners and junctions in terms of the lack of continuity of the projection matrix onto the image unit gradient. Convolution with first derivatives of oriented anisotropic gaussian kernels are used to determine the matrix singular points. The integral over all orientations of the determinant of the former wavelet transform is an energy operator on the image domain that is maximum at corners/junctions independently of their particular geometric nature. We will refer to this orientation continuity energy as orientation energy, OE, for short. The main differences with current corner characterizations are the following. By working on the image domain our operator does not require any of the preprocessing steps (Lee et al., 1995) that most 1D wavelet based corner detectors require. Because it does not rely on either the local image structure or a particular model of corner/junction, our operator characterizes both corners and junctions without any possible confusion with other image salient features (as it might happen with edges in the SUSAN (Smith and Brady, 1997) scheme and ridges/valleys in the case of maxima of curvature searchers). Besides OE increases as the corner angle becomes more acute, in contrast to operators measuring isotropy of the image (Harris et al., 1988; Noble, 1988), which response drops for angles less than 90°. Finally, by scanning the image in all possible orientations in the discrete domain, our operator is capable of detecting corners without any sort of fitting (Rohr, 1992; Rosin, 1999) or special filter design (Robbins and Owens, 1997; Simoncelli and Farid, 1996).

Still, in real images, fake responses at noisy or textured backgrounds must be suppressed. Instead of running our operator at different scales (Mokhtarian and Suomela, 1998; Lee et al., 1995), we propose using the same mechanisms that serve our visual system to ignore noise and texture. Our sensitivity to abrupt changes differs depending on the nature of the response in the surroundings: only isolated or unique salient features are taken into account. In (Marr, 1982) this biological mechanism is modelled by means of an inhibition kernel, which convolved with the image Gabor energy results in a robust edge detector (Grigorescu et al., 2003). In our case, we will apply inhibition to a representation of the image salient features obtained by means of the energy of the image wavelet transform. This map of the most significant image features is the input for OE and its inhibition serves as a noise suppression factor for IOE. Because the former strategy does not hinge upon any smoothing nor level of detail, our operator can work at the finest scale, thus ensuring maximum location accuracy.

Our landmark extractor is compared to the curvature based corner detector of (Kitchen and Rosendfeld, 1982), the Harris operator (Harris et al., 1988) and the SUSAN scheme (Smith and Brady, 1997). Following current ways of evaluating quality of corner extractions, we asses the operator performance in terms of location accuracy for different corner angulation and trade-off between true detections and fake responses. Experiments on synthetic noisy images prove IOE better performance and landmark extraction in real images shows its higher applicability and reliability.

We have structured the paper as follows. Section 2 is devoted to the description of IOE, validation on synthetic corners and landmark location in real images are presented in Section 3 and conclusions are exposed in Section 4.

2. From wavelets to corners

A usual way of describing corners is as points of maximal curvature of the image level sets. Not only is this definition ambiguous but also incorrect. Ambiguity comes from the fact that, in the case of corners of image level sets, image valleys and ridges are also characterized by maxima of curvature. Imprecision follows because, in fact, corners are points where the level curve fails to be $\mathscr{C}^1$. This fact

might hinder performance of any differential operator used to detect corners or junctions since, at these points, neither the curve curvature nor its derivatives are properly defined. A curve fails to be $\mathscr{C}^1$ at points where its tangent space is not properly defined, that is, points of discontinuity of the curve unit tangent direction. Because the image gradient is perpendicular to its level sets, we conclude that corners and junctions are characterized as discontinuity points of the image unit gradient direction, $\nabla u/|\nabla u| = (u_x/|\nabla u|, u_y/|\nabla u|)$, for $|\nabla u| = \sqrt{u_x^2 + u_y^2}$, the Euclidean norm. The image partial derivatives, (u_x, u_y), are calculated by convolving the image with the partial first derivatives of Gaussian derivatives kernels:

$$u_x = G_{x,\sigma} * u, \quad u_y = G_{y,\sigma} * u$$

where $*$ denotes convolution, and $G_{x,\sigma}$, $G_{y,\sigma}$ are the partial derivatives (in x, y, respectively) of a Gaussian kernel of variance σ:

$$G_\sigma = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

For any unitary vector field, $\xi(x,y) = (\xi_1, \xi_2)$, the matrix, P_ξ, encoding the orthogonal projection onto its vector space (Lang, 2004) is given by

$$P_\xi(x,y) = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} (\xi_1 \xi_2) = \begin{pmatrix} \xi_1^2 & \xi_1\xi_2 \\ \xi_1\xi_2 & \xi_2^2 \end{pmatrix}$$

has equal degree of differentiability than the vector direction. Thus, in a first approach we could consider that a point is a corner/junction if and only if $P_{\nabla u/|\nabla u|}$ fails to be continuous. However the projection matrix is discontinuous at sharp straight ridges (the tangent direction orientation is not well defined), and might be sensitive to image discretization. These artifacts are overcome by computing the projection matrix over the eigenvector of maximum eigenvalue of the structure tensor. This tensor is a descriptor of the image local orientation (Jahne, 1993; Kthe et al., 2003) and it is the weighted mean of the projection matrixes given by the element-wise convolution:

$$G_\rho * P_{\frac{\nabla u}{|\nabla u|}} = G_\rho * \left[\frac{1}{|\nabla u|^2} \begin{pmatrix} u_x^2 & u_x u_y \\ u_x u_y & u_y^2 \end{pmatrix} \right] \tag{1}$$

for G_ρ a 2D isotropic gaussian kernel of variance ρ. Since it is, indeed, the solution to the heat equation with initial condition $P_{\nabla u/|\nabla u|}$, the eigenvector of maximum eigenvalue, ξ, is an infinitely differentiable approximation to the image level sets normal space. It follows that it extends tangent spaces to sharp ridges/valleys and regularizes the image gradient. It follows that the definition:

Definition 2.1. A point p in an image $u = u(x,y)$ is a corner/junction if and only if the projection matrix onto the regularized image unit gradient given by the eigenvector of maximum eigenvalue of (1), $P_\xi(x,y)$, fails to be continuous.

exclusively characterizes corners and T-junctions, as, at ridges, the matrix P_ξ yields a continuous response. As example let us consider the distance map to a rectangle shown in Fig. 1(a). All level sets in its interior have four corners with the exception of the rectangle skeleton, which is a ridge of the distance map. The jump discontinuity in the arrows (representing level sets unit tangent) direction only occurs at the rectangles vertices.

Now, one of the best tools to detect discontinuities and lack of differentiability are wavelets, both from a theoretic (Mallat, 1999) and a practical point of view (Qudus and Gabbouj, 2002; Yeh, 2003). Because we are looking for discontinuities of P_u and the domain is the two dimensional image, we will use first derivatives of oriented anisotropic Gaussian kernels, namely G. Let $\sigma = (\sigma_1, \sigma_2)$, with $\sigma_1 < \sigma_2$, be the variance and $\tilde{x}$, $\tilde{y}$ the coordinates given by the rotation:

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \end{pmatrix} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Our mother wavelet will be G first derivative along its minor axis:

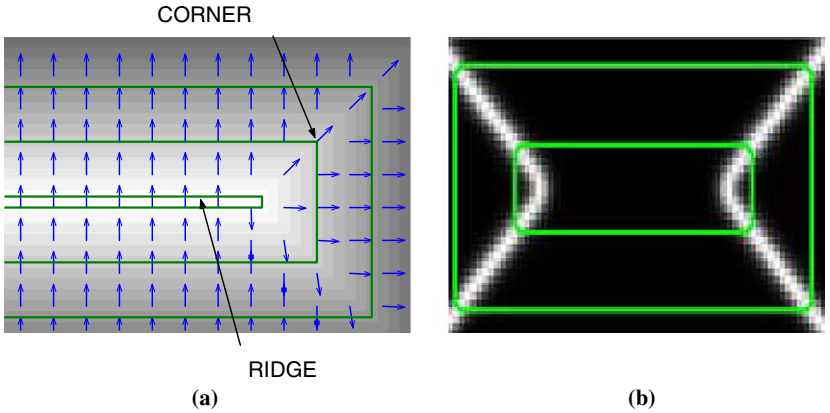

Fig. 1. Corners discontinuity in image gradient: regularized image gradient (a) and OE ideal response (b).

$$\mathrm{MH}_\theta = \partial_{\tilde{x}}(G) = \partial_{\tilde{x}}\left(\frac{1}{2\pi\sigma_1\sigma_2}\mathrm{e}^{-\frac{\tilde{x}^2}{2\sigma_1^2}-\frac{\tilde{y}^2}{2\sigma_2^2}}\right) = -\tilde{x}G/\sigma_1^2$$

Its convolution with an image $u(x,y)$ yields the oriented wavelet coefficient:

$$W_\theta u(x,y) = \int \mathrm{MH}_\theta(\tilde{x}-x,\tilde{y}-y)u(\tilde{x},\tilde{y})\,\mathrm{d}\tilde{x}\,\mathrm{d}\tilde{y}$$

And the integral over all possible orientations:

$$E(x,y) = \int (W_\theta u)^2 \mathrm{d}\theta$$

is an energy that measures the image degree of continuity at a given scale/level of detail. In the case of a matrix P_ξ, its wavelet transform is another matrix, so that E must be the integral (over all orientations) of the norm of P_ξ as a linear transformation. In our particular case the L^2 norm of the transformed matrix is given by its determinant (see Appendix A for the proof).

Proposition 2.1. *The norm $\|W_\theta P_\xi\|_2$ is given by the determinant of the matrix*

The orientation energy (OE) we suggest is

$$\begin{aligned}\mathrm{OE}(x,y) = \mathrm{OE}(u) &= \int (\det(W_\theta P_u))^2 \mathrm{d}\theta \\ \text{with } W_\theta P_u &= \int \mathrm{MH}_\theta(\tilde{x}-x,\tilde{y}-y)P_\xi(\tilde{x},\tilde{y})\mathrm{d}\tilde{x}\mathrm{d}\tilde{y}.\end{aligned} \tag{2}$$

Corners and T-junctions correspond to local maxima. Fig. 1 shows the response of OE for an ideal square corner. In Fig. 1(a) we show the eigenvector ξ and in Fig. 1(b) the resulting energy (2) for $\theta = (0, \pi/4, \pi/2, 3\pi/4)$.

Still in real images we must suppress false detections produced by textured and noisy backgrounds and a double echo at edges caused by the structure tensor (see Section 3 for further details). To avoid such false corners we propose using an inhibition kernel as in (Grigorescu et al., 2003). The main idea is to emulate the human vision system that inhibits its response at discontinuities located in areas presenting a similar singularity at all points. To such purpose, we will work with a representation of the image salient features.

2.1. The continuity of discontinuities

What captures our eyesight are abrupt changes, either in image intensity (edges, ridges) or in contours directions (corners, junctions). The first singular features serve to select those curves describing the essential shapes in an image. The second ones correspond to a punctual/point wise description of the latter, that is, landmarks. Because wavelets model human vision response, we will use a family of wavelets responding to edges and ridges to obtain a representation of the image salient features (Mallat, 1999). As in the previous section, we will consider directional derivatives of anisotropic gaussian kernels in the direction, θ, of their minor axis. That is, our wavelet filters are given by

$$\mathrm{MH}^1_\theta = \partial_{\tilde{x}}G = -\tilde{x}G/\sigma_1^2 \text{ and } \mathrm{MH}^2_\theta = \partial_{\tilde{x}\tilde{x}}G = (\tilde{x}^2-1)G/\sigma_1^2$$

If W^1_θ, W^2_θ denote the wavelet coefficients, then the energy:

$$\mathrm{MHE}(x,y) = \int (W^1_\theta u)^2 + (W^2_\theta u)^2 \mathrm{d}\theta$$

yields an image close to the representation that human perception yields. Because image edges and ridges are well outlined, this will be the input image for OE. By using MHE instead of the original input image we reduce the impact of small oscillations on level sets tangent spaces, so that the structure tensor can operate at the minimum scale. Still fake responses produced by textured backgrounds should be suppressed. Let us minimize the impact of the latter false responses by emulating the human receptive field inhibition. The human visual system rejects any salient feature in a neighborhood presenting a similar discontinuity. Roughly speaking we could say that humans check the continuity of discontinuities before processing any information. In (Marr, 1982; Grigorescu et al., 2003) this biological behavior is modelled by means of a ring-shaped inhibition kernel, IK, defined as

$$\mathrm{IK} = \frac{1}{\|H(\mathrm{DG}_\sigma)\|_2}H(\mathrm{DG}_\sigma), \quad H(z) = \begin{cases} 0 & z < 0 \\ z & z \geqslant 0 \end{cases}$$

where $\|\cdot\|_2$ denotes the L^2 norm and DG_σ is the following difference of gaussian functions:

$$\mathrm{DG}_\sigma = \frac{1}{2\pi(4\sigma)^2}\mathrm{e}^{-\frac{x^2+y^2}{2(4\sigma)^2}} - \frac{1}{2\pi\sigma^2}\mathrm{e}^{-\frac{x^2+y^2}{2\sigma^2}}$$

for the variance σ coinciding with the scale used for the wavelets filters. The inhibition kernel IK is a function of 2 variables with a ring-shaped support centered at the origin. The factor $1/\|H(\mathrm{DG}(x,y))\|_2$ is a normalization scaling to have a function of unitary L^2 norm.

The convolution of the energy MHE with IK corresponds to a weighted average in a circular ring centered at each point. It follows that if we compare this mean to the value of MHE at the point, the difference will be maximum at isolated or unique salient features and cancel response due to noise or texture. Therefore the positive response of the difference:

$$\mathrm{IA}(\mathrm{MHE}) = H(\mathrm{MHE} - \mathrm{IK} * \mathrm{MHE}) \tag{3}$$

defines an energy image defining the most representative image features as sharp maximums. The inhibition term given by (3) is added as a noise remover factor to the OE energy (2) computed on MHE to yield our inhibition orientation energy:

$$\mathrm{IOE} = \mathrm{OE}(P_{\mathrm{MHE}}) * \mathrm{IA}(\mathrm{MHE}) \tag{4}$$

In Figs. 2 and 3 we illustrate the mechanisms and results of the inhibition term on a regular area containing a well defined contour (Fig. 2(a)) and on a granulated sand area (Fig. 3(a)). The images in Figs. 2 and 3(b) and (c) are inverse (negative) energies for a better visualization (OE in boxes labelled with (b) and IOE in boxes labelled with

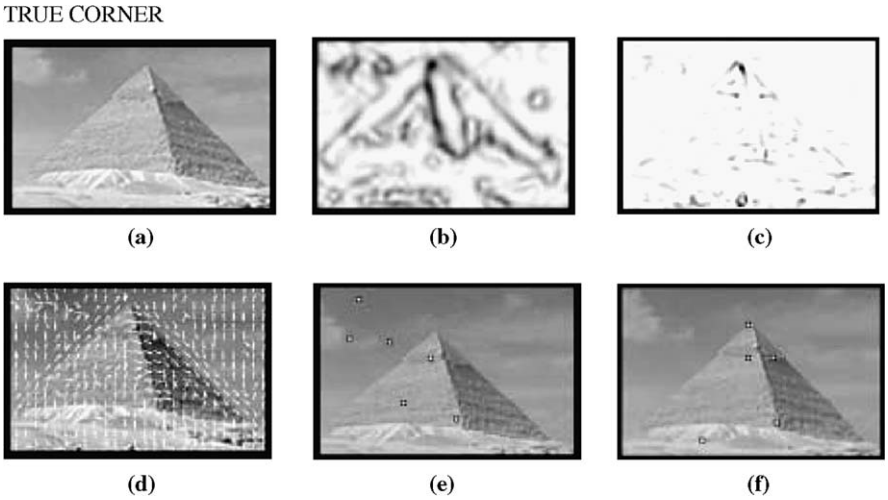

Fig. 2. Inhibition of false landmarks: regular area (a), orientation energy (OE) (b), IOE (c), vector field (d), corners without (e) and with inhibition term (f).

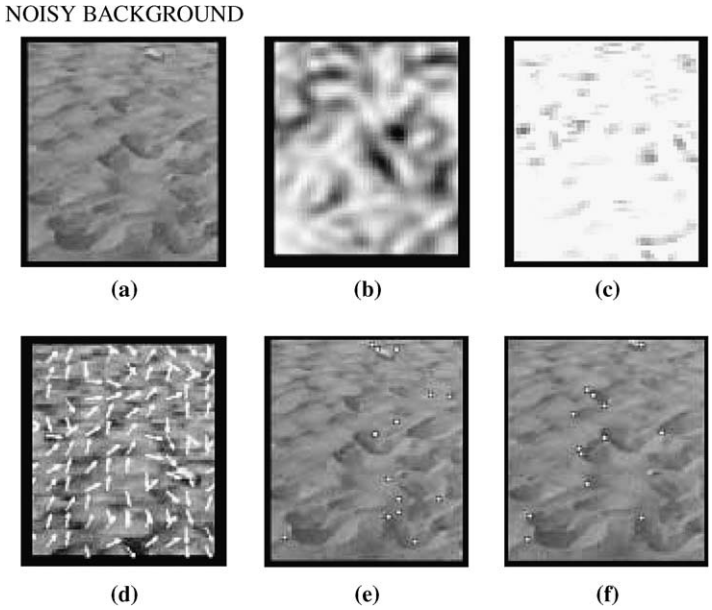

Fig. 3. Inhibition of false landmarks: regular area (a), orientation energy (OE) (b), IOE (c), vector field (d), corners without (e) and with inhibition term (f).

(c)), so that darker pixels correspond to higher energy values. Corners in Figs. 2 and 3(e) and (f) correspond to energy peaks above a given percentile for OE and IOE, respectively. On one hand, the extension of the image gradient yielded by the structure tensor (Fig. 2(b)) produces an echo in the response of OE near clean image contours. This yields fake detections in the corner map of Fig. 2(e) and omits the pyramid vertex. On the other, random orientation of the vector field in the textured background (Fig. 3(b)) produces sparse false corners. The inhibition term (Figs. 2 and 3(c)) suppresses any response outside well defined image objects. Therefore in the landmark map extracted (Figs. 2 and 3(f)) from the IOE the number of wrong detections significantly reduces, which improves the detection of the pyramid main vertex and the sand wavy shadows.

3. Experiments

We present two different experiments: assessment of performance by comparison to existing techniques and landmark extraction in real images. In all experiments, we use a scale $\{\sigma = 0.5, \rho = 0.5\}$ for the structure tensor and 6 orientations with $\sigma = 1$ for the wavelet kernels used in OE and MHE. We used two different criteria for landmark extraction:

(1) Local maxima above a given threshold of the energies characterizing landmarks. This is the criterion used for landmark extraction in synthetic images as it enables the assessment of the trade-off between false and true detections.
(2) The first N local maxima. For real images this is a more fair criterion for a visual assessment of the quality of the extracted corners, especially if any comparison to other extractors is to be made.

3.1. Synthetic corners

We have compared IOE to the curvature-based (CURV) corner detector (Kitchen and Rosendfeld, 1982), Harris (Harris et al., 1988) and SUSAN (Smith and Brady, 1997). We have applied them to a set of nine corners with angles in the range (20°, 180°) in images corrupted with gaussian noise of $\sigma = 0$, 0.1, 0.25, 0.5. The thresholds used for corner extraction have been set for each method by selecting the value that yields the best performance results. For CURV and Harris, this corresponds to the 99.5% percentile of the energy value for low noisy images ($\sigma = 0, 0.1$) and to the 99.95% percentile for the most corrupted ones. In the case of IOE we performed with a common percentile of 99.99% for all noises. For the SUSAN algorithm this percentile is given by a brightness threshold ($t \in (10, 255)$) that determines the maximum difference in gray levels between two pixels which allows them to be considered part of the same "region" in the image. Thus, the lower it is, the more corners the algorithm detects. We have set $t = 10$ for low noise images and $t = 90$ for noises with $\sigma = 0.25$, 0.5.

As in (Rohr, 2001), performance is measured in terms of location accuracy and trade-off between true and false detections. We consider a positive response is correct if it lies in a 7×7 window centered at the true corner, including the flat case (180°). Plots in Fig. 5 show the total number (for all noises) of true/false detections for each angle (abscissa axis). Images in Fig. 4 show the detections (squares on noise free angles) for an acute corner, an angle near 90° and the flat case, for $\sigma = 0.25$.

Statistics (plots in Fig. 5) for the ideal detector should yield 0 false detections, 4 good responses (1 for each noisy case) for angles less than π and no response for flat angles. Due to sensitivity to lack of differentiability, curvature-based algorithms fail to detect the most acute angles (first image in second column of Fig. 4) and produce false responses as noise increases. Harris number of detections drops as image anisotropy increases, which difficulties detection of acute corners (first image in third column of Fig. 4) and increases the number of false positives in the presence of noise (it is the worst performer with up to 4 wrong responses). SUSAN numbers would match the ideal case, if it were not for its response at flat angles (last image in 4th column of Fig. 4) and random false detections due to noise (see plot of fake detections in Fig. 5(b)). Finally, 4 right detections with 1 false one at most for angles between (40°, 140°) and null response at 180°, select IOE as the algorithm that best matches the ideal figures.

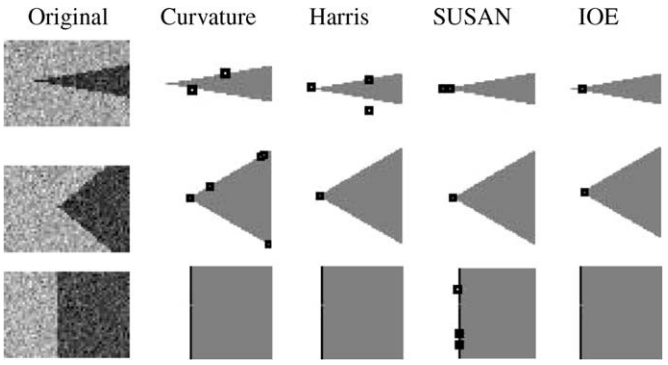

Fig. 4. Synthetic noisy corners: first row acute angle, second right and third flat.

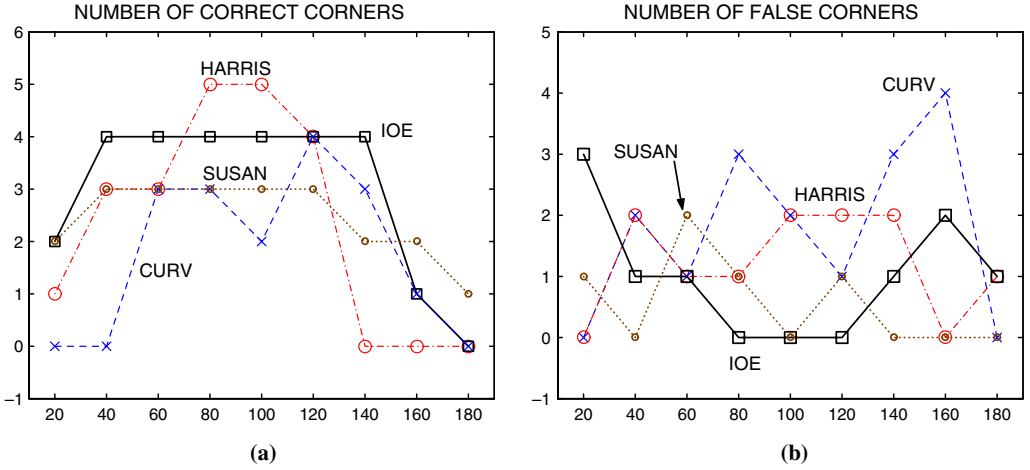

Fig. 5. Statistics for synthetic corners.

3.2. Landmark extraction in real images

Choosing the best performers of Experiment 3.1, we have applied IOE and SUSAN to landmark location on real images. The set of test images include geometric patterns (building in background of boat image in Fig. 6(a) and strips in shirt of Fig. 7(a)), faces (portrait in Fig. 7(a)) and natural scenes with texture (Fig. 9(a) and (b)). Although for a better comparison we should set each algorithm parameters to yield a given number of landmarks, there is not an explicit control on the number of landmarks extracted in the SUSAN scheme. This motivates setting the threshold parameter in the SUSAN scheme and extract as many landmarks as SUSAN detected in our case. The threshold for landmark extraction was set to $t = 40$ for the SUSAN scheme in all cases. This yields 315 points for the boat in Fig. 6, 154 for the man in stripe shirt in Fig. 7, 175 for the falcon in the first column of Fig. 9 and 186 for the birds in the second column of Fig. 9.

In general terms the number of false landmarks at edges and texture backgrounds is larger for the SUSAN scheme. Points on the border of the wooden platform and at the foreground building in Fig. 6(c) correspond to edges rather than to landmarks. In a similar fashion, SUSAN yields an over response at shirt stripes (like the line on the man shoulder in Fig. 7(c)) and branches silhouette (Fig. 9(c)), with the majority of points detected as landmarks. Textures produce some erratic false positives in water waves (right side of Fig. 6(c)), boys' hair (Fig. 7(c)) and interior of branches (Fig. 9(e)). In the case of the strong patterned nest in (Fig. 9(f)) the number of fake detections includes almost all points. Still, despite this sensitivity to corners not all

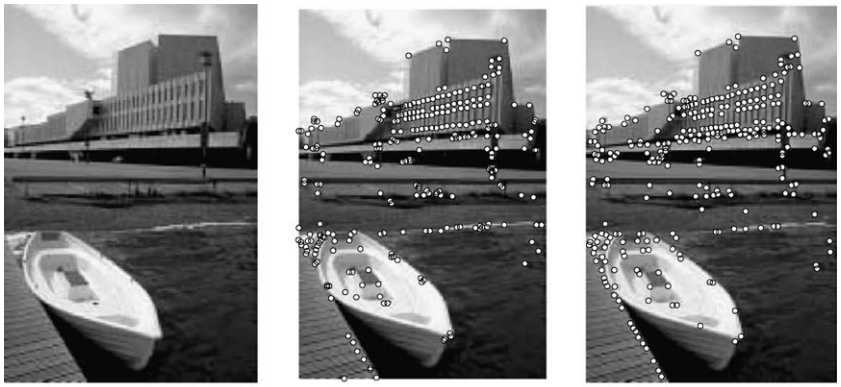

Fig. 6. Landmarks in real images I: original (a), IOE (b) and SUSAN (c).

Fig. 7. Landmarks in real images II: original (a), IOE (b) and SUSAN (c).

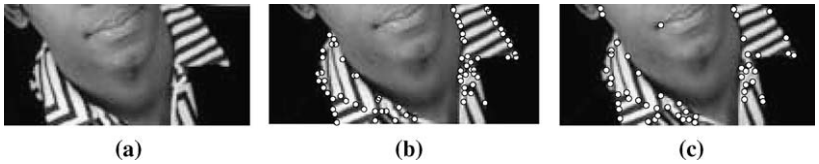

Fig. 8. Landmarks closed-up: original (a), IOE (b) and SUSAN (c).

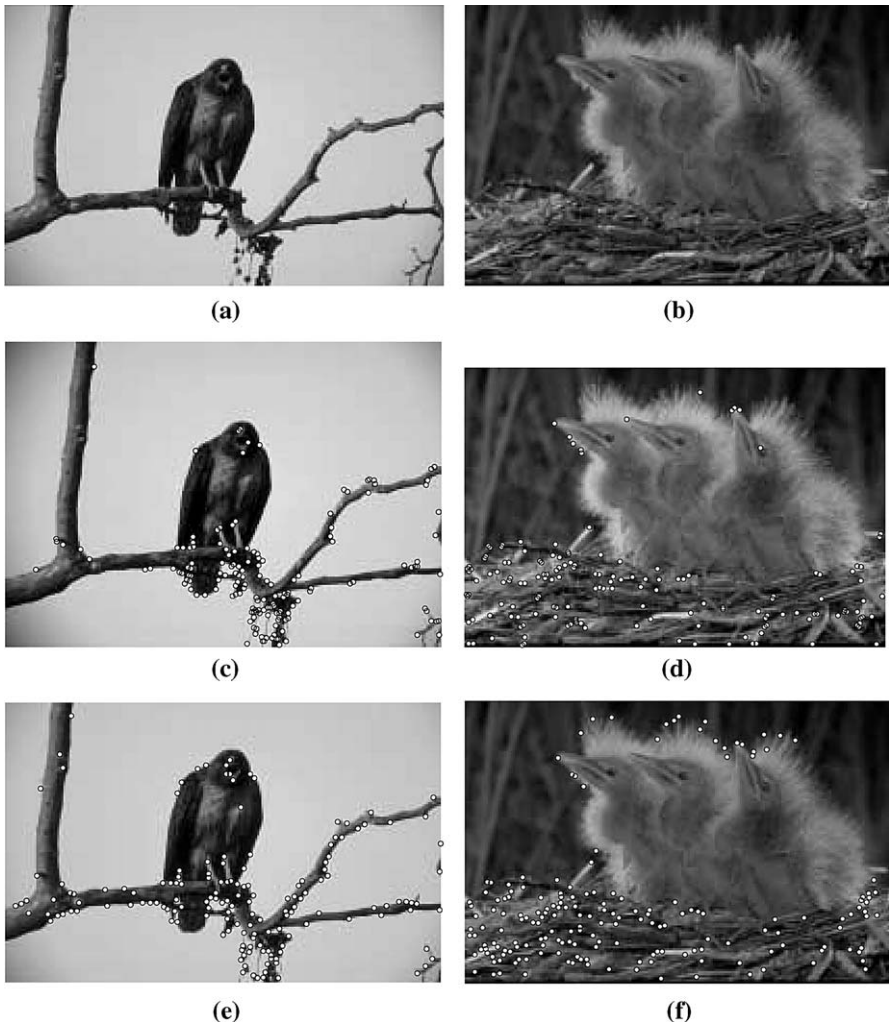

Fig. 9. Landmarks in natural scenes: original (a) and (b), IOE (c) and (d) and SUSAN (e) and (f).

stripe corners (see closed up in Fig. 8(c)) and birds beaks have been properly located by SUSAN.

On the other hand, IOE detects most geometric corners yielding optimal responses at windows and boat corners (Fig. 6(b)) and shirt stripes (Fig. 7(b)). Close-ups of the man shirt in Fig. 8 show IOE higher accuracy (Fig. 8 (b)) compared to SUSAN (Fig. 8(c)). On the other hand, IOE performance for natural scene corners and T-junctions is also competitive: branches junctions and knots in Fig. 9(c) and birds' beaks in Fig. 9(d) have been perfectly extracted. The impact of the inhibition term is determinant for IOE to yield a minimum number of fake landmarks. It lacks of response at the slightly texture induced by water waves and curly hairs and at the birds nest it gives a substantially less number of fake positives than SUSAN.

4. Conclusions

In the present paper, we define an operator that measures the continuity of the image unit gradient direction. The norm of the wavelet transform of the projection matrix onto the latter is an energy that increases with corner acuteness. In order to ensure maximum location accuracy, we propose a scale independent noise response suppressor. Basing on the ability of the visual receptive cells to inhibit their response at homogeneous noisy areas, fake responses are removed by means of the convolution of the image wavelet energy with an inhibition kernel.

Our combination of inhibition of uniform discontinuities yields a corner detector reliable enough as to robustly extract landmarks in real images. Statistics on synthetic corners and results on natural scenes prove the higher efficiency of our operator compared to usual techniques. Results on natural scenes with textured backgrounds prone to produce false responses show its applicability.

Appendix A

Proposition A.1. *The norm* $\|W_\theta P\|_2$ *is given by the determinant of the matrix*

Proof. Given a vector space with norm $\|\cdot\|$, the norm of a linear transformation (Tom, 1974), M, is given by the maximum of the transformed unitary vectors:

$$\|M\| = \max_{\|e\|=1} \|Me\|$$

We will prove that $\|W_\theta P\|_2^2$ and the determinant $|\det(W_\theta P)|$ are equivalent measures, so that we can use them indistinctly. In order to show that, it suffices (Tom, 1974) two find two positive constants $m < 1 < M$ such that

$$m|\det(W_\theta P)| \leqslant \|W_\theta P\|_2^2 \leqslant M|\det(W_\theta P)|$$

We recall that if θ is the angle between the image unit gradient and a fix axis, then P can be expressed as

$$P = \begin{pmatrix} \cos(\theta)^2 & \cos(\theta)\sin(\theta) \\ \cos(\theta)\sin(\theta) & \sin(\theta)^2 \end{pmatrix} \tag{A.1}$$

We will first prove that $W_\theta P$ is a symmetric matrix with null trace. This follows by commutativity between convolutions and derivatives:

$$\begin{aligned} W_\theta P &= \mathrm{MH}_\theta * P = \partial_{\tilde{x}} G * P = G * (\partial_{Tx} P) \\ &= G * \left[\theta_{\tilde{x}} \begin{pmatrix} -\cos(\theta)\sin(\theta) & \cos(\theta)^2 - \sin(\theta)^2 \\ \cos(\theta)^2 - \sin(\theta)^2 & \cos(\theta)\sin(\theta) \end{pmatrix} \right] \\ &= G * (\theta_{\tilde{x}} P') \end{aligned}$$

for $\theta_{\tilde{x}}$ the partial derivative with respect to the axis $\tilde{x}$ and the last equality obtained by differentiating the expression (A.1). The matrixes P' are all symmetric and have null trace. Since both symmetry and the null trace property are linear operations, they commute with the convolution. It follows that $W_\theta P$ is also symmetric with null trace. By its symmetry, we have that $W_\theta P$ diagonalizes in a basis of orthonormal eigenvectors, v, w, so it can be written as

$$W_\theta P = \begin{pmatrix} v_1 & w_1 \\ v_2 & w_2 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} v_1 & v_2 \\ w_1 & w_2 \end{pmatrix}$$

The null trace property implies opposite eigenvalues $\lambda_1 = -\lambda_2$. Since the matrix norm for square matrices is actually the maximal absolute eigenvalue (Tom, 1974), it follows that $\|M\|^2 = |\lambda|^2 = |\lambda_1 * \lambda_2| = |\det(M)|$. □

References

Grigorescu, C., Petkov, N., Westenberg, M.A., 2003. Contour perception based on non-classical receptive field inhibition. IEEE Trans. Imag. Proc. 12 (7).

Harris, C., Stephens, M., 1988. A combined Corner and Edge Detector. In: Proc. 4th Alvey Vision Conf.

Jahne, B., 1993. Spatio-temporal image processing: Theory and scientific applications. Springer-Verlag, Berlin.

Kitchen, L., Rosendfeld, A., 1982. Gray-level corner detection. Pattern Recognition Lett. 1 (95–102).

Kthe, U., Michaelis, B., Krell, G., 2003. Edge and Junction Detection with an Improved Structure Tensor. In: Proc. of 25th DAGM Symp., LNCS, vol. 2781, pp. 25–32.

Lang, S., 2004. Linear Algebra, third ed. Springer, March 9.

Lee, J.S., Sun, Y.N., Chen, C.H., 1995. Multiscale corner detection by wavelet transform. IEEE Trans. Imag. Proc. 4.

Mallat, S., 1999. A wavelet tour of signal processing. Academic Press, INC.

Marr, D., 1982. Vision. W.H. Freeman and Co.

Mokhtarian, F., Bober, M., 2003. Curvature scale space representation: Theory, applications and MPEG-7 standardization. Kluwer Academic Publishers (Springer).

Mokhtarian, F., Suomela, R., 1998. Robust image corner detection through curvature scale space. TPAMI 12 (12).

Noble, J.A., 1988. Finding corners. Imag. Vis. Comp. 6 (2).

Qudus, A., Gabbouj, M., 2002. Wavelet-based corner detection technique using optimal scale. Pattern Recognition Lett. 23.

Robbins, B., Owens, R., 1997. 2D feature detection via local energy. Imag. Vis. Comp. 15.

Rohr, K., 1992. Recognizing corners by fitting parametric models. Int. J. Comp. Vis. 9 (3).

Rohr, K., 2001. LandMark-based image analysisUsing geometric and intensity models. Kluwer Academic Publishers.

Rosin, P.L., 1999. Measuring corners properties. Comp. Vis. Imag. Unders. 73 (2).

Simoncelli, E.P., Farid, H., 1996. Steerable wedge filters for local orientation analysis. IEEE Trans. Imag. Proc. 5 (9).

Smith, S.M., Brady, J.M., 1997. SUSAN—A new approach to low level image processing. Int. J. Comp. Vis. 23 (1).

Tom, M., 1974. Apostol, second ed. Mathematical Analysis Addison-Wesley Pub. Co., January.

Yeh, C., 2003. Wavelet-based detection using eigenvectors of covariance matrices. Pattern Recognition Lett. 24.