CVGIP: GRAPHICAL MODELS AND IMAGE PROCESSING
Vol. 53, No. 1, January, pp. 40–51, 1991

Edge Detection in Multispectral Images

ALDO CUMANI

U.O. Ingegneria dei Sistemi, Istituto Elettrotecnico Nazionale "Galileo Ferraris," Torino, Italia

Communicated by Rama Chellappa

Received May 31, 1989; accepted January 24, 1990

This paper proposes an extension of edge detectors based on second-order differential operators to the case of multiple band (color) images. To this end, we define a local directional measure of multispectral contrast. A definition of edge point is then given, as the location of a directional maximum of the contrast function. Extremal contrast edges are defined, as loci of zero crossing points of the first directional derivative of the contrast function. The qualitative properties of the so defined edges are briefly discussed. An algorithm for edge location with subpixel resolution is proposed, with an example of application to actual RGB imagery. © 1991 Academic Press, Inc.

1. INTRODUCTION

Edge detection methods based upon differential operators are widely used in the early processing of one-band images, also referred to as gray-level, intensity, or monochromatic images, in contrast with multi-band, color, multispectral images. In particular, methods based on the analysis of zero-crossings of some second-order differential operator applied to image data have been extensively explored in recent years.

The two most popular operators are the Laplacian [3, 18, 20, 23, 27] and the second directional derivative in the direction of the gradient [5, 7, 11, 15, 27]; these operators share the nice property of being invariant with respect to translations and rotations in the image plane. Of the two, the second directional derivative is the more appealing, due to its connection with the extrema of the gradient magnitude; indeed, the loci of maximal gray-level gradient are a natural definition of edges in intensity images.

In contrast, differential methods seem to have received little attention in the case of multiband images. There have been attempts to define local operators for detecting changes in color images (see, e.g., [9, 13, 21, 25]); however, most of the literature on the early processing of color imagery treats the segmentation problem from the dual point of view, i.e., that of determining homogeneous regions in the image, edges being then defined as region boundaries. The algorithms proposed to this purpose range from modifications of gray-level segmentation methods, such as histogram thresholding [22, 26] or split-and-merge [1], to various kinds of clustering procedures [6, 14, 24, 28].

The inherent difficulty in extending differential methods to multi-band images stems from the fact that in this case the image function is vector-valued. To quote Machuca and Phillips [19], after the gradients of the image components are computed, "there remains the problem of how to combine them into one output."

The aim of this work is to answer that question in a way that leads to a meaningful extension of the above mentioned second-order methods. To this end, we define a local measure of directional contrast based on the gradients of the image components. This measure is maximal, at each image point, in a particular direction, which in the monochromatic case is the direction of the gray-level gradient; this leads to the definition of edge point as the location of a directional maximum of the contrast function. It should be noted that this measure of contrast was already proposed in [9]; that paper, however, did not explore the consequences of such definition in connection with edge localization.

A natural extension of second-order methods is then given, by defining extremal edges as loci of transversal zero crossings of the first derivative of the contrast function in the direction of maximal contrast. It is shown that this definition presents some nontrivial problems, connected to the sign ambiguity in the definition of the direction of maximal contrast. Conditions under which this ambiguity can be removed are given, and we discuss the qualitative properties of the resulting zero-crossing contours.

Finally, we propose an algorithm that uses the above theory for the determination of extremal edges with subpixel resolution. An example of application to actual color imagery is also given.

1049-9652/91 $3.00

2. EDGE DETECTION IN A MULTISPECTRAL IMAGE

Let an m-band image be represented by the function f: $R^2 \to R^m$ that maps a point $P = (x_1, x_2)$ in the image plane to an m-vector $\mathbf{f} = (f_1(x_1, x_2), f_2(x_1, x_2), \ldots, f_m(x_1, x_2))$. This definition clearly includes monochromatic images as a particular case ($m = 1$).

A note on notation: we shall use the notation x_i, $i = 1, 2$, for the image plane coordinates in formulas involving summations, and the more usual notation (x, y) elsewhere, it being understood that $x_1 \equiv x$ and $x_2 \equiv y$. Also, x or y as a low index will denote differentiation with respect to that variable, e.g., $g_x = \partial g/\partial x$.

We shall suppose that the components of f have been smoothed by Gaussian filtering. The strong regularizing property of Gaussian filters ensures the existence and continuity of the derivatives of $f_i(x, y)$ of any order (see, e.g., [27]). This condition could be relaxed; indeed, the results of this section will hold good provided that the image function is at least of class C^2, i.e., twice continuously differentiable.

The *image value* at a given pixel location is therefore a vector in R^m. Since we are seeking for variations in the image, we shall consider the difference in the image values at two nearby points, say P and Q, which is the m-vector

$$\Delta\mathbf{f}(P, Q) = \mathbf{f}(Q) - \mathbf{f}(P).$$

When $Q - P$ is an infinitesimal displacement $dP = (dx_1, dx_2)$ the above difference becomes the differential

$$d\mathbf{f} = \sum_{i=1}^{2} \frac{\partial \mathbf{f}}{\partial x_i}\, dx_i. \tag{1}$$

The squared norm of $d\mathbf{f}$ is

$$df^2 = \sum_{i=1}^{2}\sum_{k=1}^{2} \frac{\partial \mathbf{f}}{\partial x_i} \cdot \frac{\partial \mathbf{f}}{\partial x_k}\, dx_i dx_k = \sum_{i=1}^{2}\sum_{k=1}^{2} \gamma_{ik} dx_i dx_k, \tag{2}$$

where the dot indicates inner product of vectors in R^m. Note that speaking of "inner product" of vectors implies that some kind of *metric* has been defined on R^m [17]. This is clearly an important issue, since the choice of a particular metric may be expected to affect the output of the edge detector more or less heavily, depending upon the nature of the image being analyzed (e.g., natural color image, false-color aerial image, etc.).

A detailed discussion of this topic is, however, beyond the scope of this paper, and the reader is referred to [4, 12] for examples of metrics applicable to the case of natural color images. Although we shall use the Euclidean metric in the examples, the main results will hold good for any nonsingular Riemannian metric. In the case of the Euclidean metric, the γ_{ik} are given by

$$\gamma_{ik} = \sum_{j=1}^{m} \frac{\partial f_j}{\partial x_i} \frac{\partial f_j}{\partial x_k}.$$

Now, for constant displacement size $\|dP\|$, df^2 indicates how much the image value varies in the direction of dP. Therefore, given a unit vector $\mathbf{n} = (n_1, n_2)$ in the (x_1, x_2) plane, we define the *squared local contrast* of f at P in direction $\mathbf{n}$ as

$$S(P, \mathbf{n}) = \sum_{i=1}^{2}\sum_{k=1}^{2} \gamma_{ik} n_i n_k = En_1^2 + 2Fn_1n_2 + Gn_2^2, \tag{3}$$

where we have introduced the usual shorthand notation for the components of γ_{ik}, borrowed from differential geometry of surfaces [16]:

$$\gamma_{11} = E, \quad \gamma_{12} = \gamma_{21} = F, \quad \gamma_{22} = G.$$

It is well known that a quadratic form as (3) has, for varying $\mathbf{n}$, a maximum and a minimum value. These extreme values coincide with the eigenvalues of the 2×2 matrix $[\gamma_{ik}]$, and are attained when $\mathbf{n}$ is the corresponding eigenvector. By elementary calculations, one finds such extreme values to be

$$\lambda_\pm = (E + G \pm \sqrt{(E - G)^2 + 4F^2})/2 \tag{4}$$

and the corresponding eigenvectors are given by

$$\mathbf{n}_\pm = (\cos\theta_\pm, \sin\theta_\pm) \tag{5}$$

$$\theta_+ = \frac{1}{2}\arctan\frac{2F}{E - G} + k\pi$$

$$\theta_- = \theta_+ \pm \pi/2.$$

The above analysis shows that the quantity defined by (3) has, at each point, two extremal values attained in orthogonal directions. The meaning of this can be clarified by considering the monochromatic case ($m = 1$). In this case, letting $I(x, y) = f_1(x, y)$ be the gray value function, we have

$$E = I_x^2, \quad F = I_x I_y, \quad G = I_y^2$$

and one easily finds

$$\lambda_+ = I_x^2 + I_y^2, \quad \lambda_- = 0 \tag{6}$$

$$\theta_+ = \frac{1}{2}\arctan\frac{2I_xI_y}{I_x^2 - I_y^2} + k\pi = \arctan\frac{I_y}{I_x} + k\pi.$$

Therefore, in the mono case, the maximal squared contrast is the square of the gray-level gradient magnitude. The direction of maximal contrast is the gradient direction and the minimal contrast of zero is attained orthogonally to the gradient. This agrees well with the usual definition of contrast in one-band images as the magnitude of the gray-level gradient [7].

The form of Eqs. (4), (5) suggests that an important role is played by the two-dimensional vector function

$$\mathbf{v} = (E - G, 2F). \tag{7}$$

Indeed, we have

$$\lambda_+ - \lambda_- = \|\mathbf{v}\|, \quad \theta_+ = \tfrac{1}{2}\arg(\mathbf{v}), \tag{8}$$

from which it is seen that the maximal direction is defined wherever $\mathbf{v} \neq 0$; at points where $\mathbf{v} = 0$, the two eigenvalues of $[\gamma_{ik}]$ become equal and the eigenvector directions are undefined. Indeed, in this case $S(P, \mathbf{n})$ does not depend on $\mathbf{n}$ at P; i.e., the image variation has the same magnitude in all directions. This fact introduces a source of indeterminacy that is peculiar to the multiband case. Indeed, for $m = 1$, from (6) we see that $\mathbf{v}$ can be zero only where the gray-level gradient is zero; for $m > 1$, however, the direction of maximal contrast may be undefined even if the contrast is nonzero. We shall see in the following that the zeros of $\mathbf{v}$ have an important impact on the properties of edge lines; for the moment, however, we give a preliminary definition of edge point.

DEFINITION 1. An *edge point* of a multispectral image is a point P where λ_+, as given by (4), exhibits a *local directional maximum* in the direction $\mathbf{n}_+$ given by (5), provided that such direction is uniquely defined at P.

That Definition 1 is a reasonable one can be seen by resorting again to the monochromatic case ($m = 1$). Taking into account (6), we see that in this case Definition 1 reduces to the standard definition of edge point as the location of a maximum of the gradient magnitude in the direction of gradient.

An example of application of the above ideas is shown in Figs. 1–3. Fig. 1(a,b,c) shows the three channels of a 256 × 256 RGB image. We take as f_i the three color channels, so $m = 3$. The maximum contrast λ_+ has been computed after each channel is smoothed by a circularly symmetric Gaussian with $\sigma = 3$. This rather high value of blurring has been chosen only in order to reduce detail, so avoiding dense and hard-to-read graphs of edge contours without thresholding. Fig. 2(a) shows the values of λ_+, computed using discrete approximations to the required derivatives; we used 3 × 3 masks obtained by least-squares fitting of a quadric to pixel data as suggested in [2]. The edge points are shown in Fig. 2(b). The algorithm used to find edge points proceeds as follows:

- compute, at each pixel, the maximal contrast λ_+ and the corresponding direction $\mathbf{n}_+$, and discard pixels where $\mathbf{n}_+$ is undefined, or λ_+ is below a predefined threshold;
- at each remaining pixel (x, y), fit a least-squares quadric to the values of λ_+ at (x, y) and its eight neighbors, and test whether the fitted quadric has a maximum on a straight line through (x, y) in the direction $\mathbf{n}_+$. The pixel is marked as an edge point if this maximum is within ±0.5 pixels from (x, y).

Figure 2 shows that all visually detectable edges are correctly located by the algorithm, including the face–background edge on the left that presents a strong change in hue but almost no change in intensity. By comparison, Fig. 3 shows the result of applying the algorithm to a gray-level version of the same image (Fig. 1(d)) obtained by taking the arithmetic mean of the three color channels. The face–background edge is practically invisible in the intensity image and, as could be expected, is no longer detected by the algorithm. Also the mirror borders are localized much better in the RGB image than in the black-and-white version.

3. CONNECTIONS WITH ZERO-CROSSING METHODS

A popular second-order differential operator for edge detection in intensity images is the second directional derivative of the smoothed intensity function, denoted as D_2 for brevity [27]. Besides being rotationally invariant (as the Laplacian), D_2 is the most natural choice when edges are defined as extrema of the gradient magnitude.

Indeed, assuming the image to be represented by a sufficiently smooth gray level function, a necessary condition for the gradient magnitude to have a directional maximum at a point P is that $D_2(P)$ have a *transversal zero crossing* at P; i.e., that D_2 be zero at P while having different signs when moving away from P in the gradient direction and in the opposite one.

In order to extend such concepts to the multiband case, we state the following

DEFINITION 2. A *stationary contrast edge point* of a multispectral image is a point P where the first directional derivative of the maximal squared contrast $\lambda_+(P)$ in the direction $\mathbf{n}_+(P)$ of maximal contrast is zero.

Let us denote by $D_S(P)$ the directional derivative mentioned in Definition 2. Its value is found as the inner product of $\mathbf{n}_+$ with the (x,y) gradient of S,

$$\begin{aligned} D_S(P) &= \nabla S(P, \mathbf{n}_+) \cdot \mathbf{n}_+ \\ &= E_x n_1^3 + (2F_x + E_y) n_1^2 n_2 \\ &\quad + (G_x + 2F_y) n_1 n_2^2 + G_y n_2^3, \end{aligned} \tag{9}$$

FIG. 1. Color image. (a), (b), (c): The R, G, B channels, in that order. (d): The intensity image, $I = (R + G + B)/3$.

where the subscript + has been dropped from the components of $\mathbf{n}_+$ for simplicity.

In the derivation of (9) it may seem that we have not considered the variation of λ_+ due to the change in $\mathbf{n}_+$. Actually, using the gradient of λ_+ instead of that of S yields exactly the same result (see the Appendix).

Now, a zero of $D_S(P)$ ensures that the local directional contrast has a *stationary point* at P, but not necessarily an *extremum* (maximum or minimum). To ensure extremality of contrast, we must check whether P is a *transversal zero crossing;* i.e., we must consider the sign of D_S along a curve tangent to $\mathbf{n}_+$ at P.

This task presents some additional difficulties with respect to the analogous problem of finding zero crossings of $D_2(P)$ in the mono case. The problem is that the definition of $\mathbf{n}_+$ as an eigenvector does not uniquely specify its

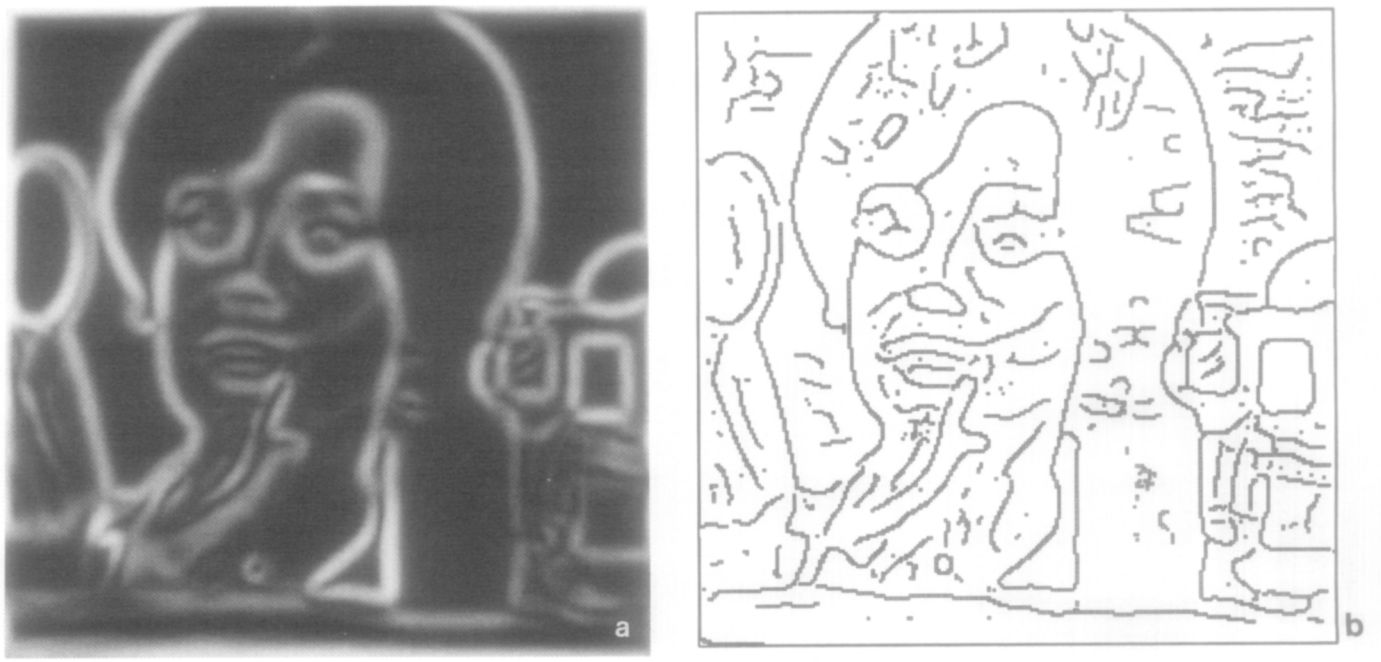

FIG. 2. (a): Maximal contrast of the filtered RGB image, Gaussian filtering with $\sigma = 3$ (log scale). (b): The detected edge points.

sign; indeed, the squared contrast defined by (3) is the same in a direction $\mathbf{n}$ and in the opposite $-\mathbf{n}$. But since D_S is *cubic* in $\mathbf{n}$, the choice of sign is not immaterial. That this problem is real can be seen by observing that θ_+ as defined by (8) exhibits a jump of $\pm\pi$ (corresponding to a change of sign in $\mathbf{n}_+$) whenever F crosses zero with $E - G < 0$.

The problem is therefore: how, and under what conditions, can we *orient* the vector $\mathbf{n}_+$, i.e., choose at any point among $+\mathbf{n}_+$ and $-\mathbf{n}_+$ in a consistent way?

A sufficient condition is given by

THEOREM 1. *The vector $\mathbf{n}_+$ can be uniquely oriented in any simply connected region V where $\mathbf{v} \neq 0$ (provided that the image function is at least C^3).*

Proof. The (x, y) gradient of θ as given by (8) is

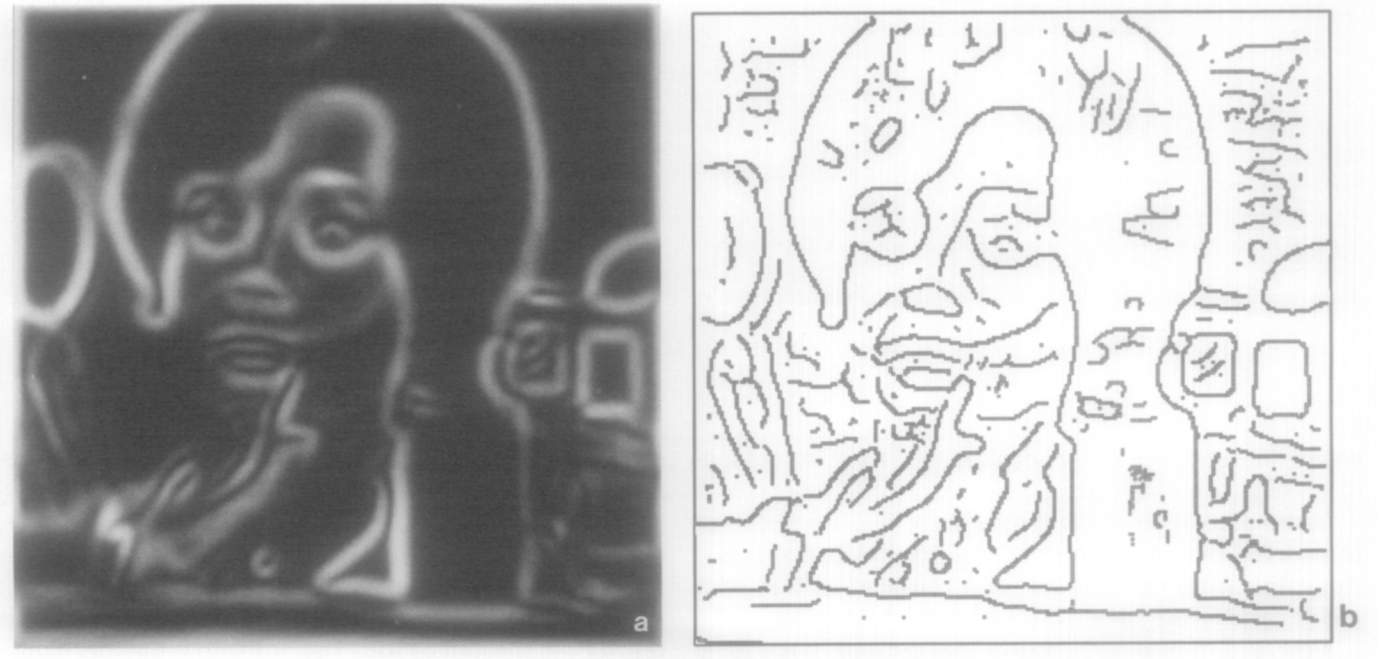

FIG. 3. (a): Maximal contrast of the filtered intensity image, Gaussian filtering with $\sigma = 3$ (log scale). (b): The detected edge points.

$$\nabla\theta = \begin{bmatrix} \partial\theta/\partial x \\ \partial\theta/\partial y \end{bmatrix} = \frac{1}{\|\mathbf{v}\|^2} \begin{bmatrix} (E-G)F_x - (E_x - G_x)F \\ (E-G)F_y - (E_y - G_y)F \end{bmatrix}. \quad (10)$$

Under the condition $f_i \in C^3$, the components of $\nabla\theta$ are continuously differentiable functions of x and y wherever $\mathbf{v} \neq 0$; therefore, from Green's theorem, the line integral of $\nabla\theta$ along any closed curve in V is zero. Given a particular point P_0 in V, we choose a particular orientation of $\mathbf{n}_+$ at P, so defining $\theta(P_0) = \theta_0$. We can then uniquely define $\theta(P)$ for all points P in V as

$$\theta(P) = \theta_0 + \int_\gamma \nabla\theta \cdot ds, \quad (11)$$

where γ is a curve in V joining P_0 to P, and from the preceding analysis, the result will not depend upon the choice of γ. ■

The proof obviously fails if V contains some point Q where $\mathbf{v} = 0$. In this case, $\nabla\theta$ is singular at Q, and its line integral along a closed curve γ surrounding Q may be nonzero. From (5), we see that this integral must be an integer multiple of π; in particular, it may be an *odd* multiple of π, which means that on moving from P_0 along γ and returning to P_0 one finds that $\mathbf{n}_+$ has reversed its sign. This means that D_S must be considered a *two-valued function* on γ, and we must make two turns on γ before achieving the same initial value.

4. ROTATIONAL PROPERTIES OF v

Theorem 1 of the previous section shows that the rotational properties of the vector function $\mathbf{v}$ are of paramount importance in the very definition of the contrast derivative D_S. Such properties are briefly analyzed in the following. We shall consider first a generic two-dimensional vector function $\mathbf{w} = (w_1, w_2)$ defined on some subset of R^2, and we shall suppose that $\mathbf{w}$ is at least of class C^1 (i.e., has continuous first derivatives). Some terminology and some results are borrowed from [19].

DEFINITION 3. The *curvature* of $\mathbf{w}$ is the vector function

$$\mathbf{k} = \frac{\mathbf{W} \cdot \mathbf{w}_\perp}{\|\mathbf{w}\|^2}, \quad (12)$$

where

$$\mathbf{w}_\perp = (-w_2, w_1),$$

so that the pair $\mathbf{w}, \mathbf{w}_\perp$ is positively oriented, and

$$\mathbf{W} = \begin{bmatrix} \partial w_1/\partial x_1 & \partial w_2/\partial x_1 \\ \partial w_1/\partial x_2 & \partial w_2/\partial x_2 \end{bmatrix} = [\nabla w_1, \nabla w_2].$$

This definition is similar to [19, 1.B.v]. Note that the curvature is defined wherever $\mathbf{w}$ is nonzero. Note also that, if $\varphi = \arg(\mathbf{w})$, i.e., the signed angle between the positive x_1 axis and $\mathbf{w}$, then

$$\mathbf{k} = \nabla\varphi. \quad (13)$$

We now consider a smooth, simple closed curve γ oriented counterclockwise and on which $\mathbf{w} \neq 0$.

DEFINITION 4 [19, 1.A.vii]. The *rotation number* of $\mathbf{w}$ on γ is

$$\eta_{\mathbf{w}}(\gamma) = \frac{1}{2\pi} \oint_\gamma \mathbf{k} \cdot ds. \quad (14)$$

THEOREM 2 [19, 1.A.viii]. *The rotation number $\eta_{\mathbf{w}}(\gamma)$ is an integer.*

The proof is almost trivial; from (13), $2\pi\eta_{\mathbf{w}}(\gamma)$ equals the variation of $\arg(\mathbf{w})$ after a complete turn on γ, which must be an integer multiple of 2π since $\mathbf{w}$ is single-valued.

DEFINITION 5. If Z is an isolated zero of $\mathbf{w}$, its *index* $I_{\mathbf{w}}(Z)$ is the rotation number of $\mathbf{w}$ on a simple closed curve γ enclosing Z and no other zero of $\mathbf{w}$.

THEOREM 3. *If $\mathbf{w} \neq 0$ on γ, and the region Γ inside γ contains only isolated zeros of $\mathbf{w}$, then $\eta_{\mathbf{w}}(\gamma)$ equals the algebraic sum of the indices of all zeros in Γ.*

Proof. It suffices to partition Γ into simple regions, each containing only one zero of $\mathbf{w}$, and observe that the integral on γ equals the algebraic sum of the integrals on the boundaries of such regions.

We now specialize the analysis to the case of $\mathbf{v} = (E - G, 2F)$ as defined by (7). We first state the following properties of $\mathbf{v}$, assuming that the components of f are entire functions, which is ensured when using two-dimensional Gaussian filtering.

THEOREM 4. *The loci of $\mathbf{v} = 0$ are either*

Z1: isolated points (almost always), or
Z2: smooth closed curves, or
Z3: smooth curves whose endpoints lie on the image boundary.

Proof. Since $\mathbf{v} = 0$ implies both $E - G = 0$ and $F = 0$, it is clear that the zeros of $\mathbf{v}$ are the intersections of zero-crossing lines of $v_1 = E - G$ and $v_2 = 2F$. Under the above assumption, v_1 and v_2 are entire functions, so their zero-crossing (ZC) loci are either isolated points or smooth curves, either closed or extending to the boundary of the considered region [27]. We may therefore exclude v_1 or v_2 being zero in a finite region unless it is zero everywhere. Therefore, we shall almost always be in case Z1 (intersection of a ZC line of v_1 with one of v_2),

while cases Z2 and Z3 may appear if the two ZC lines coincide, or if one of the components of $\mathbf{v}$ is identically zero. ■

THEOREM 5. *If Z is an isolated zero of $\mathbf{v}$, and $\det(\partial v_i/\partial x_k) \neq 0$ at Z, then $I_{\mathbf{v}}(Z) = \pm 1$.*

A formal proof of Theorem 5 can be found in the Appendix. Its importance stems from

THEOREM 6. *Let γ be a simple closed curve on which $\mathbf{v} \neq 0$. Then, $\mathbf{n}_+$ is uniquely orientable on γ iff $\eta_{\mathbf{v}}(\gamma)$ is even.*

Proof. Since $\arg(\mathbf{n}_+) = \frac{1}{2}\arg(\mathbf{v})$, if $\mathbf{n}_+$ is one-valued on γ then the rotation number of $\mathbf{v}$ is twice the rotation number of $\mathbf{n}_+$. Conversely, suppose that $\eta_{\mathbf{v}}(\gamma) = 2k$; let γ be parameterized by an arbitrary parameter t on [0, 1] so that $P(1) = P(0)$ and $P(t)$ is continuously differentiable. We choose an arbitrary determination of $\theta(P(0))$ from (5), so fixing the sign of $\mathbf{n}_+(P(0))$; using (10), we can now define $\theta(t)$ as

$$\theta(t) = \theta(P(0)) + \int_0^t \nabla\theta(P(u)) \cdot dP(u)$$

$$= \theta(P(0)) + \frac{1}{2}\int_0^t \mathbf{k}_{\mathbf{v}}(P(u)) \cdot dP(u) \qquad (15)$$

so that

$$\theta(1) = \theta(0) + \frac{1}{2}\,2\pi\eta_{\mathbf{v}}(\gamma) = \theta(0) + 2k\pi,$$

i.e., $\mathbf{n}_+(P(1)) = \mathbf{n}_+(P(0))$, and $\mathbf{n}_+$ is then uniquely defined on γ. ■

Theorems 5 and 6 show that isolated zeros of $\mathbf{v}$ have an important impact on the very definition of the contrast derivative. Their connection with extremal edges will be clarified in the next section.

5. EXTREMAL EDGES

The analysis developed in the previous sections allows us now to state more precisely the definition of an extremal edge.

DEFINITION 6. An *extremal contrast edge point* of a multispectral image is a point P where the first directional derivative of the maximal squared contrast λ_+ in the direction $\mathbf{n}_+$ of maximal contrast has a *transversal zero crossing*.

The sensibleness of the above definition can now be easily stated. Indeed, for P to be an extremal edge point we must require that $\mathbf{v}(P) \neq 0$, since otherwise $\mathbf{n}_+$ would not be defined. But the components of $\mathbf{v}$ are smooth functions of P, so we can always find a simply connected open region V containing P in its interior, and in which $\mathbf{v} \neq 0$ everywhere. From Theorem 1, $\mathbf{n}_+$ is uniquely orientable in V, so D_S is a single-valued function in V, whose sign along a curve tangent to $\mathbf{n}_+$ at P is a well-defined quantity.

It is therefore meaningful to consider *extremal edges* as the loci of extremal edge points as defined above. We have

THEOREM 7. *The extremal edges of a multispectral image are either*

E1: smooth closed curves (isolated points may be viewed as a particular case of closed curves shrunken to a single point), or

E2: smooth curve segments whose endpoints either

E21: fall on the boundary of the image, or

E22: are zeros of $\mathbf{v}$.

Proof. Assume that we have found all zeros of $\mathbf{v}$ in the image. Unless $\mathbf{v} = 0$ everywhere, from Theorem 4 we may partition the image into a set of simply connected regions, such that all zeros are either on the image boundary, or on the boundary between two (or more) regions. The directional derivative can now be uniquely defined inside each region V_i, and we may find its zero-crossing contours. From [27] we may say that the zero-crossing contours in V_i are either

- smooth closed curves contained in V_i (so we are in case E1), or
- smooth curve segments ending on the boundary δV_i of V_i.

Now, suppose that a ZC contour γ in V_i has its endpoints at P_1 and P_2 on δV_i (see Fig. 4). Considering, e.g., P_1, we have two cases:

1. $\mathbf{v}(P_1) = 0$; we are in case E22 above;
2. $\mathbf{v}(P_1) \neq 0$. If P_1 is not on the image boundary (case E21), by the argument used in the explanation of Defini-

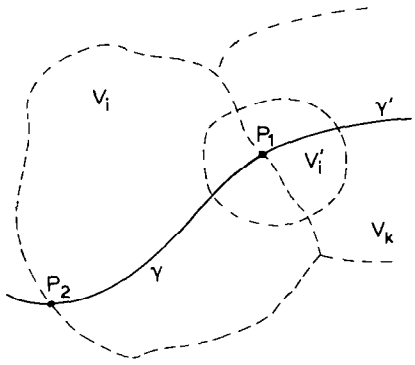

FIG. 4. Zero-crossing contours of the contrast derivative.

tion 6, we may build a new region V_i' containing P_1 and points from both V_i and from the adjacent region(s). From the above analysis, this implies that the portion of γ belonging to V_i' must be part of some smooth ZC contour of V_i' that extends beyond P_1 into some adjacent region V_k. This, in turn, implies that P_1 must be an endpoint of some contour γ' of V_k, and that γ' smoothly joins to γ at P_1.

We can repeat the above arguments on the other endpoint of γ', and so on, until we eventually arrive either at a singular point (case E22), at the image boundary (case E21), or at P_2 (closed contour, case E1). ■

Note that a transversal ZC of $D_S(P)$ does not yet ensure that P is a *maximum* of contrast. To check whether the contrast is maximum or minimum at P, we could test the sign of its second derivative, provided that the image function is sufficiently smooth (i.e., of class C^3). It is, however, more convenient to check how the sign of D_S changes when moving on a curve tangent to $\mathbf{n}_+$ at P; a change from positive to negative, in the direction of $\mathbf{n}_+$, indicates a maximum, while the opposite (negative to positive) says that P is a minimum. The fact that D_S is cubic in $\mathbf{n}_+$ in this case is of help, since reversing the sign of $\mathbf{n}_+$ reverses both the sign of D_S and the direction of travel, and the sequence of signs remains unchanged.

6. AN ALGORITHM FOR EDGE DETECTION WITH SUBPIXEL RESOLUTION

The analysis developed in the preceding section rules out the possibility of finding zero-crossing contours by computing $D_S(P)$ at each pixel and then looking for sign reversals between adjacent pixels. Indeed, whenever the image contains points where $\mathbf{v} = 0$, D_S is not single-valued and we must explicitly indicate the path on which we search for its zero crossings.

In particular, this path may be the contour of a square with vertices at four adjacent pixel locations (see Fig. 5). This observation leads to the following algorithm for finding extremal edges with subpixel resolution. The algorithm is inspired by the one proposed by [10] for the monochromatic case. For each pixel location (x, y), define $C(x, y)$ as the square having vertices at (x, y), $(x + 1, y)$, $(x + 1, y + 1)$, and $(x, y + 1)$, and choose an arbitrary direction of travel on $\delta C(x, y)$ as in Fig. 5. The algorithm then proceeds along the following lines:

A1. Compute $DS(P)$ at each pixel, assuming an arbitrary determination of $\theta(P)$ (i.e., an arbitrary sign of $\mathbf{n}_+(P)$).

A2. Try to orient each square $\delta C(x, y)$. To this end, fix the sign of $\mathbf{n}_+$ at (x, y) and the corresponding value of D_S. Then, traversing δC in the chosen direction, check whether the angle among $\mathbf{n}_+$ at each vertex and $\mathbf{n}_+$ at the preceding one is less than $\pi/2$; if so, proceed, otherwise change the sign of $\mathbf{n}_+$ and of D_S at the considered vertex. If, on return to (x, y) we get the same orientation of $\mathbf{n}$, we may go on with finding zero crossings; otherwise, we mark the square as *nonorientable* and we do not proceed further.

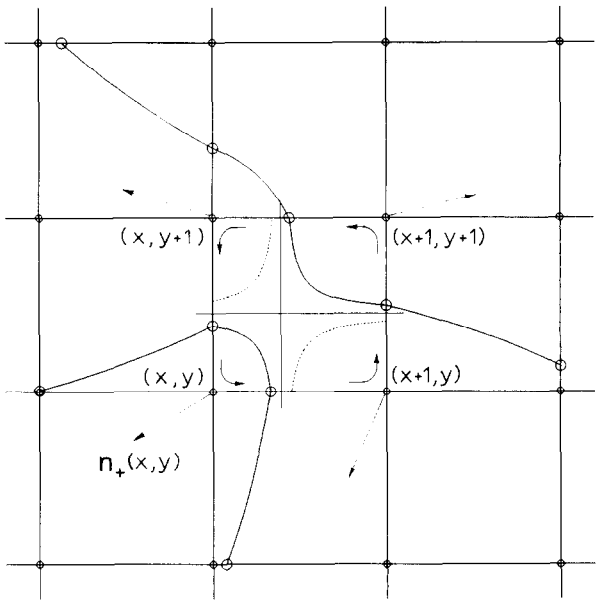

FIG. 5. Orienting the vector $\mathbf{n}_+$ on pixel-square paths and finding zero-crossing lines.

A3. For each orientable square, D_S has now a well-defined value at the corners of $C(x, y)$, so we may proceed to examine sign changes. Ruling out, for the sake of simplicity, the possibility that D_S is exactly zero on one or more of the vertices of C, we have either zero, two, or four sign changes of D_S when traversing δC. If there are sign changes, we may locate zero-crossing points on δC by using a bilinear approximation to D_S inside C:

$$\begin{aligned} D_S(x + u, y + v) \simeq\ & (1 - u)(1 - v)D_S(x, y) \\ & + (1 - u)vD_S(x, y + 1) \\ & + u(1 - v)D_S(x + 1, y) \\ & + uvD_S(x + 1, y + 1) \end{aligned} \quad (16)$$

with $0 \le u \le 1$, $0 \le v \le 1$. The bilinear approximation is useful since it yields not only the locations of ZC on δC, but also the way to connect them in the case of four ZC points, since the zeros of (16) describe an equilateral hyperbola whose orientation may be easily computed from the four corner values. In this way, each square C may yield up to two "edgels" (edge elements); such edgels can then be smoothed out (e.g., by a spline) in order to get smooth edge contours.

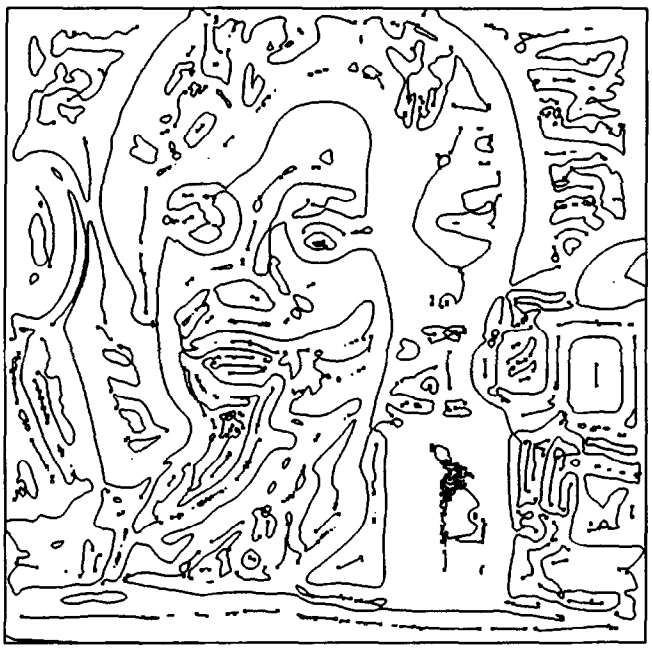

FIG. 6. Extremal edges of the filtered RGB image, drawn to subpixel resolution.

FIG. 7. Maximal edges of the filtered RGB image, drawn to subpixel resolution.

Before commenting out the above algorithm, it may be useful to look at the result of its application to the previously seen RGB image (Fig. 6). In making this figure, we have not used splines, since using straight edgels is sufficient for display purposes.

The nonorientable squares are marked by drawing their contour. Noting that nonorientable squares must contain a zero of $\mathbf{v}$, the ZC contours displayed in Fig. 6 agree with the theoretical behavior predicted by Theorem 7. Some very short contours in the figure seem to violate Theorem 7, as open curves not ending on a nonorientable square, but this is only a matter of graphic resolution; actually, they are very "thin" closed curves.

Figure 7 shows the ZC contours after nonmaximum suppression. The test for maximality is performed by checking whether D_S is decreasing or increasing when crossing the edgel in the direction of $\mathbf{n}_+$. Figure 7 should be compared with Fig. 2(b), where edge points were computed directly as maxima of λ_+ on the pixel grid.

Figure 8 offers a closer look at the area around the intersection of the face–background, hair–background, and hair–face edges on the left of the image. Nonmaximal edgels are shown as dotted lines; the figure also displays the raw vectors $\mathbf{n}_+$ and the corresponding sign of D_S at grid points (a small square on the grid point indicates a positive value of D_S).

The test of orientability used in step A2 of the algorithm can be justified by the fact that, assuming that C does not contain zeros of $\mathbf{v}$, the total rotation of $\mathbf{n}_+$ along δC must be zero. Hence, if the image is sufficiently smooth with respect to pixel size, we may expect the rotation of $\mathbf{n}_+$ on each side of C to be small, and in particular to be less than $\pi/2$. A more accurate test would require to predict the rotation of $\mathbf{n}_+$ on each side from the

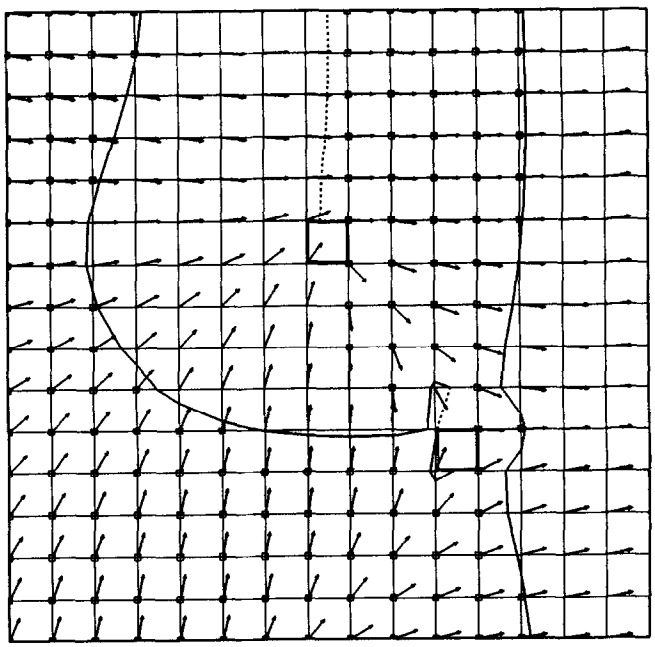

FIG. 8. The vector $\mathbf{n}_+$ and extremal edges in a small area of the image in Fig. 1.

value of $\nabla\theta$, but this would futher increase the computational load. Anyway, the proposed test, together with the bilinear approximation (16), ensures at least that for two adjacent orientable squares the ZC points (if any) on the common side coincide.

Note also that in step A2 above we said that nonorientable squares were not processed any further; this may seem too restrictive, since the non-single-valuedness of D_S along δC does not forbid us to find ZC points on δC, at least if the given square is surrounded by orientable squares. However, the bilinear approximation (16) is obviously not applicable to this case, so we do not know how to connect ZC points with edgels inside C (note that in this case we would have an odd number of ZC points on the contour).

7. CONCLUSIONS

We have tried to extend the theory of edge detection based on second-order differential operators to the multiband case, and our analysis shows that this extension is feasible.

Some results from one-band theory do not extend to this case; for example, the zero-crossing lines are no longer guaranteed to be closed. Although closedness of ZC lines seems to be considered a very useful feature in the mono case [27], from the author's point of view its loss is not too serious. When only *maximal* variations in the image are taken into account, edge lines are usually not closed even in the monochromatic case; closed edges can be obtained only by considering also *nonmaximal* zero crossings (the so called *phantom edges* [7]), whose physical significance is hardly understood.

On the other hand, the possibility of coding simultaneously the significant changes of the image in both luminance and color seems quite promising in view of a symbolic description of the image. Studies are in progress on the problem of labeling edges with quantitative attributes (e.g., the values of the gradients of the image components at the edge) suitable for a faithful description of the image, as proposed in [10, 15]. Moreover, subpixel edge detection could make this approach useful for dealing with image matching problems as those found in stereo and motion analysis.

APPENDIX

A.1. *Proof of Equation (9)*

Let, for simplicity, $\mathbf{n}_+ = (u, v)$, where $u = \cos\theta_+$ and $v = \sin\theta_+$. The directional derivative of λ_+ in the direction $\mathbf{n}_+$ is

$$\nabla\lambda_+ \cdot \mathbf{n}_+ = \frac{\partial\lambda_+}{\partial x}u + \frac{\partial\lambda_+}{\partial y}v$$

$$= \frac{1}{2}\left[E_x + G_x + \frac{(E-G)(E_x - G_x) + 4FF_x}{\sqrt{(E-G)^2 + 4F^2}}\right]u$$

$$+ \frac{1}{2}\left[E_y + G_y + \frac{(E-G)(E_y - G_y) + 4FF_y}{\sqrt{(E-G)^2 + 4F^2}}\right]v.$$

Taking into account the fact that

$$\frac{E-G}{\sqrt{(E-G)^2 + 4F^2}} = \cos 2\theta_+ = 2u^2 - 1 = 1 - 2v^2$$

and

$$\frac{2F}{\sqrt{(E-G)^2 + 4F^2}} = \sin 2\theta_+ = 2uv,$$

substituting and simplifying, one obtains

$$\Delta\lambda_+ \cdot \mathbf{n}_+ = \frac{1}{2}[2E_x u^3 + (2E_y + 4F_x)u^2 v$$

$$+ (4F_y + 2G_x)uv^2 + 2G_y v^3],$$

that is precisely Eq. (9).

A.2. *Proof of Theorem 5*

Before proving Theorem 5, we give the following rule for computing the index of an isolated zero of $\mathbf{v}$.

Let γ be a circle of radius ρ centered on an isolated zero Z of $\mathbf{v}$, and not enclosing any other zero. Then consider the zeros of the components of $\mathbf{v} = (v_1, v_2)$ on γ. Note that, due to smoothness, each zero-crossing line of v_1 and v_2 must cross γ in an even number of points; isolated zeros lying on γ may be considered as limits of closed curves, so they must be counted as double zeros.

Now, label each zero of v_1 of *odd* multiplicity with the symbol e, and each zero of v_2 of *odd* multiplicity with the symbol f. Starting from a point P_0 on γ such that $v_1(P_0) > 0$ and $v_2(P_0) > 0$, build the string s by concatenating the labels of zeros in the order in which they appear when traversing γ counterclockwise. Then simplify the string s by deleting pairs of consecutive e's and f's until s does not contain such pairs any more. After such simplifications the string s will be reduced to either

$$s = (efef)^n \quad \text{or} \quad s = (fefe)^n$$

where $n \geq 0$, and it should be clear that $I_v(Z) = n$ in the first case, and $I_v(Z) = -n$ in the second. Indeed, by considering the sequence of signs of the components of $\mathbf{v}$, one may easily convince himself that a sequence *efef* or

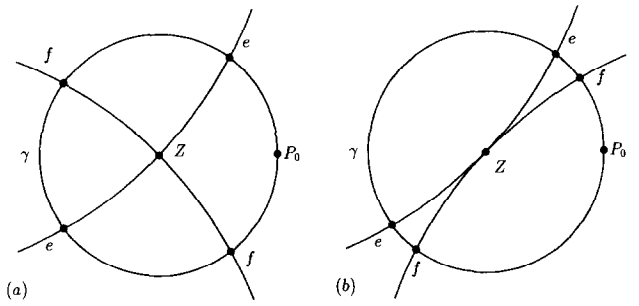

FIG. A1. Computing the index of a zero of **v**.

fefe corresponds to a *full turn* of **v** around the origin, in the positive or negative direction, respectively. Note that P_0 may not be found if one of the components of **v** is always nonpositive, but in this case it should be clear that $I_{\mathbf{v}}(Z) = 0$ since **v** makes no turn around the origin.

We now prove Theorem 5. Let Z be an isolated zero of **v**. Expanding **v** in a Taylor series around Z,

$$v_1(P) = \mathbf{w}_1^\top \mathbf{x} + \tfrac{1}{2}\mathbf{x}^\top \mathbf{H}_1 \mathbf{x} + o(\|\mathbf{x}\|^3)$$
$$v_2(P) = \mathbf{w}_2^\top \mathbf{x} + \tfrac{1}{2}\mathbf{x}^\top \mathbf{H}_2 \mathbf{x} + o(\|\mathbf{x}\|^3),$$

where $\mathbf{x} = P - Z$, $\mathbf{w}_1$ and $\mathbf{w}_2$ are the gradients of v_1 and v_2, respectively, evaluated at Z, and $\mathbf{H}_1$, $\mathbf{H}_2$ are the corresponding Hessians (again at Z). Under our hypothesis, $\mathbf{w}_1$ and $\mathbf{w}_2$ are both nonzero and nonparallel; hence, by the implicit function theorem, the zero-crossing lines of v_1 and v_2 are two smooth curves through Z, intersecting at a nonzero angle as in Fig. A1(a). Therefore, for sufficiently small ρ, the zero pattern on γ will be either *efef* or *fefe*, yielding $I_{\mathbf{v}}(Z) = \pm 1$. ■

Note that if $\mathbf{w}_1$ and $\mathbf{w}_2$ are parallel, it may happen that the situation is as depicted in Fig. A1(b), yielding $I_{\mathbf{v}}(Z) = 0$. Also, in the mono case ($m = 1$), it is easily seen that $\mathbf{w}_1$ and $\mathbf{w}_2$ are zero, since they are linear in the components of the gray level gradient **g**, which is zero at Z. In this case one could consider the quadratic terms in the Taylor expansions; this is, however, unnecessary, since we already know that in the mono case $\arg(\mathbf{v}) = 2\arg(\mathbf{g})$ so that $I_{\mathbf{v}}(Z)$ must be even.

A.3. *Computation of Image Derivatives*

The computation of the maximal squared contrast λ_+ requires the evaluation of the first derivatives of each image component; the computation of D_S, in addition, requires second derivatives. The problem of differentiation of digital images has been extensively considered in the literature. For the example shown in the paper, the required derivatives have been computed by convolving each image component f_i with a set of 3×3 masks that result from local fitting of a second degree polynomial to the data [2]. The use of a second degree polynomial is equivalent to expanding $f_i(x, y)$ in a Taylor series around the considered pixel, and truncating the series at the second derivative terms:

$$f_i(x+p, y+q) \simeq f_i(x,y) + \frac{\partial f_i}{\partial x} p + \frac{\partial f_i}{\partial y} q + \frac{1}{2}\frac{\partial^2 f_i}{\partial x^2} p^2 + \frac{\partial^2 f_i}{\partial x \partial y} p^2 + \frac{1}{2}\frac{\partial^2 f_i}{\partial y^2} q^2 = a_0 + a_1 p + a_2 q + a_3 p^2 + a_4 pq + a_5 q^2.$$

By minimizing, with respect to the a_k, the sum of squares of the differences between the above expression and the actual values of f_i over a 3×3 neighborhood of (x, y), one easily gets estimates of the image derivatives as

$$\frac{\partial f_i}{\partial x} \simeq \frac{1}{6}\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \star f_i, \quad \frac{\partial f_i}{\partial y} \simeq \frac{1}{6}\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} \star f_i,$$

$$\frac{\partial^2 f_i}{\partial x^2} \simeq \frac{1}{3}\begin{bmatrix} 1 & -2 & 1 \\ 1 & -2 & 1 \\ 1 & -2 & 1 \end{bmatrix} \star f_i, \quad \frac{\partial^2 f_i}{\partial y^2} \simeq \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ -2 & -2 & -2 \\ 1 & 1 & 1 \end{bmatrix} \star f_i,$$

$$\frac{\partial^2 f_i}{\partial x \partial y} \simeq \frac{1}{4}\begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix} \star f_i,$$

where the ★ denotes discrete convolution.

ACKNOWLEDGMENTS

The author thanks his colleagues of the Gruppo di Visione Artificiale of IEN and CNR for helpful discussions. This work was partly supported by PROMETHEUS (PRO-ART) project.

REFERENCES

1. M. Amadasun and R. A. King, Low-level segmentation of multispectral images via agglomerative clustering of uniform neighbourhoods, *Pattern Recognition* **21,** 1988, 261–268.
2. P. Beaudet, Rotationally invariant image operators, in *Proceedings, Int. Joint Conf. on Pattern Recognition, 1987,* pp. 579–583.
3. V. Berzins, Accuracy of laplacian edge detectors, *Comput. Vision Graphics Image Process.* **27,** 1984, 195–210.
4. F. Bumbaca and K. C. Smith, Design and implementation of a colour vision model for computer vision applications, *Comput. Vision Graphics Image Process.* **39,** 1987, 226–245.
5. J. Canny, A computational approach to edge detection, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-8,** 1986, 679–698.
6. M. Celenk, A recursive clustering technique for color picture segmentation, in *Proceedings, IEEE Conference on Computer Vision and Pattern Recognition, Ann Arbor, MI, 1988,* pp. 437–444.

7. J. J. Clark, Authenticating edges produced by zero crossing algorithms, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-11,** 1989, 43–57.
8. A. Cumani, *Edge Detection in Multispectral Images,* Technical Report #373, Istituto Elettrotecnico Nazionale G. Ferraris, Turin, 1989.
9. S. Di Zenzo, A note on the gradient of a multi-image, *Comput. Vision. Graphics Image Process.* **33,** 1986, 116–125.
10. P. Grattoni and A. Guiducci, *Image Recovery from Contours,* Technical Report #372, Istituto Elettrotecnico Nazionale G. Ferraris, Turin, 1989.
11. R. M. Haralick, Zero-crossings of second directional derivative edge operator, in *SPIE Proceedings Robot Vision, 1982.*
12. G. Healey and T. O. Binford, A color metric for computer vision, in *Proceedings IEEE Conference on Computer Vision and Pattern Recognition, Ann Arbor, MI 1988,* pp. 10–17.
13. T. L. Huntsberger, C. Rangarajan, and S. N. Jayaramamurthy, Representation of uncertainty in computer vision using fuzzy sets, *IEEE Trans. Computers* **C-35,** 1986, 145–156.
14. G. J. Klinker, S. A. Shafer, and T. Kanade, Color image analysis with an intrinsic reflection model, in *Proceedings, 2nd Int. Conf. on Comp. Vision, Tampa, FL, 1988,* pp. 292–296.
15. A. F. Korn, Toward a symbolic representation of intensity changes in images, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-10,** 1988, 610–625.
16. M. M. Lipschutz, *Differential Geometry,* McGraw–Hill, New York, 1969.
17. D. Lovelock and H. Rund, *Tensors, Differential Forms and Variational Principles,* Wiley, New York, 1975.
18. W. H. H. J. Lunscher and M. P. Beddoes, Optimal edge detector design. I. Parameter selection and noise effects, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-8,** 1986, 164–177.
19. R. Machuca and K. Phillips, Applications of vector fields to image processing, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-5,** 1983, 316–329.
20. D. C. Marr and E. C. Hildreth, Theory of edge detection, *Proc. Roy. Soc. London Ser. B* **207,** 1980, 187–217.
21. R. Nevatia, A color edge detector and its use in scene segmentation, *IEEE Trans. Systems Man Cybernet.* **SMC-7,** 1977, 820–826.
22. Y. Ohta, T. Kanade and T. Sakai, Color information for region segmentation, *Comput. Vision Graphics Image Process.* **13,** 1980, 222–241.
23. J. A. Reimer and P. D. Lawrence, Characterizing $\nabla^2 G$ filtered images by their zero crossings, in *Proceedings ICASSP87, Dallas, TX 1987,* pp. 253–256.
24. A. Sarabi and J. K. Aggarwal, Segmentation of chromatic images, *Pattern Recognition* **13,** 1981, 417–427.
25. A. Shiozaki, Edge extraction using entropy operator, *Comput. Vision, Graphics Image Porcess.* **36,** 1986, 1–9.
26. S. Tominaga, Color image segmentation using three perceptual attributes, in *Proceedings, IEEE Conference on Computer Vision and Pattern Recognition, Miami Beach, FL, 1986,* pp. 628–630.
27. V. Torre and T. Poggio, On edge detection, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-8,** 1986, 147–163.
28. M. M. Trivedi and J. C. Bezdek, Low-level segmentation of aerial images with fuzzy clustering, *IEEE Trans. Systems Man Cybernet.* **SMC-16,** 1986, 589–597.