

Scale-invariant filtering design and analysis for edge detection

Sasan Mahmoodi

Proc. R. Soc. A 2011 **467**, 1719-1738 first published online 19 January 2011
doi: 10.1098/rspa.2010.0475

References

[This article cites 16 articles](#)

<http://rspa.royalsocietypublishing.org/content/467/2130/1719.full.html#ref-list-1>

Subject collections

Articles on similar topics can be found in the following collections

[image processing](#) (11 articles)

[computer vision](#) (3 articles)

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click [here](#)

To subscribe to *Proc. R. Soc. A* go to: <http://rspa.royalsocietypublishing.org/subscriptions>

Scale-invariant filtering design and analysis for edge detection

BY SASAN MAHMOODI*

*School of Electronics and Computer Science, Southampton University,
Building 1, Southampton SO17 1BJ, UK*

Existing edge detection filters work well on straight edges but make significant errors near sharp corners by producing rounded corners. This is due to the fact that the edge maps produced by these filters are scale variant. We enhance Canny's optimality criteria to incorporate detection performance near corners as an explicit design objective. The resulting optimal filter, termed 'Bessel integral filter', can be derived analytically and exhibits superior performance over recent alternatives, both in terms of numerical accuracy and experimental fidelity. A noise-free localization index is also derived here to account for the detection accuracy of discontinuities forming sharp corners in the absence of noise. We prove here that edges detected by the filters that are not optimal with respect to this noise-free localization index are scale variant. However, the Bessel integral filter proposed here is optimal with respect to the noise-free localization index and therefore it is a scale-invariant filter.

Keywords: edge detection; Canny-like criteria; Bessel integral filter; scale-space smoothing; two-dimensional filtering design; scale-invariant filters

1. Introduction

One of the popular approaches in edge detection is to detect discontinuities by convolving a linear filter with an image. It is a common practice in the literature to design a linear filter in one dimension for signals by modelling the edge as a one-dimensional Heaviside function and then extend it for a two-dimensional image by simply adding one dimension to the calculated one-dimensional linear filter. Such an approach in filter design leads to edge detection algorithms, which do not account for two-dimensional features in edges, such as sharp corners and curved discontinuities and therefore may demonstrate a behaviour known as scale-space smoothing causing edges to deform, distort and smooth. Such a behaviour is a consequence of the fact that these filters are scale variant (Lindeberg 1998*a,b*). The amount of feature distortion depends on the size (scale) of the filter. This smoothing effect distorts the original shape of objects in images. It is, therefore, important to design filters to detect edges regardless of the filter size (scale). Our contribution in this paper is to introduce a new framework to design scale-invariant filters for images with two-dimensional features. To this end, we exploit

*sm3@ecs.soton.ac.uk

some basic principles similar to the ones introduced by Canny (1986). A two-dimensional ideal wedge is defined here to model various types of corner-shaped discontinuities. The problem is then reduced to design a linear filter to detect any point on sharp corners forming discontinuities regardless of the wedge angle and orientation. In his seminal paper, Canny (1986) proposes a framework for computational edge detection by introducing certain criteria such as maximum signal-to-noise ratio (SNR) and localization and least multiplicity of response to derive an optimal filter approximated by a derivative of Gaussian (*DroG*) for edge detection. Deriche proposes a computationally more efficient recursive filter with a constant execution time with infinite support to implement Canny's *DroG* in Deriche (1987). Tagare & deFigueiredo (1990) exploit the theory of stochastic processes to model the distribution of zero-crossings to calculate an optimal localization index for *DroG*. An effective width is defined for an indefinite length filter in Sarkar & Boyer (1991) to implement a filter that outperforms *DroG* and is computationally more efficient. Canny-like criteria are employed by Petrou & Kittler (1991) and Petrou (1995) to propose an optimal filter for ramp edges by arguing that edges in real world images are closer to ramps. Infinite symmetric exponential filter is also proposed by Shen & Castan (1992) to demonstrate that such a filter has a better localization property than *DroG* filter. A Wiener filter is proposed in Rao & Ben-Arie (1994) by optimizing discriminative signal-to-noise ratio (DSNR) instead of SNR to demonstrate superior performance in terms of Pratt's figure of merit (Pratt 2007). DSNR is also employed by Wang *et al.* (1996) to propose an optimal filter for bipolar ramp edges demonstrating less distortion and higher figures of merit than the optimal filters for ramp edges proposed in Petrou (1995) and Petrou & Kittler (1991). Demigny (2002) employs Canny's criteria for the optimal design of filters for edge detection in a discrete setting. Canny-like criteria are employed directly in a two dimensional domain by Jacob & Unser (2004) to design optimal steerable filters initially introduced in Freeman & Adelson (1991) by using Gaussian window functions. The technique of scale multiplication is proposed in Bao & Wu (2005) to employ multiple scales to enhance the localization criterion in the *DroG* filter for edge detection. The rest of the paper consists of the following sections. In §2, we initially define the Canny-like criteria, then derive an optimal filter in a two-dimensional domain and finally discuss its mathematical properties. Section 3 deals with numerical evaluations of Canny-like criteria for the optimal filter proposed here and some popular filters in the literature. Experimental results are presented in §4 and finally the paper concludes in §5.

2. Two-dimensional optimal filter and its mathematical properties

Our objective here is to design a filter to detect image discontinuities with numerical cost as low as that of a linear convolution. In an image, such discontinuities may appear as edges along straight lines, edges forming sharp corners and curved edges. In order to account for all kinds of aforementioned edges, let us consider a piecewise constant image (an ideal wedge) shown in figure 1 forming a sharp corner in the centre of coordinate system, point $O(\varphi \neq \pi)$. In this figure, without loss of generality, we can assume that the x -axis is one of the wedge sides to simplify our calculations. Point O can also be considered on

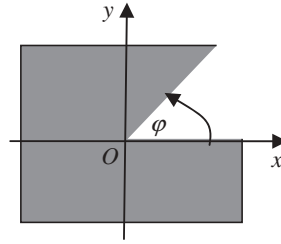


Figure 1. The input image used for the optimal filter calculation.

discontinuity along a straight line for $\varphi = \pi$. It can also be assumed to be on a curved edge for $\varphi = \pi$ by considering point O to be on an infinitesimal line on the curve. Therefore, an optimal filter is sought to detect the discontinuity at point O regardless of the wedge angle and orientation.

(a) Filter derivation

In this paper, by considering an ideal wedge as an input signal shown in figure 1, the first Canny's (1986) assumption that two-dimensional edges locally have a constant cross section in some direction, is violated (e.g. see point O in figure 1). This wedge is defined as a piecewise constant image conveniently described in polar coordinates as

$$g(r, \theta) = (H(\theta) - H(\theta - \varphi)), \quad (2.1)$$

where H , r and θ are the Heaviside step function, radial and angular coordinates, respectively. A two-dimensional image g given in equation (2.1) (or figure 1) is assumed to be the input signal modelling discontinuity with any arbitrary angle and orientation. Let us also assume that the input image is contaminated with a zero-mean Gaussian noise. The detection performance and localization indices (Canny 1986) are then optimized by using the method of Lagrange multipliers to find the optimal filter.

Let us now start by calculating the detection performance and localization indices for point O at the centre of coordinate system $(0,0)$ in figure 1. Our first assumption is that the desired filter is linear and space invariant. We further assume the impulse response is isotropic, i.e. in polar coordinates our desired filter is a function of only r (radial coordinate). In order to calculate the detection performance, the response of the isotropic filter to the discontinuity at point O is calculated by a convolution in Cartesian coordinates evaluated in polar coordinates as

$$U_g|_O = h * g = \int_0^{+\infty} \int_0^{2\pi} g(r, \theta + \pi) h(r) r \, dr \, d\theta, \quad (2.2)$$

where $*$, U_g , g and h are the two-dimensional convolution operation, the convolved image, the ideal wedge defined in equation (2.1) and the desired filter, respectively. To detect the edge in the sharp corner O , we consider the directional derivative of U_g along the unit vector \mathbf{n} with respect to which axis x (in Cartesian coordinates) has a fixed angle θ_n so that $\theta_n \in (0, \varphi)$. This unit vector \mathbf{n} is represented in polar coordinates as $\mathbf{n} = \cos(\theta - \theta_n) \hat{e}_r - \sin(\theta - \theta_n) \hat{e}_\theta$, where \hat{e}_r and \hat{e}_θ are radial and

angular unit vectors in polar coordinates, respectively. The gradient magnitude of the convolved image U_g is examined to find pixels representing discontinuities. Since the relationship between the gradient magnitude and the directional derivative along the unit vector \mathbf{n} at point O is $\partial U_g / \partial n = \nabla U_g \cdot \mathbf{n} = |\nabla U_g| \cos(\alpha)$ where α is a constant angle between \mathbf{n} and $\nabla U_g|_O$, the directional derivative of the convolved image is proportional to its gradient magnitude. The directional derivative of U_g along the unit vector \mathbf{n} at point O in figure 1 is therefore calculated as

$$H_g(0, \theta_n) = \frac{\partial U_g}{\partial n}(0, \theta_n) = \int_0^{+\infty} \int_0^{\vartheta+\varphi+\pi} \frac{\partial h}{\partial n} r \, dr \, d\theta, \quad (2.3)$$

where $\partial h / \partial n = \nabla h \cdot \mathbf{n} = (dh/dr) \cos(\theta - \theta_n)$. By integrating with respect to θ , $H_g(0, \theta_n)$ can be written as

$$H_g(0, \theta_n) = (\sin(\theta_n - \varphi) - \sin(\theta_n)) \int_0^{+\infty} r \frac{dh}{dr} dr. \quad (2.4)$$

Let us now assume that the ideal wedge shown in figure 1 is contaminated with zero-mean Gaussian noise $N(r, \theta)$ with variance N_0^2 . The filter response to noise is therefore written as

$$H_n(0, \theta_n) = \int_0^{+\infty} \int_0^{2\pi} N(r, \theta + \pi) \frac{dh}{dr} \cos(\theta - \theta_n) r \, dr \, d\theta. \quad (2.5)$$

Therefore, the mean-squared response of filter $\partial h / \partial n$ to the noise at point O is $E(H_n^2) = N_0^2 (\int_0^{+\infty} \int_0^{2\pi} r (dh/dr)^2 \cos^2(\theta - \theta_n) dr \, d\theta)$, where $E(\cdot)$ is the statistical expectation. By integrating with respect to θ , the above equation is rewritten as

$$E(H_n^2) = N_0^2 \pi \left(\int_0^{+\infty} r \left(\frac{dh}{dr} \right)^2 dr \right). \quad (2.6)$$

Using equations (2.4) and (2.6), the detection performance is written as

$$\Sigma = \frac{|\int_0^{+\infty} r (dh/dr) dr|}{\sqrt{\int_0^{+\infty} r (dh/dr)^2 dr}}. \quad (2.7)$$

To calculate the localization index, let us now assume that $(\partial H_g / \partial n)(0, \theta_n) = 0$ where \mathbf{n} is the unit vector that has an θ_n angle ($\theta_n \in (0, \varphi)$) with respect to axis x (in Cartesian coordinates). Generally the condition $(\partial H_g / \partial n)(0, \theta_n) = 0$ may not hold for any given isotropic filter. In §2*b*, it is demonstrated that filters which do not meet the abovementioned condition, produce scale-variant edges. For now, we assume that $(\partial H_g / \partial n)(0, \theta_n) = 0$, since we seek to detect point O in figure 1. Therefore, if H_n is the response of the filter $\partial h / \partial n$ to the noise, then the filter response at the presence of noise, should have a local maximum in a point with coordinates (r_0, θ_n) , i.e.

$$\left. \frac{\partial H_n}{\partial n} \right|_{(r_0, \theta_n)} + \left. \frac{\partial H_g}{\partial n} \right|_{(r_0, \theta_n)} = 0. \quad (2.8)$$

Let us now evaluate the individual terms in equation (2.8). The McLaurin's series expansion along the unit vector \mathbf{n} is employed to calculate $(\partial H_g/\partial n)|_{r_0, \theta_n}$ for a small displacement r_0 :

$$\left. \frac{\partial H_g}{\partial n} \right|_{(r_0, \theta_n)} = \left. \frac{\partial H_g}{\partial n} \right|_{(0, \theta_n)} + r_0 \left. \frac{\partial^2 H_g}{\partial n^2} \right|_{(0, \theta_n)} + O(r_0^2). \quad (2.9)$$

As indicated before, we assume $(\partial H_g/\partial n)|_{(0, \theta_n)} = 0$ therefore equation (2.9) for a small displacement r_0 is written as

$$\left. \frac{\partial H_g}{\partial n} \right|_{(r_0, \theta_n)} = r_0 \left. \frac{\partial^2 H_g}{\partial n^2} \right|_{(0, \theta_n)}. \quad (2.10)$$

The terms with higher orders of r_0 are ignored, since the displacement r_0 is assumed small. Equation (2.8) is, therefore, written as

$$\left. \frac{\partial H_n}{\partial n} \right|_{(r_0, \theta_n)} = -r_0 \left(\left. \frac{\partial^2 H_g}{\partial n^2} \right|_{(0, \theta_n)} \right). \quad (2.11)$$

Owing to the presence of zero-mean Gaussian noise, it is reasonable to assume that r_0 is a zero-mean random variable whose variance is calculated as the expectation of r_0^2 by using equation (2.11), i.e.

$$E(r_0^2) = \frac{E[(\partial H_n/\partial n)|_{(r_0, \theta_n)}]^2}{[(\partial^2 H_g/\partial n^2)|_{(0, \theta_n)}]^2}. \quad (2.12)$$

Let us now evaluate the variance of the Gaussian random variable $(\partial H_n/\partial n)|_{(r_0, \theta_n)}$ in polar coordinates. Similar to the calculation method in equation (2.6), the variance of the random variable $(\partial H_n/\partial n)|_{(r_0, \theta_n)}$ can be written as

$$E \left[\left(\left. \frac{\partial H_n}{\partial n} \right|_{(r_0, \theta_n)} \right)^2 \right] = \left(\frac{3\pi}{4} \right) \int_0^{+\infty} \left(r \left(\frac{d^2 h}{dr^2} \right)^2 + \frac{1}{r} \left(\frac{\partial h}{\partial r} \right)^2 \right) dr. \quad (2.13)$$

On the other hand, the denominator in equation (2.12) can be calculated by considering the boundary conditions, $(dh/dr)|_{r=0} = (dh/dr)|_{r=+\infty} = 0$, and therefore

$$\left. \frac{\partial^2 H_g}{\partial n^2} \right|_{(0, \theta_n)} = (\sin^3(\theta_n - \varphi) - \sin^3(\theta_n)) \int_0^{+\infty} \frac{1}{r} \frac{dh}{dr} dr. \quad (2.14)$$

The localization index is the reciprocal of equation (2.12). By using equations (2.12)–(2.14), the localization index can, therefore, be written as equation (2.15)

$$\mathcal{A} = \frac{|\int_0^{+\infty} (1/r)(dh/dr)dr|}{\sqrt{\int_0^{+\infty} (r(d^2 h/dr^2)^2 + (1/r)(dh/dr)^2)dr}}. \quad (2.15)$$

The problem of finding a filter h maximizing the product $\Sigma \mathcal{A}$ evaluated in equations (2.7) and (2.15) can be reduced to extremizing one of the integral terms and keeping the other terms constant in equations (2.7) and (2.15) by using the method of Lagrange multipliers. We also note that the terms in equations (2.7)

and (2.15) do not depend on h itself and consist of only the first and second derivatives of h . To find the optimal filter in the space of admissible functions, the functional to extremize can, therefore, be written with respect to $P = dh/dr$, i.e.

$$E(P, P_r) = \int_0^{+\infty} \Gamma(P, P_r, r) dr = \int_0^{+\infty} \left(\gamma rP + \alpha rP^2 + \beta \left(rP_r^2 + \frac{P^2}{r} \right) + \frac{\omega}{r} P \right) dr, \quad (2.16)$$

where $\Gamma(P, P_r, r)$ is the Lagrangian, $\gamma, \alpha, \beta, \omega$ are constant coefficients and $P_r = dP/dr$. We therefore arrive at the Euler–Lagrange equation:

$$rP_{rr} + P_r - \left(\alpha' r + \frac{1}{r} \right) P - \gamma' r - \frac{\omega'}{r} = 0, \quad (2.17)$$

where $\alpha' = \alpha/\beta$, $\gamma' = \gamma/\beta$ and $\omega' = \omega/\beta$. By assuming that $r \neq 0$, the above equation is rewritten as

$$r^2 P_{rr} + rP_r - (\alpha' r^2 + 1)P - \gamma' r^2 - \omega' = 0. \quad (2.18)$$

Equation (2.18) is of the type of modified Bessel differential equation (Polyanin & Zaitsev 2003). Solutions for differential equation (2.18) exist, only if $(\gamma/\alpha) = (\omega/\beta)$. The necessary condition for functional (2.16) to have a minimizer is that its second variation is positive. The second variation for functional (2.16) can then be calculated as

$$\delta^2 E = \int_0^{+\infty} (\Gamma_{PP} q^2 + \Gamma_{P_r P_r} q_r^2) dr = \int_0^{+\infty} \left(\left(2\alpha r + \frac{2\beta}{r} \right) q^2 + 2r\beta q_r^2 \right) dr, \quad (2.19)$$

where $\Gamma_{PP} = (\partial^2 \Gamma / \partial P^2)$, $\Gamma_{P_r P_r} = (\partial^2 \Gamma / \partial P_r^2)$, $\Gamma_{PP_r} = (\partial^2 \Gamma / \partial P \partial P_r) = 0$, $q(r) = \delta P(r)$ and $q_r = dq/dr$ are from the space of admissible functions. The second variation calculated above is positive if inequalities (2.20) are satisfied for $\forall r \in R^+$:

$$\alpha r + \frac{\beta}{r} > 0 \quad \text{and} \quad 2r\beta > 0. \quad (2.20)$$

In order for inequalities (2.20) to be true for any $r \in R^+$, it is required that $\alpha > 0$, $\beta > 0$. The general solution of equation (2.18) is therefore written as $P = AK_1(r\sqrt{\alpha'}) + BI_1(r\sqrt{\alpha'}) + C$, where $I_1(r)$ and $K_1(r)$ are modified Bessel functions of the first and second kind and degree one (Polyanin & Zaitsev 2003) and A, B and C are constants, respectively. It is noted that $\alpha' = \alpha/\beta$ is positive since both α and β are positive. This solution is subject to the boundary conditions $P(+\infty) = 0$ and $P(0) = 0$. We also note that $P(r)$ is required to be asymmetric along the unit vector \mathbf{n} . Therefore, $P(r)$ along the unit vector \mathbf{n} is written as

$$P(r) = \begin{cases} -K_1(r\sqrt{\alpha'}) & r > 0 \\ K_1(r\sqrt{\alpha'}) & r < 0. \\ 0 & r = 0 \end{cases} \quad (2.21)$$

Figure 2a depicts $P(r)$ for $\alpha' = 0.2$ along the unit vector \mathbf{n} . It is clear that as $r \rightarrow 0$, $P(r)$ approaches to infinity. For implementation purposes, this is not computationally tractable. Therefore, a regularization procedure is proposed here

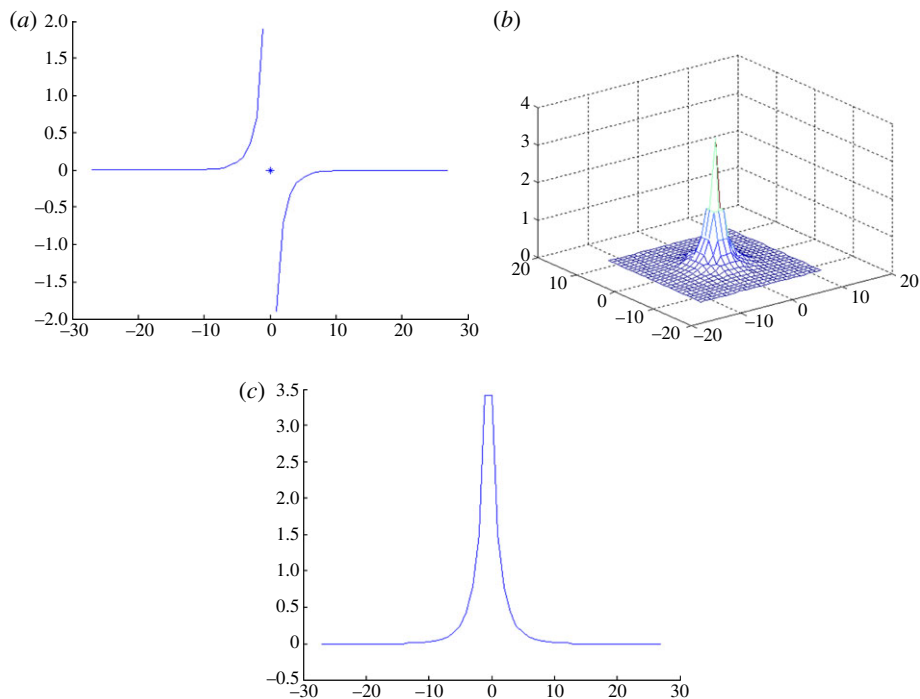


Figure 2. (a) The Bessel function $P(r)$ along the unit vector; (b) a three-dimensional view of the optimal isotropic filter h ; (c) the cross section of the filter. (Online version in colour.)

as follows:

$$P_{\varepsilon}(r) = \begin{cases} -K_1(|r|\sqrt{\alpha'}) & r \geq \varepsilon \\ K_1(|r|\sqrt{\alpha'}) & r \leq -\varepsilon. \\ 0 & |r| < \varepsilon \end{cases} \quad (2.22)$$

The isotropic filter h is then calculated by integrating the above function with respect to r .

$$h_{\varepsilon}(r) = \int_{-\infty}^r P_{\varepsilon}(q) dq. \quad (2.23)$$

In this paper, we refer to isotropic filter (2.23) as *Bessel Integral* filter. Figure 2*b,c* shows a three-dimensional view of the isotropic filter h and its cross section for $\alpha' = 0.2$ and $\varepsilon = \Delta = 1$, respectively, where Δ is the sampling distance.

(b) Mathematical analysis for edge detection filters and algorithms

In the first part of this section, we present some mathematical properties of the Bessel Integral filter. In the second part of §2*b*, we describe general properties of scale-variant and invariant filters to demonstrate that the Bessel integral filter belongs to a family of scale-invariant filters. In the following theorem, we initially prove that the local extrema of the gradient magnitude of the convolved image calculated in equation (2.2) correspond to the discontinuities of the piecewise constant image g .

Theorem 2.1. *Let $U_g: \Omega \rightarrow R^+$ (Ω is the image domain) be the image obtained in equation (2.2) by convolving the regularized function h_ε in equation (2.23) and a piecewise constant image g shown in figure 1. In a neighbourhood of point O , if the gradient magnitude of U_g has a local extremum in (r_0, θ_0) so that r_0 is a small displacement from point O and $\theta_0 \neq 0, \varphi$, then this local extremum approaches point O , as $\varepsilon \rightarrow 0$.*

A proof for theorem 2.1 is provided in appendix A. In theorem 2.2 presented below, we prove that the gradient magnitude of the image convolved by the input image and the isotropic filter h is infinity on the discontinuities of the original image when the regularizing parameter ε approaches zero.

Theorem 2.2. *The gradient magnitude of the convolved image $U_g: \Omega \rightarrow R^+$ between the regularized isotropic function h_ε calculated in equation (2.23) and a piecewise constant image g approaches infinity on the discontinuities of the input image g , if $\varepsilon \rightarrow 0$.*

A proof for the above theorem is presented in appendix B. Proposition 2.3 presented below readily follows from theorems 2.1 and 2.2. This proposition is the basis for an edge detection algorithm by finding the local maxima of the gradient magnitude of the input image convolved with the Bessel integral filter.

Proposition 2.3. *The gradient magnitude of the convolved image $U_g: \Omega \rightarrow R^+$ between the regularized isotropic filter h_ε calculated in equation (2.23) and a piecewise constant image g has local maxima on the discontinuities of the input image g , if $\varepsilon \rightarrow 0$.*

Given that the gradient magnitude is always positive, the proof of proposition 2.3 is straightforward and is therefore skipped here. According to proposition 2.3, if filter (2.23) is convolved with any input image, the local maxima of the gradient magnitude of the convolved image correspond to discontinuities of the input image. Proposition 2.3 is therefore a fundamental statement for our discontinuity detection algorithm proposed in this paper. The result obtained from theorem 2.1 leads also to proposition 2.4. This proposition rationalizes the assumption made to arrive at equation (2.10) from equation (2.9) for the Bessel integral filter.

Proposition 2.4. *For the Bessel integral filter and the input image g shown in figure 1, equation (2.9) approaches equation (2.10) as $\varepsilon \rightarrow 0$ if r_0 is a small displacement.*

The proof of this proposition is straightforward and immediate from the result obtained from theorem 2.1 and is therefore skipped. According to proposition 2.4, localization calculated in equation (2.15) is only true for the Bessel integral filter. However, the localization index for any other isotropic edge detection filter is calculated by using equations (2.8) and (2.9). From these equations, the reciprocal of the square root of variance of r_0 being the localization index is calculated as

$$A = \frac{|(\partial^2 H_g / \partial n^2)_{(0, \theta_n)}|}{\sqrt{E[(\partial H_n / \partial n)_{(r_0, \theta_n)}]^2 + ((\partial H_g / \partial n)_{(0, \theta_n)})^2}}. \quad (2.24)$$

In the absence of noise, a noise-free localization index is derived as

$$A_{\text{NF}} = \left| \frac{(\partial^2 H_g / \partial n^2)|_{(0, \theta_n)}}{(\partial H_g / \partial n)|_{(0, \theta_n)}} \right|. \quad (2.25)$$

Equations (2.24) and (2.25), theorem 2.1 and proposition 2.4 lead to lemma 2.5.

Lemma 2.5. *Let the function $h(r) : R^+ \rightarrow R$ continuous in $(0, +\infty)$, be a two-dimensional isotropic edge detection filter in a polar coordinate system. If the central value of such a filter $h(0)$ is bounded, the two-dimensional noise-free localization index calculated in equation (2.25) for the wedge angle $\varphi \neq \pi$, has a bounded value.*

Lemma 2.5 is proved in appendix C. Lemma 2.5 in fact indicates what kind of filters produces scale-variant edges and it therefore leads to lemma 2.6 describing what property a filter should possess to be able to produce scale-invariant edges. In the proof of lemma 2.5, it is demonstrated that equation (2.25) for an isotropic filter $h(r)$ (r is the radial coordinate in the polar coordinate system) can be written as

$$A_{\text{NF}} = \left| \frac{(\sin^3(\theta_n - \varphi) - \sin^3(\theta_n)) \int_0^{+\infty} (1/r)(dh/dr) dr}{h(0)(\sin(2(\varphi - \theta_n)) + \sin(2\theta_n))} \right|. \quad (2.26)$$

It can also be verified that for a Gaussian filter and in the absence of noise, the localization index calculated in equation (2.26) is proportional to the reciprocal of the standard deviation (scale) of the Gaussian filter, i.e. if the scale of the Gaussian filter increases, the noise-free localization of the detected edge corresponding to the corner point O in figure 1 decreases. This phenomenon observed in the case of images with discontinuities containing sharp corners such as the one shown in figure 1 is related to the scale variance property of a Gaussian filter. The conditions under which an isotropic filter detects scale-invariant edges are indicated in lemma 2.6 concluded from the results obtained from theorem 2.1 and lemma 2.5.

Lemma 2.6. *Let the function $h(r) : R^+ \rightarrow R$ continuous in $(0, +\infty)$, be a two-dimensional isotropic edge detection filter in a polar coordinate system. The localization index in the absence of noise for such a filter is infinity regardless of the value of wedge angle and its orientation as shown in figure 1, if $h(r)$ approaches infinity as $r \rightarrow 0$.*

The proof of this lemma is similar to the proof of theorem 2.1 and can be given by evaluating the localization index calculated in equation (2.26) in lemma 2.5 as $r \rightarrow 0$ and it is therefore skipped in this paper. Lemma 2.6 can be considered as a direct consequence of proposition 2.3 for the Bessel integral filter as a member of a family of filters for which this lemma is applicable.

3. Numerical evaluations of detection performance and localization

In this section, the detection performance and localization indices are numerically evaluated with respect to the parameters of the Bessel integral filter derived in the previous section. These indices are evaluated here to compare with some existing

filters in the literature. The filter $P_\varepsilon(r)$ regularized in equation (2.22) is employed to numerically evaluate indices with respect to the regularizing parameter ε and $\mu = 1/\alpha'$. The parameter μ is used here instead of α' for convenience. The detection performance and localization indices are calculated according to Canny (1986). Figure 3 depicts the detection performance and localization indices and their product with respect to μ and the regularizing parameter ε . As shown in this figure, by lowering the regularizing parameter ε , the detection performance index decreases as depicted in figure 3*a*. The localization index greatly improves as ε is lowered (figure 3*b*). The product of these two indices improves as ε is lowered (figure 3*c*). We also notice that the ability of the filter for noise removal increases by raising the values for μ (and as a result the filter size). The localization, on the other hand, almost remains unchanged, if μ is increased (or decreased) as demonstrated in figure 3*b*. We now need to determine the window size according to which the regularized Bessel integral filter is truncated, since this filter in spatial domain is not band limited. Therefore, we define the truncated filter as

$$P_\varepsilon(r, d) = \begin{cases} 0 & 0 < r < \varepsilon \\ K_1\left(\frac{r}{\sqrt{\mu}}\right) & \varepsilon < r < d. \\ 0 & r > d \end{cases} \quad (3.1)$$

The normalized truncation error $E(d)$ defined in equation (3.2) as

$$E(d) = \frac{\int_0^{+\infty} |P_\varepsilon(r) - P_\varepsilon(r, d)| dr}{\max(\int_0^{+\infty} |P_\varepsilon(r) - P_\varepsilon(r, d)| dr)}. \quad (3.2)$$

We notice that the normalized truncation error $E(d)$ for $\varepsilon \in [0.01, 1]$ depends only on $d/\sqrt{\mu}$. For $d/\sqrt{\mu} = 4$, the error term E becomes negligible ($E = 0.0265$). Therefore, in this paper, the window size of the filter proposed here is set to $W = 8\sqrt{\mu} \times 8\sqrt{\mu}$. Figure 4 depicts the detection performance, localization indices and their product with respect to the filter size for *DroG*, Demigny and Bessel integral filters with $\varepsilon = 0.01, 0.1$ and 1 . As shown in the figure, the localization of the filter proposed here increases and its detection performance decreases, as lower values for ε are selected, which confirms the result shown in figure 3. The localization of the Bessel integral filter for all values of ε is superior to the localization of the other two filters (figure 4*b*). Finally, figure 4*c* depicts the product of the detection performance and localization. The Bessel integral filter with $\varepsilon = 0.01$ has the best overall performance (product of indices) over the whole range of filter sizes.

4. Experimental results

Before the experimental results are discussed in this section, there are two implementation issues that require more clarification. Firstly, if the regularizing parameter less than unity (e.g. as low as $\varepsilon = 0.01$) is required anywhere in this paper, the filter is initially constructed in a higher resolution (e.g. $\Delta = 0.01$) with the required ε and then down-sampled to the normal resolution $\Delta = 1$. Secondly,

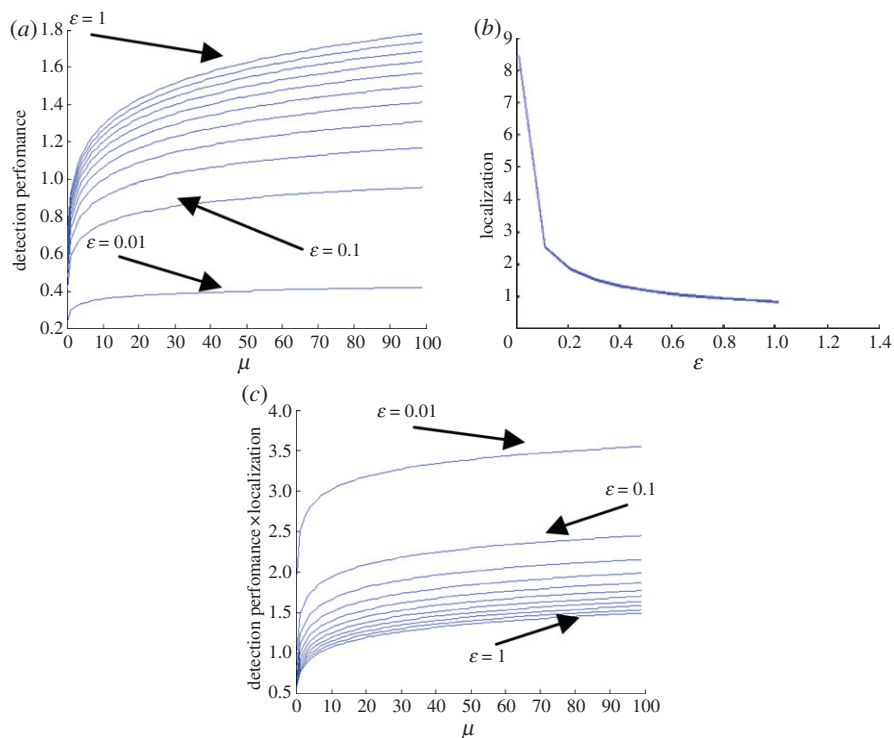


Figure 3. The detection performance and localization with respect to a range of values for μ and regularizing parameter ϵ . (a) Detection performance index; (b) localization index; (c) the product of these two indices. (Online version in colour.)

similar to other filters, the Bessel integral filter is also convolved with the original image, and then the non-maximum suppression algorithm is applied on the gradient magnitude of the convolved image to detect discontinuities. Let us now compare the algorithm proposed here with *DroG* (Canny 1986), Demigny (2002) and Jacob & Unser (2004) edge detection algorithms. The three Jacob–Unser filters employed in this paper are a wedge filter with wedge angle $\pi/2$, and two third-order edge detectors with filter smoothing parameters 0.09 and 0.2 (named as edge 1 and edge 2 filters, respectively, in this paper) (see Jacob & Unser (2004) table 1 for more details). The six filters are applied on the synthetic star image shown in figure 6 to compare the accuracy (localization) of the detected edges. To measure the distance between each detected edge pixel from the corresponding real edge pixel on the original image, we use signed distance function (SDF) (Sethian 1996) of the original image of figure 6. To compute an error term representing the average error between the true and the detected edge pixels, let us assume that $z = \phi(x, y)$ is the SDF of the synthetic star image calculated by using the fast marching algorithm (Sethian 1996) and x_i and y_i are the x and y coordinates of the i th edge pixel of an edge map produced by one of the algorithms investigated here. The average error term associated with the edge map containing N_{edge} pixels is measured by $1/N \sum_{i=1}^N |\phi(x_i, y_i)|$. Same filter size

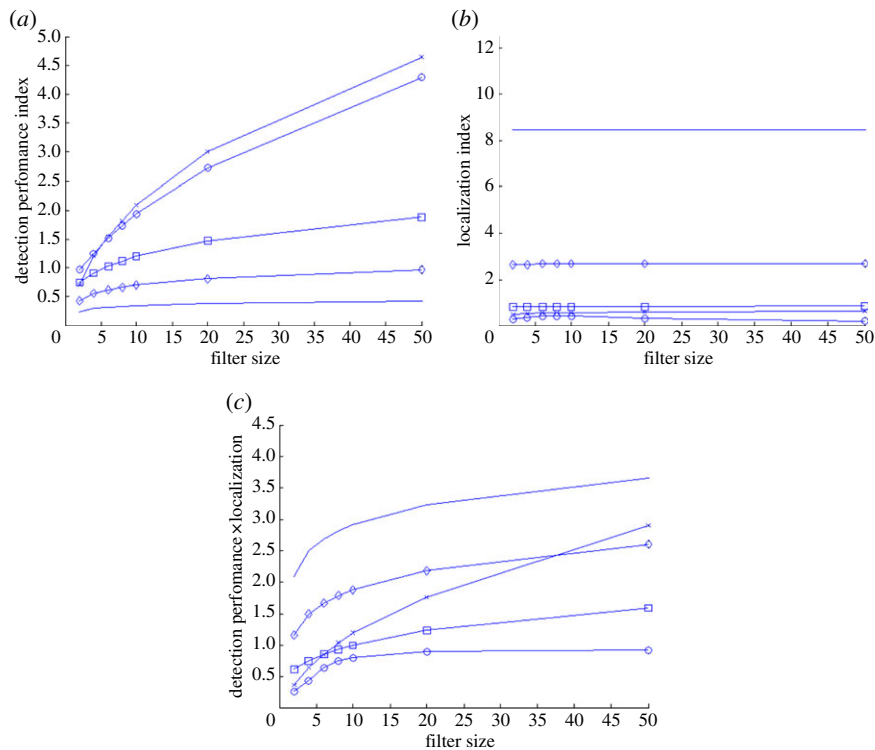


Figure 4. Detection performance and localization indices with respect to various filter sizes for Demigny, *DroG* and Bessel integral filters with $\varepsilon = 0.01, 0.1$, and 1 . (a) Detection performance index, (b) localization index and (c) product of these indices. (a–c) Thin line, Bessel $\varepsilon = 0.01$; diamonds with solid line, Bessel $\varepsilon = 0.1$; squares with solid line, Bessel $\varepsilon = 1$; crosses with solid line, Demigny; circles with solid line, Canny. (Online version in colour.)

is used for all filters in each measurement. Figure 5 shows the average error term for each algorithm for filter sizes 5, 9, 13, 17, 21, 41 and 101. By changing the filter sizes for each measurement, other filter parameters are also changed accordingly. For example, for *DroG* and Jacob–Unser algorithms, standard deviation = filter size/8, and for the Bessel integral filter $\mu = (\text{filter size})^2/64$ are calculated. As depicted in figure 5, the average error term increases for all scale-variant filters. However, the average error of edge detection for the Bessel integral filter is unchanged and lowest as implied from theorems 2.1, 2.2, proposition 2.3 and lemma 2.6 proved in §2. The edge maps detected by the six filters for the filter size = 41 is depicted in figure 6. As shown in this figure, the sharp corners are detected properly only with the Bessel integral filter regardless of filter size and wedge angle. The other filters with the same size have smoothed the sharp corners. We notice that the central processing unit (CPU) time required to produce edge maps is the same for all filters with the same size. Because all edge detection algorithms investigated here are based on the calculation of the gradient of the convolved image with a linear filter. For example, it takes around 0.1 s for a PC with a 2.6 GHz CPU to run every algorithm discussed in the

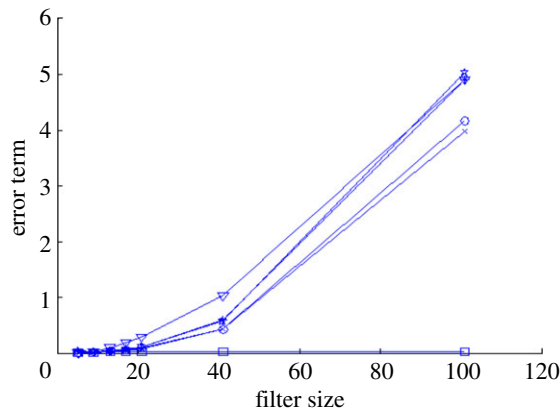


Figure 5. The average error between the detected edges and real edges in the ‘Star’ image with respect to filter size for the filters discussed in this paper (see the text for more details). Squares with solid line, Bessel $\epsilon = 1$; crosses with solid line, Demigny; circles with solid line, Canny; inverted triangles with solid line, Jacob–Unser edge 1; stars with solid line, Jacob–Unser wedge; asterisks with solid line, Jacob–Unser edge 2. (Online version in colour.)

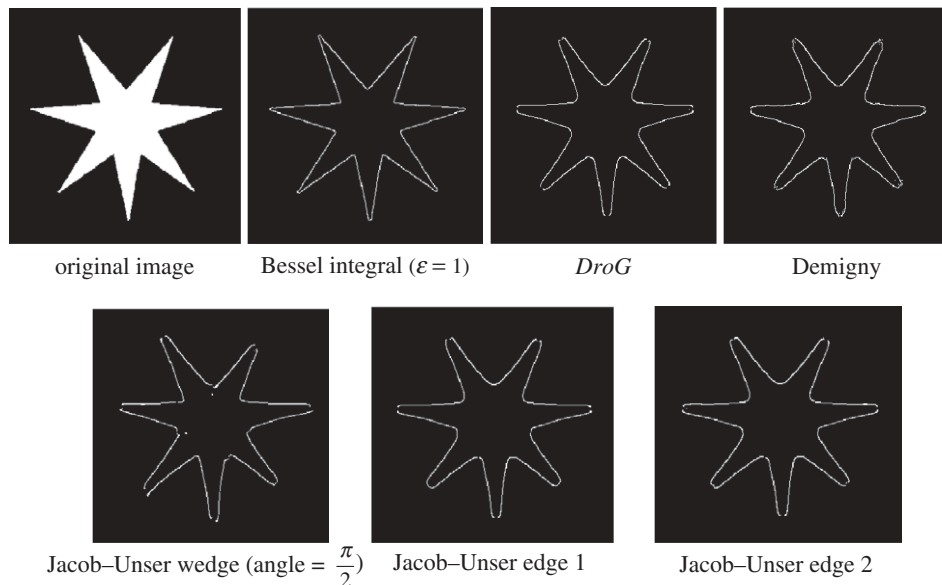


Figure 6. The edge maps of the ‘Star’ image (320×320) detected by the filters investigated in this paper with filter size = 41×41 .

section with filter size 41×41 in a MATLAB environment (v. 7.3) on the original image of figure 6 for edge detection. In order to choose equivalent threshold levels in all algorithms investigated in this paper for a fair comparison, a strip of logarithmically increasing grey scales is added to the bottom of the noisy and real world images used here. This strip of grey scales named as threshold scaling strip

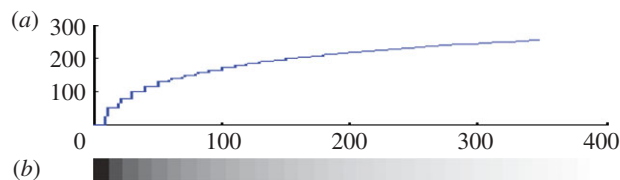


Figure 7. Threshold scaling strip, (a) a cross section of the strip; (b) the strip in grey scale. (Online version in colour.)

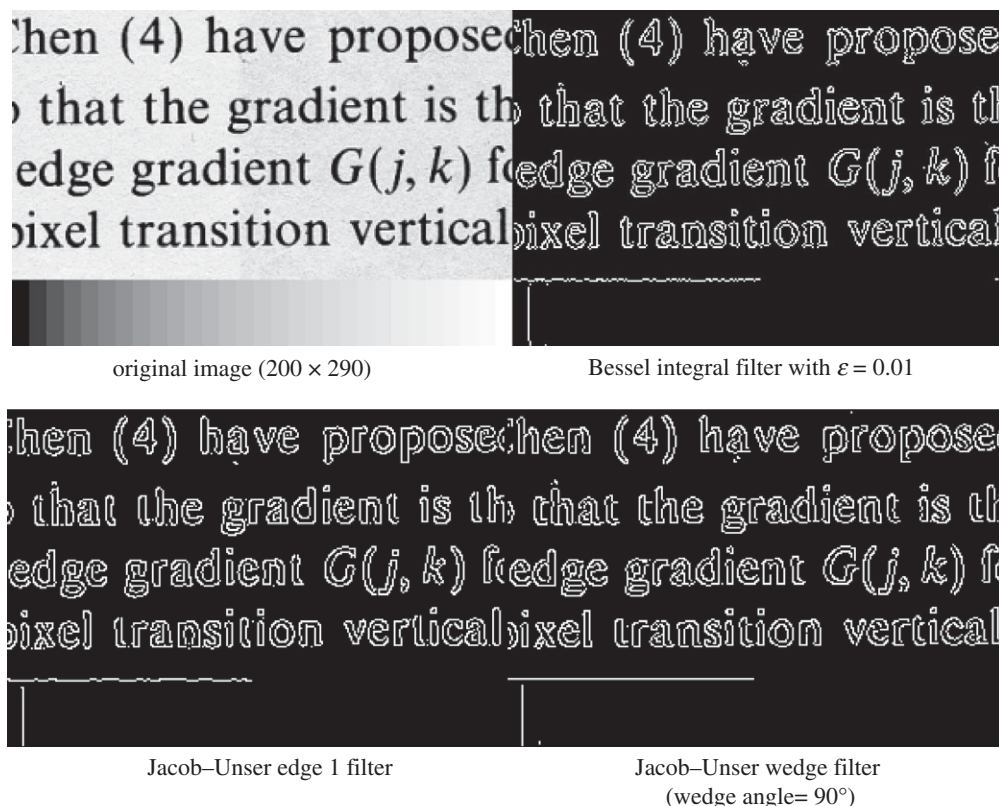


Figure 8. Edge map produced by Bessel integral filter with $\varepsilon = 0.01$ and Jacob–Unser edge 1 filter and Jacob–Unser wedge filter with wedge angle = 90° (filter size = 9×9).

in this paper and its cross section are shown in figure 7. The threshold levels in all algorithms are adjusted so that the same number of discontinuities in the strip is detected by all algorithms. Figure 8 shows a real world image of a text book. Noise is introduced to this image in the data acquisition process. The Bessel integral, Jacob–Unser edge 1 and Jacob–Unser wedge filters (size = 9×9) are applied to the image of this figure. The threshold levels for all algorithms are adjusted so that the first edge in the threshold scaling strip is detected. A close inspection of this figure indicates that the edge map produced by the Bessel integral filter has the best accuracy (e.g. see ‘t’ in ‘that’, ‘transition’ and ‘gradient’, ‘r’ in ‘gradient’

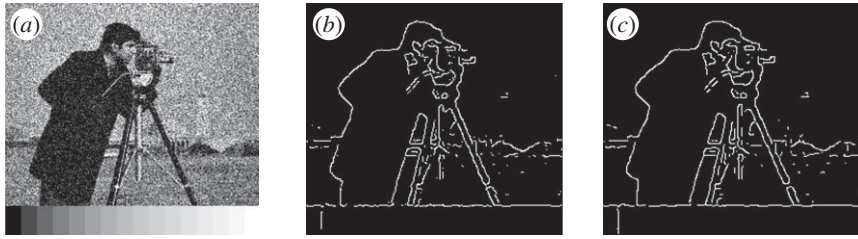


Figure 9. Comparison of edge detection performance in a noisy environment. (a) Original noisy image of Cameraman (256×256), (b) Bessel integral filter ($\epsilon = 3.4$) and (c) Jacob–Unser edge 1 filter (filter size = 13×13).

and ‘propose’, and so on). As the filter size increases, the scale-space distortion of Jacob–Unser filters becomes more severe. However, the Bessel integral filter experiences less scale-space distortion. The Bessel integral and Jacob–Unser edge 1 filters (filter size = 13×13) are applied to the noisy image of Cameraman as shown in figure 9a. This image is contaminated with the zero-mean Gaussian noise with standard deviation 50. The edge maps produced by the Bessel integral and Jacob–Unser filters are depicted in figure 9b and c, respectively. As shown in this figure, the Bessel integral filter produces competitive results in comparison with Jacob–Unser filter in the presence of excessive noise. Soft edge maps, as suggested in Martin *et al.* (2004), are produced by applying Bessel integral and Jacob–Unser edge 1 filters (with size 13) on a real world image from Berkeley segmentation data base (Martin *et al.* 2004), as shown in figure 10a. There are many details that are detected with higher accuracy by Bessel integral filter. For example, number ‘96’ on the front car and the signs ‘CASTROL’ in the background (figure 10a) are completely readable in figure 10b. However, it is very difficult (if not impossible) to read these details in the soft edge map produced by Jacob–Unser filter as shown in figure 10c. The source code of the edge detection algorithm proposed in this paper is written in MATLAB (v. 7.3) and is available in <http://users.ecs.soton.ac.uk/sm3/curricul2.htm>.

5. Conclusion

A scale-invariant filter termed as Bessel integral filter is derived here by defining the Canny-like criteria in a two-dimensional image domain. The Bessel integral filter proposed here is singular at the centre. A regularization method for the filter is therefore proposed here to improve the robustness of this edge detection method in the presence of noise. It is demonstrated that the product of the detection performance and localization indices of the filter proposed here is higher than some recent filters in the literature. It is also analytically proved in this paper that the gradient amplitude of the filtered image has local maxima on the discontinuities of the input image, when the regularizing parameter of the Bessel integral filter approaches zero. This property of the Bessel integral filter is the main reason for its scale-invariance property in edge detection. The condition under which a filter may be scale variant or invariant is presented here. Two

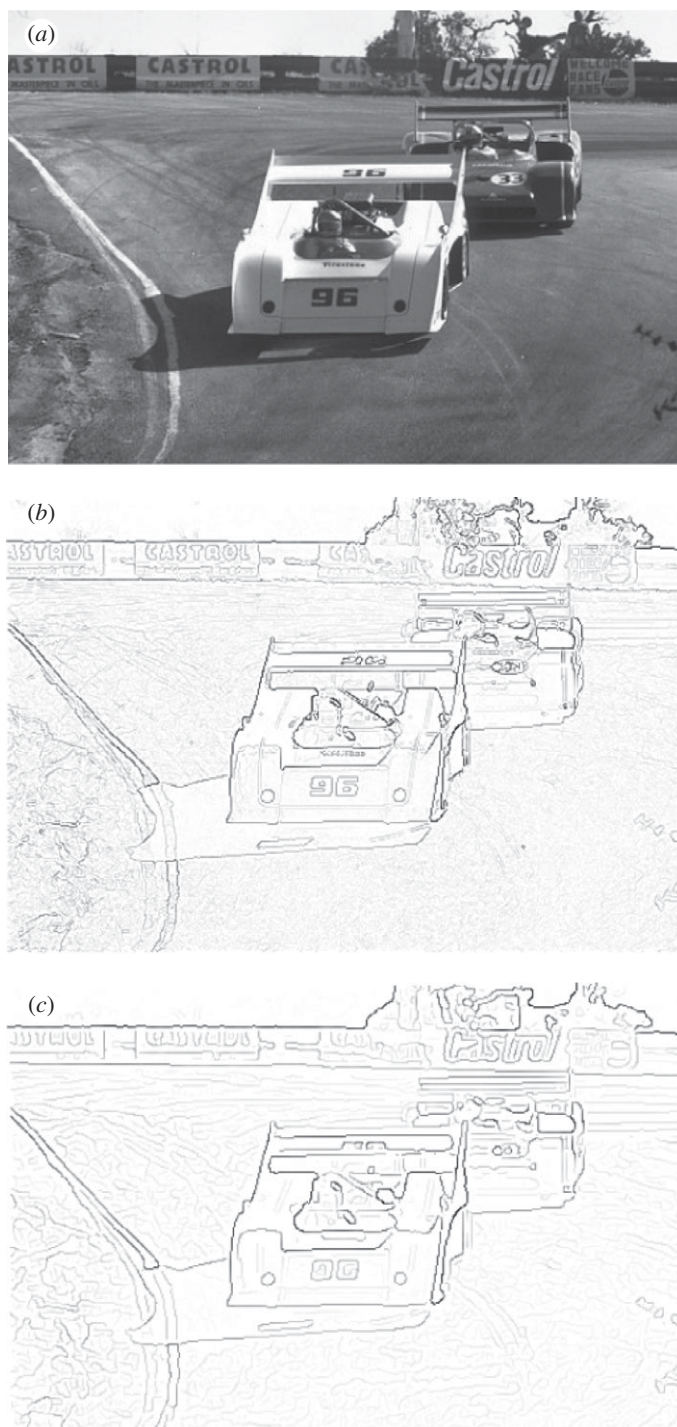


Figure 10. Bessel and Jacob–Unser edge 1 filters (with filter size 13) are employed to produce soft edge maps for a real world image from Berkeley University data base. (a) Original real world image (480×320). (b) Soft edge map produced by Bessel filter with $\varepsilon = 0.01$. (c) Soft edge map produced by Jacob–Unser edge 1 filter.

issues, not discussed in this paper, are open for further research: (i) the possibility of employing the filter proposed here in a class of steerable filters as in Jacob & Unser (2004); (ii) the filter proposed here is designed in a continuous domain. A discrete treatment as in Demigny (2002) for the filtering design introduced here would lead to a discrete filter equivalent to the Bessel integral filter derived here in a continuous domain.

This work was supported in part by the IST program of the European Community, under the PASCAL2 Network of Excellence, the IST-2007-216886 and PinView project with grant number 216529.

Appendix A

Proof for theorem 2.1. Let us consider the piecewise constant image g as shown in figure 1. Let us further assume that the local extremum of $H_g = \partial U_g / \partial n$ corresponding to point O in the original image is at coordinate (r_0, θ_0) , where r_0 is a small displacement, $\theta_0 \neq 0, \varphi$ and $\mathbf{n} = \cos(\theta - \theta_n)\hat{e}_r - \sin(\theta - \theta_n)\hat{e}_\theta$. By using the McLaurin's series, we can write:

$$\begin{aligned} \left. \frac{\partial H}{\partial n} \right|_{(r_0, \theta_0)} &= \left. \frac{\partial H}{\partial n} \right|_{(0, \theta_n)} + r_0 \left. \frac{\partial^2 H}{\partial n^2} \right|_{(0, \theta_n)} + r_0(\theta_0 - \theta_n) \left. \frac{\partial^2 H}{\partial n_\perp \partial n} \right|_{(0, \theta_n)} \\ &+ \text{higher order terms,} \end{aligned} \quad (\text{A } 1)$$

where $\mathbf{n}_\perp = \sin(\theta - \theta_n)\hat{e}_r + \cos(\theta - \theta_n)\hat{e}_\theta$ is perpendicular to \mathbf{n} . The higher order terms are ignored owing to the fact that r_0 is a small displacement. Since $H_g(r_0, \theta_0)$ is a local extremum, $(\partial H_g / \partial n)|_{(r_0, \theta_0)} = 0$. Equation (A 1) can therefore be used to calculate r_0 the distance between the local extremum and point O , i.e.

$$r_0 = \frac{-(\partial H / \partial n)|_{(0, \theta_n)}}{(\partial^2 H / \partial n^2)|_{(0, \theta_n)} + (\theta_0 - \theta_n)(\partial^2 H / \partial n_\perp \partial n)|_{(0, \theta_n)}}. \quad (\text{A } 2)$$

From equation (2.2), straightforward calculations show that

$$\left. \frac{\partial H_g}{\partial n} \right|_{(0, \theta_n)} = \int_0^{+\infty} \int_0^{2\pi} \left(r \frac{d^2 h}{dr^2} \cos^2(\theta - \theta_n) + \frac{dh}{dr} \sin^2(\theta - \theta_n) \right) g dr d\theta. \quad (\text{A } 3)$$

where g is defined in equation (2.1). By integrating the above equation with respect to θ , and then integrating by parts with respect to r for the first term in the above equation and finally considering the asymptotic behaviour of the Bessel function, we can write:

$$\begin{aligned} \left. \frac{\partial H_g}{\partial n} \right|_{(0, \theta_n)} &= -(\sin(2(\varphi - \theta_n)) + \sin(2\theta_n)) \int_0^{+\infty} \left(\frac{dh_\varepsilon}{dr} \right) dr \\ &= (\sin(2(\varphi - \theta_n)) + \sin(2\theta_n)) h(\varepsilon). \end{aligned} \quad (\text{A } 4)$$

It is also straightforward to show that

$$\left. \frac{\partial^2 H_g}{\partial n_\perp \partial n} \right|_{(0, \theta_n)} = (\cos(\theta_n) \sin^2(\theta_n) - \cos(\varphi - \theta_n) \sin^2(\varphi - \theta_n)) \int_0^{+\infty} \frac{1}{r} h_\varepsilon(r) dr. \quad (\text{A } 5)$$

By substituting equations (A 4), (A 5) and (2.14) into equation (A 2), we can write:

$$r_0 = \frac{(\sin(2(\varphi - \theta_n)) + \sin(2\theta_n))h(\varepsilon)}{\lambda \int_\varepsilon^{+\infty} (1/r)h(r) dr}, \quad (\text{A } 6)$$

where $\lambda = (\cos(\theta_n) \sin^2(\theta_n) - \cos(\varphi - \theta_n) \sin^2(\varphi - \theta_n))(\theta_0 - \theta_n) + (\sin^3(\theta_n - \varphi) - \sin^3(\theta_n))$. Equation (A 6) is indeterminate as $\varepsilon \rightarrow 0$. By using l'Hôpital rule for intermediate terms, we can write equation (A 6) as

$$\begin{aligned} r_0 &= \lim_{\varepsilon \rightarrow 0} \frac{(\sin(2(\varphi - \theta_n)) + \sin(2\theta_n))h(\varepsilon)}{(\sin^3(\theta_n - \varphi) - \sin^3(\theta_n)) \int_\varepsilon^{+\infty} (1/r)(dh/dr)dr} \\ &= \lim_{\varepsilon \rightarrow 0} \frac{(\sin(2(\varphi - \theta_n)) + \sin(2\theta_n))(dh/d\varepsilon)}{(\sin^3(\theta_n - \varphi) - \sin^3(\theta_n))(1/\varepsilon)(dh/d\varepsilon)} = 0. \end{aligned} \quad (\text{A } 7)$$

Therefore, the distance between point O and the local extremum $H_g(r_0, \theta_0)$ approaches zero, as $\varepsilon \rightarrow 0$, i.e. $H_g(0, \theta_0)$ is a local extremum as $\varepsilon \rightarrow 0$. Therefore $|\nabla U_g|$ has a local extremum in point O . The above argument is also correct for edges along straight lines for $\varphi = \pi$. Any edge point along curved boundaries can be considered a point on an infinitesimal line for which the above argument also applies. ■

Appendix B

Proof for theorem 2.2. Let us consider the piecewise constant image shown in figure 1 described by equation (2.1) in polar coordinates as an input image g in equation (2.2). In figure 1, point O is placed on the image discontinuity forming a sharp corner. We aim to show that $|\nabla U_g|_O$ approaches infinity as $\varepsilon \rightarrow 0$. By using equation (2.2), let us evaluate $H_g = (\partial U_g / \partial n)|_{(0, \theta_n)}$, where $\mathbf{n} = \cos(\theta - \theta_n)\hat{e}_r - \sin(\theta - \theta_n)\hat{e}_\theta$, and θ_n , \hat{e}_r and \hat{e}_θ are any arbitrary angle between 0 and φ , and the unit vectors (radial and angular unit vectors) in polar coordinates, respectively. The Cartesian convolution evaluated in polar coordinates is written as

$$\frac{\partial U_g}{\partial n}(0, \theta_n) = \lim_{\varepsilon \rightarrow 0} \int_0^{+\infty} \int_{\vartheta+\pi}^{\vartheta+\varphi+\pi} \frac{\partial h_\varepsilon}{\partial n} r dr d\theta. \quad (\text{B } 1)$$

On the other hand, $\partial h / \partial n$ can be calculated as

$$\frac{\partial h_\varepsilon}{\partial n} = \nabla h_\varepsilon \cdot \mathbf{n} = \frac{\partial h_\varepsilon}{\partial r} \cos(\theta - \theta_n) = P_\varepsilon(r) \cos(\theta - \theta_n), \quad (\text{B } 2)$$

where $P_\varepsilon(r)$ is defined in equation (2.22). By substituting equation (B 2) into equation (B 1), then integrating by parts with respect to r and finally integrating

with respect to θ , we can write:

$$\frac{\partial U_g}{\partial n}(0, \theta_n) = (\sin(\theta_n - \varphi) - \sin(\theta_n)) \lim_{\varepsilon \rightarrow 0} \left[(rh_\varepsilon(r))|_0^{+\infty} - \int_0^{+\infty} h_\varepsilon(r) dr \right], \quad (\text{B } 3)$$

where $h_\varepsilon(r) = \int_{-\infty}^r P_\varepsilon(x) dx$. Using the asymptotic behaviour of $K_1(r)$ as $r \rightarrow 0$ and $r \rightarrow +\infty$ (e.g. Polyanin & Zaitsev 2003; Arfken & Weber 2005) and l'Hôpital's rule for indeterminate terms, equation (B 3) can be rewritten as

$$\frac{\partial U_g}{\partial n}(0, \theta_n) = (\sin(\theta_n) - \sin(\theta_n - \varphi)) \lim_{\varepsilon \rightarrow 0} \left(\int_0^{+\infty} h_\varepsilon(r) dr \right). \quad (\text{B } 4)$$

As $\varepsilon \rightarrow 0$, the integration term in the above equation approaches infinity. Since $(\partial U_g / \partial n)(0, 0) = \nabla U_g|_O \cdot \mathbf{n}$, then $\lim_{\varepsilon \rightarrow 0} |\nabla U_g|_O = +\infty$. We notice that for $\varphi = \pi$, the wedge shown in figure 1 becomes an edge along a straight line on which point O can be regarded as an arbitrary point. It is also noted that if point O is placed on some part of the edge (contour) that can be regarded as a regular curve, then the contour at point O can be approximated as an infinitesimal line passing through point O for both of which the above argument applies. ■

Appendix C

Proof for lemma 2.5. In the absence of noise, the noise-free localization index is calculated according to equation (2.25). We would like to evaluate the noise-free localization by calculating the terms $(\partial H_g / \partial n)|_{(0, \theta_n)}$ and $(\partial^2 H_g / \partial n^2)|_{(0, \theta_n)}$:

$$\left. \frac{\partial H_g}{\partial n} \right|_{(0, \theta_n)} = h(0)(\sin(2(\varphi - \theta_n)) + \sin(2\theta_n)) \quad (\text{C } 1)$$

and

$$\left. \frac{\partial^2 H_g}{\partial n^2} \right|_{(0, \theta_n)} = (\sin^3(\theta_n - \varphi) - \sin^3(\theta_n)) \int_0^{+\infty} \frac{1}{r} \frac{dh}{dr} dr, \quad (\text{C } 2)$$

where $h(r)$ is only a function of radial coordinate in a polar coordinate system. By replacing (C 1) and (C 2) into equation (2.25), the noise-free localization index is calculated as

$$A_{\text{NF}} = \left| \frac{(\sin^3(\theta_n - \varphi) - \sin^3(\theta_n)) \int_0^{+\infty} (1/r)(dh/dr) dr}{(\sin(2(\varphi - \theta_n)) + \sin(2\theta_n))h(0)} \right|. \quad (\text{C } 3)$$

Since $h(0)$ is bounded, $(dh/dr)|_{(r=0)} = 0$, and the wedge angle $\varphi \neq \pi$, then the noise-free localization index evaluated in (C 3) is bounded. ■

References

- Arfken, G. B. & Weber, H. J. 2005 *Mathematical methods for physicists*. San Diego, CA: Harcourt Academic Press.
- Bao, P. & Wu, X. 2005 Canny edge detection enhancement by scale multiplication. *IEEE Trans. Pattern Anal. Mach. Intell.* **27**, 1485–1490. (doi:10.1109/TPAMI.2005.173)

- Canny, J. F. 1986 A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **8**, 679–698. (doi:10.1109/TPAMI.1986.4767851)
- Demigny, D. 2002 On optimal linear filtering for edge detection. *IEEE Trans. Image Process.* **11**, 728–737. (doi:10.1109/TIP.2002.800887)
- Deriche, R. 1987 Using Canny's criteria to derive a recursively implemented optimal edge detector. *Int. J. Comput. Vis.* **1**, 728–737. (doi:10.1007/BF00123164)
- Freeman, W. T. & Adelson, H. 1991 The design and use of steerable filters. *IEEE Trans. Pattern Anal. Mach. Intell.* **13**, 891–906. (doi:10.1109/34.93808)
- Jacob, M. & Unser, M. 2004 Design of steerable filters for feature detection using canny-like criteria. *IEEE Trans. Pattern Anal. Mach. Intell.* **26**, 1007–1019. (doi:10.1109/TPAMI.2004.44)
- Lindeberg, T. 1998*a* Edge detection and ridge detection with automatic scale selection. *Int. J. Comput. Vis.* **30**, 117–154. (doi:10.1023/A:1008097225773)
- Lindeberg, T. 1998*b* Feature detection with automatic scale selection. *Int. J. Comput. Vis.* **30**, 79–116. (doi:10.1023/A:1008045108935)
- Martin, D. R., Fowlkes, C. C. & Malik, J. 2004 Learning to detect natural image boundaries using local brightness, color, and texture cues. *IEEE Trans. Pattern Anal. Mach. Intell.* **26**, 530–549. (doi:10.1109/TPAMI.2004.1273918)
- Petrou, M. 1995 Separable 2D filters for the detection of ramp edges. *IEE Proc. Vis. Image Signal Process.* **142**, 228–231. (doi:10.1049/ip-vis:19952090)
- Petrou, M. & Kittler, J. 1991 Optimal edge detectors for ramp edges. *IEEE Trans. Pattern Anal. Mach. Intell.* **13**, 483–491. (doi:10.1109/34.134047)
- Polyanin, A. D. & Zaitsev, V. F. 2003 *Handbook of exact solutions for ordinary differential equations*. London, UK: Chapman & Hall.
- Pratt, W. K. 2007 *Digital image processing*. New York, NY: John Wiley & Sons Inc.
- Rao, K. R. & Ben-Arie, J. 1994 Optimal edge detection using expansion matching and restoration. *IEEE Trans. Pattern Anal. Mach. Intell.* **16**, 1169–1177. (doi:10.1109/34.387490)
- Sarkar, S. & Boyer, K. L. 1991 On optimal infinite impulse response detection filters. *IEEE Trans. Pattern Anal. Mach. Intell.* **13**, 1154–1171. (doi:10.1109/34.103275)
- Sethian, J. A. 1996 *Level set methods: evolving interfaces in geometry, fluid mechanics, computer vision and material science*. Cambridge, UK: Cambridge University Press.
- Shen, J. & Castan, S. 1992 An optimal linear operator for step edge detection. *CVGIP Graph. Models Image Process.* **54**, 112–133. (doi:10.1016/1049-9652(92)90060-B)
- Tagare, H. D. & deFigueiredo, R. J. P. 1990 On localization performance measure and optimal edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **12**, 1186–1190. (doi:10.1109/34.62607)
- Wang, Z., Rao, K. R. & Ben-Arie, J. 1996 Optimal ramp edge detection using expansion matching. *IEEE Trans. Pattern Anal. Mach. Intell.* **18**, 1092–1097. (doi:10.1109/34.544078)