

# Scale-space Properties of Quadratic Edge Detectors

Paul Kube<sup>1</sup> and Pietro Perona<sup>2,3</sup>

<sup>1</sup> Computer Science and Engineering, University of California, San Diego  
La Jolla, California 92093, U.S.A.      kube@cs.ucsd.edu

<sup>2</sup> Department of Electrical Engineering, California Institute of Technology  
Pasadena, California 91125, U.S.A.      perona@systems.caltech.edu

<sup>3</sup> Università di Padova, Dipartimento di Elettronica ed Informatica, Padova, Italy

**Abstract.** Edge detectors which use a quadratic nonlinearity in the filtering stage are attracting interest in machine vision applications because of several advantages they enjoy over linear edge detectors. However, many important properties of these quadratic or “energy” edge detectors remain unknown. In this paper, we investigate the behavior of quadratic edge detectors under scaling. We consider two cases important in practice: quadratic detectors with constituent filters related by the Hilbert transform, and with constituent filters related by the first spatial derivative. We prove that in one dimension, Hilbert-pair detectors with Gaussian scaling permit the creation of new features as scale is increased, but such causality failures cannot generically occur with derivative-pair detectors. In addition, we report experiments that show the effects of these properties in practice. Thus at least one class of quadratic edge detectors can have the same desirable scaling property as detectors based on linear differential filtering.

## 1 Introduction

The process of detecting image features across a range of scales is important in many machine vision applications, and dates at least from Rosenfeld [14] and Marr [9]. In practical systems using multiscale techniques, features detected at a coarse scale can determine processing at finer scales. It is important in this context that features detected at a given resolution were not created gratuitously at that scale, but rather are “grounded” in image detail at a finer resolution. When a multiscale feature detection method never introduces features as the scale is coarsened, the method has the desirable property of *causality* [6].

It is known that edge detectors which mark edges at zeros or extrema in the output of linear differential filters acting on the image have the causality property if scale is selected by convolution of the filters or image with a Gaussian [2, 15], and these results have been extended to scale selection by anisotropic diffusion [12]. More recently, quadratic nonlinear filters for feature detection have been proposed as having advantages over linear filters, particularly in their ability to localize edges with complex structure [5,7,10,11,13]. However, the question

whether these quadratic or “energy” filters permit a causal scale selection technique has remained open.

In this paper, we address this question for quadratic edge detectors in the context of images containing one-dimensional edges scaled by convolution. We concentrate on quadratic detectors with two constituent filters; this coincides with existing practice [10,11] and theory [7]. We consider two special cases of practical interest: constituent filters related by the first spatial derivative, and constituent filters related by the Hilbert transform. We show that in the case of derivative-pair filters, Gaussian scaling is causal, whereas Gaussian scaling with Hilbert-pair filters permits noncausal edge creation at any scale. Thus we show that in the one dimensional case, at least one class of quadratic edge detectors has the same desirable scaling property as the more familiar detectors based on linear differential filtering.

## 2 Causality, Catastrophe, and Edge Detection

Let  $h(x)$  be a real-valued signal. In this paper we assume  $x \in \mathbb{R}$ , giving direct application to time signals, images with unidirectionally oriented edges, and images with edges that are one-dimensional on a scale comparable to the filter size.

A quadratic or “energy” edge detector marks an edge at scale  $\sigma$  wherever there is a local maximum with respect to  $x$  of the nonlinearly filtered signal

$$E(x, \sigma) = \sum_{j=1}^M [(f^j * g^\sigma * h)(x)]^2 . \quad (1)$$

Here  $*$  denotes convolution. The  $f^j$  are the impulse responses of the “constituent filters” of the quadratic edge detector; we will concentrate on the case of detectors with two constituent filters, one even and one odd, with impulse responses  $f^e$  and  $f^o$  respectively. The function  $g^\sigma$  is the impulse response of a scaling filter with scale parameter  $\sigma$ ; it has the form

$$g^\sigma(x) = g(x/\sigma)/\sigma \quad (2)$$

with  $g$  even, so increasing  $\sigma$  corresponds to coarser scales.

We assume that  $f^e$ ,  $f^o$ ,  $h$ , and  $g^\sigma$  are such that the energy signal  $E$  is smooth. We are interested in whether some choices of  $f^e$ ,  $f^o$ , and  $g^\sigma$  may guarantee that for any  $h$  no new maxima are introduced in  $E$  as  $\sigma$  increases, i.e. that  $E$  has the causality property. We will pay particular attention to the Gaussian  $g^\sigma(x) = \exp(-x^2/2\sigma^2)/\sigma$ , and approach this question by considering the conditions under which the degenerate critical points of  $E$  may be noncausal. A discussion of an approach from the point of view of systems of partial differential equations satisfying a maximum principle can be found in [8].

## 2.1 Quadratic Filters and Scalespace Singularities

Suppose a maximum of the energy  $E(x, \sigma)$  is created with increasing  $\sigma$  in the interior of a bounded domain, say at  $x = x_0, \sigma = \sigma_0$ . Then  $(x_0, \sigma_0)$  is a degenerate critical point of  $E$ , and generically the set of critical points  $E_x = 0$  in a neighborhood of  $(x_0, \sigma_0)$  form an upward opening parabola (a fold catastrophe) [1, 4]. Conditions accompanying such a generic causality failure can be expressed in terms of partial derivatives of  $E$  at the degenerate critical point  $x = x_0, \sigma = \sigma_0$ , viz (cf. [15])

$$\begin{aligned} E_x(x_0, \sigma_0) &= 0 & (3) \\ E_{xx}(x_0, \sigma_0) &= 0 \\ E_{x\sigma}(x_0, \sigma_0)E_{xxx}(x_0, \sigma_0) &< 0 . \end{aligned}$$

We will be interested in whether these conditions can exist in (1) for some choice of constituent filters  $f^j$ , scaling function  $g^\sigma$ , and image signal  $h$ . If they can, then  $E$  is not causal for that choice of constituent and scaling filters. If no such  $h$  exists for some choice of  $f^j$  and  $g^\sigma$ , then no signal will produce generic causality failures for that scalable quadratic filter design. We do not consider the case of nongeneric causality failures.

## 2.2 Toward Constructing Causality Failures

We consider the occurrence of causality failures at  $x_0 = 0$ ; since  $E$  is shift invariant with respect to the image  $h$ , this is without loss of generality. Then partial derivatives of  $E$  in  $\sigma$  and  $x$  involved in (3) will contain terms of the form

$$\frac{\partial^{n+m}}{\partial x^n \partial \sigma^m} (f^j * g^\sigma * h)(0) \quad (4)$$

with  $j \in \{e, o\}$ . The conditions (3) can be expressed in a form which reduces the question of noncausality to the question of solving a constrained system of linear equations. We adopt the following notation.

We suppose that  $h(x)$  is the sum of  $k$  sinusoidal components;  $k$  may be large as desired, and so  $h$  approximates in the  $L_2$  sense any image function as closely as desired. Let  $\{u_1, \dots, u_k\}$  be the (positive) frequencies of the sinusoidal components defining  $h$ . Further, let  $F^e, F^o, G$ , and  $H$  be the Fourier transforms of  $f^e, f^o, g$ , and  $h$  respectively, and let  $G'$  be the derivative of  $G$ . We introduce the real  $k$ -vectors  $\mathbf{f}^e, \mathbf{f}^o, \mathbf{f}^{e'}, \mathbf{f}^{o'}, \mathbf{h}^e, \mathbf{h}^o$ , with the  $l$ th component of each specified as

$$\mathbf{f}^e_l = F^e(u_l)G(\sigma u_l) \quad (5)$$

$$\mathbf{f}^o_l = F^o(u_l)G(\sigma u_l) \quad (6)$$

$$\mathbf{f}^{e'}_l = F^e(u_l)G'(\sigma u_l) \quad (7)$$

$$\mathbf{f}^{o'}_l = F^o(u_l)G'(\sigma u_l) \quad (8)$$

$$\mathbf{h}^e_l = 2 \operatorname{Re} H(u_l) \quad (9)$$

$$\mathbf{h}^o_l = \frac{2}{i} \operatorname{Im} H(u_l) . \quad (10)$$

Further, let  $\mathbf{U}$  be the  $k \times k$  matrix  $\text{diag}(u_1, \dots, u_k)$ .

Now, satisfying (3) at  $(0, \sigma)$  is equivalent to satisfying the system

$$\begin{pmatrix} \mathbf{f}^{eT} \\ (\mathbf{U}\mathbf{f}^o)^T \\ (\mathbf{U}^2\mathbf{f}^e)^T \\ (\mathbf{U}^3\mathbf{f}^o)^T \\ (\mathbf{U}\mathbf{f}^{e'})^T \\ (\mathbf{U}^2\mathbf{f}^{o'})^T \end{pmatrix} \mathbf{h}^e = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v'_2 \\ v'_3 \end{pmatrix}, \quad \begin{pmatrix} \mathbf{f}^{oT} \\ (\mathbf{U}\mathbf{f}^e)^T \\ (\mathbf{U}^2\mathbf{f}^o)^T \\ (\mathbf{U}^3\mathbf{f}^e)^T \\ (\mathbf{U}\mathbf{f}^{o'})^T \\ (\mathbf{U}^2\mathbf{f}^{e'})^T \end{pmatrix} \mathbf{h}^o = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w'_2 \\ w'_3 \end{pmatrix} \quad (11)$$

$$E_x(0, \sigma) = 4\pi(v_2w_1 - v_1w_2) = 0$$

$$E_{xx}(0, \sigma) = 8\pi^2(v_2^2 - v_1v_3 + w_2^2 - w_1w_3) = 0$$

$$E_{xxx}(0, \sigma)E_{x\sigma}(0, \sigma) = 64\pi^4(3v_3w_2 + v_1w_4 - 3v_2w_3 - v_4w_1) \\ \times (-v'_2w_2 - v_1w'_3 + v_2w'_2 + v'_3w_1) < 0 .$$

In the following sections, we consider the existence of solutions to this constrained system.

### 3 Hilbert Pair Filters and Gaussian Scaling

Quadratic edge detectors with an even and an odd constituent filter in quadrature, i.e. such that one is the Hilbert transform of the other, are perhaps the most common in practice [10,11]. This case gives a particularly simple form for the system (11), since when  $f^e$  and  $f^o$  are Hilbert transforms of each other we have  $\mathbf{f}^e = \mathbf{f}^o = \mathbf{f}$ . A further simplification is possible when Gaussian scaling is used; then  $G(u\sigma) = \exp(-u^2\sigma^2/2)$  and  $\mathbf{f}^{e'} = \mathbf{f}^{o'} = -\sigma\mathbf{U}\mathbf{f}$ . We consider this case, and obtain

**Proposition 1.** *Assume a quadratic edge detector  $E$  with constituent filters  $f^e$  and  $f^o$ , even and odd respectively, such that one is the Hilbert transform of the other, and with Gaussian scaling function. Assume the Fourier transform of  $f^e$  is nonzero at at least four distinct positive frequencies. Then  $E$  is not causal.*

*Proof.* The proof is constructive; given any filter of the stated type, and any scale  $\sigma_0$ , we produce a signal  $h$  which satisfies (11) and so produces a causality failure in  $E(x, \sigma)$  at  $x = 0$   $\sigma = \sigma_0$ . See [8] for details.  $\square$

### 4 Derivative Pair Filters and Gaussian Scaling

The case of quadratic edge detectors with an even and an odd constituent filter such that one is the derivative of the other have been studied by Kube [7]. Here we consider this case in which  $f^e = df^o/dx$ , which in the present notation implies  $\mathbf{f}^e = \mathbf{U}\mathbf{f}^o$  and  $\mathbf{f}^{e'} = -\sigma\mathbf{U}\mathbf{f}^e$ . As in the previous section, these assumptions together with Gaussian scaling lead to simplifications in the form of the constrained system (11). Here however the result is more favorable:

**Proposition 2.** *Assume a quadratic edge detector  $E$  with constituent filters  $f^e$  and  $f^o$ , even and odd respectively, such that one is the derivative (with respect to the spatial coordinate  $x$ ) of the other, and with a Gaussian scaling function. Then generic causality failures cannot occur in  $E$ .*

*Proof.* The proof proceeds by cases to show that under the stated conditions the system (11) has no solutions; see [8] for details.  $\square$

## 5 Results of Experiments

Sections 3 and 4 showed that, in principle, causality failures can occur in quadratic edge detector schemes which use Hilbert-pair filters and Gaussian scaling, but cannot generically occur if derivative-pair filters are used with Gaussian scaling. However, these results leave open the questions whether causality failures occur often in practice with Hilbert-pair quadratic detectors, and whether nongeneric causality failures occur in practice with derivative-pair quadratic detectors. We have attempted to address these questions with experiments we report here.

Five images were obtained of scenes consisting of matte surfaced right rectangular prisms and cylinders placed randomly on a surface. The camera geometry and illumination was such that all edges in the images were approximately vertical. A typical row of 512 pixels was taken from each image; this one-dimensional signal was subjected to quadratic edge detection as defined in Section 2, using Gaussian scaling with  $\sigma$  in the range 1 to 64 pixels. In one set of experiments, the quadratic detector had as constituent filters the first-derivative operator and its Hilbert transform; in the other set of experiments, the quadratic detector had as constituent filters the first-derivative operator and its first derivative (i.e. the second derivative operator). Impulse responses of the constituent filters at a scale of 8 pixels are shown in Figures 1 and 2 respectively. In each case, local maxima were detected without thresholding at each scale and the resulting scale space representations were examined for causality failures.

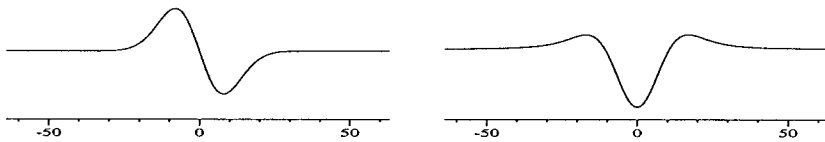


Fig. 1.

Scaled impulse responses of the constituent filters  $f^o * g^\sigma$ ,  $f^e * g^\sigma$  for the Hilbert-pair quadratic edge detector discussed in the text. Left, the first derivative of a Gaussian with standard deviation 8 pixels; right, its Hilbert transform.

For each of the five images, the Hilbert-pair quadratic detector generated a few noncausal edges, some of which were quite robust and extended over a range

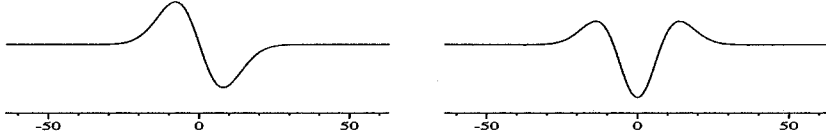


Fig. 2.

Scaled impulse responses of the constituent filters  $f^o * g^\sigma$ ,  $f^e * g^\sigma$  for the derivative-pair quadratic edge detector discussed in the text. Left, the first derivative of a Gaussian with standard deviation 8 pixels; right, its derivative.

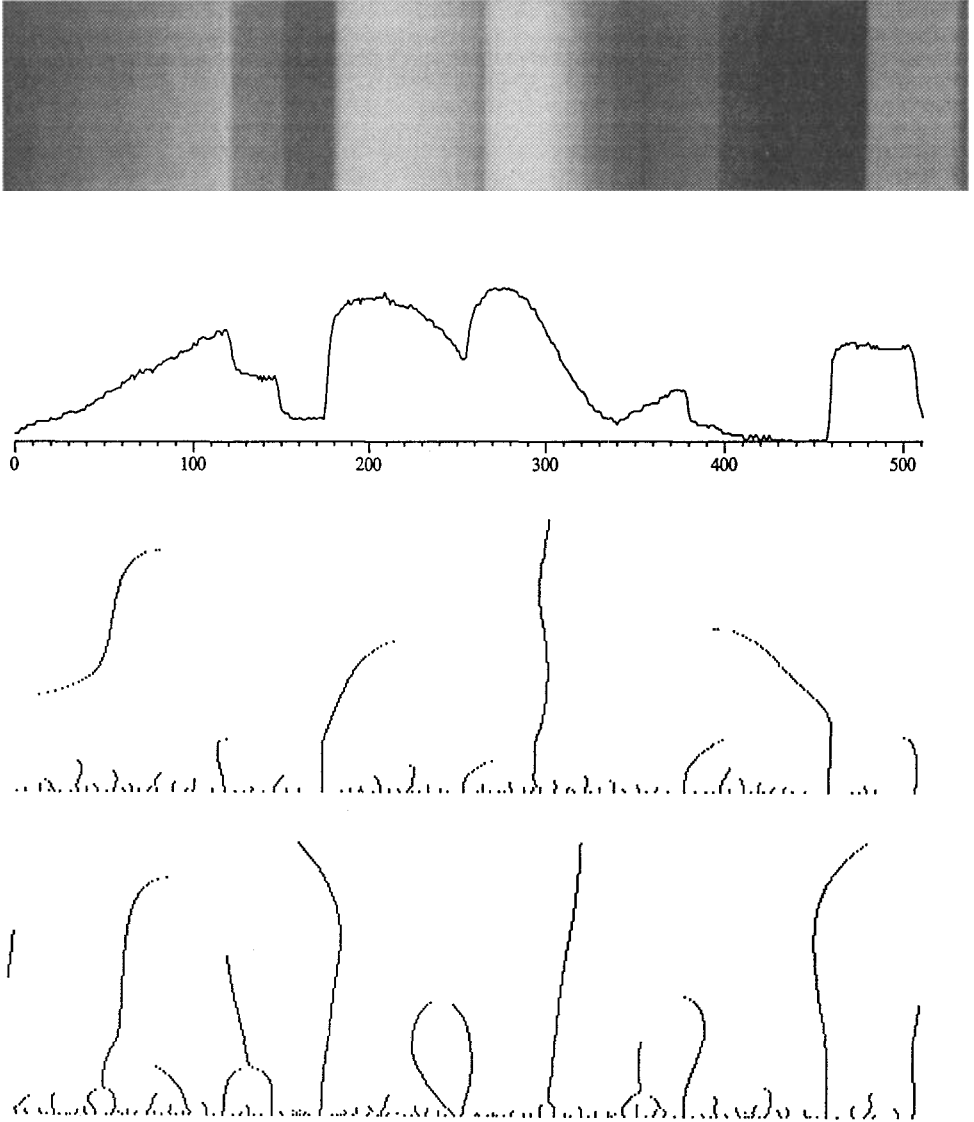
of scales. The derivative-pair quadratic detector generated no noncausal edges. Results from a typical image are shown in in Figure 3; for others, see [8]. It is perhaps somewhat surprising that filters with such similar impulse responses produce exhibit such qualitatively different performance. These experimental results are, of course, consistent with the theoretical results of previous sections in this paper.

## 6 Summary

The scale-space properties of quadratic edge detectors being of potential interest, we have investigated whether quadratic detectors admit of causal scaling. Considering the cases of detectors with constituent filters related by the Hilbert transform and with constituent filters related by the first spatial derivative, we have stated theorems to the effect that Gaussian scaling leads to causality failures in the first case but not in the second, and we have supported these results with experimental observations on real images. Thus we have shown that, in the one dimensional case at least, a class of quadratic edge detectors has the same desirable causal scaling property as the more well known types of edge detectors based on linear differential filtering.

We have not shown that the Gaussian is unique in providing causal scaling in the derivative-pair case, nor that no scaling filter is causal in the Hilbert-pair case. While we believe sharper results are obtainable along these lines, general statements about solutions to the constrained system (11) are equivalent to feasibility testing for fairly general systems of quadratic inequalities and so entail difficult problems [3].

In addition, we have worked here with the assumption of one-dimensional edges. The results are thus applicable to feature detection in time signals, images with edges oriented in the same direction, or images with edges that are one-dimensional on a scale comparable to the filter size. The extension to higher dimensions appears feasible and is the subject of ongoing work.



**Fig. 3.**

From top to bottom: An image with one-dimensional edges; the graph of image intensity for the top row of pixels in the image; edge scale-space generated from that 1D signal with a Hilbert-pair quadratic edge detector; edge scale-space generated from that 1D signal with a derivative-pair quadratic edge detector. Gaussian scaling and periodic convolution are used in each case. Edge scale-space graphs have spatial coordinates 0-511 pixels in  $x$  and scale coordinates 1-64 pixels in  $\sigma$ . The Hilbert-pair detector exhibits a robust causality failure at  $(x, \sigma) \approx (17, 22)$ . Smaller causality failures can be observed near  $(93, 19)$  and  $(103, 18)$ . The derivative-pair filter exhibits no causality failures. See text for discussion.

## 7 Acknowledgements

This research was in part sponsored by NSF Research Initiation grant IRI 9211651, and by NSF grant IRI 9306155, and by ONR grant NOOO14-93-1-0990. P.P. gratefully acknowledges the Newton Institute for Mathematical Sciences of Cambridge, UK, where he conducted part of this research.

## References

- [1] V. I. Arnold, *Singularity Theory*, Cambridge University Press, 1981.
- [2] J. Babaud, A. Witkin, M. Baudin & R. Duda, "Uniqueness of the Gaussian kernel for scale-space filtering," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8 (1986), 26–33.
- [3] A. I. Barvinok, "Feasibility testing for systems of real quadratic equations," *STOC* (1992).
- [4] J. J. Clark, "Singularity theory and phantom edges in scale space," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 10 (1988), 720–726.
- [5] W. T. Freeman & E. H. Adelson, "The design and use of steerable filters," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13 (1991), 891–906.
- [6] J. Koenderink, "The structure of images," *Biological Cybernetics* 50 (1984), 363–370.
- [7] P. Kube, "Properties of energy edge detectors," *IEEE Computer Science Society Conference on Computer Vision and Pattern Recognition* (1992).
- [8] P. Kube & P. Perona, "Scale-space properties of quadratic edge detectors," California Institute of Technology, Computation and Neural Systems Technical Report 31, October, 1993.
- [9] D. Marr & E. Hildreth, "Theory of edge detection," *Proc. R. Soc. Lond. B* 207 (1980), 187–217.
- [10] M. C. Morrone & R. A. Owens, "Feature detection from local energy," *Pattern Recognition Letters* 6 (1987), 303–313.
- [11] P. Perona & J. Malik, "Detecting and localizing edges composed of steps, peaks and roofs," *Proceedings of the Third International Conference of Computer Vision* (1990).
- [12] P. Perona & J. Malik, "Scale space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12 (1990), 629–639.
- [13] C. Ronse, "On idempotence and related requirements in edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15 (1993), 484–491.
- [14] A. Rosenfeld & M. Thurston, "Edge and curve detection for visual scene analysis," *IEEE Trans. Comput.* 20 (1971), 562–569.
- [15] A. L. Yuille & T. A. Poggio, "Scaling theorems for zero crossings," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8 (1986), 15–25.