Lecture notes on edge detection

R. Brémond

Université Paris Est, LEPSiS, INRETS-LCPC

Abstract

Abstracts of reference papers in the field of edge detection are proposed: (Marr and Hildreth, 1980) and (Perona and Malik, 1990), as well as surveys on edge detection (Ziou and Tabbone, 1998; Basu, 2002) and old forgotten papers (Fram and Deutsch, 1975). Personal remarks are sometimes included.

Keywords: edge detection, visual perception, computer vision

1. On the quantitative evaluation of edge detection schemes and their comparison with human performance

(Fram and Deutsch, 1975)

This is the first published paper to propose a methodology in order to rate the respective merit of Edge Detectors (ED). Of course the tested ED are not up-to-date, but the framework is still interesting.

ED are supposed to find and locate the "boundaries of objects". Even in 1975, a number of operators were available, and a comparision framework was needed, as well as a ground truth in terms of human visual performance. A number of edge properties would benefit from such an evaluation framework, such as edge orientation estimation, detection of curved edges, ability to discriminate close edges, etc. The following of the paper focuses on straight edge detection in the presence of noise. Three ED are compared: those from Rosenfeld and Thurston (1971), McLeod (1972) and Hueckel (1973). In short, Hueckel (1973) detects edges thanks to a set of ideal edge lines (templates). McLeod (1972) uses a convolution with an oriented kernel: it looks for an edge in a given direction. Rosenfeld and Thurston (1971) proposed a 2-steps operator, fitting a very basic model of the adjacent regions in various directions. The resulting edges are then thinned.

For the operator's evaluation, an image database was build. Each image contained one vertical edge, separating two adjacent regions. Then, some noise was added. Each image was described by the (Gaussian) noise rate and by the "edge's strength" (intensity difference) between the two regions.

The outputs from the ED produced edge weights, and sometimes additional informations. This was normalized, so as to get the same level of information out of any ED: a binary image. The number of edge points was estimated *a priori*, which allowed to threshold the edge weights. For each image and each ED, 2 indexes were computed: $P_1 = TP/(TP + FP)$ (TP: True Positives, FP: False Positives), and $P_2 = TP/(TP + M)$ (M: Misses).

Then a psycho-visual experiment allowed to compare the detector's performances to a mean human performance (mean over 5 observers). The psychovisual task did not mimic the detector's task: people were asked the orientation of a displayed edge, among four possible orientations. This led to some discussions about how the two performance indexes could be compared. Finally, the authors computed an index, thresholding P_1 and P_2 (which includes the hazardous step of setting the thresholds values), and compared it to human perfomances.

2. Theory of edge detection

(Marr and Hildreth, 1980)

Overview. This paper was not published in a computer vision or image processing journal, but in the proceedings of the Royal Society (London). It is a starting point for edge detection algorithms, where vision science was first introduced. Several important ideas are proposed, some of them have become popular in computer vision, such as the zero-crossings idea. As a whole, Marr an Hildreth's paper is still cited, but their approach is mostly considered as a dead-end in the edge detection community[1].

One tough issue in edge detection is that, as stated by Marr and Hildreth, "the concept of an edge has a partly visual and partly physical meaning". Their paper was an effort to make this double bind clear.

The paper is divided in two parts: a single scale edge detector is followed by a multiscale edge detector. The basic tool is the Laplacian of Gaussian (LoG), applied to image I: the edges are expected to be located at the zero-crossings of the LoG, for a given scale. Then, a multiscale representation of the image is proposed, where Gaussian filters at various scales have to be combined in order to find out the "true" edges. Finally, the proposed edge detection theory is discussed in the context of psychophysical knowledge.

The starting point of Marr and Hildreth's paper was to discuss previous results in vision science about early visual processing. First, Hubel and Wiesel (1962) found the so-called *simple cells* in V1, with bar or edge-shaped receptive fields. Second, Campbell and Robson (1968) showed that visual processing includes parallel spatial frequency tuned channels, which results in some kind of a Fourier analyzer. These findings are not easy to mix up. Marr's feeling is that some notions about the task devoted to early vision would help. His proposal is that a *primal sketch* is built: a "primitive but rich description of the image that is to be used to determine the reflectance and illumination of the visible surfaces[2], and their orientation and distance relative to the viewer" (p. 188). The article deals with the computation of the *raw* primal sketch, which describes the image in terms of edges, bars, blobs and terminations[3].

One paradoxal idea is to detect edges, which is directional if anythink is, with an isotropic spatial filter. The authors anounce a future paper, which addresses directional selectivity (Marr and Ullman, 1981).

The single scale problem. The first issue addressed in the paper is the selection of an optimal smoothing filter, in order to detect intensity changes at a given scale. The proposal is to design a bandpass filter with a smooth spectrum in the frequency domain; in other words, the variance in the frequency domain, $\Delta\omega$, should be small. The second constraint is that the edge should be accurately localized, as "the things in the world that give rise to intensity changes" are spatially localized. This is expressed with Δx, the spatial variance of the filter, which should be small. The choice of a linear filter is not discussed. The point is, an uncertainty principle states that $\Delta\omega\Delta x \leq \pi/4$. The optimal distribution in terms of uncertainty was found to be the Gaussian distribution (Leipnik, 1960).

Given a smoothing filter, where are the edges at the selected scale? Intensity changes correspond to a peak in the first derivative, that is, a zero-crossing (maximum) in the second derivative. For a 2D Gaussian filter, edges are found where this second derivative crosses zero. Marr and Hildreth note that this second derivative "closely resembles" the Difference of Gaussians (DoG) found by Wilson and Giese (1977) in vision science (this similarity is discussed in an Appendix: the second derivative of the Gaussian is the limit of the DoG when... the size of the two Gaussians tend to one another!).

The next issue is to compute the edge direction, when an edge is found. That is, in what direction

[1] I guess it is not.

[2] This idea may be related to the Retinex theory by Land and McCann (1971).

[3] Marr and Hildreth's approach seems purely bottom-up, following the first trend of cognitive science, that is, the information theory in the 1950'.

should we compute the second derivative? The answer is that the edge direction is given by the local zero-crossings. Assuming that intensity variations along the line of zero-crossing are (locally) linear, the solution is given by the Laplacian ∇^2. Finally, the proposed filter is a Laplacian of Gaussian (LoG, or $\nabla^2 G$). Thus, the edge orientation is the one at which the zero-crossing has the maximum slope.

Marr and Hildreth define edges in the image as linear segments (in the u direction) of zero-crossings of $\partial^2 G/\partial v^2$, where $u \perp v$. An amplitude ν is associated with each segment: it is the slope of this second derivative along the segment (supposed to be almost constant). Finally, the set of these segments is the raw primal sketch mentioned above, at the selected scale.

The multispectral problem. The main proposal for combining results from different channels was called the *spatial coincidence assumption.* A given intensity change is spatialy localized, and shoud be detected at various scales with (more or less) the same size and (more or less) the same orientation and (more or less) the same localization. From this statement, 3 cases are discussed: (1) isolated edges, (2) bars, (3) blobs and terminations. The purpose of this discussion is to build an informative so-called raw primal sketch; that is, to describe the relevant properties of the detected edges, along with their localization. For instance, the edge amplitude and width may be computed. Bars are pairs of parallel edges; they may reveal the presence of a thick edge, that is, a single edge in a higher spatial scale[4].

Psychophysics of edge detection. The proposed model is compared to Wilson's models (Wilson and Giese, 1977; Wilson and Bergen, 1979) for orientation-dependent, spatial frequency-tuned channels. The authors discuss the idea that Hubel and Wiesel's simple cells (Hubel and Wiesel, 1962) may simply measure second directional derivatives (they guess not). Then, they argue that the LoG and DoG are not so different, meaning that biological models including DoG may as well be implemented with LoG: it is not so easy to discriminate between these models on the basis of experimental data.

Marr and Hildreth propose that $\nabla^2 G \star I$ may be coded in the lateral geniculate nucleus (LGN) of the thalamus, while some "simple cells" of V1 (in the sense of Hubel and Wiesel) may code segments of zero-crossings in the previous map. They argue that the DoG have been proposed as a model for LGN cells (Rodieck and Stone, 1965; Enroth-Cugell and Robson, 1966), and may implement their LoG in the sustained cells (carrying either the positive or the negative part of $\nabla^2 G \star I$). Then, the zero-crossings detection in V1 could be an `AND` operator, combining the positive and negative components of the input signal from the LGN. The amplitude of the edge detector should be linked with ν in some way.

3. Scale-space and edge detection using anisotropic diffusion

(Perona and Malik, 1990)

One nice formalism for a multiscale image description is the scale-space filtering (Witkin, 1983; Koendrink, 1984). An image I_0 is embedded in a family of derived images:

$$I(t) = I_0 \star G(t) \tag{1}$$

where G is a Gaussian kernel with variance t. Then, Koendrink (1984) noticed that the image family may be viewed as the solution of the heat conduction (with t the time): $I(t) = \Delta I$ (with initial condition $I(0) = I_0$). He proposed two properties for the kernel: "causality" (a new feature [edge] should not appear at a coarse level), and "Isotropy" (the kernel should be isotropic). The last criterion was chosen foro the sake of simplicity, and Perona and Malik (1990) show it is not necessary.

Anisotropic diffusion is proposed as a solution to several weaknesses of the scale-space paradigm: boundaries are shifted at coarse scales, and some junctions disapear. The authors replace Koendrink's "Isotropy" criterion with two more criteria: "immediate localization" (boudaries are well localized at all

[4] Note that due to the author's goal of building a semi-semantic, low level primal sketch, some discussions and questions do not adress edge detection, but edge description.

scales) and "piecewise smoothing" (smoothing apply to the regions rather than to the edges). The proposed solution is that the diffusion coefficient may vary spatially, along with an estimation of the edge presence (local contrast). The heat diffusion equation becomes:

$$I(t) = div(c.\nabla I) = c(t)\Delta I + \nabla c.\nabla I \qquad (2)$$

with ∇ and Δ the gradient and Laplacian operators, respectively. It reduces to $I(t) = c\Delta I$ if c is constant over space and time. The idea is to smooth the image within regions, not across boundaries. The optimal solution is $c = 1$ inside regions, and $c = 0$ on edges, but at this stage the solution is not known! Thus, the best edge estimator is used instdead[5].

For practical application, various values for c were tested of the form $c = g(\|\nabla I\|)$, based on the idea that $\|\nabla I\|$ is an estimation of the presence of an edge. g should be monotonic and decrease between 1 and 0. The authors did not report a big difference depending on g. One proposal was the Gaussian function[6]:

$$g(x) = e^{-x^2/K^2} \qquad (3)$$

Then, it is shown that using adiabatic boundary conditions ($c = 0$ on the image boundaries), the maximum principle is followed: no new maxima emerge at coarse scales. Another nice property is that edges are enhanced by the diffusion process, while noise is reduced (this is demonstrated analytically, under some reasonable hypotheses). One interesting result is that if $\phi(I_x) = c.I_x$ denotes the flux along x, one sees at the same time forward diffusion almost everywhere ($\phi'(I_x) > 0$), and backward diffusion around the edges ($\phi'(I_x) < 0$). Backwards diffusion is said to quickly shrink, so that the process is stable.

[5]It should be noted that unlike many papers on edge detection, edges are considered as region boundaries, and the issue of edge detection is seen as a joint issue, together with regions segmentation.

[6]The setting of the K parameter is not much discussed: it is either fixed by hand, or using Canny's noise estimator (Canny, 1986). The idea is that K should be set to the "typical contrast value", and thus may vary across the image if this typical value strongly varies.

A practical implementation is proposed, on a 4-connexity lattice, where the nodes code for intensity (I), and the arcs code for diffusion (c). The author prove that no maximum can emerge: $I(t+1)$ is always between the local `Min` and `Max` in the neighbourhood of $I(t)$.

This implementation is compared to other edge detectors, which are classified into two classes: fixed-neighbourhood (such as Canny (1986)) and energy/probability global schemes, such as Geman and Geman (1984). Compared to Canny, in addition to the nice property of avoiding localization errors at coarse scales, the combination of result images at various scales is avoided. Moreover, edge thining and linking becomes almost unnecessary. They also note that the proposed algorithm is massively parallel.

Anisotropic diffusion is also compared to energy minimization approaches (Geman and Geman, 1984), which is similar to maximizing a Markov probability distribution function. In this approach, the Energy U is the sum of an *a priori* term (containing prior knowledge about the image space) and an *a posteriori* term (which depends on the available information, that is, the image data). An interesting result is that the proposed anisotropic diffusion may be seen as a gradient descent of the *a priori* part V of the energy function.

4. Edge detection techniques. An overview.

(Ziou and Tabbone, 1998)

The first sentance of this paper may be quoted: "In computer vision, edge detection is a process which attemps to capture the significant properties of *objects* in the image" (emphasized by myself). these properties give rise to variations in the image (step edges, line edges, junctions). The purpose of edge detection is to localize them, "and to identify the physical phenomena which produces them".

The basic tool is the computation of image derivatives, which is sensible to noise. Smoothing removes some noise, at the cost of lost information and spatial uncertainty. Edge detection algorithms have tried to cope with these two constraints.

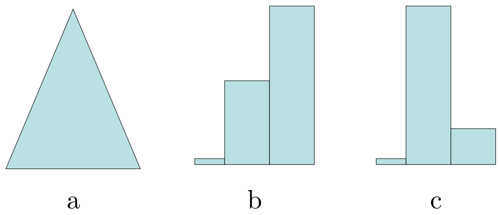

Figure 1: Line, staircase and pulse profiles.

Edge definition. Physical edges correspond to discontinuities in the topological, geometrical or photometrical properties of the objects. They inform about variations in reflectance, illumination (Stockham, 1972), orientation or depth, to name some.

A *line* results from mutual illumination between two objects in contact (Fig. 4.a).

The "usual" edge is called a *step*: it is a fronteer between two regions. Due to the camera point spread function, to noise, to the pixel's size, to quantification, and so on, the image is not identical to the scene[7]. Still, step edges are localized where the first order derivatives reach local extrema (that is, zero-crossings of the second order derivative).

Most edge definitions do not consider the spatial distribution of edge points: instead, the edge detectors detect "edge points" rather than edges. The basic approach is to look for double step edges, that is, pairs of inflexion points near each other. Depending on their relative sign, they are denoted "staircase" or "pulse" (Fig. 4.b and 4.c). T-junctions are also encountered sometimes.

Properties of edge detectors. An edge detector uses an image as input, and produces an edge map as output. Additional information may be provided, such as the segmentation, strength, orientation and scale of the edges. Edge Detectors (ED) may be classified depending on their use of prior information. General-purpose ED (without prior information) are local operators, while contextual ED use prior information about the edges, or about the scene. Most published ED use no prior information. Both kind of operators use the same 3-step scheme: differentiation (computing the derivatives), smoothing (removing noise) and labeling (localization, and removing "false" edges). Differentiation may happen befor or after the smoothing.

Smoothing. Smoothing is a tradeoff between information loss and noise reduction[8]. Thus, a smoothing may be said to be "optimal" if one can quantify both the information loss (edge conservation) and the noise reduction. This is the aim of the regularization theory.

To make the problem easier, one may choose a function family (e.g. Gaussian filters, cubic splines, Green functions, etc.). Then, a scale parameter (regularization parameter) may be choosen in order to optimize something. It makes people happy, and they say the problem is well-posed, because an analytical computation is made possible, providing that we know something about the "noise".

Non-linear filters perform better than linear filters, especially with random white noise (Pitas and Venetsanopoulos (1986), see also Perona and Malik (1990)). Still, most ED use linear filters, and rotationnal invariance is preferred. Using linear filters, it is te same to smooth the image before or after the differenciation. Using non-linear filters, the smoothing must precede the differentiation.

Differentiation. The differentiation is the computation of the image derivatives, in order to localize the edges. The most commonly used operators are the gradient $\vec{\nabla}$, the Laplacian ∇^2 and second order directional derivatives (this last one is neither linear nor invariant to rotations):

$$\vec{\nabla} I = \{\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}\} \tag{4}$$

$$\nabla^2 I = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \tag{5}$$

$$\frac{\partial^2 I}{\partial \vec{n}^2} \tag{6}$$

[7] The map is not the territory (Korzybski, 1931).

[8] Thus, some notions about the noise's properties are needed.

where $\vec{n}$ is the gradient direction and ψ the corresponding angle. One key assumption is that the gradient is perpendicular ot the edge orientation.

Edge labeling. The basic idea is to threshold a plausibility index on the output of the previous two steps. The plausibility is often taken as the gradient modulus. As this may produce thick edges, a skeletization may be required. Canny also proposed to find local maxima along $\vec{n}$.

For zero-crossing algorithms, one checks whether there are positive and negative values in the neighbourhood of a pixel. Then the question arises: *what* value should be thresholded? some authors use the gradient (Clark, 1989), other prefer the slope a the zero-crossing location (Marr and Hildreth, 1980; Ziou and Tabbone, 1993), which is quite noise-sensitive [9].

False edge do not only originate from noise (they may be due to surface texture or image acquisition). Most operators use a threshold to select the edges, which breaks the edges. Hysteresis thresholding, then, is an improvement (Canny, 1986). Another kind of false edge is the "phantom edge", which may appear for thick edges when one looks at it at the wrong scale.

One unsolved problem is the threshold selection, which may be local or global. Most authors use a trial-error process (needing a ground truth database).

Multiscale approaches. Marr and Hildreth (1980), after Rosenfeld and Thurston (1971), apply edge detectors at various scales, and then combine the resulting edges. The scale selection issue is replaced by a combination problem. Moreover, in multiscale approaches, the smoothing are computed *after* the differenciation.

One key issue is to label the same edge, detected at various scales, with the same label. Fine-to-coarse and coarse-to-fine strategies have been proposed (Canny, 1986; Bergholm, 1987). One original approach is the one from Ziou and Tabbone (1993), which is neither coarse-to-fine nor fine-to-coarse. The behaviour of a given edge is compared to 4 edge models (ideal, blurred, pulse, staircase) in scale space, and two scales are selected for each edge.

Evaluation of detectors. A good ED produces primitives (chaine, straight lines, circles, splines, etc.) from which an object may be found with little computation. Thus, it is a first step in a specific strategy for object segmentation, and the evaluation should address the object segmentation task. However, experimental evaluations of edge detectors have focused on their specific failures in the subtask of "true edge" detection.

Subjective judgments by humans "cannot be used to measure the performance of detectors, but only to establish their failure". This is because human judgments are unaccurate and depend on many strange factors[10] (however, see Heath et al. (1997)). Objective evaluation is much better, it just needs a *subjective* labelling of true edges (ground truth)[11]. Quantitative indexes are proposed, for instance, in Kitchen and Rosenfeld (1981) and Kanungo et al. (1995). The main idea is to look for various *a priori* error types, such as false detection, non-detection, double edges, localization error, lack of continuity, etc.

Survey of edge detectors. The first operators were limited to the differentiation (without filtering), using 3×3 masks, such as Δ_x and $\Delta_y = \Delta_x^T$:

$$\Delta_x = \begin{bmatrix} -1 & 0 & 1 \\ -a & 0 & a \\ -1 & 0 & 1 \end{bmatrix}$$

where $a = 1$ for Prewitt's mask, $a = 2$ for Sobel's mask. The Laplacian is implemented with:

$$\nabla = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

The idea of parametric fitting is to use an *a priori* model of what an edge is (Hueckel, 1971; Baker

[9] Using the gradient module and the zero-crossing image to compute the edges has something to do with the joint bilateral filtering idea of Petschnigg et al. (2004).

[10] For instance, Webb and Pervin (1984) found that subjective contours do not correspond to local discontinuities!

[11] :-)

et al., 1998). A pixel's neighbourhood is compared to this model, and classified as an edge or not. Models may be available for step edges, corners, etc. Haralick proposed to use a function family (e.g. Tchebychev's polynomials) instead of a single model for the local fitting (Haralick, 1984). One good point in parametric fitting is that the images are considered as sample data, and the fit is in $\mathbb{R}^2$.

The alternative to parametric fitting is optimal enhancement: (1) performance criterion definition, and (2) optimal filter selection (Marr and Hildreth, 1980; Canny, 1986). The performance criterion should have a mathematical expression, so as to select the filter with an analytic derivation of the criterion's constraints. For instance, Canny's performance is related to the detector's performance in the presence of Gaussian white noise. Many authors followed Canny's path (Deriche, 1987; Sarkar and Boyer, 1991; Pettrou and Kittler, 1991). Still, one weakness of most ED is that they do not detect 2D edges, only 0D edge points.

Lines. A line is a local extremum of a grey level image. They are detected by thinning algorithms, in binary (Smith, 1987) and grey-level images (Dyer and Rosenfeld, 1979). Haralick's technique of polynomial fitting is also relevant here. Giraudon (1991) proposed that a line is a negative local maximum of the second derivative of the image (rather than a zero-crossing in Marr and Hildreth (1980)). To some extend, we agree with this.

5. Gaussian-Based Edge-Detection Methods. A Survey.

(Basu, 2002)

Edge detection and Gaussian filters. This pessimistic paper reviews the main developments in edge-detection techniques, using Gaussian filters, after Marr and Hildreth's seminal paper in 1980 (Marr and Hildreth, 1980). The edge-detection problem is defined, as usual, as having nothing to do with edges: " In a gray level image, an edge may be defined as a sharp change in intensity. Edge detection is the process which detects the presence and location of these intensity transitions". The ambiguity may be related to the various applications of edge detection, from image compression to visibility estimation. However, the author claim that the main purpose of edge detection is to mimic the Human Vision System (HVS): it is supposed to be a low-level image processing step, on the way to object detection.

Old operators, such as Sobel's gradient, are very sensitive to noise. Since then, the main trend use linear operators which are derivatives of some sort of smoothing filters, among which the Gaussian filter is the most popular. These operators are described in the present paper, while a more general overview of edge detection can be found in Ziou and Tabbone (1998).

The theoretical framework of edge detection seems to be the signal detection theory: "true edges" are presents in an image, and the challenge is to detect them (and localize at their "true position") without missing any, and without false detections (that is, locate "intensity changes where edges do not exist"). Thus, one key issue in edge detection is the sensibility to noise (robustness), where noise is defined as an intensity transition which is not an edge. In short, noise is a false edge[12].

Based on the signal/noise framework, a problem appears. The noise is expected to be removed by bluring filters, at the cost of localization errors for the remaining edges. Anyways, a consensus seems to appear on the fact that smoothing an image on a single scale is not a good idea. The key issue in multi-scale edge detection, still, is how the response of each filter are combined.

The Gaussian filter was proposed by Marr and Hildreth, because they share a number of intersting properties. First, it is a good candidate to mimic some aspects of image processing in the HVS (Enroth-Cugell and Robson, 1966) (see also Wandell (1995)), which may include Difference of Gaussian (DoG) operators. Marr and Hildreth demonstrated

[12] To my opinion, the edge/noise distinction is due to an ill-defined problem. One way to make this clear is to split between edges in the image and edges in the scene: edge-detection should detect the object's edges in the scene, not in the image.

that under some hypothesis, a Gaussian filter followed by a Laplacien is close to a DoG.

Starting from the edge-detection as an optimization problem, Canny derived an optimal edge detector which turned out to be a first derivative of a Gaussian function (Canny, 1986). One fine property of Gaussian filters is that they do not create new zero-crossing when moving from fine-to-coarse scales (together with the Laplacian, see Yuille and Poggio (1986)). Another fine property of the Gaussian filter is that it is the only one to fulfill Eq. 7:

$$\Delta x \Delta \omega \geq \frac{1}{2} \tag{7}$$

where Δx and $\Delta \omega$ are the variance of the filter in the spatial and frequency domains, respectively.

Marr and Hildreth. The first insight from Marr and Hildreth is the spatial coincidence: a relevant edge should be detected at several spatial scales. This was the rationale for a multiscale analysis in edge detection. They also argue that the Gaussian filter is a good candidate for edge detection, thanks to Eq. 7. Their basic proposal Marr and Hildreth (1980) was that the edges produce zero-crossings in the second derivative of the Gaussian filter. Given a Gaussian filter G_σ:

$$G_\sigma(x, y) = \frac{1}{2\pi\sigma^2} e^{-r^2/2\sigma^2} \tag{8}$$

the Laplacian of Gaussian (LoG) is:

$$\nabla^2 = \frac{\partial^2 G}{\partial x^2} + \frac{\partial^2 G}{\partial y^2} = \frac{r^2 - 2\sigma^2}{\sigma^4} G \tag{9}$$

One problem with such a detector is spatial accuracy, especially with corners and curves.

Marr and Hildreth claim that their operator have support from biological vision, namely, from the DoG model of Enroth-Cugell and Robson (1966)[13]. This was another insight: rely edge detection in computer vision to edge detection in human vision.

[13]However, vision science results support the DoG model rather than the zero-crossing proposal.

The problem when equating edges with zero-crossings is that it fails when the edges are close to each other, and when the Signal-to-Noise Ratio (SNR) is low. The classical problems are the shift of the edge location (related to σ), missed edges, and the detection of "false" edges (noise artifacts). In addition, combining LoG at different scales raise specific problems:

- a given edge produce zero-crossings only in a limited range of scales;
- the zero-crossing of a given edge moves (spatially) from one scale to another;
- noise also produces zero-crossings at small scales.

Three trends emerge from Marr's seminal paper:

1. improving Marr's approach, under linear constraints. The key paper here is Canny (1986);
2. improving Marr's approach, focusing on biological vision. The key paper here is Kennedy and Basu (1997);
3. improving Marr's approach, in the non-linear domain. The key paper here is Perona and Malik (1990).

Linear approach. Canny proposed an edge-detection operator from an optimization point of view, and happened to find out that the first derivative of a Gaussian filter was this optimal operator. His idea was that an optimal edge-detector should be a good detector, with good localization, and should give only one detection for a single edge. This was derived from an information theory point of view, computing the SNR, detection and localization for a given edge, noise and detector (local filter). For step edges, Canny's optimal operator was similar to Marr's LoG. Canny's proposal for the multi-resolution problem was called *feature synthesis*, and is fine-to-coarse. Two thresholds are included in an hysteresis threshold; their value depend on the noise estimation.

Witkin (1983) initiated intensive studies about scale-space localisation of zero-crossings. For instance, Bergholm (1987) proposed a coarse-to-fine method, called *edge focusing.* A strong Gaussian smoothing detects edges, and these edges are tracked at finer scales in the vicinity of the higher level edges. Thus, their localisation is more and more accurate, and the final edges are not due to noise. Of course the start and end scales have to be chosen; moreover, a threshold is needed. Instead of zero-crossings, Goshtasby records the sign of the pixels, after filtering with a LoG operator (Goshtasby, 1994), which avoids disconnected edges.The image analysis compares the sign of a given pixel at consecutive scales. There is no need for arbirary scale choice, the main problem is the amount of memory needed. Instead of a full multiscale approach, Jeong and Kim (1992) compute, for each location, the optimal size for a Gaussian filter in terms energy miminization. They proposed a definition of energy such that σ is large in flat regions, and small in sharp regions. Unfortunately, the computation is slow.

The LoG zero-crossings may be considered as a guide to the true edges (Quian and Huang, 1996). This was done using a parametric edge model, with a SNR optimization[14]. In this approach, edge segments are combined in a fine-to-coarse strategy. Then, Lindeberg (1998) proposed to compute two measures of the *edge strength*: the magnitude of the gradient, and a second index, to check if the gradient magnitude is maximal in the gradient direction. A parameter γ is added, to make scale selection dependent on the diffuseness (sharpness) of the edge. Following this approach, Elder and Zucker (1998) include the sensor properties, and make the scale a function of the second moment of the sensor noise. The authors argue that the classification between important and unimportant edges should not happen at such a low level of image processing. Basu's reply is that it should happen somewhere.

Non-linear approach. The starting point of non-linar approaches of edge detection with Gaussian filter is the similarity between this filter and the heat equation; only, the spatial scale σ in the Gaussian filter coresponds to time t in the diffusion equation. Based on this analogy, Perona and Malik (1990) proposed a multislcae image representation, based on anisotropic diffusion. The heat diffusion coefficient is made dependent on the image gradient. The main instability of this method is the *stairacse effect*: an edge is split in linear segments separated by jumps (Nitzberg and Shiota, 1992); however it is rarely observed in practical situations.

Human vision approach. Basu (1994), followed by Kennedy and Basu (1997), introduced a Line-Weight Function (LWF) to enhance the edges. It is a combination of zero and second order Hermite functions, which is equivalent to a Gaussian and its second derivative. The operator was consistent with spatial vision knowledge about edge detection (Young, 1987), as well as a mathematical derivation of contrast sensitivity in the HVS (Stewart and Pinkham, 1991).

[14] We call it the Tarel top-down approach.

References

Baker, S., Nayar, S. K., Murase, H., 1998. Parametric feature detection. International Journal of Computer Vision 27 (1), 27–50.

Basu, M., 1994. A Gaussian derivative model for edge enhancement. Pattern Recognition 27, 1451–1461.

Basu, M., 2002. Gaussian-based edge-detection methods. A survey. IEEE Trans. on Systems, Man and Cybernetics, Part C 32 (3), 252–260.

Bergholm, F., 1987. Edge focusing. IEEE Trans. on Pattern Analysis and Machine Intelligence 9, 726–741.

Campbell, F. W., Robson, J. G., 1968. Application of Fourier analysis to the visibility of gratings. Journal of Physiology 197, 551–566.

Canny, J., 1986. A computational approach to edge detection. IEEE Trans. on Pattern Analysis and Machine Intelligence 8, 679–698.

Clark, J. J., 1989. Authenticating edges produced by zero-crossing algorithms. IEEE Transactions Pattern Analysis and Machine Intelligence 11 (1), 43–57.

Deriche, R., 1987. Using Canny's criteria to derive an optimal edge detector recursively implemented. International Journal on Computer Vision 2 (1).

Dyer, C. R., Rosenfeld, A., 1979. Thinning algorithms for gray-scale pictures. IEEE Trans. on pattern Analysis and Machine Intelligence 1 (1), 88–89.

Elder, J. H., Zucker, S. W., 1998. Local scale control for edge detection and blur estimation. IEEE Transactions on Pattern Analysis and Machine Intelligence 20, 699–716.

Enroth-Cugell, C., Robson, J. G., 1966. The contrast sensitivity of retinal ganglions cells of the cat. J. Physiol. 187, 517–522.

Fram, J. R., Deutsch, E. S., 1975. On the quantitative evaluation of edge detection schemes and their comparison with human performance. IEEE Trans. on Computers 26 (6), 616–628.

Geman, S., Geman, D., 1984. Stochastic relaxation, Gibbs distrbutions, and teh Bayesian restoration of images. IEEE Trans. on Pattern Analysis and Machine Intelligence 6, 721–741.

Giraudon, G., 1991. Edge detection from local negative maximum of second derivative. In: Proc. IEEE CVPR. pp. 643–645.

Goshtasby, A., 1994. On edge focusing. Image and Vision Computing 12, 247–256.

Haralick, R. M., 1984. Digital step edges from zero crossing of second directional derivatives. IEEE Trans. on Pattern Analysis and Machine Intelligence 6 (1), 58–68.

Heath, M., Sarkar, S., Sanocki, T., Bowyer, K., 1997. A robust visual method for assessing the relative performance of edge-detection algorithms. IEEE Trans. on Pattern Analysis and Machine Intelligence 19 (12), 1338–1359.

Hubel, D. H., Wiesel, T. H., 1962. Receptive field, binocular interaction and functional architecture in the cat's visual cortex. J. Physiology 160, 106–154.

Hueckel, M. H., 1971. An operator which locates edges in digitalized pictures. Journal of the ACM 19 (1), 113–125.

Hueckel, M. H., 1973. A local visual operator which recognizes edges and lines. Journal of the ACM 20, 634–647.

Jeong, H., Kim, C. I., 1992. Adaptative determination of filter scales for edge detection. IEEE Trans. on Pattern Analysis and Machine Intelligence 14, 579–585.

Kanungo, T., Jaisimba, M. Y., Palmer, J., Haralick, R. M., 1995. A methodology for quantitative performance evaluation of detection algorithm. IEEE Trans. on Image Processing 12 (4), 1667–1674.

Kennedy, L. M., Basu, M., 1997. Image enhancement using a human visual system model. Pattern Recognition 30 (12), 2001–2014.

Kitchen, L., Rosenfeld, A., 1981. Edge evaluation using local edge coherence. IEEE Trans. on Systems, Man and Cybernetics 11 (9), 597–605.

Koendrink, J., 1984. The structure of images. Biological Cybernetics 50, 363–370.

Korzybski, A., 1931. A non-aristotelian system and its necessity for rigour in mathematics and physics. In: Proc. American Association for the Advancement of Science. American Mathematical Society, New Orleans (USA), reprinted in Science and Sanity, 1933, p. 747–61.

Land, E. H., McCann, J. J., 1971. Lightness and retinex theory. Journal of the Optical Society of America 61 (1), 1–11.

Leipnik, R., 1960. The extended entropy uncertainty principle. Information and Control 3 (1), 18–25.

Lindeberg, T., 1998. Edge detection and ridge detection with automatic scale selection. International Journal of Computer Vision, 117–156.

Marr, D., Hildreth, E., 1980. Theory of edge detection. In: Proceedings Royal Society London B. Vol. 207. pp. 187–217.

Marr, D., Ullman, S., 1981. Directional selectivity and its use in early visual processing. In: Proceedings of the Royal Society of London B. Vol. 211. pp. 151–180.

McLeod, I. D. G., 1972. Comment on "techniques for edge detection". In: Proc. IEEE (Lett.). Vol. 60. p. 344.

Nitzberg, M., Shiota, T., 1992. Nonlinear image filtering with edge and corner enhancement. IEEE Trans. on Pattern Analysis and Machine Intelligence 14, 826–833.

Perona, P., Malik, J., 1990. Scale-space and edge detection using anisotropic diffusion. IEEE Trans. on Pattern Analysis and Machine Intelligence 12, 629–639.

Petschnigg, G., Agrawala, M., Hoppe, H., Szeliski, R., Cohen, M., Toyama, K., 2004. Digital photography with flash and no-flash image pairs. In: Proc. SIGGRAPH. pp. 664–672.

Pettrou, M., Kittler, J., 1991. Optimal edge detector for ramp edges. IEEE Trans. on Pattern Analysis and machine Intelligence 13 (5), 483–491.

Pitas, I., Venetsanopoulos, N., 1986. Edge detector based on nonlinear filters. IEEE Trans. on Pattern Analysis and Machine Intelligence 8 (4), 538–550.

Quian, R. J., Huang, T. S., 1996. Optimal edge edtection in two-dimensional images. IEEE Trans. on Image Processing 5, 1215–1220.

Rodieck, R. W., Stone, J., 1965. Analysis of receptive fields of cat retinal ganglion cells. J. Neurophysiology 28, 833–849.

Rosenfeld, A., Thurston, M., 1971. Edge and curve detection for visual scene analysis. IEEE Transactions on Computer 20 (5), 562–569.

Sarkar, S., Boyer, K. L., 1991. On optimal infinite impulse response edge detection filters. IEEE Trans. on Pattern Analysis and machine Intelligence 13 (11), 1154–1171.

Smith, R. W., 1987. Computer processing of line images: A survey. Pattern Recognition 20 (1), 7–15.

Stewart, A. L., Pinkham, R., 1991. A spatial-variant differential operator for visual sensitivity. Biological Cybernetics 64, 373–379.

Stockham, T. G., 1972. Image processing in the context of a visual model. IEEE 60 (7), 828–841.

Wandell, B., 1995. Foundations of vision. Sinauer associates, Sunderland, MA, USA.

Webb, J. A., Pervin, E., 1984. The shape of subjective contours. In: AAAI. pp. 340–343.

Wilson, H. R., Bergen, J. R., 1979. A four mechanisms model for spatial vision. Vision Research 19, 19–32.

Wilson, H. R., Giese, S. C., 1977. Threshold visibility of frequency gradient patterns. Vision Research 17, 1177–1190.

Witkin, A., 1983. Scale-space filtering. In: Proc. Int. Joint. Conf. Artificial Intelligence. Vol. 2. pp. 1019–1022.

Young, R. A., 1987. The Gaussian derivative model for spatial vision: Part I, Retinal mechanisms. Spatial Vision 2 (4), 273–293.

Yuille, A. L., Poggio, T. A., 1986. Scaling theorems for zero-crossings. IEEE Trans. on Pattern Analysis and Machine Intelligence 8, 15–25.

Ziou, D., Tabbone, S., 1993. A multi-scale edge detector. Pattern Recognition 26 (9), 1305–1314.

Ziou, D., Tabbone, S., 1998. Edge detection technique. An overview. Journal of Pattern Recognition and Image Analysis 8, 537–559.