# Generalized antisymmetric filters for edge detection

Nicolás Madrid
Institute for Research and Applications of Fuzzy Modeling, University of Ostrava
Ostrava, Czech Republic
Email: nicolas.madrid@osu.cz

Carlos Lopez-Molina
Dpto. Automatica y Computacion
Universidad Publica de Navarra
Pamplona, Spain
Email: carlos.lopez@unavarra.es

Bernard De Baets
Dept. of Mathematical Modeling, Statistics and Bioinformatics, Ghent University
Gent, Belgium
Email: Bernard.DeBaets@UGent.be

***Abstract*—A large number of filters has been proposed to compute local gradients in grayscale images, usually having as goal the adequate characterization of edges. A significant portion of such filters are antisymmetric with respect to the origin. In this work we propose to generalize those filters by incorporating an explicit evaluation of the tonal difference. More specifically, we propose to apply restricted dissimilarity functions to appropriately measure the tonal differences. We present the mathematical developments, as well as quantitative experiments that indicate that our proposal offers a clear option to improve the performance of classical edge detection filters.**

***Index Terms*—Image processing; edge detection; restricted dissimilarity function; convolution filter.**

## I. Introduction

Edge detection is assumed to be a multiphase process, as enunciated by Torre and Poggio in [1], where the authors stated that *"A traditional belief in computational vision is that [...] at least two separate stages are required"*. Different authors have proposed breakdown structures to specify the role and goal of each phase of such process. Examples are the proposals by Prager (*preprocessing*, *edge representation*, *grouping* and *postprocessing* [2]) or Law *et al.* (*filtering*, *detection* and *tracing* [3]), although the most comprehensive work is that by Bezdek et al. [4]. In [4] four phases were defined, namely *conditioning*, *feature extraction*, *blending* and *scaling*. The *feature extraction* phase has been the one receiving more attention from the scientific community. In this phase local edge cues and indicators are gathered, so that a decision about the edges can be made at subsequent phases. In a sense, the information in the conditioned (enhanced) image is projected to a feature space for a further classification of the pixels into edge or non-edges.

Most of the proposals for feature extraction for edge detection are based on convolution filters, which are used to estimate the first or second derivatives of an image. This is due to the fact that the derivatives of an image contain most of the information about the local intensity changes, and can be consequently used to discern the transition between objects. In the early years of digital image processing discrete convolution filters were presented, being the most famous those by Sobel [5], Prewitt [6] and Roberts [7]. In the 80s, different works were grounded on notions from signal processing to develop filters that were optimal to detect edges under some restrictive conditions [8], [9], [10]. These works, which include the widespread proposals by Marr and Hildreth [11] and Canny [12], [13], have been often revisited. Although their mathematical grounds have been critiziced (or even shown to be debatable [14], [15]), they usually stand as the reference for comparison. Moreover, despite a large number of edge detection methods has been presented thereafter, convolution filters still stand as one of the most common ways to compute edge features.

The study of convolution filters for gradient computation has been rather still in the past 20 years. Despite remarkable works have been presented, most of the novelties consist of introducing enhanced constraints leading to the redefinition of the concept of optimality and, consequently, producing new convolution filters [10], [15]. A significant breakthrough has been the multiscale methods, which propose to generate scale-spaces from which a richer interpretation of the partial derivatives can be extracted. Although already suggested by Canny (who proposed the so-called *feature synthesis* to combine the results of convolution filters of different size and orientation), the study of multiscale methods and their corresponding scale-spaces was boosted after the works by Witkin [16], Perona and Malik [17] or Lindeberg [18].

In this work we revisit the convolution filters for gradient computation. More specifically, we propose to incorporate a more elaborated treatment of the tonal information in the decision about the significance of a given tone transition. We propose a generalization that can be applied to most of such filters, and has as ultimate goal the enhancement of their adaptability to specific problems. This generalization is partially inspired by bilateral filtering [19], in the sense that it aims to incorporate into the spatial filtering (provided by the convolution filter) an explicit evaluation of the tonal differences.

In Section II we recall some concepts that are applied in the development of this work. Section III contains our proposal, which is backed up by quantitative, experimental results in Section IV. To conclude, in Section V we summarize the work and discuss some future lines of research.

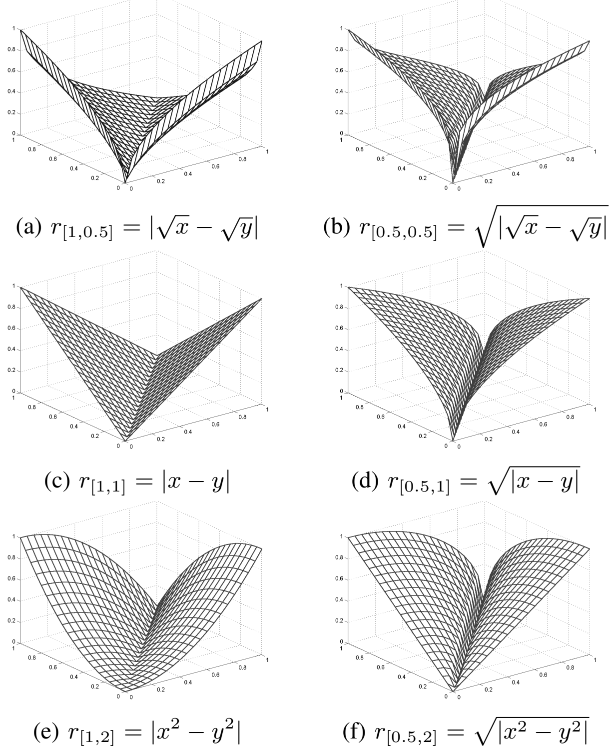
## II. Preliminaries

In this work we make use of Restricted Dissimilarity Functions (RDFs) [20].

*Definition 1:* [20] A *restricted dissimilarity function* is a mapping $r : [0,1]^2 \to [0,1]$ satisfying

(R1) $r(x,y) = r(y,x)$ for all $x, y \in [0,1]$;

978-1-4799-3400-3/13/$31.00 



(a) $r_{[1,0.5]} = |\sqrt{x} - \sqrt{y}|$ (b) $r_{[0.5,0.5]} = \sqrt{|\sqrt{x} - \sqrt{y}|}$

(c) $r_{[1,1]} = |x - y|$ (d) $r_{[0.5,1]} = \sqrt{|x - y|}$

(e) $r_{[1,2]} = |x^2 - y^2|$ (f) $r_{[0.5,2]} = \sqrt{|x^2 - y^2|}$

Fig. 1. Restricted dissimilarity functions constructed as in Proposition 1.

(R2) $r(x, y) = 1$ if and only if $(x, y) = (0, 1)$ or $(x, y) = (1, 0)$:

(R3) $r(x, y) = 0$ if and only if $x = y$;

(R4) for all $x, y, z \in [0, 1]$, if $x \leq y \leq z$, then $r(x, y) \leq r(x, z)$ and $r(y, z) \leq r(x, z)$.

The following proposition shows a way to construct RDFs from automorphisms.

*Proposition 1:* Let $\varphi_1$ and $\varphi_2$ be two automorphisms of the unit interval, then the mapping $r : [0, 1]^2 \to [0, 1]$ defined by

$$r(x, y) = \varphi_1(|\varphi_2(x) - \varphi_2(y)|) \tag{1}$$

is a restricted dissimilarity function.

In order to narrow down the scope of this work, we will only consider the RDFs constructed as in Proposition 1. More specifically, we refer with $r_{[\alpha,\beta]}$, where $\alpha, \beta \in \mathbb{R}$, to the RDF constructed with $\varphi_1(x) = x^\alpha$ and $\varphi_2(x) = x^\beta$. That is,

$$r_{[\alpha,\beta]}(x, y) = (|x^\beta - y^\beta|)^\alpha \ . \tag{2}$$

Figure 1 includes some RDFs constructed as in Eq. (2). Note that $r_{[1,1]}$ corresponds to the absolute difference.

In the remainder of this work we use lower case for scalar data ($a$, $b$), while the vectorial data takes up boldface symbols ($\vec{a}$,$\vec{b}$).

## III. Proposal

In this section we present a reformulation of the convolution operator for partial derivative approximation. This reformulation is grounded on the assumption that the filters used for such approximations are antisymmetric. Our proposal results in a generalization of the classical convolution in which the measurement of the tonal differences is carried out by a problem-specific funtion. Furthermore, we describe how to apply RDFs in the generation of such problem-specific funtions and relate our proposal with other approaches in the literature.

### A. Computing partial derivatives on digital images

It is well known that the concept of derivative roots on the variation of a function. Hence, in a discrete setting, the local difference could be considered as a possible generalization of the derivative. In this way, the partial derivatives of an image $f$ at each pixel could be computed as its difference with the *next* one, i.e.:

$$\begin{aligned} \frac{\partial f}{\partial x}(i, j) &= f(i + 1, j) - f(i, j) \text{ and} \\ \frac{\partial f}{\partial y}(i, j) &= f(i, j + 1) - f(i, j) \,. \end{aligned} \tag{3}$$

However, in practical applications this renders in deceptive results, due to the contamination of the data, which makes this methodology extremely unreliable. An example of the unreliability of the estimations based on Eq. (3) can be seen in Fig. 2. In the top row of this figure we see two signals, the second being a contaminated version of the first. In this signal four edges coexist, two of them related to each of the objects (plateaus) in the signal. Note that two of such edges should produce positive peaks (maxima) of the first derivative, while the remaining two should manifest as local minima. In the second row of Fig. 2 we see the approximation of the derivative of such signals using Laligant's filter [21], which corresponds to Eq. (3). We observe that, although accurate in the original signal, it becomes extremely sensitive to the noise in the contaminated one. Consequently, different methodologies have been developed to compute partial derivatives in a discrete setting able to cope with contamination, mainly based on convolution filters.

There exist very different filters for the computation of the partial derivatives of an image, either discrete (such as those by Sobel [5] and Prewitt [6]) or continuous (such as the well-known Canny filter [12] and its subsequent evolutions [15], [22]). Any of these filters is used for convolving the image. Specifically, the partial derivative of the image is computed as the convolution of the image with the filter $g$.

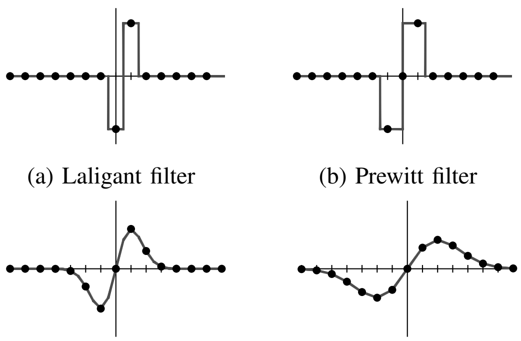
Let us recall that the convolution of an image $f$ and a filter $g$ is defined as

$$f * g(\mathbf{x}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(\xi) \cdot f(\mathbf{x} + \xi) \cdot d\xi \,. \tag{4}$$

In order to approximate a partial derivative of an image by a convolution it is necessary to choose an appropriate filter. Most of the filters used for such task satisfy antisymmetry[1,2] $g(\xi) =$

[1] Note that from a pure mathematical point of view, $g$ is an odd function. However, we have opted to follow here the terminology generally used in image processing.

[2] In general, this property does not hold for those filters used to estimate the second derivative, since these filters are usually symmetric [9], [11].



(a) Laligant filter (b) Prewitt filter

(c) Canny filter ($\sigma = 1.0$) (d) Canny filter ($\sigma = 2.0$)

Fig. 3. One-dimensional convolution filters used in Fig. 2. The dots account for the discrete representation of the filter.

$-g(-\xi)$. Thus, by considering an antisymmetric filter $g$ we have

$$\begin{aligned}
f * g(\mathbf{x}) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(\xi)\cdot f(\mathbf{x}+\xi)\cdot d\xi = \\
&= \int_{-\infty}^{\infty}\int_{0}^{\infty} g(\xi) f(\mathbf{x}+\xi) d\xi + \int_{-\infty}^{\infty}\int_{-\infty}^{0} g(\xi) f(\mathbf{x}+\xi) d\xi \\
&= \int_{-\infty}^{\infty}\int_{0}^{\infty} g(\xi) f(\mathbf{x}+\xi) d\xi + \int_{-\infty}^{\infty}\int_{0}^{\infty} g(-\xi) f(\mathbf{x}-\xi) d\xi \\
&= \int_{-\infty}^{\infty}\int_{0}^{\infty} g(\xi) f(\mathbf{x}+\xi) d\xi - \int_{-\infty}^{\infty}\int_{0}^{\infty} g(\xi) f(\mathbf{x}-\xi) d\xi \\
&= \int_{-\infty}^{\infty}\int_{0}^{\infty} g(\xi)(f(\mathbf{x}+\xi) - f(\mathbf{x}-\xi)) d\xi \,.
\end{aligned}$$

Finally, note that from a practical point of view, we can assume a finite support of $g$. In that case, the formula above can be reformulated as:

$$f * g(\mathbf{x}) = \int_{-k}^{k}\int_{0}^{k} g(\xi)(f(\mathbf{x}+\xi) - f(\mathbf{x}-\xi)) d\xi \,. \quad (5)$$

where $[-k,k]^2$ is the support of the filter $g$.

From the previous equation, it is clear that the estimation of the partial derivative of the signal $f$ by a convolution involves:

- the difference between the intensity on both sides of the pixel (i.e. the expression $f(\mathbf{x}+\xi) - f(\mathbf{x}-\xi)$)
- weighted by the value of the filter $g$.

Hence, instead of measuring such a difference by subtraction, we can consider a mapping $\epsilon\colon [0,1]^2 \to [-1,1]$ in charge of measuring the perceived dissimilarity in the transition from the tone $f(\mathbf{x}-\xi)$ to $f(\mathbf{x}+\xi)$ and, subsequently, reformulate Eq. (5) as

$$f * g(\mathbf{x}) = \int_{-k}^{k}\int_{0}^{k} g(\xi)\epsilon(f(\mathbf{x}+\xi), f(\mathbf{x}-\xi)) d\xi \,. \quad (6)$$

In this way, the computation of the partial derivative of an image is dependent upon the spatial (domain) filter $f$ and a certain tonal (range) filter $\epsilon$. We refer to $g(\xi)$ as the *spatial term*, while $\epsilon(f(\mathbf{x}+\xi), f(\mathbf{x}-\xi))$ is called *tonal term.*

### B. Using restricted dissimilarity functions

We propose to replace the arithmetic difference in Eq. (6) by a function $\epsilon$ able to evaluate the perceived difference between two tones. This function should produce large absolute values when the perceived tonal dissimilarities are large, as well as near-zero values when the tones are perceptually similar. In this work we propose to generate such function $\epsilon$ using RDFs, more specifically to use as tonal evaluation term the expression

$$\epsilon_r(y,z) = \operatorname{sign}(z-y)\cdot r(y,z)\,, \quad (7)$$

where $r$ is a restricted dissimilarity function.

There is a practical motivation behind the use of RDFs for measuring the perceived dissimilarity between two tones. We have that the properties of RDFs (see Definition 1) are perfectly suited for our purposes. By using Eq. (7) the functions $\epsilon$ satisfy a minimal set of properties, namely:

(E1) $\epsilon_r(y,z) = -\epsilon_r(z,y)$:
(E2) $\epsilon_r(y,z) = 0$ if and only if $y = z$;
(E3) $\epsilon_r(y,z) = 1$ if and only if $(y,z) = (1,0)$;
and, $\epsilon_r(y,z) = -1$ if and only if $(y,z) = (1,0)$;
(E4) for any $t,y,z \in [0,1]$, $t \le y \le z$, we have that
$\epsilon_r(t,y) \le \epsilon_r(t,z)$ and $\epsilon_r(y,z) \le \epsilon_r(z,t)$ .

These properties guarantee the expected behaviour for the tonal evaluation. For example, because of (E2), we know that Eq. (6) will only produce a zero response if and only if $f(\mathbf{x}+\xi) = f(\mathbf{x}-\xi)$ for every position within the support of the filter.

In addition to the good behaviour guaranteed by the properties (E1)-(E4), constructing functions $\epsilon$ from RDFs has the advantage of providing a connection between a novel proposal, such as ours, with a well established field of research. In this way, we can aim to transfer the knowledge gathered in previous developments to our filtering proposal.

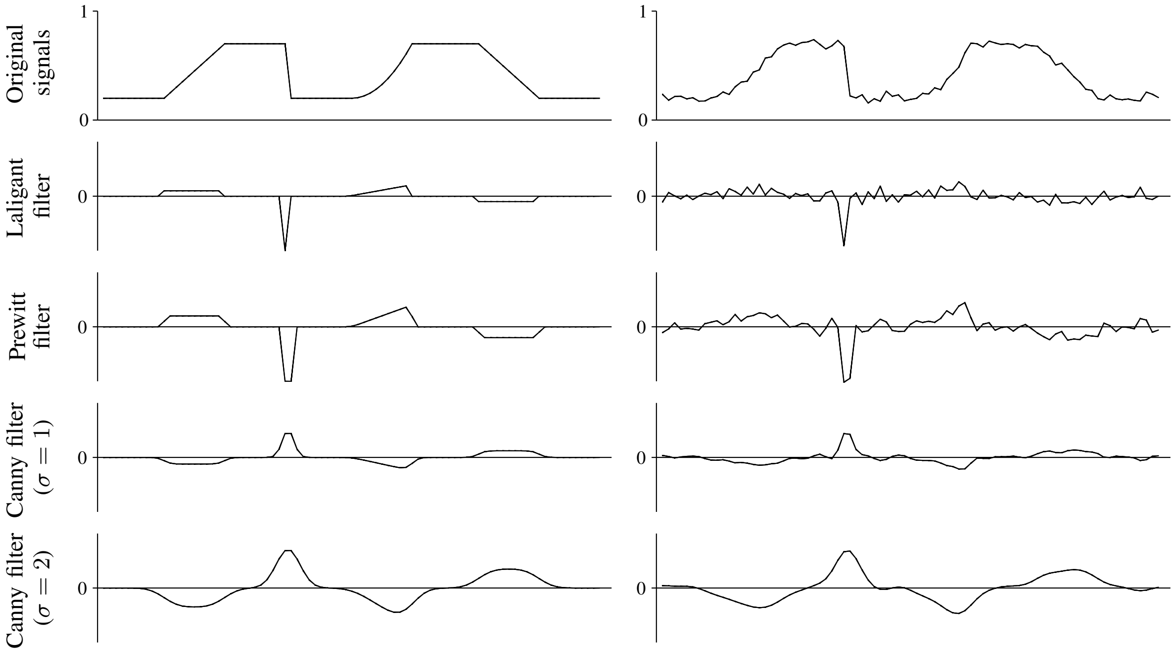
### C. Relationship with other filtering paradigms in the literature

This work is not the first proposal to incorporate explicit tonal analysis into classical filtering. In fact, any Content-Aware (CA) filtering technique somehow integrates the tonal analysis (to design the filter) with the spatial analysis (to apply the filter). However, both aspects of the problem are approached in two separated and consecutive steps.

The most relevant contribution integrating the treatment of tonal and spatial information in an explicit way is bilateral filtering [19]. In [19] the authors propose to combine filters in the tonal and spatial domain for CA regularization. More specifically, they propose to filter a signal $f$ as

$$f^*(\mathbf{x}) = \frac{1}{k_d}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\mathbf{x}+\xi)\cdot g(\xi)\cdot h(\|f(\xi) - f(\mathbf{x}+\xi)\|) d\xi$$

where $k_d$ is a normalization factor, $\|\cdot\|$ represents the Euclidean norm, $g$ is a classical low-pass filter and $h$ is a unary, decreasing function. Thus, the value of each translation $\mathbf{x}+\xi$ contributes to the regularized value of $f^*(\mathbf{x})$ as long as it is close in the spatial (filter $g(\xi)$) and in the tonal (function $h(\|f(\xi) - f(\mathbf{x}+\xi)\|)$) domains.




(a) Synthetic signal

(b) Contaminated signal

Fig. 2. Estimations of the first derivative of a signal computed using different 1D convolution operators. The contaminated signal has been obtained after adding to the synthetic signal zero-mean Gaussian noise with standard deviation 0.02. The convolution operators are those by Laligant [21], Prewitt [6] and Canny [13].

Bilateral filters have been thereafter applied to different problems, but they have been mostly dedicated to CA regularization with specific goal, such as denoising or de-texturing, among others [23]. Other authors have focused on developing mathematical models connecting bilateral filtering, with other well-known CA regularization techniques, such as anisotropic diffusion or mean-shift [24], [25], [26].

Our proposal, although conceptually similar to that in [19], renders in a different formulation. The most relevant change is that we consider tonal evaluation based on the values of symmetric positions of each pixel neighbourhood, while in bilateral filtering the tonal evaluation is performed taking into account the value of the neighbor and that of the central pixel (more specifically, the difference between them). The other critical difference is that the scope of the bilateral filter involves the whole neighbourhood around a each pixel, while we only consider half of it (the other is implicitly taken into account in the tonal evaluation).

## IV. Experimental validation

The proposal in Section III is based on the assumption that the arithmetic difference is not necessarily the best descriptor of the tonal dissimilarity between two pixels. However, there is a need for qualitatively and quantitatively validating this fact. Since qualitative experiments can be seen as mere examples, and given the limited space, we present in this section a quantitative experiment measuring the performance obtained by our proposal.

### A. Edge extraction technique

The procedure used to produce binary edge images is described as follows:

1. Smooth the image with a Gaussian filter with standard deviation $\sigma = 1$;
2. Compute the gradients as the combination of the partial derivatives, which are obtained according to Eq. (6) with the settings described below.
3. Compute the edginess at each pixel as the Euclidean magnitude of the components of its gradient.
4. Binarize the edge image using non-maxima suppresion [27] in combination with hysteresis [12], whose thresholds are set using the technique by Medina-Carnicer *et al.* [28].

In this experiment, we have used the following settings for our proposal (see Eq. (6)):

- $\epsilon$ as in Eq. (7) constructed from RDFs $r_{[\alpha,\beta]}$, with $\alpha, \beta \in \{0.5, 1, \ldots, 3\}$;
- $g_h$ and $g_v$ functions based on the Canny filter, i.e

$$g_h(\xi) = \frac{-\xi_x}{\sigma^2} e^{-\frac{\xi_x^2+\xi_y^2}{2\sigma^2}} \quad \text{and} \quad g_v(\xi) = \frac{-\xi_y}{\sigma^2} e^{-\frac{\xi_x^2+\xi_y^2}{2\sigma^2}}, \tag{8}$$

where $\sigma = 1$.

Note that, when using $r_{[1,1]}$, the algorithm becomes computationally equivalent to the Canny filters. Hence, by including $(\alpha, \beta) = (1, 1)$ in the comparison, we compared the results by different RDFs with those obtained by the Canny method, and, consequently we can evaluate whether the replacement of the arithmetic difference by other dissimilarity measures can result in quantitative improvements.

### B. Quantification of the quality of the results

In this experiment we adhere to the most common experimental setup in the literature, which is based on computing the $F$-measure [29], [30] on the BSDS dataset [31]. The details of our procedure for quantitative evaluation are analogous to those in [32].

### C. Experimental results

In Table I we list the results obtained in our experiments. First, in Table I(a) we list the average performance of each combination $(\alpha, \beta)$. Second, in Table I(b) and I(c) we display number of images for which each combination of $(\alpha, \beta)$ produces the best and worst result. That is, the number of images for which each combination leads to the best (or worst) possible results, in terms of $F$-measure.

We can observe that the combination of $(\alpha, \beta) = (1, 1)$ (which recovers the Canny method) is not the best performer in the set of candidates. Several other settings produce better results, both in terms of average performance and in number of best/worst results Hence, from these results we conclude that $r_{[0.5,\beta]}$ and $r_{[1,\beta]}$ with $\beta \in \{1.5, 2, 2.5\}$ often lead to results better than those of $r_{[1,1]}$. This conclusion holds for the average performance, but also for the number of best/worst results. We consider noteworthy the case of $r_{[0.5,1.5]}$ and $r_{[0.5,2]}$, since they are assessed an average performance better than that of $r_{[1,1]}$ and, at the same time, have the best ratio of best/worst results in the comparison. Although more extensive experiments should be carried out, our results point out that the arithmetic difference might not be the best function to express the dissimilarity between tones.

| | | $\beta$ | | | | | |
|---|---|---|---|---|---|---|---|
| | | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 |
| $\alpha$ | 0.5 | .46 | .56 | .59 | .59 | .57 | .55 |
| | 1.0 | .47 | .56 | .58 | .57 | .54 | .51 |
| | 1.5 | .48 | .56 | .56 | .53 | .50 | .48 |
| | 2.0 | .48 | .54 | .52 | .50 | .47 | .44 |
| | 2.5 | .49 | .51 | .50 | .46 | .44 | .41 |
| | 3.0 | .48 | .49 | .47 | .44 | .41 | .38 |

(a) Average $F$-measure for each combination of $(\alpha, \beta)$.

| | | $\beta$ | | | | | |
|---|---|---|---|---|---|---|---|
| | | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 |
| $\alpha$ | 0.5 | 3 | 5 | 13 | 13 | 7 | 9 |
| | 1.0 | 3 | 6 | 8 | 4 | 1 | 1 |
| | 1.5 | - | 3 | 2 | - | - | 2 |
| | 2.0 | 1 | 4 | 3 | 1 | 1 | - |
| | 2.5 | 1 | 2 | 2 | 1 | - | - |
| | 3.0 | 1 | 1 | - | 1 | 1 | - |

(b) Number of images in the dataset for which each combination of $(\alpha, \beta)$ produces the best result.

| | | $\beta$ | | | | | |
|---|---|---|---|---|---|---|---|
| | | 0.5 | 1.0 | 1.5 | 2.0 | 2.5 | 3.0 |
| $\alpha$ | 0.5 | 14 | - | 1 | - | - | 2 |
| | 1.0 | 5 | - | - | - | - | - |
| | 1.5 | 5 | - | - | - | 1 | - |
| | 2.0 | - | - | - | - | - | 2 |
| | 2.5 | 2 | - | - | - | - | 4 |
| | 3.0 | 6 | - | 2 | 3 | 5 | 48 |

(c) Number of images in the dataset for which each combination of $(\alpha, \beta)$ produces the worst result.

TABLE I

QUANTITATIVE RESULTS OBTAINED WITH OUR PROPOSAL ON THE BSDS TEST SET.

## V. CONCLUSIONS

In this work we have presented a novel model to generalize antisymmetric filters for gradient computation with application to edge detection. Our model proposes to use an explicit tonal term in the filtering of the image to be able to adapt the perceived dissimilarity between tones. To this end, we have proposed to use restricted dissimilarity functions, which provide a minimal set of desired properties, while also offering high flexibility. We have illustrated, using quantitative experiments, that our proposal is able to improve the results by classical filters for edge feature extraction. We consider that the results are promising and, although they should be backed up by more extensive experiments, they indicate that tonal evaluation is a potential source of improvements for classical, convolution-based edge detection methods. Moreover, we highlight the fact that the explicit tonal term offers new possibilities for the training of specialized functions in the tonal space, which can adapt the filters to specific applications.

## ACKNOWLEDGMENT

This work has been partially funded by the Spanish Ministry of Science (Project TIN2010-15055) and by the European Regional Development Fund in the IT4Innovations Centre of Excellence (Projects CZ.1.05/1.1.00/02.0070 and CZ.1.07/2.3.00/30.0010).

## REFERENCES

[1] V. Torre and T. Poggio, "On edge detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 8, no. 2, pp. 147–163, 1984.

[2] J. M. Prager, "Extracting and labeling boundary segments in natural scenes," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 2, no. 1, pp. 16–27, 1980.

[3] T. Law, H. Itoh, and H. Seki, "Image filtering, edge detection, and edge tracing using fuzzy reasoning," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 18, no. 5, pp. 481–491, 1996.

[4] J. Bezdek, R. Chandrasekhar, and Y. Attikouzel, "A geometric approach to edge detection," *IEEE Trans. on Fuzzy Systems*, vol. 6, no. 1, pp. 52–75, 1998.

[5] I. Sobel and G. Feldman, "A 3x3 isotropic gradient operator for image processing," 1968, presented at a talk at the Stanford Artificial Intelligence Project.

[6] J. M. S. Prewitt, *Object enhancement and extraction*, ser. Picture Processing and Psychopictorics. Academic Press, 1970, pp. 75–149.

[7] L. G. Roberts, "Machine perception of three-dimensional solids," Ph.D. dissertation, Massachusetts Institute of Technology, 1963.

[8] R. M. Haralick, “Digital step edges from zero crossing of second directional derivatives,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 6, no. 1, pp. 58–68, 1984.
[9] M. Basu, “Gaussian-based edge-detection methods- A survey,” *IEEE Trans. on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 32, no. 3, pp. 252–260, 2002.
[10] S. Mahmoodi, “Edge detection filter based on Mumford–Shah Green function,” *SIAM Journal on Imaging Sciences*, vol. 5, no. 1, pp. 343–365, 2012.
[11] D. Marr and E. Hildreth, “Theory of edge detection,” *Proceedings of the Royal Society of London*, vol. 207, no. 1167, pp. 187–217, 1980.
[12] J. Canny, “Finding edges and lines in images,” Massachussets Institute of Technology, Cambridge, MA, USA, Tech. Rep., 1983.
[13] ——, “A computational approach to edge detection,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.
[14] H. Tagare and R. deFigueiredo, “On the localization performance measure and optimal edge detection,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 12, no. 12, pp. 1186–1190, 1990.
[15] W. McIlhagga, “The Canny edge detector revisited,” *International Journal of Computer Vision*, vol. 91, pp. 251–261, 2011.
[16] A. P. Witkin, “Scale-space filtering,” in *Proc. of the International Joint Conference on Artificial Intelligence*, vol. 2, 1983, pp. 1019–1022.
[17] P. Perona and J. Malik, “Scale-space and edge detection using anisotropic diffusion,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.
[18] T. Lindeberg, “Edge detection and ridge detection with automatic scale selection,” *International Journal of Computer Vision*, vol. 30, no. 2, pp. 117–156, 1998.
[19] C. Tomasi and R. Manduchi, “Bilateral filtering for gray and color images,” in *Proc. of the International Conference on Computer Vision*, 1998, pp. 838–846.
[20] H. Bustince, E. Barrenechea, and M. Pagola, “Relationship between restricted dissimilarity functions, restricted equivalence functions and normal EN-functions: Image thresholding invariant,” *Pattern Recognition Letters*, vol. 29, no. 4, pp. 525–536, 2008.
[21] O. Laligant and F. Truchetet, “A nonlinear derivative scheme applied to edge detection,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 32, no. 2, pp. 242–257, 2010.
[22] P.-L. Shui and W.-C. Zhang, “Noise-robust edge detector combining isotropic and anisotropic Gaussian kernels,” *Pattern Recognition*, vol. 45, no. 2, pp. 806–820, 2012.
[23] S. Paris, P. Kornprobst, J. Tumblin, and F. Durand, “Bilateral filtering: Theory and applications,” *Foundations and Trends in Computer Graphics and Vision*, vol. 4, no. 1, pp. 1–73, 2008.
[24] M. Elad, “On the origin of the bilateral filter and ways to improve it,” *IEEE Trans. on Image Processing*, vol. 11, no. 10, pp. 1141–1151, 2002.
[25] D. Barash, “A fundamental relationship between bilateral filtering, adaptive smoothing, and the nonlinear diffusion equation,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 844–847, 2002.
[26] D. Comaniciu and P. Meer, “Mean shift: a robust approach toward feature space analysis,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.
[27] A. Rosenfeld and M. Thurston, “Edge and curve detection for visual scene anaiysis,” *IEEE Trans. on Computers*, vol. 20, no. 5, pp. 562–569, 1971.
[28] R. Medina-Carnicer, F. Madrid-Cuevas, A. Carmona-Poyato, and R. Muñoz-Salinas, “On candidates selection for hysteresis thresholds in edge detection,” *Pattern Recognition*, vol. 42, no. 7, pp. 1284–1296, 2009.
[29] D. Martin, C. Fowlkes, and J. Malik, “Learning to detect natural image boundaries using local brightness, color, and texture cues,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 26, no. 5, pp. 530–549, 2004.
[30] C. Lopez-Molina, B. De Baets, and H. Bustince, “Quantitative error measures for edge detection,” *Pattern Recognition*, vol. 46, no. 4, pp. 1125–1139, 2013.
[31] D. Martin, C. Fowlkes, D. Tal, and J. Malik, “A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics,” in *Proc. of the 8th International Conference on Computer Vision*, vol. 2, 2001, pp. 416–423.
[32] C. Lopez-Molina, B. De Baets, H. Bustince, J. Sanz, and E. Barrenechea, “Multiscale edge detection based on Gaussian smoothing and edge tracking,” *Knowledge-Based Systems*, vol. 44, pp. 101–111, 2013.