

Regularization, Scale-Space, and Edge Detection Filters

Mads Nielsen¹, Luc Florack¹, and Rachid Deriche²

¹ DIKU, Universitetsparken 1, DK-2100 Copenhagen, Denmark

² INRIA, 2004 Route des Lucioles, BP 93, F-06902 Sophia Antipolis, France

Abstract. Computational vision often needs to deal with derivatives of digital images. Such derivatives are not intrinsic properties of digital data; a paradigm is required to make them well-defined. Normally, a linear filtering is applied. This can be formulated in terms of scale-space, functional minimization, or edge detection filters. The main emphasis of this paper is to connect these theories in order to gain insight in their similarities and differences. We take regularization (or functional minimization) as a starting point, and show that it boils down to Gaussian scale-space if we require scale invariance and a semi-group constraint to be satisfied. This regularization implies the minimization of a functional containing terms up to infinite order of differentiation. If the functional is truncated at second order, the Canny-Deriche filter arises.

1 Introduction

Given a digital signal in one or more dimensions, we want to define its derivatives in a well-posed way. This can be done in a distributional sense [1] by convolving the signal with a smooth test function. Instead of taking the derivatives of the signal, we take the derivatives of the convolved signal by differentiation of the smooth filter prior to convolution. In this way, the derivatives of any integrable signal are operationally defined and well-posed.

Distributional differentiation has been implemented in various conceptually different ways. In computational vision, Gaussian scale-space, regularization, or edge detection filters are typically applied. Anyway, the method must be linear since differentiation by its very definition is linear. In the following, we describe these three methods, their motivation, and discuss their relations.

Gaussian scale-space [2] can be motivated from different points of views. Koenderink [3] introduced it as a one parameter family of blurred images satisfying the causality criterion: every isophote in scale-space must be upwards convex. This requirement expresses in a precise sense that coarse scale details must have a cause at finer scales and essentially singles out the normalized Gaussian filter. Florack [4] defines a visual front-end to be linear, spatially isotropic, spatially homogeneous, separable and scale invariant, leading to the Gaussian as well. Since the Gaussian is smooth, all derivatives of the images are well-defined in Gaussian scale-space.

In Tikhonov regularization theory [5,6], we search for a differentiable function, which in some precise sense is closest to our signal. The measure of difference between the solution and the signal is a functional of the solution. In this way, the problem of well-posedness of differentiation is turned into a functional minimization problem. Using regularization we can directly specify up to which order the solution must be differentiable.

In edge detection many methods have been used for finding robust image derivatives. One method is to define a sample edge as well as criteria for an optimal detection. Canny [7] used a noisy step edge to find the optimal linear detection filter according to criteria of signal-to-noise-ratio, localization, and uniqueness of detection. Deriche [8] used these criteria on an infinite domain to find an optimal filter.

In the following sections, regularization is reviewed and the relation to Gaussian scale-space is derived. Multi-dimensional regularization under the constraint of Cartesian invariance is considered. Furthermore, it is shown that the Canny-optimal Deriche-filter corresponds to a regularization. Finally, it is shown that regularization can be implemented in an efficient way using recursive filtering without introducing further approximations.

2 Linear regularization

Regularization of a signal $g \in \mathcal{L}^2(\mathbb{R})$ can be formulated [5] as the minimization with respect to the regularized solution f of an energy functional E .

Definition 1. (Tikhonov regularization) The Tikhonov regularized solution f of the signal $g \in \mathcal{L}^2(\mathbb{R})$ minimizes the energy functional

$$E[f] \equiv \frac{1}{2} \int dx \left((f - g)^2 + \sum_{i=1}^{\infty} \lambda_i \left(\frac{\partial^i}{\partial x^i} f \right)^2 \right) \quad (1)$$

with nonnegative λ_i .

When $\lambda_n \neq 0$ and $\lambda_i = 0$ for all $i > n$ we talk about n th order regularization. The functional is convex and the minimum exists, implying a uniquely defined minimum. It can be found by linear filtering:

Proposition 2. (Regularization by convolution) *Linear convolution of the signal $g \in \mathcal{L}^2(\mathbb{R})$ by the filter h , having the Fourier transform*

$$\hat{h} = \frac{1}{\sum_{i=0}^{\infty} \lambda_i \omega^{2i}} \quad (2)$$

yields the solution of the regularization of Definition 1. By definition $\lambda_0 \equiv 1$.

Proof In the Fourier domain the energy functional (1) yields (according to Parsevals Theorem)

$$E[\hat{f}] = \frac{1}{2} \int d\omega \left((\hat{f} - \hat{g})^2 + \sum_{i=1}^{\infty} \lambda_i \omega^{2i} \hat{f}^2 \right)$$

A necessary (and sufficient) condition for E to be minimized with respect to \hat{f} is that the variation of E is zero:

$$0 = \frac{\delta E}{\delta \hat{f}} = (\hat{f} - \hat{g}) + \sum_{i=1}^{\infty} \lambda_i \omega^{2i} \hat{f} \quad \Leftrightarrow \quad \hat{f} = \frac{1}{1 + \sum_{i=1}^{\infty} \lambda_i \omega^{2i}} \hat{g}$$

The optimal \hat{f} can thus be found by a linear filtering of the initial signal g with a filter defined by the constants λ_i . Defining $\lambda_0 \equiv 1$ we obtain the linear filter h given by (2). \square

Thus any regularization using only sums of quadratic terms of the derivatives of the solution can be reformulated as a linear convolution with the Fourier inverse (provided it exists) of \hat{h} as given by (2). The differentiation of a regularized signal can be performed by a differentiation of the filter h prior to convolution. We make the following proposition:

Proposition 3. (Completeness of regularization) *Regularization on the form of (1) can implement any real, normalized, even, and analytic filter.*

Proof Regularization implements the linear filtering with \hat{h} of (2). Since the denominator of \hat{h} can be the Taylor series expansion of any analytic, even, and real function which is unity in zero, the filter \hat{h} itself can be any analytic, even, and real function which is unity in zero. Hence, the spatial filter h can be any even, real, analytic, and normalized filter. \square

Regularization can be formulated as a mapping of the initial function into a function space, in which all functions have the desired regularity properties. In Tikhonov regularization a square integrable function is mapped to a function within a Sobolev space of specified order. A Sobolev space of order N is the space of all functions being square integrable, and having derivatives up to order N all well-defined and square integrable.

2.1 Scale invariant regularization

For an action to be scale invariant it must be dimensionless, i.e. without units [9]. Any physical law is in general scale invariant. In order for a filter to be physically meaningful it must thus be dimensionless. In the following we use dimensional analysis to insure scale invariance of the filters. The notation $[\cdot]$ denotes “the dimension of”.

A filtering h is scale invariant if it commutes with the scaling operation $S : x \mapsto x\gamma$ on the spatial domain. A scale invariant regularization other than the identity regularization cannot be constructed. A regularization parametrized by a scale parameter s so that

$$S(h(s, g)) = h(s', S(g))$$

can be constructed. Here $s' = \phi(s, S)$ is an identifiable function. We denote this scale invariant regularization.

Proposition 4. (Scale invariant regularization) *A regularization implemented by a filter $\hat{h}(\omega, s)$ is invariant to a scaling $S : x \mapsto x\gamma$ when*

$$\hat{h}(\omega, s) = \frac{1}{\sum_0^\infty \lambda_i \omega^{2i}} \quad \text{where} \quad \lambda_i(s) = \frac{1}{i!} a_i s^i$$

and $a_i \in \mathbb{R}_+$ are scale independent (dimensionless) constants, $s \propto \gamma^2$, and the factorial $i!$ is included for later algebraic simplicity.

Proof If every term of the filter is dimensionless, then the filter is scale invariant. Since $[\omega] = \text{length}^{-1}$, then $[\lambda_i]$ must be length^{2i} . This is ensured since $[s] = \text{length}^2$. \square

The above construction does not limit the class of possible regularization. It expresses how the parameters of the regularization must transform if a change of units of the spatial domain is taking place.

2.2 Semi-group constraint

We now add constraints on the filter h stating that the initial signal g should not play a privileged role, but must be treated on equal footing with any reconstructed signal f . This is called a *recursivity constraint*: the regularization of a regularized signal must be expressible as a single regularization of the signal. This constraint is most intuitive when the signal g is given as a digitized signal. In this case, the digitizer has already performed a filtering of g . If g then plays a privileged role, a commitment in the image processing is made to the digitizer. We can perceive the recursivity constraint as a necessary condition to unconfound the image processing from the grid details. The recursivity constraint ensures that a filtered signal smoothed to some degree (eventually by the digitizer) can be smoothed to a higher degree by a filter from the same family [4,10]. The recursivity constraint can be expressed as the *semi-group constraint*.

Definition 5. (Semi-group property) A *semi-group* G is a set of compositions: $G \times G \rightarrow G$ which is associative.

Because regularizations can be implemented as linear filters, and linear filters are associative, so are regularizations. Regularization forms a semi-group, since it also satisfies the requirement $G \times G \rightarrow G$, which is the recursivity constraint.

If we embed h into a 1-parameter family $h(s)$, this family is a subset of all regularizations and this subset may not satisfy the semi-group constraint. We may formulate the *semi-group* property of the 1-parameter family $h(s)$ as:

Definition 6. (Semi-group property of filter family) The semi-group property of the 1-parameter family of filters $h(s)$ is

$$\forall s, t \in \mathbb{R}_+ : h(s \oplus_p t) = h(s) * h(t)$$

where the parameter-concatenation is the p -norm addition ³

$$s \oplus_p t = (s^p + t^p)^{1/p}$$

³ This is only the most general parameter concatenation when the filters must be dimensionless [14].

We call the filter-parameter the *scale parameter*, since it resembles the scale parameter in scale-space theory.

Proposition 7. (Scale invariant semi-group regularization) *If a one-parameter regularization filter family $h(s)$, $s \in \mathbb{R}_+$ is a scale invariant filter (Prop. 4) and the filter fulfills the semi-group constraint using the p -norm addition, then the filter family can be written as*

$$\hat{h}_p(\omega, t) = \frac{1}{\sum_{i=0}^{\infty} \frac{t^{ip}}{ip!} \omega^{2ip}} = e^{-(\omega^2 t)^p}$$

here given in the Fourier domain, where $t \propto s$.

Proof According to proposition 4, we can write the filter constants

$$\lambda_i(s) = \frac{1}{i!} a_i s^i$$

The semi-group property yields in conjunction with the actual form of h (given by (2))

$$\forall s, t \in \mathbb{R}_+ : \sum_{i=0}^{\infty} \lambda_i(s \oplus_p t) \omega^{2i} = \left(\sum_{i=0}^{\infty} \lambda_i(s) \omega^{2i} \right) \left(\sum_{i=0}^{\infty} \lambda_i(t) \omega^{2i} \right)$$

By separation into terms of equal power of ω :

$$\forall s, t \in \mathbb{R}_+^2 : \lambda_i(s \oplus_p t) = \sum_{l=0}^i \lambda_l(s) \lambda_{i-l}(t)$$

For clarity we now analyse the situation of $p = 1$. Choose $a_1 \equiv \tau$. By induction we find $a_i = \tau^i$, resulting in a filter of the form ($t = \tau s$)

$$\hat{h}(\omega, t) = \frac{1}{\sum_{i=0}^{\infty} \frac{t^i}{i!} \omega^{2i}} = e^{-\omega^2 t}$$

which is the well-known Gaussian. We formulate this as the result:

Result 1 (Gaussian regularization) *Scale invariance and the semi-group constraint applied to Tikhonov regularization, yields Gaussian scale-space if the 1-norm addition is used for concatenation of scales associated with the dimension of length squared.*

Applying the general p -norm scale-concatenation results in a broader class of filters: The derivation can be found in [14].

$$\hat{h}_p(\omega, t) = \frac{1}{\sum_{i=0}^{\infty} \frac{t^{ip}}{ip!} \omega^{2ip}} = e^{-(\omega^2 t)^p}$$

This is a class of filters all having the semi-group property for different scale-concatenation norms p . □

The above filters were derived from the class of regularization filters, but constitute the full class of all analytic scale invariant semi-group filters as it was derived by Pauwels et al. [11]. Here the existence of an infinitesimal generator is the starting point. Combined with scale invariance the above class of filters arises. Since the semi-group constraint is implied by the existence of an infinitesimal generator the above approach and Pauwels et al. [11] are quite similar.

All \hat{h}_p , $p \in \mathbb{N}$ are Fourier-invertible. For $p = 1$ it is the Gaussian, for p increasing towards infinity, it converges towards the ideal low-pass filter with cut-off frequency $\omega_c = 1/\sqrt{t}$.

2.3 Evolution equations

In the above we adjusted the weighting of the derivatives in Tikhonov regularization to implement Gaussian filtering. We can also reverse the argumentation: Gaussian scale-space can be implemented as the minimization of an energy functional (using a smoothness term of infinite order). The corresponding energy functional is

$$E[f] = \frac{1}{2} \int dx (f - g)^2 + \sum_{i=1}^{\infty} \frac{t^i}{i!} \left[\frac{\partial^i}{\partial x^i} f \right]^2$$

A necessary condition for f to minimize this is that the first variation yields zero. This is the Euler-Lagrange formulation and yields for the above functional

$$0 = (f - g) + \sum_{i=1}^{\infty} \frac{(-t)^i}{i!} \frac{\partial^{2i}}{\partial x^{2i}} f \equiv e^{-t\Delta} f - g$$

where Δ is the Laplacean and its exponentiation is defined by its Taylor series expansion. This differential equation is an equation using only spatial derivatives, but having the same solutions as the Heat Equation. The solution can formally (only formally since g is not necessarily differentiable) be written as

$$f = e^{t\Delta} g$$

By differentiation with respect to t , we obtain the Heat Equation:

$$\begin{cases} \frac{\partial}{\partial t} f = \Delta f \\ f|_{t=0} = g \end{cases}$$

The same derivation can be made for the other scale invariant semi-group filters under the mapping $\Delta \mapsto \Delta^p$, where Δ^p denotes the Laplacean applied p times.

2.4 Scale-space interpretation of regularization

We have seen that infinite order regularization with appropriate weights of the different orders boils down to linear scale-space. We can reformulate this result as follows: as time evolves, the signal governed by the Heat Equation travels through the minima of a one-parameter functional. In general, we would like to

be able to construct a partial differential equation which travels through the minima of a given 1-parameter functional. Here, we look into the case of Tikhonov regularization. We want to find the function f minimizing the functional

$$E[f] = \int dx \left((f - g)^2 + \lambda \sum_{i=1}^n \lambda_i \left(\frac{\partial^i}{\partial x^i} f \right)^2 \right)$$

where $\lambda_i \geq 0$ are arbitrary. We find the Euler-Lagrange equation by setting the first variation of this to zero

$$(f - g) + \lambda \sum_{i=1}^n \lambda_i (-\Delta)^i f = 0$$

where $(-\Delta)^i$ denotes the Laplacean applied i times and negated if i is odd. We perceive of f as being a function of space x and evolution parameter λ and find by differentiation with respect to λ the following differential equation:

$$\left(1 + \lambda \sum_{i=1}^n \lambda_i (-\Delta)^i \right) \frac{\partial}{\partial \lambda} f = - \sum_{i=1}^n \lambda_i (-\Delta)^i f$$

This is a differential equation governing f , so that it travels through the minima of the above functional. In Fourier space it yields

$$\left(1 + \lambda \sum_{i=1}^n \lambda_i \omega^{2i} \right) \frac{\partial}{\partial \lambda} \hat{f} = \sum_{i=1}^n \omega^{2i} \hat{f}$$

In case of first order regularization (i.e. $\lambda_1 = 1$ and all other $\lambda_i = 0$) this is the Heat Equation, but with the following remapping of scale

$$d\lambda = (1 + \lambda\omega^2) ds \quad \Leftrightarrow \quad s(\lambda; \omega) = \frac{\log(1 + \lambda\omega^2)}{\omega^2}$$

In this way first order regularization can be interpreted as Gaussian scale-space with a frequency-dependent remapping of scale. Note that this remapping of scale is not Fourier invertible. Hence, a first order regularization cannot be constructed as a Gaussian scale-space with a spatially varying scale.

So far, we have discussed 1-dimensional regularization. In more dimensions mixed derivatives show up complicating the picture. However, imposing Cartesian invariance [12] (translation, rotation, and the scaling given implicit by the scaling nature of regularization) on the functional E leads to:

Proposition 8. (High-dimensional Cartesian invariant regularization)

The most general Cartesian invariant functional used for regularization in D dimensions has only one independent parameter λ_i per order of differentiation.

The proof is given in [14]. Using this translationally and rotationally invariant regularization, we can easily generalize the results from the previous sections to higher dimensions. All what is needed is to substitute ω by $|\omega|$ in the filters. Also the semi-group property using the 1-norm addition leads trivially to the Gaussian filter.

3 Truncated smoothness operators

To be consistent with scale-space, Tikhonov regularization must be of infinite order. Nevertheless, low-order regularization is often performed [13]. We might perceive of low-order regularization as a regularization using a truncated Taylor-series of the smoothness functional. Here we list the linear filters h_z found by Fourier inversion. The notation h_z indicates the filter where all λ_i are zero except those mentioned by their index in the set z (E.g. in $h_{1,2}$ all λ s are zero except λ_1 and λ_2).

$$h_1(x) = \frac{1}{\sqrt{2\lambda}} e^{-\frac{|x|}{\sqrt{\lambda}}} \quad (3)$$

$$h_2(x) = \frac{\pi}{\lambda^{1/4}} \cos\left(\frac{\sqrt{2}|x|}{\lambda^{1/4}}\right) e^{-\frac{|x|}{\sqrt{2}\lambda^{1/4}}} \quad (4)$$

In the case of mixed first and second order regularization with $\lambda = \lambda_1$ and $\lambda_2 = \lambda_1^2/2$ corresponding to the truncation to second order of the energy functional implying Gaussian filtering, we find

$$h_{1,2}(x) = \frac{\pi(|x| + \sqrt{\lambda})}{2\lambda} e^{-\frac{|x|}{\sqrt{\lambda}}}$$

We notice that the first order regularization filter is always positive, while the second order filter is similar apart from a multiplication with an oscillating term. The latter explains why second order regularization might give inadequate and oscillating results (so-called overshooting or ringing effects). The second order truncated Gaussian filter $h_{1,2}$ is always positive, and will in general give more sensible results than one without the first order term. Higher order non-oscillating regularization filters can be constructed by successive first order regularizations. An analysis of positiveness of regularization filters can be found in [14].

4 Canny optimality

Canny proposed three criteria of optimality of a feature detection filter. The feature is detected in the maxima of the linear filtering. In the case of a general feature $e(x)$ with uncorrelated Gaussian noise, the measures of signal-to-noise ratio Σ , localization A , and uniqueness of zero-crossing Υ are:

$$\Sigma[f] = \frac{|\int dx e(x)f(x)|}{(\int dx f^2(x))^{1/2}}, \quad A[f] = \frac{|\int dx e'(x)f'(x)|}{(\int dx f'^2(x))^{1/2}}, \quad \Upsilon[f] = \frac{(\int dx f'^2(x))^{1/2}}{(\int dx f''^2(x))^{1/2}}$$

where all integrals are taken over the real axis. We now *define* the feature to be a symmetric step edge:

$$e(x) = \int_{-\infty}^x dt \delta_0(t)$$

The symbol δ_0 denotes the Dirac delta functional. Canny tries first simultaneously to maximize Σ and A and finds the box filter as the optimal solution on a

finite domain. To avoid the box-filter he then introduces the uniqueness measure. After this, a simultaneous optimization of all three measures using Lagrange multipliers is performed. Deriche [8] finds the optimal solution on a infinite domain. In the following we show that the uniqueness criterion can be omitted on the infinite domain leading to a conceptually simpler result, but lacking simplicity of the edges. Simplicity means they are zero-crossings of a differentiable function.

In order to find the optimal smoothing filter h to be differentiated so as to give the optimal step edge detector, we substitute $h'(x) = f(x)$ in the measures and try to find h . Assume, h is symmetric and normalized. We use one of the factors as optimality criterion and the others as constraints to find a composite functional Ψ . To obtain symmetry in this formulation we multiply all factors by an arbitrary Lagrange multiplier λ_i .

$$\Psi[h] = \int dx \lambda_1 e h' + \lambda_2 h'^2 + \lambda_3 e' h'' + \lambda_4 h''^2 + \lambda_5 h''^2$$

Omitting the uniqueness constraint corresponds to $\lambda_5 = 0$. A necessary condition for an optimal filter is a zero variation:

$$\frac{\delta\Psi}{\delta h} = -\lambda_1 \delta_0 - 2\lambda_2 h^{(2)} + \lambda_3 \delta_0^{(2)} + 2\lambda_4 h^{(4)} - 2\lambda_5 h^{(6)} = 0$$

Here, a parenthesized superscript denotes the order of differentiation. By Fourier transformation we find

$$\hat{h}(\omega) = \frac{\lambda_1 + \lambda_3 \omega^2}{2(\lambda_2 \omega^2 + \lambda_4 \omega^4 + \lambda_5 \omega^6)}$$

Normalized in the spatial domain implies $\lambda_1 = 0$ and $\lambda_3 = 2\lambda_2$:

$$\hat{h}(\omega) = \frac{1}{1 + \alpha \omega^2 + \beta \omega^4}$$

where $\alpha = 2\lambda_4/\lambda_3$ and $\beta = 2\lambda_5/\lambda_3$. This is a second order regularization filter. Deriche [8] chooses $\alpha = 2\sqrt{\beta}$, and fixes in this way a one-parameter family of optimal filters. This choice corresponds well with a dimensional analysis and is the second order truncated Gaussian $h_{1,2}$.

Another way to select a one-parameter family is to select $\beta = 0$. By omitting the uniqueness criteria we find the optimal step edge detection filter to be the derivative of

$$\hat{h}(\omega) = \frac{1}{1 + \alpha \omega^2}$$

This is a first order regularization filter. It seems intuitively correct that the optimal first order filter should project into a Sobolev space of first order. We do not in general require the second derivative of the signal to exist when we consider only first order characteristics. In the spatial domain, we find the filter

$$h'(x) = \frac{\pi}{\sqrt{\alpha}} \frac{x}{|x|} e^{-|x|/\sqrt{\alpha}}$$

This is the edge detection filter proposed by Castan et al. [16]. The derivative of this filter is not well-defined, and we see that we indeed have projected into a Sobolev space of first order.

The above results can be generalized to optimal detection of higher order image features [14]. This also implies regularization but of higher order.

5 Implementation issues

The regularization using quadratic stabilizers can be implemented in various ways. It can be implemented by a gradient descent algorithm. This has the advantage that the natural boundary conditions can be implemented directly, leading to sensible results near the boundaries. The disadvantage is slow convergence (especially for large λs).

Implementation can be done by convolution in the spatial domain. In this case the boundaries can be handled by cutting off the filters and renormalizing. The computational complexity will be $O(MN\lambda^2n)$, where the size of the image is $M \times N$ pixels and n is the order of regularization.

The filtering can also be implemented in the Fourier domain as a multiplication, using the Fast Fourier Transform (FFT). In this case the image is assumed to be cyclic, which may imply strange phenomena near the boundaries. The computational complexity is $O(MN \log M \log N)$ independently of the order of regularization and the λs .

Finally, the regularization can be implemented as a recursive filtering. In this case, the boundaries can be handled by cutting off like in the case of convolution in the spatial domain. The computational complexity is $O(MNn)$. In most practical cases, this will be the fastest implementation.

In order to deal with the recursive system, we need to reformulate the regularization on a discrete grid. We define the energy

$$E(f) \equiv \sum_x \left((f - g)^2 + \sum_{i=1}^N \lambda_i (d_i * f)^2 \right)$$

where d_i is the i th order difference filter. We recall the definition of the z -transform

$$\hat{h} = \sum_{x=-\infty}^{\infty} h(x)z^{-x}$$

and make the following proposition:

Proposition 9. (Recursive filtering) *The discrete regularization can be implemented by recursive filtering using no more than $2N$ multiplications and additions per output element, where N is the order of regularization.*

Proof By following the argumentation in the continuous case (and substituting the Fourier transform by the z -transform) or the discrete formulation given by

Unser et al. [17], we find that the minimization is implemented by convolution with the filter

$$\hat{h}(z) = \frac{1}{1 + \sum_{i=1}^N \lambda_i \hat{d}_i(z) \hat{d}_i(z^{-1})}$$

where the hat indicates the z -transform. The transform of the difference operator is given by

$$\hat{d}_i(z) = z^{-i/2}(1 - z)^i$$

implying that $\hat{d}_i(z)\hat{d}_i(z^{-1})$ is a N order polynomial in z multiplied by a N order polynomial in z^{-1} . Because the transfer function is symmetric in z and z^{-1} the $2N$ roots in the denominator will appear in pairs of z and z^{-1} . The transfer function can be decomposed as $\hat{h}(z) = \hat{h}^+(z)\hat{h}^+(z^{-1})$, where

$$\hat{h}^+(z) = \frac{c}{\prod_{i=1}^N (1 - z_i z^{-1})} = \frac{c}{1 - a_1 z^{-1} - \dots - a_N z^{-N}}$$

and z_i are the roots in the denominator ordered by length. This means that the regularization can be implemented by the forward and backward recursive filters with identical coefficients:

$$\begin{aligned} f^+(x) &= g(x) + a_1 f^+(x-1) + \dots + a_N f^+(x-N) \\ f^-(x) &= f^+(x) + a_1 f^-(x+1) + \dots + a_N f^-(x+N) \\ f(x) &= c^2 f^-(x) \end{aligned}$$

□

We have here proven that regularization in the discrete 1D case can be implemented (without any other approximation than the discrete implementation) as recursive filtering. The proof follows the lines of the proof given by Unser et al. [17] for the case, where a single order stabilizer is used.

6 Summary

We have presented a formulation of the Heat Equation as a minimization of a functional corresponding to a Tikhonov regularization. Furthermore, we have developed an infinite series of regularization filters all possessing the semi-group property. The simplest of these is the well-known Gaussian. The series of semi-group filters converges towards an ideal low-pass filtering, when the scale addition convention converges towards the infinity norm.

We have shown that the Canny-Deriche filter is a second order regularization filter, when we choose coefficients corresponding to the truncation of the functional leading to Gaussian scale-space. The Canny-Deriche filter is optimized for detection of zero order step edges.

Furthermore, we have shown that regularization in a higher dimensional space in its most general form is a rotation of the 1D regularization, when Cartesian invariance is imposed. This means that all of the above results can be generalized to any dimension, under the constraint that the functional is a Cartesian invariant.

The above work is a concentrate of the technical report [14], where the omitted proofs and further results can be found. We have only treated linear problems: Linear convolution, linear scale-space, linear biased diffusion. In future a major subject for studies will be to find the resemblance between non-linear biased and non-biased diffusion. Can we transform a non-linear diffusion equation [18] into a non-linear biased diffusion equation [19] (and thereby capture minimization of non-quadratic functionals) and vice versa?

References

1. L. Schwartz, "*Théorie des Distributions*", Hermann, 2nd edition, Paris, 1966.
2. A. P. Witkin, "*Scale-Space Filtering*", Proc. of IJCAI, Karlsruhe, Germany, 1983.
3. J. J. Koenderink, "*The Structure of Images*", Biol. Cybern., vol. 50, pp. 363-370, 1984.
4. L. M. J. Florack, "*The Syntactical Structure of Scalar Images*", PhD-thesis, University of Utrecht, 1993.
5. A. N. Tikhonov and V. Y. Arseninn, "*Solution of ill-posed problems*" V.H Winston & Sons, John Wiley, 1977.
6. T. Poggio, H. Voorhees and A. Yuille, "*A Regularized Solution to Edge Detection*", A. I. Memo 883, M.I.T, May 1985.
7. J. F. Canny, "*A Computational Approach to Edge Detection*", IEEE PAMI, Vol. 8, No. 6, Nov 1986.
8. R. Deriche, "*Using Canny's Criteria to Derive a Recursively Implemented Optimal Edge Detector*", Int. Jour. of Computer Vision, pp. 167-187, 1987.
9. J. Fourier, "*The Analytic Theory of Heat*", New York Dover publications, 1995, Original "*Théori Analytique de la Chaleur*", Paris, 1822.
10. T. Lindeberg, "*Scale-Space Theory in Computer Vision*", Kluwer Academic Publishers, Netherlands, 1994.
11. E. J. Pauwels, P. Fiddelaers, T. Moons, and L. J. V. Gool, "*An extended Class of Scale-Invariant and Recursive Scale-Space Filters*", Katholieke Universiteit Leuven, Tech. Rep. KUL/ESAT/M12/9316, 1994.
12. D. Hilbert, *Ueber die vollen Invariantsystemen*, Math. Annalen, vol. 42, pp. 313-373, 1893.
13. W. E. L. Grimson, "*From Images to Surfaces*", MIT Press, Cambridge, Massachusetts, 1981.
14. M. Nielsen, L. M. J. Florack, and R. Deriche, "*Regularization and Scale-Space*", RR-2352, INRIA, France, 1994.
15. R. Deriche, "*Fast Algorithms for Low-level Vision*", IEEE PAMI, Vol. 12, No. 1, Jan 1990.
16. S. Castan, J. Zhao, and j. Shen, "*Optimal Filter for Edge Detection Methods and Results*", ECCV 90, Antibes, France, May, 1990.
17. M. Unser, A. Aldroubi, and M. Eden, "*Recursive Regularization Filters: Design, Properties and Applications*", IEEE PAMI, vol. 13, no. 3, March 1991.
18. P. Perona and J. Malik, "*Scale-Space and Edge Detection using Anisotropic Diffusion*" IEEE PAMI, vol. 12, pp. 429-439, 1990.
19. N. Nordström, "*Biased Anisotropic Diffusion - a Unified Regularization and Diffusion Approach to Edge Detection*", Image and Visual Computing, no. 8, pp. 318-327, 1990.