In J.-F. Aujol, M. Nikolova, N. Papadakis (Eds.):
Scale Space and Variational Methods in Computer Vision.
Lecture Notes in Computer Science, Vol. 9087, pp. 425–436, Springer, Berlin, 2015.
The final publication is available at link.springer.com.

A Linear Scale-Space Theory for Continuous Nonlocal Evolutions

Giovanno Marcelo Cárdenas, Joachim Weickert, and Sarah Schäffer

Mathematical Image Analysis Group, Faculty of Mathematics and Computer Science,
Campus E1.7, Saarland University, 66041 Saarbrücken, Germany
{cardenas, weickert, schaeffer}@mia.uni-saarland.de

Abstract. Most scale-space evolutions are described in terms of partial differential equations. In recent years, however, nonlocal processes have become an important research topic in image analysis. The goal of our paper is to establish well-posedness and scale-space properties for a class of nonlocal evolutions. They are given by linear integro-differential equations with measures. In analogy to Weickert's diffusion theory (1998), we prove existence and uniqueness, preservation of the average grey value, a maximum–minimum principle, image simplification properties in terms of Lyapunov functionals, and we establish convergence to a constant steady state. We show that our nonlocal scale-space theory covers nonlocal variants of linear diffusion. Moreover, by choosing specific discrete measures, the classical semidiscrete diffusion framework is identified as a special case of our continuous theory. Last but not least, we introduce two modifications of bilateral filtering. In contrast to previous bilateral filters, our variants create nonlocal scale-spaces that preserve the average grey value and that can be highly robust under noise. While these filters are linear, they can achieve a similar performance as nonlinear and even anisotropic diffusion equations.

Keywords: nonlocal processes, scale-space, diffusion, integro-differential equations, well-posedness, bilateral filtering

1 Introduction

Starting with Iijima's pioneering work in 1962 [1] and its western counterparts by Witkin [2] and Koenderink [3] two decades later, the scale-space concept has become an integral part of many image processing and computer vision methods. For example, it is the backbone of the widely used SIFT detector for feature matching [4].

Scale-spaces embed an original image f into a family $\{T_t f \,|\, t \geq 0\}$ such that

$T_0 f = f$ and larger values of t correspond to simpler representations of f. Numerous attempts have been made to formalise this idea and supplement it with additional assumptions in order to restrict the scale-space evolution to a specific class of processes, or even single out a unique scale-space in an axiomatic way. Such evolutions include linear processes such as Gaussian scale-space [1–3, 5–7], the Poisson scale-space [8] and its generalisation to α-scale-spaces [9]. Typical representatives of nonlinear scale-spaces are given by nonlinear diffusion scale-spaces [10], the morphological equivalent of Gaussian scale-space [11], and curvature-driven evolutions such as the affine morphological scale-space [12]. Moreover, also spatio-temporal scale-spaces have been considered [13, 14], and regularisation methods have been identified as scale-spaces [15].

Many of these processes exhibit a local behaviour and can be described in terms of partial differential equations (PDEs) or pseudodifferential equations. More recently, however, nonlocal processes have become very popular in research. For instance, bilateral filters [16, 17] and patch-based methods [18, 19] are widely-used in image processing applications, and classical PDEs and variational methods have been generalised to nonlocal evolutions [20]. However, less is known about scale-space theory for nonlocal processes. Related work can be found in [25], where the authors develop nonlocal morphological scale-spaces as an extension of [12].

The goal of our paper is to address this issue from the point of view of diffusion processes. By restricting ourselves to a class of nonlocal evolutions that are given by linear integro-differential equations, we establish well-posedness and scale-space results that are in analogy to the diffusion framework by Weickert [10]. This includes existence and uniqueness, preservation of the average grey value, an extremum principle, a large class of Lyapunov functionals, and convergence to a flat steady state. We show that our framework covers nonlocal generalisations of Gaussian scale-space as well as space-discrete diffusion scale-spaces. Moreover, we introduce two modifications of bilateral filtering that are in accordance with our theory and can be much more robust under noise.

Our paper is organised as follows. In Section 2 we derive theoretical results on well-posedness and scale-space properties. The third section discusses examples and presents experiments. Our paper is concluded with a summary in Section 4.

2 Theoretical Results

Let us begin by giving a precise formulation of the problem we are concerned with. For this matter let $\Omega \subset \mathbb{R}^N$ be a bounded N-dimensional image domain, and let μ be a locally finite Borel measure in $\mathbb{R}^N$. We consider the following linear evolution process:

$$\partial_t u(\mathbf{x},t) = \int_\Omega K(\mathbf{x},\mathbf{y})\,(u(\mathbf{y},t) - u(\mathbf{x},t))\,d\mu(\mathbf{y}) \quad \text{in} \quad \bar{\Omega} \times [0, t_0]\,, \tag{1}$$

$$u(\mathbf{x},0) = f(\mathbf{x}) \quad \text{in} \quad \bar{\Omega}\,, \tag{2}$$

with the subsequent assumptions:

(NL1) Regularity: $K \in C(\bar{\Omega} \times \bar{\Omega})$ and $f \in C(\bar{\Omega})$.
(NL2) Symmetry: $K(\mathbf{x}, \mathbf{y}) = K(\mathbf{y}, \mathbf{x})$ in $\bar{\Omega} \times \bar{\Omega}$.
(NL3) Nonnegativity: $K(\mathbf{x}, \mathbf{y}) \geq 0$ in $\bar{\Omega} \times \bar{\Omega}$.
(NL4) Irreducibility: There exists a finite family of μ-measurable sets $\mathcal{F} := \{B_i \subset \Omega \,:\, 1 \leq i \leq p\}$, such that:
  (i) There exists a constant $c > 0$ such that $K(\mathbf{x}, \mathbf{y}) \geq c$ whenever $B \in \mathcal{F}$ and $\mathbf{x}, \mathbf{y} \in B$.
  (ii) $\Omega = \bigcup_{i=1}^{p} B_i$ and $\mu(B_i \cap B_{i+1}) > 0$ for $1 \leq i \leq p-1$.

We will see that not all of these assumptions will be necessary for our results: (NL1) is needed for establishing well-posedness, the proof of a maximum–minimum principle involves (NL1) together with (NL3), while preservation of the average grey value uses (NL1) and (NL2). The existence of Lyapunov functionals and the convergence to a constant steady state require (NL1)–(NL4).

2.1 Well-Posedness

Let us first define a solution concept for (1)–(2).

Definition 1. We say that $u \in C(\bar{\Omega} \times [0, t_0])$ is a *solution* of problem (1)–(2) if

$$u(\mathbf{x}, t) = f(\mathbf{x}) + \int_0^t \int_\Omega K(\mathbf{x}, \mathbf{y})(u(\mathbf{y}, s) - u(\mathbf{x}, s))\, d\mu(\mathbf{y})\, ds\,, \qquad 0 \leq t \leq t_0\,. \quad (3)$$

This definition allows us to prove the following result.

Proposition 1 (Existence and Uniqueness). There exists a solution of (1)–(2), and this solution is unique.

Proof. The proof is similar to the one given in [21], where the authors considered this type of processes with a special function $K(\mathbf{x}, \mathbf{y}) = J(\mathbf{x} - \mathbf{y})$, for some continuous radial and symmetric function J and with μ being equal to the Lebesgue measure. The fixed point arguments in [21] also work under the conditions of our more general framework. In our case, we consider the operator given by the r.h.s. of (3) defined on the space $C(\bar{\Omega} \times [0, t_0])$. □

Remark 1. We will interpret $\partial_t u$ as the right and left derivative when $t = 0$ and $t = t_0$, respectively. In this case, the unique solution u in the sense of Definition 1 satisfies (1)–(2). In fact, it is not hard to verify that the expression $\int_\Omega K(\mathbf{x}, \mathbf{y})\,(u(\mathbf{y}, s) - u(\mathbf{x}, s))\, d\mu(\mathbf{y})$ is continuous with respect to the variable s. Thus, (3) and the fundamental theorem of integral calculus imply that $\partial_t u$ exists and that (1)–(2) hold.

2.2 Scale-Space Properties

Now that we have established the existence of a unique solution, we are in a position to prove a number of scale-space results that are in analogy to the ones for anisotropic diffusion [10].

Proposition 2 (Preservation of the Average Grey Value). The solution of (1)–(2) preserves the average grey value:

$$\frac{1}{\mu(\Omega)} \int_{\Omega} u(\mathbf{x},t)\, d\mu(\mathbf{x}) = \frac{1}{\mu(\Omega)} \int_{\Omega} f(\mathbf{x})\, d\mu(\mathbf{x}) \qquad \text{for} \quad 0 \leq t \leq t_0\,. \tag{4}$$

Proof. Integrating (3) over Ω with respect to μ and applying the Tonelli-Fubini theorem together with (NL2), we obtain

$$\begin{aligned} \int_{\Omega} (u(\mathbf{x},t) - f(\mathbf{x}))\, d\mu(\mathbf{x}) &= \int_0^t \int_{\Omega} \int_{\Omega} K(\mathbf{x},\mathbf{y})(u(\mathbf{y},s) - u(\mathbf{x},s))\, d\mu(\mathbf{y})\; d\mu(\mathbf{x})\; ds \\ &= -\int_0^t \int_{\Omega} \int_{\Omega} K(\mathbf{x},\mathbf{y})(u(\mathbf{y},s) - u(\mathbf{x},s))\, d\mu(\mathbf{y})\; d\mu(\mathbf{x})\; ds\,. \end{aligned} \tag{5}$$

This implies the result. □

Proposition 3 (Preservation of Nonnegativity). If u is a solution of (1)–(2) with $f(\mathbf{x}) \geq 0$, then

$$\min_{(\mathbf{x},t) \in \bar{\Omega} \times [0,t_0]} u(\mathbf{x},t) \geq 0\,. \tag{6}$$

Proof. Assume that $f(\mathbf{x}) \geq 0$ and that $\min_{(\mathbf{x},t) \in \bar{\Omega} \times [0,t_0]} u(\mathbf{x},t) < 0$. Then there exists an $\epsilon > 0$ such that the function $v := u + t\epsilon$ has a strictly negative minimum in some point $(\mathbf{x_m}, t_m) \in \bar{\Omega} \times]0, t_0]$. However,

$$0 = \partial_t v(\mathbf{x_m}, t_m) = \epsilon + \int_{\Omega} K(\mathbf{x_m},\mathbf{y})(u(\mathbf{y},t_m) - u(\mathbf{x_m},t_m))\, d\mu(\mathbf{y}) > 0\,. \tag{7}$$

This is a contradiction. □

From this last proposition we get the following maximum–minimum principle.

Proposition 4 (Maximum–Minimum Principle). If u is a solution of (1)–(2), then

$$\min_{\mathbf{z} \in \bar{\Omega}} f(\mathbf{z}) \leq u(\mathbf{x},t) \leq \max_{\mathbf{z} \in \bar{\Omega}} f(\mathbf{z}) \qquad \forall\ (\mathbf{x},t) \in \bar{\Omega} \times [0,t_0]\,. \tag{8}$$

Proof. To prove the first inequality, we apply Proposition 3 to problem (1)–(2) with f replaced by $v_0 := f - \min_{\mathbf{x} \in \bar{\Omega}} f$. In fact, $v_0 \geq 0$, and it follows that the solution v of this problem should satisfy $v \geq 0$. However, from the linearity of (1)–(2), we also know that $v = u - \min_{\mathbf{x} \in \bar{\Omega}} f$, which gives the result. The second inequality can be proven in a similar way, applying Proposition 3 to problem (1)–(2) with f replaced by $\max_{\mathbf{x}\in\bar{\Omega}}(f) - f$. □

Our next goal is to analyse the behaviour of the solution of (1)–(2) as $t_0 \to \infty$. We will need the following lemma.

Lemma 1. Let $r : \mathbb{R} \to \mathbb{R}$ be a convex C^2 function. If u is a solution of (1)–(2), then $\frac{d}{dt}\int_\Omega r(u(\mathbf{x},t))\,d\mu(\mathbf{x})$ exists for $t \in [0,t_0]$ (here we mean the right and left derivative for $t = 0$ and $t = t_0$, respectively). Moreover, this expression is equal to $\int_\Omega r'(u(\mathbf{x},t))\,\partial_t u(\mathbf{x},t)\,d\mu(\mathbf{x})$.

Proof. Let $t \in [0,t_0]$ and define $F_h(\mathbf{x}) := \frac{1}{h}\left(r(u(\mathbf{x},t+h)) - r(u(\mathbf{x},t))\right)$, for $t, t+h \in [0,t_0]$. Then, since $r \in C^2$, we obtain from Remark 1 that $\partial_t r(u(\mathbf{x},t))$ exists for every $(\mathbf{x},t) \in \bar{\Omega} \times [0,t_0]$ and is equal to $r'(u(\mathbf{x},t))\partial_t u(\mathbf{x},t) = \lim_{h\to 0} F_h(\mathbf{x})$. On the other hand, since $u \in C(\bar{\Omega} \times [0,t_0])$, we also know that

$$|F_h(\mathbf{x})| = |r'(u(\mathbf{x},t_x))| \tag{9}$$

for some $t_x \in [0,t_0]$ such that $|t_x - t| < h$. Thus, we may bound $F_h(\mathbf{x})$ with a constant $M > 0$ which is independent of x and h. This allows us to apply Lebesgue's convergence theorem to obtain that

$$\lim_{h\to 0}\int_\Omega F_h(\mathbf{x})\,d\mu(x) = \int_\Omega \lim_{h\to 0} F_h(\mathbf{x})\,d\mu(\mathbf{x})\,, \tag{10}$$

as wanted. □

In what follows we will denote the constant function that is equal to the average grey value of f by

$$\tilde{u}(\mathbf{x}) := \frac{1}{\mu(\Omega)}\int_\Omega f(\mathbf{z})\,d\mu(\mathbf{z}) \qquad \forall\, \mathbf{x} \in \bar{\Omega}\,. \tag{11}$$

With this notation we can state the following result.

Proposition 5 (Lyapunov Functionals). Let u be the solution of (1)–(2). For any convex C^2 function $r : \mathbb{R} \to \mathbb{R}$, the expression

$$V(t) = \Phi(u(.,t)) := \int_\Omega r(u(\mathbf{x},t))\,d\mu(\mathbf{x}) \tag{12}$$

is a Lyapunov functional, i.e.

(i) $\Phi(u(.,t)) \geq \Phi(\tilde{u})$ for all $t \geq 0$.
(ii) $V \in C^1[0,\infty[$ and $V'(t) \leq 0$ for all $t \geq 0$.

Moreover, if $r'' > 0$, then $V(t)$ is even a strict Lyapunov functional, i.e.

(iii) For all $t \geq 0$ we have that $\Phi(u(.,t)) = \Phi(\tilde{u})$, if and only if $u(.,t) = \tilde{u}$ μ-a.e. in Ω.

(iv) If $t \geq 0$, then $V'(t) = 0$, if and only if $u(t) = \tilde{u}$ μ-a.e. in Ω.
(v) $V(0) = V(T)$ for $T > 0$, if and only if $\forall t \in [0,T] : u(\mathbf{x},t) = \tilde{u}$ μ-a.e. in Ω.

Proof.

(i) From Jensen's inequality and the preservation of the average grey value we obtain that

$$\Phi(u(.,t)) = \int_\Omega r(u(\mathbf{z},t))\, d\mu(\mathbf{z}) \geq \int_\Omega r\left(\int_\Omega \frac{u(\mathbf{z},t)}{\mu(\Omega)} d\mu(\mathbf{z})\right) d\mu(\mathbf{y}) = \Phi(\tilde{u})\,. \tag{13}$$

(ii) From Lemma 1 we know that

$$V'(t) = \int_\Omega r'(u(\mathbf{x},t)) \frac{d}{dt} u(\mathbf{x},t)\, d\mu(\mathbf{x})\,. \tag{14}$$

Then, from (1)–(2) we obtain that

$$\begin{aligned} 2V'(t) &= 2\int_\Omega \int_\Omega K(\mathbf{y},\mathbf{x})\, r'(u(\mathbf{x},t))(u(\mathbf{y},t) - u(\mathbf{x},t))\, d\mu(\mathbf{y})\; d\mu(\mathbf{x}) \\ &= \int_\Omega \int_\Omega K(\mathbf{y},\mathbf{x})(r'(u(\mathbf{x},t)) - r'(u(\mathbf{y},t))) \cdot \\ &\qquad\qquad \cdot (u(\mathbf{y},t) - u(\mathbf{x},t))\, d\mu(\mathbf{y})\; d\mu(\mathbf{x}) \end{aligned} \tag{15}$$

where we used (NL2) for the second equality. Since r is convex we know that r' is nondecreasing. Therefore, the quantity $(r'(u(\mathbf{x},t)) - r'(u(\mathbf{y},t))) \cdot \cdot (u(\mathbf{y},t) - u(\mathbf{x},t))$ is allways nonpositive and it follows from the nonnegativity of K (NL3) that $V'(t) \leq 0$. Continuity of $V(t)$ and $V'(t)$ follows from the uniform continuity of u in $\bar{\Omega} \times [0,t_0]$ and (15).

(iii) If we assume that r is strictly convex, then we obtain from the strict Jensen's inequality that $\Phi(u(.,t)) = \Phi(\tilde{u})$ if and only if $u(\mathbf{x},t) = C$ μ-a.e. for some constant C. However, from the preservation of the average grey value (Proposition 2) we conclude that the only possibility is $C = \tilde{u}$, as wanted.

(iv) From (15) and the irreducibility condition (NL4) we obtain that u is μ-a.e. equal to a constant. However, this constant can only be $\tilde{u}$ because of the preservation of the average grey value. This proves the result.

(v) We use the fact that V is nonincreasing together with (iv). □

As explained in [10], Lyapunov functionals guarantee that a scale-space acts image simplifying in many ways. By choosing specific strictly convex functions for r, it follows that the scale-space evolution reduces all L^p norms for $p \geq 2$, all even central moments, and it increases the entropy of the image. Last but not least, Lyapunov functionals are also useful for proving the following convergence result.

Proposition 6 (Convergence) Let u be a solution of (1)–(2). Then

$$\lim_{t\to\infty} ||u(t) - \tilde{u}||_{L^2(\Omega,\mu)} = 0\,. \tag{16}$$

Proof. Let $v = u - \tilde{u}$ be the solution of (1)–(2) when f is replaced by $f - \tilde{u}$. If we consider the Lyapunov functional of Proposition 5 for the solution v, with the particular choice $r(x) = x^2$ in the definition (12) of V, we get that

$$\lim_{t\to\infty} ||u(.,t) - \tilde{u}||_{L^2(\Omega,\mu)} = \ell\,, \tag{17}$$

for some finite value $\ell \geq 0$, as a consequence of (i) and (ii) of Proposition 5. Moreover, we know that

$$\int_0^\infty |V'(t)|dt \leq V(0) - \lim_{t\to\infty} V(t) < \infty\,. \tag{18}$$

This implies that there exists a sequence t_i such that $\lim_{i\to\infty} t_i = \infty$ and $\lim_{i\to\infty} V'(t_i) = 0$, or equivalently,

$$\lim_{i\to\infty} \int_\Omega \int_\Omega K(\mathbf{y},\mathbf{x})(u(\mathbf{y},t_i) - u(\mathbf{x},t_i))^2\, d\mu(\mathbf{y})\, d\mu(\mathbf{x}) = 0\,. \tag{19}$$

Now, for every B in the family $\mathcal{F}$ of condition (NL4) we may apply the Cauchy-Schwartz inequality to obtain

$$\begin{aligned} &\int_B \left| u(\mathbf{x},t) - \frac{1}{\mu\,(B)} \int_B u(\mathbf{y},t)\, d\mu(\mathbf{y}) \right|^2 d\mu(\mathbf{x}) \\ &\leq \frac{1}{\mu\,(B)} \int_B \int_B |u(\mathbf{y},t) - u(\mathbf{x},t)|^2\; d\mu(\mathbf{x})\; d\mu(\mathbf{y}) \\ &\leq \frac{1}{\mu\,(B)} \int_B \int_B \frac{K(\mathbf{x},\mathbf{y})}{c} |u(\mathbf{y},t) - u(\mathbf{x},t)|^2\; d\mu(\mathbf{x})\; d\mu(\mathbf{y})\,, \end{aligned} \tag{20}$$

where $c > 0$ is the lower bound for K in condition (NL4). Let us denote by $h_k(t)$ the constant function defined on each $B_k \in \mathcal{F}$ of condition (NL4) that is equal to $\frac{1}{\mu(B_k)} \int_{B_k} u(\mathbf{x},t) d\mu(x)$ for $1 \leq k \leq p$. The last inequality and (19) imply that

$$||u(.,t_i) - h_k(t_i)||_{L^2(B_k,\mu)} \to 0\,. \tag{21}$$

Moreover, from (17) we know that $u(.,t_i)$ is bounded in $L^2(\Omega,\mu)$. Therefore, also $h_k(t_i)$ is bounded. We may choose a subsequence of t_i which we continue to denote the same way, such that $\lim_{i\to\infty} h_k(t_i)$ exists and is finite for $1 \leq k \leq p$. Furthermore, the quantity $\gamma := \min\{\mu(B_k \cap B_{k+1}) :\; 1 \leq k \leq p-1\}$ is positive because of (NL4). Therefore, we obtain that

$$\begin{aligned} \gamma|h_k(t) - h_{k+1}(t)|^2 &\leq |h_k(t) - h_{k+1}(t)|^2 \int_{B_k\cap B_{k+1}} d\mu(\mathbf{x}) \\ &= ||h_k(t) - h_{k+1}(t)||_{L^2(B_k\cap B_{k+1},\mu)} \\ &\leq ||u(.,t) - h_k(t)||_{L^2(B_k,\mu)} + ||u(.,t) - h_{k+1}(t)||_{L^2(B_{k+1},\mu)} \end{aligned} \tag{22}$$

for $0 \leq k \leq p-1$. These inequalities, together with (NL4) and (21), allow us to conclude that all $h_k(t_i)$ converge to the same constant. Hence, it follows that u converges in $L^2(\Omega,\mu)$ to a constant. This implies that the value ℓ in (17) has to be 0 as wanted. □

3 Examples

3.1 Continuous Setting with Shift-Invariant Kernels

Let us consider the specific problem

$$\partial_t u(\mathbf{x},t) = \int_\Omega J(\mathbf{x}-\mathbf{y})\,(u(\mathbf{y},t)-u(\mathbf{x},t))\,d\mathbf{y}, \quad \text{in} \quad]0,t_0]\times\Omega\,, \tag{23}$$

$$u(\mathbf{x},0) = f(\mathbf{x}) \quad \text{in} \quad \Omega\,, \tag{24}$$

with $f \in L^1(\Omega)$ and some nonnegative radial function $J \in C(\mathbb{R}^N,\mathbb{R})$ such that $J(0) > 0$ and $\int_{\mathbb{R}^N} J(\mathbf{x})\,d\mathbf{x} = 1$. Notice that since we restrict ourselves to continuous initial data, i.e. $f \in C(\bar{\Omega})$, it is not difficult to check that (23) satisfies all conditions (NL1)–(NL4). Thus, we may apply the results of the previous section, for μ equal to the Lebesgue measure of $\mathbb{R}^N$.
Interestingly, this process was studied also in [21]. The authors proved that the family of solutions u_ϵ of (23) with J replaced by an appropriate rescaled version J_ϵ, approximates the solution of the usual Neumann problem for homogeneous diffusion. More precisely, if v is a solution of

$$\partial_t u = \Delta u \quad \text{in} \quad]0,t_0]\times\Omega\,, \tag{25}$$

$$\frac{du}{d\nu} = 0 \quad \text{on} \quad \partial\Omega\,, \tag{26}$$

$$u(\mathbf{x},0) = f(\mathbf{x}) \quad \text{in} \quad \Omega\,, \tag{27}$$

where ν denotes the outer normal vector to $\partial\Omega$, then

$$\lim_{\epsilon\to 0} \|u_\epsilon - v\|_{L^\infty(\Omega\times[0,t_0])} = 0\,. \tag{28}$$

For this reason, the nonlocal problem (23) solves a diffusion problem. In other words, observe that using any kernel J as specified above will always lead us to Gaussian scale-space. This statement has its stochastic counterpart in the central limit theorem, which tells us that an iterated application of a smoothing kernel converges to a Gaussian. This motivates us to consider more general filters below, where the kernel can be a space-variant function of the initial image f.

3.2 Discrete Setting

Now we discuss the case when the measure μ is a discrete measure concentrated on a finite subset of Ω. We will focus on the one-dimensional case since the extension to higher dimension is straightforward.

Let $\Omega =]0,1[$ and let $h = \frac{1}{M}$ for some fixed integer $M > 1$. Moreover, we define μ as the restriction to Ω of the discrete measure that is concentrated on the set $\mathcal{Z}_h := \left\{h(z-\frac{1}{2})\ ;\ z\in\mathbb{Z}\right\}$. In what follows we set $k_{i,j} := K((i-\frac{1}{2})h,(j-\frac{1}{2})h)$ and

$u_i(t) := u((i-\frac{1}{2})h, t)$ for $1 \le i, j \le M$. With these choices, the problem (1)–(2) becomes

$$\frac{d}{dt} u_i = \sum_{j=1}^{M} k_{i,j}(u_j - u_i) \qquad (1 \le i \le M)\,, \tag{29}$$

$$u_i(0) = f_i \qquad (1 \le i \le M)\,. \tag{30}$$

This is a semidiscrete evolution process for the vector $\mathbf{u} := (u_1, u_2, ..., u_M)^\top$. Conditions (NL1)-(NL4) imply that the matrix $\mathbf{K} = (k_{i,j})_{i,j=1}^{M}$ is symmetric, nonnegative, and irreducible. Notice that (29) can be written as

$$\frac{d}{dt}\mathbf{u}(t) = \mathbf{A}\mathbf{u}(t)\,, \tag{31}$$

$$\mathbf{u}(0) = \mathbf{f}\,, \tag{32}$$

where $\mathbf{f} = (f_1, f_2, ..., f_M)^\top$ and $\mathbf{A} = (a_{i,j})_{i,j=1}^{M}$ is the matrix with entries

$$a_{i,j} = \begin{cases} k_{i,j} & (i \neq j)\,, \\ -\sum\limits_{n \neq i} k_{i,n} & (i = j)\,. \end{cases} \tag{33}$$

This process satisfies all the properties of the semidiscrete framework for anisotropic diffusion considered in [10]. In fact, since $\mathbf{K}$ is a matrix, the corresponding linear operator is Lipschitz-continuous. Moreover, $\mathbf{K}$ is symmetric, has nonnegative entries, and is irreducible. Thus, it follows that $\mathbf{A}$ is Lipschitz-continuous, symmetric, has nonegative off-diagonal entries, and is irreducible. Moreover, (33) implies that $\mathbf{A}$ has zero row sums. These are the conditions required in [10].

Remarks 2.

(a) Note that the fact that in the linear case, the semidiscrete diffusion framework is covered by our nonlocal continuous framework is a benefit of our formulation in terms of measures.
(b) This also shows that Weickert's semidiscrete diffusion theory is more general than his continuous one, which requires local processes in terms of PDEs.
(c) Extensions to higher dimensions can be obtained by choosing $\Omega \subset \mathbb{R}^N$ and the measure $\mu_n := \mu \times \mu \times .... \times \mu$, where the product is taken n times, and μ is a discrete measure concentrated on $\mathcal{Z}_h := \left\{ h(z - \frac{1}{2})\;;\; z \in \mathbb{Z} \right\}$ as above.

3.3 A Scale-Space Variant of Bilateral Filtering

Bilateral filtering goes back to Aurich and Weule [16] and became popular by a paper of Tomasi and Manduchi [17]. In a continuous notation, it filters a greyscale image $f : \Omega \to \mathbb{R}$ by means of spatial and tonal averaging with Gaussian weights:

$$u(\mathbf{x}) = \frac{\int_\Omega g_\lambda(|f(\mathbf{y}) - f(\mathbf{x})|)\; g_\rho(|\mathbf{y} - \mathbf{x}|)\; f(\mathbf{y})\; d\mathbf{y}}{\int_\Omega g_\lambda(|f(\mathbf{y}) - f(\mathbf{x})|)\; g_\rho(|\mathbf{y} - \mathbf{x}|)\; d\mathbf{y}}\,, \tag{34}$$

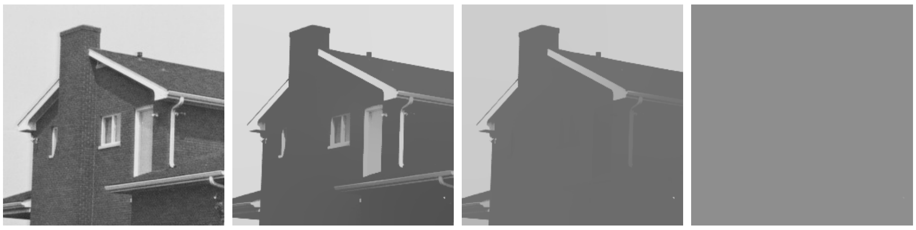

Fig. 1. Bilateral scale-space evolution (35) of a test image (256 $\times$ 256 pixels, $\rho = 5$, $\lambda = 10$). **From left to right:** $t = 0$, 500, 10000, and 200000.

where $g_\rho(s) := \exp(-s^2/(2\rho^2))$. While bilateral filtering is a nonlocal process, it does not preserve the average grey value. Moreover, it is typically applied in a noniterative way.
We propose the following modification that leads to an evolution equation:

$$\partial_t u(\mathbf{x}, t) = \frac{1}{c} \int_\Omega g_\lambda(|f(\mathbf{y}) - f(\mathbf{x})|)\, g_\rho(|\mathbf{y} - \mathbf{x}|)\, \big(u(\mathbf{y}, t) - u(\mathbf{x}, t)\big)\, d\mathbf{y}\,, \tag{35}$$

where $c := \int_\Omega g_\rho(\mathbf{y})\, d\mathbf{y}$ performs a normalisation of the spatial weighting. In our terminology, this is a nonlocal linear scale-space with the specific kernel $K(\mathbf{x}, \mathbf{y}) = \frac{1}{c}\, g_\lambda(|f(\mathbf{y}) - f(\mathbf{x})|)\, g_\rho(|\mathbf{y} - \mathbf{x}|)$ and the Lebesgue measure μ. It is straightforward to check that it satisfies the requirements (NL1)–(NL4) of our theory. This implies e.g. that it preserves the average grey value.
Figure 1 illustrates such a scale-space evolution. It has been obtained with an explicit finite difference scheme. As predicted by the theory, we observe that the image is gradually simplified. For $t \to \infty$, it converges to a flat steady state with the same average grey value as the initial image. It is remarkable how well the localisation of edges is preserved.

3.4 Robustified Bilateral Scale-Space

While our bilateral integro-differential equation (35) gives an interesting scale-space evolution, its performance under noise is less favourable. The reason is easy to understand: Noise creates large values for $|f(\mathbf{y}) - f(\mathbf{x})|$, such that the corresponding tonal weight $g_\lambda(|f(\mathbf{y}) - f(\mathbf{x})|)$ becomes very small. As a result, noisy structures are rewarded by a longer lifetime in scale-space. Similar problems are also well-known for the Perona–Malik diffusion filter [22]. Therefore, we can also use a similar strategy to overcome this problem: Following Catté et al. [23], we replace the image f in the argument of the tonal weight g_λ by a Gaussian-smoothed variant f_σ, where σ denotes the standard deviation of the Gaussian. Hence, our robustified bilateral evolution is given by

$$\partial_t u(\mathbf{x}, t) = \frac{1}{c} \int_\Omega g_\lambda(|f_\sigma(\mathbf{y}) - f_\sigma(\mathbf{x})|)\, g_\rho(|\mathbf{y} - \mathbf{x}|)\, \big(u(\mathbf{y}, t) - u(\mathbf{x}, t)\big)\, d\mathbf{y}\,. \tag{36}$$

[b]

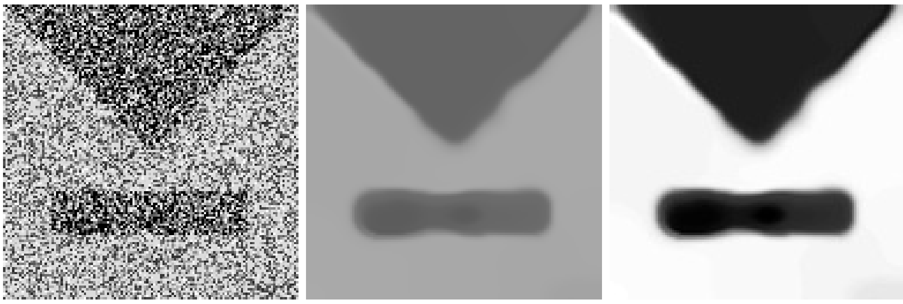

Fig. 2. (a) Left: Noisy test image, 128×128 pixels. **(b) Middle:** After processing with the robustified bilateral process (36) with $\sigma = 2$, $\rho = 5$, $\lambda = 1.4$, and $t = 500$. **(c) Right:** After rescaling the filtered result from (b) to the greyscale interval $[0, 255]$.

Its behaviour is illustrated in Fig. 2. We observe that this process is well-suited for removing even a large amount of noise, while keeping the semantically important edge structures. Its performance is comparable to the edge-enhancing anisotropic nonlinear diffusion filter from [24]. However, this is achieved with a linear process, that does not require to struggle with the numerical challenges of implementing anisotropic filters with a diffusion tensor.

4 Conclusions

In our paper we have established a nonlocal scale-space theory. To this end, we have studied a general type of linear nonlocal problems and have proven scale-space properties. We have shown that some existing diffusion methods can be interpreted within this general formulation. More importantly we have also introduced two modifications of bilateral filtering that satisfy our nonlocal scale-space requirements and can be highly robust under noise.
In our future work we intend to generalise our theory from the linear to the nonlinear setting, such that its applicability is further broadened.

References

1. Iijima, T.: Basic Theory on Normalization of Pattern (in Case of Typical One- Dimensional Pattern). Bulletin of the Electrotechnical Laboratory 26, 368–388, (1962) In Japanese.
2. Witkin, A.P.: Scale-Space Filtering. In: Proc. Eighth International Joint Conference on Artificial Intelligence Volume 2, p.p. 945–951. Karlsruhe, West Germany (August 1983)
3. Koenderink, J.J.: The Structure of Images. Biological Cybernetics 50, 363–370 (1984)
4. Lowe, D.L.: Distinctive Image Features From Scale-Invariant Keypoints. International Journal of Computer Vision 60(2), 91–110 (2004)

5. Lindeberg, T.: Scale-Space Theory in Computer Vision. Kluwer, Boston (1994)
6. Florack, L.: Image Structure. Volume 10 of Computational Imaging and Vision. Kluwer, Dordrecht (1997)
7. Weickert, J., Ishikawa, S., Imiya, A.: Linear Scale-Space has First been Proposed in Japan. Journal of Mathematical Imaging and Vision 10(3), 237–252 (May 1999)
8. Felsberg, M., Sommer, G.: Scale-Adaptive Filtering Derived From the Laplace Equation. In Radig, B., Florczyk, S., eds.: Pattern Recognition. Volume 2032 of Lecture Notes in Computer Science, p.p. 95–106. Springer, Berlin (2001)
9. Duits, R., Florack, L., de Graaf, J., ter Haar Romeny, B.: On the Axioms of Scale Space Theory. Journal of Mathematical Imaging and Vision 20, 267–298 (2004)
10. Weickert, J.: Anisotropic Diffusion in Image Processing. Teubner, Stuttgart (1998)
11. van den Boomgaard, R.: The Morphological Equivalent of the Gauss Convolution. Nieuw Archief Voor Wiskunde 10(3), 219–236 (November 1992)
12. Alvarez, L., Guichard, F., Lions, P.L., Morel, J.M.: Axioms and Fundamental Equations in Image Processing. Archive for Rational Mechanics and Analysis 123, 199–257 (1993)
13. Guichard, F.: A Morphological, Affine, and Galilean Invariant Scale-Space for movies. IEEE Transactions on Image Processing 7(3), 444–456 (March 1998)
14. Lindeberg, T.: Generalized Gaussian Scale-Space Axiomatics Comprising Linear Scale-Space, Affine Scale-Space and Spatio-Temporal Scale-Space. Journal of Mathe- matical Imaging and Vision 40, 36–81 (2011)
15. Scherzer, O., Weickert, J.: Relations Between Regularization and Diffusion Filtering. Journal of Mathematical Imaging and Vision 12(1), 43–63 (February 2000)
16. Aurich, V., Weule, J.: Non-Linear Gaussian Filters Performing Edge Preserving Diffu- sion. In: Sagerer, G., Posch, S., Kummert, F., eds.: Mustererkennung 1995, p.p. 538–545. Springer, Berlin (1995)
17. Tomasi, C., Manduchi, R.: Bilateral Filtering for Gray and Color Images. In: Proc. Sixth International Conference on Computer Vision, p.p. 839–846. Narosa Publishing House, Bombay, India (January 1998)
18. Buades, A., Coll, B., Morel, J.M.: A Review of Image Denoising Algorithms, with a New One. Multiscale Modeling and Simulation 4(2), 490–530 (2005)
19. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Image Denoising by Sparse 3D Transform-Domain Collaborative Filtering. IEEE Transactions on Image Processing 16(8), 2080–2095 (August 2007)
20. Gilboa, G., Osher, S.: Nonlocal Operators with Applications to Image Processing. Multiscale Modeling and Simulation 7, 1005–1028 (2008)
21. Andreu-Vaillo, F., Maz on, J.M., Rossi, J.D.: Nonlocal Diffusion Problems. American Mathematical Society, Providence (2010)
22. Perona, P., Malik, J.: Scale-Space and Edge Detection using Anisotropic diffusion. IEEE Transactions on Pattern Analysis and Machine Intelligence 12, 629–639 (1990)
23. Catt e, F., Lions, P.L., Morel, J.M., Coll, T.: Image Selective Smoothing and Edge Detection by Nonlinear Diffusion. SIAM Journal on Numerical Analysis 32, 1895–1909 (1992)
24. Weickert, J.: Theoretical Foundations of Anisotropic Diffusion in Image Processing. Computing Supplement 11, 221–236 (1996)
25. Ballester, C., Calderero, F., Caselles, V., Facciolo, G.: Multiscale Analysis of Similarities between Images on Riemannian Manifolds. Multiscale Modeling and Simulation 12(2), 616–649 (2014)