Anisotropic Diffusion in Image Processing

Joachim Weickert

Vom Fachbereich Mathematik der Universität Kaiserslautern zur Verleihung des akademischen Grades Doktor der Naturwissenschaften (Doctor rerum naturalium, Dr. rer. nat.) genehmigte Dissertation

1. Gutachter: Prof. Dr. H. Neunzert

2. Gutachter: Prof. Dr. P.-L. Lions

Vollzug der Promotion: 29. 01. 1996

D 386

Meinen Eltern

Preface

Through many centuries physics has been one of the most fruitful sources of inspiration for mathematics. As a consequence, mathematics has become an economic language providing a few basic principles which allow to explain a large variety of physical phenomena. Many of them are described in terms of partial differential equations (PDEs).

In recent years, however, mathematics also has been stimulated by other novel fields such as image processing. Goals like image segmentation, multiscale image representation, or image restoration cause a lot of challenging mathematical questions. Nevertheless, these problems frequently have been tackled with a pool of heuristical recipes. Since the treatment of digital images requires very much computing power, these methods had to be fairly simple. With the tremendous advances in computer technology in the last decade, it has become possible to apply more sophisticated techniques such as PDE-based methods which have been inspired by physical processes.

Among these techniques, parabolic PDEs have found a lot of attention for smoothing and restoration purposes, see e.g. [113]. To restore images these equations frequently arise from gradient descent methods applied to variational problems. Image smoothing by parabolic PDEs is closely related to the scale-space concept where one embeds the original image into a family of subsequently simpler, more global representations of it. This idea plays a fundamental role for extracting semantically important information. The pioneering work of Alvarez, Guichard, Lions and Morel [11] has demonstrated that all scale-spaces fulfilling a few fairly natural axioms are governed by parabolic PDEs with the original image as initial condition. Within this framework, two classes can be justified in a rigorous way as scale-spaces: the linear diffusion equation with constant diffusivity and nonlinear so-called morphological PDEs. All these methods satisfy a monotony axiom as smoothing requirement which states that, if one image is brighter than another, then this order is preserved during the entire scale-space evolution.

An interesting class of parabolic equations which pursue both scale-space and restoration intentions is given by nonlinear diffusion filters. Methods of this type have been proposed for the first time by Perona and Malik in 1987 [190]. In order to smooth the image and to simultaneously enhance semantically important features such as edges, they apply a diffusion process whose diffusivity is steered by local image properties. These filters are difficult to analyse mathematically, as they may act locally like a backward diffusion process. This gives rise to well-posedness questions. Moreover, since they violate the monotony axiom, they cannot be treated within the abovementioned scale-space axiomatic. On the other hand, nonlinear diffusion filters are frequently applied with very impressive results; so there appears the need for a theoretical foundation.

The goal of the present work is to develop results in this direction: We shall investigate if there is a scale-space interpretation beyond the monotony requirement, and in which sense restoration properties and scale-space requirements are compatible.

To this end, we consider a general class of nonlinear diffusion processes. They utilize a diffusion tensor depending on the local image structure via the so-called structure tensor (second-moment matrix), a well-established tool for texture analysis. This class comprises linear diffusion filters as well as spatial regularizations of the Perona–Malik process. Since it reveals a diffusion tensor instead of a scalar diffusivity, the diffusive flux does not have to be parallel to the grey value gradient: the filter may become anisotropic. The use of diffusion tensors allows to design novel diffusion filters which outperform isotropic ones with respect to certain applications such as denoising of highly degraded edges or enhancing coherent flow-like images by closing interrupted one-dimensional structures. In order to establish well-posedness and scale-space properties for this class, we shall investigate existence, uniqueness, stability, maximum–minimum principles, Lyapunov functionals, and invariances.

In practice, however, one has to deal with digital images which are sampled on a grid. For this reason it is important to know if the results for the continuous framework carry over to the discrete setting. These questions shall be addressed in the present work as well. A general characterization of semidiscrete and fully discrete filters, which reveal similar properties as their continuous diffusion counterparts, is presented.

Due to the interdisciplinary character of image processing, this book is written in a style which should be understandable for researches with a very different scientific background. Therefore, the intention is to keep it as self-consistent as possible and to require only basic knowledge in partial differential equations and image processing. The author apologizes if some sections may reiterate topics that are already known to some of the readers.

As a first step to provide a common basic knowledge, Chapter 1 surveys the fundamental ideas behind PDE-based smoothing and restoration methods. The discussion of limitations and open questions of these techniques puts us in a position to define the goals of the present work more precisely. These goals and a detailed outline of what is discussed in the subsequent chapters can be found in Section 1.7.

It is a pleasure to take this opportunity to express my gratitude to all people who have contributed to the success of this work:

- Professor Helmut Neunzert for drawing my interest to diffusion processes in image processing and for providing the possibility to carry out this work at the Laboratory of Technomathematics.

- Professor Pierre–Louis Lions (CEREMADE, University Paris IX) for inviting me to the CEREMADE and for giving me the honour to get, as an invited speaker at the conference "Multiscale Analysis in Image Processing", into contact to many researchers in this field.

- Professor Jean–Michel Morel (also CEREMADE) for fruitful discussions and for his support.

- Professor Bernd Jähne (IWR, University of Heidelberg) for stimulating my interest in structure tensor analysis.

- Dr. Bart ter Haar Romeny (Imaging Center Utrecht, Utrecht University) for his encouragement and for introducing me to the members of the ESPRIT–NSF project "Geometry-Driven Diffusion".

- Dr. Martin Reißel for undertaking the hard job of proofreading, for valuable hints, and for many discussions about mathematics and life in general.

- Andrea Bechtold for her thorough assistance in all kinds of difficulties with the English language.

- All colleagues at the Laboratory of Technomathematics who helped me whenever mathematical or software problems occured.

- The Center for Applied Mathematics Darmstadt–Kaiserslautern and the Stiftung Rheinland–Pfalz für Innovation for their financial support.

- Last but not least, I owe a lot of thanks to my parents Gerda and Norbert for giving me all kinds of help throughout my whole life.

Without the aid of all of them this work would not have been possible.

Kaiserslautern, November 1995 Joachim Weickert

Contents

Chapter 1

Image smoothing and restoration by PDEs

PDE-based methods appear in a large variety of image processing areas ranging from shape-from-shading, active contour models and optical flow to stereo vision and even histogramme modification.

This chapter reviews their main application, namely image smoothing, image restoration and edge detection. It is written in an informal style and refers to a large amount of original literature, where proofs and full mathematical details can be found.

The goal is to give an introduction to PDE methods that are related to our nonlinear diffusion approach. On the one hand, this should make the reader sensitive to the similarities, differences and problems of all these methods, on the other hand it shows how our work relates to them and motivates the reader to study how some of these problems will be solved later on.

For each class of methods we shall sketch the basic ideas as well as the theoretical background, numerical aspects and applications. Many of these ideas are borrowed from physical phenomena such as wave propagation or transport of heat and mass. Nevertheless, also gas dynamics, crack propagation, grassfire flow, the study of salinity profiles in oceanography, or mechanisms of the retina and the brain are closely related to some of these approaches. Although a detailed discussion of these connections would be far beyond the scope of this work, we shall mention them wherever they appear, in order to allow the interested reader to pursue these ideas.

The outline of this chapter is as follows:

We start with reviewing the physical ideas behind diffusion processes. This helps us to understand the next sections which are concerned with the properties of linear and nonlinear diffusion filters in image processing. The subsequent study of image enhancement methods of diffusion–reaction type relates diffusion filters to variational image restoration techniques. Afterwards we investigate morphological filters, a topic which seems at first glance fairly different to the diffusion approach. Nevertheless, it reveals some interesting relations when it is interpreted within a PDE framework. This becomes especially evident when considering curvature-based morphological PDEs. Finally we shall discuss total variation image restoration techniques which permit discontinuous solutions. The last section summarizes the advantages and shortcomings of the main methods and gives an outline of the questions we are concerned with in the subsequent chapters.

1.1 Linear diffusion filtering

The simplest and best investigated PDE method for smoothing images is to apply a linear diffusion process. After having sketched the physical background of diffusion processes, we shall focus on the relation between linear diffusion filtering and the convolution with a Gaussian, analyse its smoothing properties for the image as well as its derivatives, and review the fundamental properties of the Gaussian scale-space induced by linear diffusion filtering. Afterwards we shall give a survey on discrete aspects and discuss shortcomings of the linear diffusion paradigm. The section is concluded by sketching directed diffusion processes, a method for introducing bias into the process without renouncing linearity.

1.1.1 Physical background of diffusion processes

Most people have an intuitive impression of diffusion as a physical process that equilibrates concentration differences without creating or destroying mass. This physical observation can be easily cast in a mathematical formulation.

The equilibration property is expressed by *Fick's law*

$$j = -D \cdot \nabla u. \tag{1.1}$$

This equation states that a concentration gradient ∇u causes a flux j which aims to compensate for this gradient. The relation between ∇u and j is described by the *diffusion tensor* D, a positive definite symmetric matrix. The case where j and ∇u are parallel is called *isotropic*. Then we may replace the diffusion tensor by a positive scalar-valued *diffusivity* g. In the general *anisotropic* case, j and ∇u are not parallel.

The observation that diffusion does only transport mass without destroying it or creating new mass is expressed by the *continuity equation*

$$\partial_t u = -\mathrm{div}\ j \tag{1.2}$$

where t denotes the time.

If we plug in Fick's law into the continuity equation we end up with the *diffusion equation*

$$\partial_t u = \operatorname{div}(D \cdot \nabla u). \tag{1.3}$$

This equation appears in many physical transport processes. In the context of heat transfer it is called *heat equation*.

In image processing we may interpret the concentration as a grey value at a certain point. The diffusion tensor does not have to be constant: frequently it is advantageous to choose it as a function of the local image structure. This leads us to *nonlinear diffusion filters*. Three cases are relevant for image processing:

(a) linear isotropic diffusion filters using a constant diffusivity,

(b) nonlinear isotropic diffusion filters with diffusivities being adapted to the local image structure,

(c) nonlinear anisotropic diffusion filters with diffusion tensors being adapted to the local image structure.

The linear case (a) is treated below, and the nonlinear cases (b) and (c) will be discussed in section 1.2.

1.1.2 Foundations of linear diffusion filtering

Gaussian smoothing

Let a grey-scale image f be represented by a real-valued mapping $f \in L^1(\mathbb{R}^2)$. A widely-used way to smooth f is by calculating the convolution

$$(K_\sigma * f)(x) := \int_{\mathbb{R}^2} K_\sigma(x-y)\, f(y)\, dy \tag{1.4}$$

where K_σ denotes the two-dimensional Gaussian of width (standard deviation) $\sigma > 0$:

$$K_\sigma(x) := \frac{1}{2\pi\sigma^2} \cdot \exp\left(-\frac{|x|^2}{2\sigma^2}\right). \tag{1.5}$$

There are multiple reasons for the excellent smoothing properties of this method:

First we observe that since $K_\sigma \in C^\infty(\mathbb{R}^2)$ we get $K_\sigma * f \in C^\infty(\mathbb{R}^2)$, even if f is only absolutely integrable.

Next, let us investigate the behaviour in the frequency domain. When defining the Fourier transformation $\mathcal{F}$ by

$$(\mathcal{F}f)(\omega) := \int_{\mathbb{R}^2} f(x)\, \exp(-i\langle \omega, x\rangle)\, dx \tag{1.6}$$

we obtain by the convolution theorem that

$$(\mathcal{F}(K_\sigma * f))(\omega) \;=\; (\mathcal{F}K_\sigma)(\omega)\cdot(\mathcal{F}f)(\omega). \tag{1.7}$$

Since the Fourier transform of a Gaussian is again Gaussian-shaped,

$$(\mathcal{F}K_\sigma)(\omega) \;=\; 2\pi\cdot\exp\left(-\frac{|\omega|^2}{2/\sigma^2}\right), \tag{1.8}$$

we observe that (1.4) is a low-pass filter that damps high frequencies in a monotone way.

Interestingly, the smoothing behaviour can also be understood in the context of a PDE interpretation.

Equivalence to linear diffusion filtering

It is a classical result (cf. e.g. [195, pp. 267–271] and [120, pp. 43–56]) that for any bounded $f \in C(\mathbb{R}^2)$ the linear diffusion process

$$\partial_t u \;=\; \Delta u, \tag{1.9}$$

$$u(x,0) \;=\; f(x) \tag{1.10}$$

possesses the unique solution

$$u(x,t) = \begin{cases} f(x) & (t=0) \\ (K_{\sqrt{2t}} * f)(x) & (t>0). \end{cases} \tag{1.11}$$

This solution is unique, provided we restrict ourselves to functions satisfying

$$|u(x,t)| \;\le\; M\cdot\exp\left(a|x|^2\right) \qquad (M, a > 0). \tag{1.12}$$

It depends continuously on the initial image f with respect to $\|.\|_{L^\infty(\mathbb{R}^2)}$, and it fulfils the maximum–minimum principle

$$\inf_{\mathbb{R}^2} f \;\le\; u(x,t) \;\le\; \sup_{\mathbb{R}^2} f \quad \text{on} \quad \mathbb{R}^2\times[0,\infty). \tag{1.13}$$

From (1.11) we observe that the time t corresponds with the spatial width $\sqrt{2t}$ of the Gaussian. Hence, smoothing structures of order σ requires to stop the diffusion process at time

$$T = \tfrac{1}{2}\,\sigma^2. \tag{1.14}$$

Gaussian derivatives

In order to understand the structure of an image we have to analyse grey value fluctuations within a neighbourhood of each image point, that is to say, we need information about its derivatives. However, differentiation is an ill-posed problem in the sense of Hadamard [239], as small perturbations in the original image can lead to arbitrary large fluctuations in the derivatives. Hence, the need for regularizing methods arises.

One possibility to regularize is to convolve the image with a Gaussian prior to differentiation [239]. By the equality

$$\partial_{x_1}^n \partial_{x_2}^m (K_\sigma * f) \; = \; K_\sigma * (\partial_{x_1}^n \partial_{x_2}^m f) \; = \; (\partial_{x_1}^n \partial_{x_2}^m K_\sigma) * f \qquad (1.15)$$

we observe that all derivatives undergo the same Gaussian smoothing process as the image itself and this process is equivalent to convolving the image with derivatives of a Gaussian.

The resulting *Gaussian derivatives* can be successfully applied to the deblurring of images [115]. Moreover, they can be combined (often in a nonlinear way) to expressions *(differential invariants)* that are invariant under transformations such as rotations, for instance $|\nabla K_\sigma * u|$ or $\Delta K_\sigma * u$.

Differential invariants are useful for the detection of features such as edges, ridges, junctions, and blobs (see [153] for an overview). To illustrate this, we focus on two applications for detecting edges.

A frequently used method is the *Canny edge detector* [51]. It is based on calculating the first derivatives of the Gaussian-smoothed image. After applying sophisticated thinning and linking mechanisms (*non-maxima suppression* and *hysteresis thresholding*), edges are identified as locations where the gradient magnitude has a maximum. This method is acknowledged to be the best linear edge detector [58], and it has become almost a standard in edge detection.

Another interesting edge detector is the *Marr–Hildreth operator* [165], which uses the *Laplacian-of-Gaussian (LoG)* ΔK_σ as convolution kernel. Edges of f are identified as zero-crossings of $\Delta K_\sigma * f$. This needs no further postprocessing and always gives closed contours. There is some evidence that LoGs and especially their approximation by *differences-of-Gaussians (DoGs)* play an important role in the visual system of mammals, see [165] and the references therein.

If one investigates the temporal evolution of the zero-crossings of an image filtered by linear diffusion, one observes an interesting phenomenon: When increasing the smoothing scale σ, no new zero-crossings can be created which are not present at finer scales [263]. This *evolution property* is closely connected to the maximum–minimum principle of certain parabolic operators [124]. Attempts to reconstruct the original image from the temporal evolution of the zero-crossings of the Laplacian have been carried out [125] with the result that this is practically unstable unless additional information is provided.

Nevertheless, the evolution property of the zero-crossings was one of the key investigations leading to the scale-space idea [263], which we shall describe below.

1.1.3 Scale-space properties

The general scale-space concept

It is a well-known fact that images usually contain structures at a large variety of scales. In those cases where it is not clear in advance which is the right scale for the depicted information it is desirable to have an image representation at multiple scales. Moreover, by comparing the structures at different scales, one obtains a hierarchy of image structures, which eases a subsequent image interpretation.

A *scale-space* is an image representation at a continuum of scales, embedding the image f into a family $\{T_t f \mid t \geq 0\}$ of gradually simplified versions of it, provided that it fulfils certain architectural, smoothing and invariance requirements which we shall sketch briefly below. For more details the reader is referred to [11] from which most of the following nomenclature is borrowed.

An important architectural assumption is *recursivity*, i.e. for $t = 0$, the scale-space representation gives the original image f, and the filtering may be split into a sequence of filter banks:

$$T_0 f = f, \tag{1.16}$$

$$T_{t+s} f = T_t(T_s f) \qquad \forall\, s, t \geq 0. \tag{1.17}$$

This property is also sometimes referred to as the *semigroup property*. Other architectural axioms require regularity properties of T_t and local behaviour as t tends to 0.

Information reduction arises from the wish that the smoothing transformation should not create artefacts when passing from fine to coarse representation. Thus, at a coarse scale, we must not have additional structures which are caused by the filtering method itself and not by underlying structures at finer scales. This property is specified by numerous authors in different ways, using concepts such as no creation of new level curves [143, 124, 150], nonenhancement of local extrema [24, 150], decreasing number of local extrema [150], maximum–minimum principle [124, 192], and comparison principle [11].

We may regard an image as a representative of an eqivalence class containing all images that depict the same object. Two images of this class differ e.g. by grey-level shifts, translations and rotations or even more complicated transformations such as affine mappings. This makes the requirement plausible that the scale-space analysis should be invariant to as many of these transformations as possible, in order to analyse only the depicted object [15].

The work of Alvarez, Guichard, Lions and Morel ([11], see also [8]) shows that every scale-space fulfilling some fairly natural architectural, information-reducing and invariance properties is governed by a PDE with the original image as initial condition.

Gaussian scale-space

The best investigated example of a scale-space is the *Gaussian scale-space*, which is obtained via convolution with Gaussians of increasing variance [263], or – equivalently – by linear diffusion filtering according to (1.9), (1.10). Numerous theoretical results indicate that this is the only "reasonable" way to define a *linear* scale-space:

- Koenderink [143] derived the Gaussian scale-space in one and two dimensions as the unique transformation that (a) does not create new level curves when increasing the scale parameter and (b) treats all spatial points and scale levels equally. He named the first requirement *causality* and the second one *homogeneity and isotropy*. Yuille and Poggio [268] came to similar results based on assumptions concerning the zero-crossings of the Laplacian.

- In the one-dimensional case, Babaud et al. [24] showed that the Gaussian is the unique kernel fulfilling some natural constraints such as symmetry, normalization, and nonenhancement of local extrema.

- Lindeberg [150] derived the one-dimensional Gaussian scale-space by combining the requirement of not introducing new local extrema with the semigroup property and normalization and symmmetry assumptions on the kernel.

- The uniqueness of the Gaussian kernel and the Gaussian derivatives can also be established from linearity, isotropy and spatial shift invariance constraints in connection with scale invariance, see ter Haar Romeny et al. [114]. This holds for any dimension.

- Alvarez et al. [11] proved that the Gaussian scale-space is the only linear transformation fulfilling a semigroup, locality and regularity axiom as well as a monotony requirement (comparison principle) and invariance with respect to grey level shifts and Euclidean transformations.

A very detailed treatment of linear scale-space theory can be found in the monograph by Lindeberg [151].

1.1.4 Numerical aspects

The preceding theory is entirely continuous. However, in practical problems, the image is sampled at the nodes *(pixels)* of a fixed equidistant grid. Thus, the diffusion filter has to be discretized.

By virtue of the equivalence of solving the linear diffusion equation and convolving with a Gaussian, we can either approximate the convolution process or the diffusion equation.

Let us start with convolution-based methods. When restricting the image to a finite domain and applying the Fast Fourier Transformation (FFT), convolution in the spatial domain can be reduced to multiplication in the frequency domain, cf. (1.7). Due to the overhead of transformation and backtransformation, this proceeding is especially efficient if the kernel size σ is large.

When carrying out the convolution in the spatial domain, the Gaussian kernel has to be sampled and truncated. Although this technique is frequently used, it has the drawback that it does not preserve the semigroup property of the continuous Gaussian scale-space [150].

Lindeberg [150] has developed a linear scale-space theory for the *semidiscrete*[1] case. He proved that the discrete analogue of the Gaussian kernel is given in terms of modified Bessel functions of integer order. Since this scale-space family arises naturally from a semidiscretized version of the diffusion equation, he concludes that approximating the diffusion equation should be preferred to discretizing the convolution integral.

Among the numerous numerical possibilities to approximate the linear diffusion equation, finite difference (FD) schemes dominate the field. Except for a few implicit approaches [108, 48, 49], explicit schemes are mainly used. A very efficient approximation of the Gaussian scale-space results from applying multigrid ideas. The *Gaussian pyramid* [47] gives a multilevel representation at finitely many scales of different resolution. By subsequently smoothing the image with an explicit scheme for the diffusion equation and restricting the result to a coarser grid, one obtains a simplified image representation at the next coarser grid. Due to their simplicity and efficiency, pyramid decompositions are meanwhile integrated into commercially available hardware ([151], p. 38).

1.1.5 Shortcomings

In spite of its multitude of properties that make linear diffusion filtering unique and easy to handle, it reveals several drawbacks:

(a) An obvious disadvantage of Gaussian smoothing is the fact that it does not only smooth noise, but also blurs important features such as edges

[1] By semidiscrete we mean discrete in space and continuous in time throughout this work.

and makes them harder to identify. This equation cannot comprise any additional information on structures which are worth being preserved (or even enhanced).

(b) Linear diffusion filtering dislocates edges when moving from finer to coarser scales, see e.g. Witkin [263]. So structures which are identified at a coarse scale do not give the right location and have to be traced back to the original image [263, 29]. In practise, relating dislocated information obtained at different scales is difficult and gives rise to instabilities, unless the step size between subsequent scales is very small. These coarse-to-fine tracking difficulties are generally denoted as the *correspondence problem.*

(c) Some causality properties of Gaussian scale-space do not hold for dimensions larger than one: A closed zero-crossing contour can split into two as the scale increases [268], and it is generally not true that the number of local extrema is nonincreasing, see [149, 150] for illustrative counterexamples. A deep mathematical analysis of such phenomena has been carried out by Damon [69].

Due to the uniqueness of the Gaussian scale-space within a linear framework we know that any modification in order to overcome these problems will either renounce linearity or some scale-space properties. We shall see that the shortcomings (a) and (b) can be avoided by nonlinear diffusion processes, while (c) requires morphological equations.

Before we turn our attention to those processes, let us first investigate a linear modification which has no scale-space interpretation anymore, but allows to include additional information into the evolution.

1.1.6 Directed diffusion

Provided we are given some background information in form of a smooth image b, it has been shown by Illner and Neunzert [127], that under some technical requirements and suitable boundary conditions the classical solution u of

$$\partial_t u = b\,\Delta u - u\,\Delta b, \tag{1.18}$$

$$u(x,0) = f(x) \tag{1.19}$$

converges to b along a path where the relative entropy with respect to b increases in a monotone way. Numerical experiments have been carried out by Giuliani [101], and an analysis in terms of nonsmooth b and weak solutions is due to Illner and Tie [128].

Such a *directed diffusion process* requires to specify an entire image as background information in advance; in many applications it would be desirable to include a priori knowledge in a less specific way, e.g. by prescribing that features within a certain contrast and scale range are considered to be semantically important and processed differently. Such demands can be satisfied by nonlinear diffusion filters.

1.2 Nonlinear diffusion filtering

Adaptive smoothing methods are based on the idea to apply a process which itself depends on local properties of the image. Although this concept is well-known in the image processing community (see [209] and the references therein for an overview), a corresponding PDE formulation was first given by Perona and Malik [190] in 1987. We shall discuss this model in detail, especially its ill-posedness aspects. This gives rise to study regularizations. These techniques can be extended to true anisotropic processes which make use of an adapted diffusion tensor instead of a scalar diffusivity.

1.2.1 The Perona–Malik model

Basic idea

Perona and Malik propose a nonlinear diffusion method for avoiding the blurring and localization problems of linear diffusion filtering [190, 192]. They apply a nonuniform process (which they name anisotropic) that reduces the diffusivity at those locations having a larger likelihood to be edges. This likelihood can be measured by $|\nabla u|^2$. Hence, they use[2]

$$\partial_t u = \operatorname{div}\left(g(|\nabla u|^2)\ \nabla u\right). \tag{1.20}$$

Among the diffusivities they propose is

$$g(s^2) = \frac{1}{1 + s^2/\lambda^2} \qquad (\lambda > 0). \tag{1.21}$$

It should be noted that – in our terminology – the Perona–Malik filter is regarded as an isotropic model, since it utilizes a scalar-valued diffusivity and not a diffusion tensor.

Interestingly, there exists a relation between (1.20) and the neural dynamics of brightness perception: In 1984 Cohen and Grossberg [63] proposed a model of the primary visual cortex with similar inhibition effects as in the Perona–Malik model.

[2] Since we want the edge estimator to depend smoothly on ∇u, we consider $|\nabla u|^2$ instead of $|\nabla u|$.

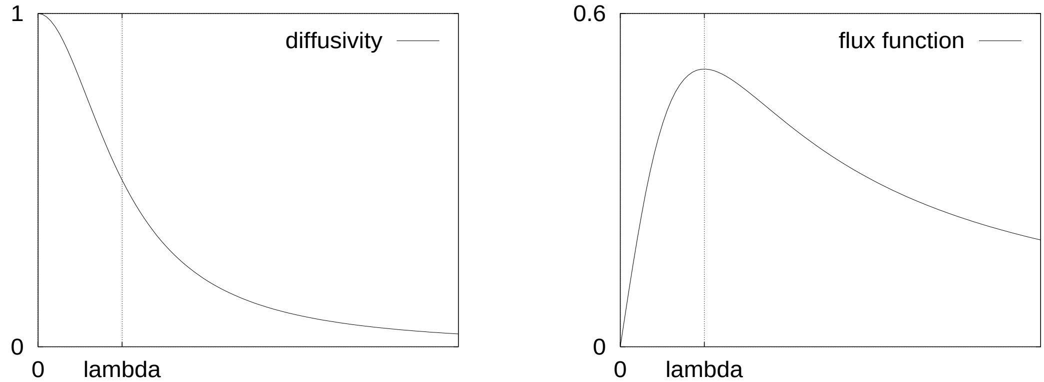

Figure 1.1: (a) LEFT: Diffusivity $g(s^2) = \frac{1}{1+s^2/\lambda^2}$. (b) RIGHT: Flux function $\Phi(s) = \frac{s}{1+s^2/\lambda^2}$.

The experiments of Perona and Malik were visually very impressive: edges remained stable over a very long time. It was demonstrated [192] that edge detection based on this process clearly outperforms the linear Canny edge detector, even without applying non-maximum suppression and hysteresis thresholding. This is due to the fact that diffusion and edge detection interact in one single process instead of being treated as two independent processes which are to be applied subsequently.

Moreover, there is another reason for the impressive behaviour at edges, which we shall discuss next.

Edge enhancement

To study the behaviour of the Perona–Malik filter at edges, let us for a moment restrict ourselves to the one-dimensional case. This simplifies the notation and illustrates the main behaviour, since near a straight edge a two-dimensional image approximates a function of one variable.

For the diffusivity (1.21) it follows that the *flux function* $\Phi(s) := sg(s^2)$ satisfies $\Phi'(s) \geq 0$ for $|s| \leq \lambda$, and $\Phi'(s) < 0$ for $|s| > \lambda$, see Figure 1.1. Since (1.20) can be rewritten as

$$\partial_t u = \Phi'(u_x) u_{xx}, \tag{1.22}$$

we observe that – in spite of its nonnegative diffusivity – the Perona–Malik model is of *forward parabolic type* for $|u_x| \leq \lambda$, and of *backward parabolic type* for $|u_x| > \lambda$. Hence, λ plays the role of a *contrast parameter* separating forward (low contrast) from backward (high contrast) diffusion areas.

It is not hard to verify that the Perona–Malik filter increases the slope at inflection points of edges within a backward area: If there exists a sufficiently smooth solution u it satisfies

$$\partial_t(u_x^2) \;=\; 2u_x\partial_x(u_t) \;=\; 2\Phi''(u_x)u_x u_{xx}^2 + 2\Phi'(u_x)u_x u_{xxx}. \tag{1.23}$$

A location x_0 where u_x^2 is maximal at some time t is characterized by $u_x u_{xx} = 0$ and $u_x u_{xxx} \le 0$. Therefore,

$$\left(\partial_t(u_x^2)\right)(x_0, t) \;\ge\; 0 \quad \text{for} \quad |u_x(x_0,t)| > \lambda \tag{1.24}$$

with strict inequality for $u_x u_{xxx} < 0$.

In the two-dimensional case, (1.22) is replaced by [11, 267]

$$\partial_t u \;=\; \Phi'(\nabla u)u_{\eta\eta} + g(|\nabla u|^2)u_{\xi\xi} \tag{1.25}$$

where the *gauge coordinates* ξ and η denote the directions perpendicular and parallel to ∇u, respectively. Hence, we have forward diffusion along *isophotes* (i.e. lines of constant grey value) combined with forward–backward diffusion along *flowlines* (lines of maximal grey value variation).

We observe that the forward–backward diffusion behaviour is not only restricted to the special diffusivity (1.21), it appears for all diffusivities $g(s^2)$ whose rapid decay causes non-monotone flux functions $\Phi(s) = sg(s^2)$. Such diffusivities are explicitly intended in the Perona–Malik method, as they give the desirable result of blurring small fluctuations and sharpening edges. Therefore, they are the main reason for the visually impressive results of this restoration technique. On the other hand, they lead to severe theoretical problems, which we shall discuss now.

Ill-posedness

Although there is no general theory for nonlinear parabolic processes, there exist certain frameworks which allow to establish well-posedness results for a large class of equations. Let us recall three examples:

- Let $S(N)$ denote the set of symmetric $N \times N$ matrices and $\text{Hess}(u)$ the Hessian of u. Classical differential inequality techniques [244] based on the Nagumo–Westphal lemma require that the underlying nonlinear evolution equation

  $$\partial_t u \;=\; F(t, x, u, \nabla u, \text{Hess}(u)) \tag{1.26}$$

  satisfies

  $$F(t, x, r, p, Y) \;\ge\; F(t, x, r, p, X) \tag{1.27}$$

  for all $X, Y \in S(2)$ where $Y - X$ is positive semidefinite.

- The same requirement is needed for applying the theory of viscosity solutions [68].

- Let H be a Hilbert space with scalar product $(.,.)$ and $A : H \to H$. In order to apply the concept of maximal monotone operators [39] to the problem

$$\frac{du}{dt} + Au \;=\; 0, \tag{1.28}$$

$$u(0) \;=\; f \tag{1.29}$$

  one has to ensure that A is *monotone*, i.e.

$$(Au - Av, u - v) \;\geq\; 0 \qquad \forall\, u, v \in H. \tag{1.30}$$

We observe that the nonmonotone flux function of the Perona–Malik process implies that neither (1.27) is satisfied nor A defined by $Au := -\mathrm{div}\,(g(|\nabla u|^2)\,\nabla u)$ is monotone. Therefore, none of these frameworks is applicable to ensure well-posedness results.

Although for general smooth nonmonotone Φ the existence question is not settled [122], Höllig [121] has proved that if Φ is piecewise linear, decreasing in a bounded interval, and increasing elsewhere, there exist initial functions for which the corresponding initial value problem has infinitely many solutions.

The current understanding of the Perona–Malik process makes it rather unlikely that it possesses smooth solutions, but it seems to be possible that this equation admits weak solutions. Uniqueness and stability with respect to the initial image should not be expected [58, 193].

Interestingly, such forward–backward diffusion equations are not as unnatural as they look at first glance: they have been proposed as a mathematical model for heat and mass transfer in a stably stratified turbulent shear flow to explain the evolution of stepwise constant temperature or salinity profiles in the ocean (although the well-posedness question was usually circumvented). Related equations also play a role in population dynamics and viscoelasiticity, see [27] and the references therein.

In the context of oceanography, numerical experiments were carried out by Posmentier [196] in 1977. Starting from a smoothly increasing initial distribution he reported the creation of perturbations which led to a stepwise constant salinity profile after some time. He also observed instabilities, a first experimental hint to the ill-posedness of this equation. Instabilities were also reported later on by Dzhu Magaziewa [78].

In image processing, numerical simulations on the ill-posedness of the one-dimensional Perona–Malik filter were performed by Nitzberg and Shiota [181], Fröhlich and Weickert [92], and Benhamouda [28]. All results point in the same direction: the solution depends strongly on the regularizing effect of the discretization. Finer discretizations are less regularizing and reveal a larger danger of *staircasing effects*, where the number of inflection points of a smoothed step edge increases.

Scale-space interpretation

Perona and Malik renounced the assumption of Koenderink's linear scale-space axiomatic [143] that the smoothing treats all spatial points and scale levels equally. Instead of this, they required that region boundaries should be sharp and should coincide with the semantically meaningful boundaries at each resolution level *(immediate localization)* and that intra-region smoothing should be preferred to inter-region smoothing *(piecewise smoothing)*. These properties are of significant practical interest, as they guarantee that structures can be detected easily and correspondence problems can be neglected. Their experiments demonstrated that the Perona–Malik filter satisfies these requirements fairly well [192].

In order to establish a smoothing scale-space property for this nonlinear diffusion process, it was argued that the Perona–Malik filter satisfies a maximum–minimum principle, provided that it possesses C^2 solutions [192]. However, this scale-space reasoning is not very satisfactory since it is unlikely that the Perona–Malik filter exhibits smooth solutions.

As an alternative to construct nonlinear diffusion scale-spaces, Salden [210], Florack [88, 90] and Eberly [79] propose to carry over the linear scale-space theory to the nonlinear case by considering nonlinear diffusion processes which result from special rescalings of the linear one.[3] Unfortunately, the Perona–Malik filter turns out not to belong to this class [88, 90].

Alvarez, Guichard, Lions and Morel [11] have developed a nonlinear scale-space axiomatic which comprises the linear scale-space theory as well as nonlinear morphological processes (which we will discuss in 1.4 and 1.5). Their smoothing axiom is a monotony assumption *(comparison principle)* requiring that the scale-space is order-preserving:

$$f \;\le\; g \quad\Longrightarrow\quad T_t f \;\le\; T_t g \quad \forall\, t \ge 0. \tag{1.31}$$

This property is closely related to a maximum–minimum principle and to L^∞-stability of the solution [11, 155]. However, the Perona–Malik model does not fit into this framework as well since its nonmonotonic flux function prevents it from satisfying the comparison principle (see also [193]).

[3] A comparison of the scale-space reasonings of Florack and Alvarez et al. can be found in [117].

1.2.2 Regularized nonlinear models

In spite of the abovementioned theoretical problems, it has been observed that practical implementations of the Perona–Malik process work frequently very well. This suggests that the numerical schemes provide implicit regularizations which stabilize the process. Hence, it seems natural to introduce the regularization directly into the continuous equation in order to become more independent of the numerical implementation [58, 181].

Höllig's nonuniqueness result [121] shows that the dynamics of the solution may critically depend on the sort of regularization. Hence, the regularization should be specifically adjusted to the desired goal of the forward–backward heat equation [27].

One can apply spatial or temporal regularization (and of course, a combination of both). Below we shall discuss three examples which illustrate the variety of possibilities and their tailoring towards a specific task.

(a) The first spatial regularization attempt is probably due to Posmentier who observed numerically the stabilizing effect of averaging the gradient within the diffusivity [196].

A mathematically sound formulation of this idea is given by Catté, Lions, Morel and Coll [58]. By replacing the diffusivity $g(|\nabla u|^2)$ of the Perona–Malik model by $g(|\nabla u_\sigma|^2)$ with $u_\sigma := K_\sigma * u$ they end up with

$$\partial_t u = \operatorname{div}\,(g(|\nabla u_\sigma|^2)\;\nabla u). \tag{1.32}$$

In [58] existence, uniqueness and regularity of a solution for $\sigma > 0$ has been established.

Other spatial regularizations of equations of Perona–Malik type have been proposed by Weickert [248, 250] and will be described in 1.2.3. Whitaker and Pizer [255] have suggested that the regularization parameter σ should be a decreasing function in t. In addition, Li and Chen [148] have proposed to subsequently decrease the contrast parameter λ.

We shall see that these spatial regularizations belong to a class of well-posed problems which possess a large class of Lyapunov functionals which guarantee that the solution converges to a constant steady-state.

From a practical point of view, spatial regularizations offer the advantage that they make the filter insensitive to noise at scales smaller than σ. Therefore, when regarding (1.32) as an image restoration equation, it exhibits besides the contrast parameter λ an additional *scale-parameter* σ. This avoids a shortcoming of the genuine Perona–Malik process which misinterprets strong oscillations due to noise as edges which should be preserved or even enhanced.

(b) P.-L. Lions proved in a private communication to Mumford that the one-dimensional process

$$\begin{aligned} \partial_t u &= \partial_x \left(g(v) \, \partial_x u \right), && (1.33)\\ \partial_t v &= \tfrac{1}{\tau} \left(|\partial_x u|^2 - v \right) && (1.34) \end{aligned}$$

leads to a well-posed filter (cf. [193]). We observe that v is intended as a time-delay regularization of $|\partial_x u|^2$ where the parameter $\tau > 0$ determines the delay. These equations arise as a special case of the spatio-temporal regularizations of Nitzberg and Shiota [181], when neglecting any spatial regularization. Mumford conjectures that this model gives piecewise constant steady states [180]. In this case, the steady-state solution would solve a segmentation problem.

(c) In the context of shear flows, Barenblatt et al. [27] regularized the one-dimensional forward–backward heat equation by considering the third-order equation

$$\partial_t u = \partial_x(\Phi(u_x)) + \tau \partial_{xt}(\Psi(u_x)) \qquad (1.35)$$

where Ψ is strictly increasing and uniformly bounded in $\mathbb{R}$, and $|\Phi'(s)| = O(\Psi'(s))$ as $s \to \pm\infty$. This regularization was physically motivated by introducing a relaxation time τ into the diffusivity.

For the corresponding initial boundary value problem with homogeneous Neumann boundary conditions they proved the existence of a unique generalized solution. They also showed that smooth solutions may become discontinuous within *finite* time, before they finally converge to a piecewise constant steady state.

1.2.3 Anisotropic nonlinear models

All nonlinear diffusion filters that we have investigated so far utilize a scalar-valued diffusivity g which is adapted to the underlying image structure. Therefore, they are isotropic and the flux $j = -g \nabla u$ is always parallel to ∇u. Nevertheless, in certain applications it would be desirable to rotate the flux towards the orientation of interesting features. These requirements cannot be satisfied by a scalar diffusivity anymore, a diffusion tensor leading to anisotropic diffusion filters has to be introduced.

First anisotropic ideas in image processing date back to Graham [109] in 1962, followed by Newman and Dirilten [176], Lev, Zucker and Rosenfeld [146], and Nagao and Matsuyama [174]. They used convolution masks that depended on the underlying image structure. Related statistical approaches were proposed by Knutsson, Wilson and Granlund [141, 262].

Convolution with shape-adapted anisotropic Gaussian kernels is studied by Nitzberg and Shiota [181], Lindeberg and Gårding ([152], cf. also [153]), and Yang et al. [266]. Unlike in the linear case, this is no more equivalent to suitable diffusion processes. If one wants to relate these ideas to diffusion filtering, one has to iterate the convolution and carry out sophisticated scaling limits, cf. [181].

Anisotropic diffusion filters usually apply spatial regularization strategies (an exception is Cottet's time-delay regularization [67]). Below we study two typical representatives of anisotropic diffusion processes. The first one offers advantages at noisy edges, whereas the second one is well-adapted to the processing of one-dimensional features.

(a) *Anisotropic regularization of the Perona–Malik process*
In the interior of a segment the nonlinear isotropic diffusion equation (1.32) behaves almost like the linear diffusion filter (1.9), but at edges diffusion is inhibited. Therefore, noise at edges cannot be eliminated successfully by this process.

To overcome this problem, a desirable method should prefer diffusion along edges to diffusion perpendicular to them.

Anisotropic models do not only take into account the modulus of the edge detector ∇u_σ, but also its direction. To this end, we construct the orthonormal system of eigenvectors v_1, v_2 of the diffusion tensor D such that they reflect the estimated edge structure:

$$v_1 \parallel \nabla u_\sigma, \qquad v_2 \perp \nabla u_\sigma. \tag{1.36}$$

In order to prefer smoothing along the edge to smoothing across it, Weickert [247, 250] proposed to choose the corresponding eigenvalues λ_1 and λ_2 as

$$\lambda_1(\nabla u_\sigma) := g(|\nabla u_\sigma|^2), \tag{1.37}$$

$$\lambda_2(\nabla u_\sigma) := 1. \tag{1.38}$$

In general, ∇u does not coincide with one of the eigenvectors of D as long as $\sigma > 0$. Hence, this model behaves really anisotropic. If we let the regularization parameter σ tend to 0, we end up with the isotropic Perona–Malik process.

Another anisotropic model which can be regarded as a regularization of an isotropic nonlinear diffusion filter has been described in [248].

(b) *Anisotropic models for smoothing one-dimensional objects*
A second motivation for introducing anisotropy into diffusion processes arises from the wish to process one-dimensional features such as line-like

structures. To this end, Cottet and Germain [66] constructed a diffusion tensor with eigenvectors as in (1.36) and corresponding eigenvalues

$$\lambda_1(\nabla u_\sigma) := 0, \tag{1.39}$$

$$\lambda_2(\nabla u_\sigma) := \frac{\eta|\nabla u_\sigma|^2}{1+(|\nabla u_\sigma|/\sigma)^2} \qquad (\eta > 0). \tag{1.40}$$

This is a process diffusing solely perpendicular to ∇u_σ. For $\sigma \to 0$, we observe that ∇u becomes an eigenvector of D with correponding eigenvalue 0. Therefore, the process stops completely. In this sense, the Cottet–Germain model is not intended as an anisotropic regularization of the Perona–Malik equation.

In [66] the existence of weak solutions for this process was proved.

Since the Cottet–Germain model diffuses only in one direction, its success critically depends on a correct estimate of the appropriate direction. We shall see that this process can be improved when replacing $\nabla u_\sigma^\perp$ by a more robust descriptor of local orientation, the structure tensor.

1.2.4 Generalizations

Higher dimensions. It is easily seen that many of the previous results can be generalized to higher dimensions. This may be useful when considering e.g. computerized tomography (CT) or magnetic resonance tomography (MRT) image sequences arising from medical applications or when applying diffusion filters to the postprocessing of fluctuating higher-dimensional numerical data. Three-dimensional diffusion filters have been investigated by Gerig et al. [99] and Rambaux and Garçon [202].

More sophisticated structure descriptors. The edge detector ∇u_σ enables us to adapt the diffusion to magnitude and direction of edges, but it cannot distinguish between edges and corners or gives a reliable measure of local orientation. To this end, one can steer the smoothing process by more advanced structure descriptors such as higher-order derivatives [82] or tensor-valued expressions of first-order derivatives [181, 179, 252]. The theoretical analysis in the present work shall comprise the second possibility.

Vector-valued models. Vector-valued images can arise either from devices measuring multiple physical properties or from a feature analysis of one single image. Examples for the first category are colour images, multi-spectral Landsat exposures and multi-spin echo MRT images, whereas representatives of the second class are given by statistical moments or the jet space induced by the image itself and its partial derivatives up to a given order. Feature vectors play an important role for tasks like texture segmentation.

The simplest idea how to apply diffusion filtering to multichannel images would be to diffuse all channels separately and independently from each other. This leads to the undesirable effect that edges may be formed at different locations for each channel. In order to avoid this, one should use a common diffusivity which combines information from all channels. Such isotropic vector-valued diffusion models were proposed by Gerig, Kübler, Kikinis, Jolesz [99] and Whitaker [257, 258] in the context of medical imagery.

1.2.5 Numerical aspects

In contrast to linear diffusion filtering, no satisfactory semidiscrete or discrete scale-space interpretation for nonlinear diffusion filters has been available up to now. Chapter 3 and 4 will provide such a theoretical framework.

Nevertheless, numerous numerical methods for nonlinear diffusion filtering have been applied:

Fröhlich and Weickert [92] have compared three schemes for a one-dimensional regularized nonlinear diffusion filter: a wavelet method of Petrov–Galerkin type, a spectral method and a finite-difference (FD) scheme. It turned out that – especially for large σ – all results were fairly similar. Since the computational effort is of a comparable order of magnitude, it seems to be a matter of taste which scheme is preferred.

Of course, other numerical methods are applicable as well, e.g. finite elements [220]. Neural network realizations of nonlinear diffusion filters are investigated by Cottet [65, 67]. Perona and Malik [191] and Schnörr [223] propose hardware realizations by means of analogue VLSI networks with nonlinear resistors.

In most applications of nonlinear diffusion filters, finite differences are preferred, since they are easy to handle and the pixel structure of a real digital image already provides a natural discretization on a fixed rectangular grid. Explicit schemes are very simple to implement and therefore they are used almost exclusively [23, 66, 71, 99, 192, 189, 236, 256]. Due to their local behaviour, they are well-suited for parallel architectures [192, 209]. Nevertheless, they suffer from the fact that fairly small time step sizes are needed in order to ensure stability. Semi-implicit schemes (which approximate the diffusivity or the diffusion tensor in an explicit way and the rest implicitly) are considered by [58, 246]. They possess much better stability properties. Further speed-up can be achieved by using splitting techniques [246]. A fast multigrid technique using a pyramid algorithm for the Perona–Malik filter has been studied by Acton et al. ([3], see also [209] for related ideas).

1.2.6 Applications

Nonlinear diffusion filters have been applied for postprocessing fluctuating data [159, 250] and for visualizing quality-relevant features in computer aided quality control [248, 250, 252]. They have proved to be useful for texture segmentation [257, 260, 261], for image sequence analysis [221, 222], and for edge detection [4] and segmentation [3] of remotely sensed images. Most applications, however, are concerned with the filtering of medical images [99, 23, 65, 66, 67, 189, 258, 157, 236, 250, 260]. Some of these applications will be investigated in more detail in Chapter 5.

Besides such specific problem solutions, nonlinear diffusion filtering is currently integrated into commercially available software packages, for instance the image processing tool Heurisko.[4]

1.3 Methods of diffusion–reaction type

This section investigates variational frameworks, in which diffusion–reaction equations or coupled systems of them are interpreted as steepest descent minimizers of suitable energy functionals. This idea connects diffusion methods to edge detection and segmentation ideas.[5]

Besides the variational interpretation there exist other theoretical frameworks for diffusion filters such as the Markov random field and mean field annealing context [96, 235], and deterministic interactive particle models [166]. However, their discussion would lead us beyond the scope of this work.

1.3.1 Single diffusion–reaction equations

Nordström [182, 183] has suggested to obtain a reconstruction u of a degraded image f by minimizing the energy functional

$$E_f(u,w) := \int_\Omega \Big(\beta\cdot(u-f)^2 \;+\; w\cdot|\nabla u|^2 \;+\; \lambda^2\cdot(w-\ln w)\Big)\, dx. \qquad (1.41)$$

The parameters β and λ are positive weights and $w : \Omega \to [0,1]$ gives a fuzzy edge representation: in the interior of a region, w approaches 1 while at edges, w is close to 0 (as we shall see below).

The first summand of E punishes deviations of u from f *(deviation cost)*, the second term detects unsmoothness of u within each region *(stabilizing cost)*, and

[4] Heurisko is a registered trademark of AEON Verlag & Studio, Fraunhoferstr. 51B, 63454 Hanau, Germany.

[5] Diffusion–reaction methods with constant diffusivities can also be used for local contrast normalization in images, see [194].

the last one measures the extend of edges *(edge cost)*. Cost terms of these three types are typical for variational image restoration methods.

The corresponding Euler equations to this energy functional are given by

$$0 = \beta\cdot(u-f) - \operatorname{div}(w\nabla u), \tag{1.42}$$

$$0 = \lambda^2\cdot(1-\tfrac{1}{w}) + |\nabla u|^2, \tag{1.43}$$

equipped with a homogeneous Neumann boundary condition for u.

Solving (1.43) for w gives

$$w = \frac{1}{1+|\nabla u|^2/\lambda^2}. \tag{1.44}$$

We recognize that w is identical with the Perona–Malik diffusivity $g(|\nabla u|^2)$ introduced in (1.21). Therefore, (1.42) can be regarded as the steady-state equation of

$$\partial_t u = \operatorname{div}(g(|\nabla u|^2)\,\nabla u) + \beta(f-u). \tag{1.45}$$

This equation can also be obtained directly as the descent method of the functional

$$F_f(u) := \int_\Omega \left(\beta\cdot(u-f)^2 + \lambda^2\cdot\ln\left(1+\tfrac{|\nabla u|^2}{\lambda^2}\right)\right) dx. \tag{1.46}$$

The diffusion–reaction equation (1.45) consists of the Perona–Malik process with an additional bias term $\beta\cdot(f-u)$. Nordström claims that this modification frees the user from the difficulty of specifying an appropriate stopping time for the Perona–Malik process. Moreover, by regarding it as a method for minimizing the energy functional (1.41), he argues that this gives nonlinear diffusion filtering an appealing sense of optimality and establishes it as a method for solving a well-defined mathematical problem.

However, it is obvious that the Nordström model just shifts the problem of specifying a stopping time T to the problem of determining β (although this question was circumvented by setting $\beta := 1$).

Besides the fact that this method suffers from the same ill-posedness problems as the underlying Perona–Malik equation, it is easily seen that the energy functional (1.46) is nonconvex. Therefore, it may possess numerous local minima, and the process (1.45) with f as initial condition does not necessarily converge to a global minimum.

As a remedy, Schnörr [220] renounces edge-enhancing diffusivities in order to end up with (nonquadratic) convex functionals. In this case the classical theory is applicable ensuring well-posedness results and stability of a standard finite-element approximation.

For other diffusion–reaction methods such as [96], the convergence problem has usually not been addressed. One exception is the paper of Cottet and Germain [66]. They show that their filter allows the possibility of having almost steady piecewise constant states and prove their attractivity. The stability of these solutions is indicated by showing that piecewise constant states are almost steady-state solutions of a semidiscretized version where the diffusion tensor is constant with respect to time.

1.3.2 Coupled sytems of diffusion–reaction equations

Mumford and Shah [172, 173] have proposed to obtain a segmented image u from f by minimizing the functional

$$E_f(u,K) = \beta \int_{\Omega} (u-f)^2 \, dx \; + \int_{\Omega \setminus K} |\nabla u|^2 \, dx \; + \; \alpha |K| \tag{1.47}$$

with nonnegative parameters α and β. The discontinuity set K consists of the edges, and its one-dimensional Hausdorff measure $|K|$ gives the total edge length. Like the Nordström functional (1.41), this expression consists of three cost terms: the first one is the deviation cost, the second one gives the stabilizing cost, and the third one represents the edge cost.

The Mumford–Shah functional can be regarded as a continuous version of the Markov random field method by Geman and Geman [98] and the weak membrane model of Blake and Zisserman [32]. Related approaches are also used to model materials with two phases and a free interface (cf. [70]).

The fact that (1.47) leads to a free discontinuity problem causes many challenging theoretical questions [145]. The book of Morel and Solimini [171] covers a very detailed analysis of this functional. Although the existence of a global minimizer with a closed edge set K has been established [70, 17], uniqueness is in general not true [160, pp. 197–198]. Regularity results for K have recently been obtained [37].

The concept of energy functionals for segmenting images offers the practical advantage that it provides a framework for comparing the quality of two segmentations. On the other hand, (1.47) exhibits also some shortcomings, e.g. the problem that flat edges produce multiple segmentation boundaries *(oversegmentation)* [229]. Another drawback results from the fact that the Mumford–Shah functional allows only singularities which are typical for minimal surfaces: Corners or T-junctions are not possible and segments meet at triple points with 120^o angle [173]. In order to avoid such problems, modifications of the Mumford–Shah functional have been proposed by Ambrosio and Mantegazza (cf. [205]) and Shah [231]. An affine invariant version of (1.47) is investigated in [26, 25].

Since many algorithms in image processing can be restated as versions of the Mumford–Shah functional [171] and since it is a prototype of a free discontinuity problem it is interesting to study this variational problem in more detail.

Numerical complications arise from the fact that the Mumford–Shah functional has numerous local minima. Global minimizers such as the simulated annealing method used by Geman and Geman [98] are extremely slow. Hence, one searches for fast (suboptimal) deterministic strategies, e.g. pyramidal region-growing algorithms [2, 142].

Another interesting class of numerical methods is based on the idea to approximate the discontinuity set K by a smooth function w, which is close to 0 near edges of u and which approximates 1 elsewhere.

We may for instance study the functional

$$F_f(u,w) := \int\limits_{\Omega} \Big(\beta\cdot(u-f)^2 \,+\, w^2\cdot|\nabla u|^2 \,+\, \alpha\cdot\Big(c|\nabla u|^2 + \tfrac{(1-w)^2}{4c}\Big)\Big)\, dx \tag{1.48}$$

with a positive parameter c specifying the "edge width". Ambrosio and Tortorelli proved that this functional converges to the Mumford–Shah functional for $c \to 0$ (in the sense of Γ-convergence, see [19] for more details).

Minimizing F_f corresponds to the gradient descent equations

$$\partial_t u \;=\; \operatorname{div}(w^2 \nabla u) \,+\, \beta\cdot(f-u), \tag{1.49}$$

$$\partial_t w \;=\; c\Delta w \,-\, \tfrac{2w}{\alpha}|\nabla u|^2 \,-\, \tfrac{(1-w)}{2c} \tag{1.50}$$

with homogeneous Neumann boundary conditions. Equations of this type are investigated by Richardson and Mitter [205]. Since (1.49) is very similar to the Nordström process (1.45), similar problems arise: The functional F_f is not jointly convex in u and v, so it may have many local minima and a gradient descent algorithm may get trapped in a poor local minimum. Well-posedness results for this system have not been obtained up to now, but a maximum–minimum principle and a local stability proof have been established.

Another diffusion–reaction system is studied by Shah [227, 228]. He replaces the functional (1.47) by two coupled convex energy functionals and applies gradient descent. This results in finding an equilibrium state between two competing processes. Experiments indicate that it converges to a stable solution. Proesmans et al. [200, 199] observed that this solution looks fairly blurred since the equations contain diffusion terms such as Δu. They obtained pronounced edges by replacing such a term by its Perona–Malik counterpart $\operatorname{div}(g(|\nabla u|^2)\,\nabla u)$. Of course, this approach gives rise to the same theoretical questions as (1.49), (1.50).

The system of Richardson and Mitter is used for edge detection [205]. Shah investigates diffusion–reaction systems for matching stereo images [230, 232], while Proesmans et al. apply coupled diffusion-reaction equations to image sequence analysis, vector-valued images and stereo vision [199, 201]. Their finite difference algorithms run on a parallel transputer network.

It should also be mentioned that there exist reaction–diffusion systems which have been applied to image restoration [197, 198] or texture generation [240, 264] and which are not connected to the Perona–Malik or Mumford–Shah ideas.

1.4 Classic morphological processes

Morphology is an approach to image analysis based on shapes. Its mathematical formalization goes back to the group around Matheron and Serra, both working at ENS des Mines de Paris in Fontainebleau. The theory had first been developed for binary images, afterwards it was extended to grey tone images by regarding level sets as shapes. Its applications cover biology, medical diagnostics, histology, quality control, radar and remote sensing, science of material, mineralogy, and many others.

Morphology is usually described in terms of algebraic set theory, see e.g. [167, 225, 118, 119] for an overview. Nevertheless, recently PDE formulations for classic morphological processes have been discovered by Brockett and Maragos [41, 42], van den Boomgaard [34], Arehart et al. [22] and Alvarez et al. [11].

This section surveys the basic ideas and elementary operations of binary and grey-scale morphology, presents its PDE representations for images and curves, and summarizes the results concerning well-posedness and scale-space properties. It ends up with discussing numerical aspects of the PDE formulation of these processes.

1.4.1 Binary and grey-scale morphology

Binary morphology considers *shapes (silhouettes)*, i.e. closed sets $X \subset \mathbb{R}^2$ whose boundaries are Jordan curves [15]. Henceforth, we identify a shape X with its characteristic function

$$\chi(x) := \begin{cases} 1 & \text{if } x \in X, \\ 0 & \text{else.} \end{cases} \tag{1.51}$$

Binary morphological operations affect only the boundary curve of the shape and, therefore, they can be viewed as curve or shape deformations.

Grey-scale morphology generalizes these ideas [163] by decomposing an image f into its level sets $\{X_a f,\ a \in \mathbb{R}\}$, where

$$X_a f := \{x \in \mathbb{R}^2,\ f(x) \geq a\}. \tag{1.52}$$

A binary morphological operation A can be extended to some grey scale image f by defining

$$X_a(Af) := A(X_a f) \quad \forall\, a \in \mathbb{R}. \tag{1.53}$$

We observe that for morphological operations only grey-level sets matter. As a consequence, they are invariant under monotone grey-level rescalings. This is a very desirable property in all cases where brightness changes of the illumination occur or where one wants to be independent of the specific contrast range of the camera. On the other hand, for applications like edge detection or image restoration, contrast provides important informations which should be taken into account. Moreover, in some cases isolines may give inadequate information about the depicted physical object boundaries.

1.4.2 Basic operations

In order to analyse a shape, classic morphology uses a so-called *structuring element*, a bounded set $B \subset \mathbb{R}^2$. Typical shapes for B are discs, squares, or ellipses.

The two basic morphological operations, *dilation* and *erosion* with a structuring element B, are defined for a grey-scale image $f \in L^\infty(\mathbb{R}^2)$ by [42]

$$\text{dilation:} \qquad (f \oplus B)\,(x) \;:=\; \sup\,\{f(x-y),\ y \in B\}, \tag{1.54}$$
$$\text{erosion:} \qquad (f \ominus B)\,(x) \;:=\; \inf\,\{f(x+y),\ y \in B\}. \tag{1.55}$$

These names can be easily motivated when considering a shape in a binary image and a disc structuring element. In this case dilation blows up its boundaries, while erosion shrinks them.

Dilation and erosion form the basis for constructing other morphological processes, for instance *opening* and *closing:*

$$\text{opening:} \qquad (f \circ B)\,(x) \;:=\; ((f \ominus B) \oplus B)\,(x), \tag{1.56}$$
$$\text{closing:} \qquad (f \bullet B)\,(x) \;:=\; ((f \oplus B) \ominus B)\,(x). \tag{1.57}$$

In the preceding shape interpretation opening smoothes the shape by breaking narrow isthmuses and eliminating small islands, while closing smoothes by eliminating small holes, fusing narrow breaks and filling gaps on the contours [118].

1.4.3 Continuous-scale morphology

Let us consider a convex structuring element tB with a scaling parameter $t > 0$. Then, calculating $u(t) = f \oplus tB$ and $u(t) = f \ominus tB$, respectively[6], can be shown to be equivalent to solving

$$\partial_t u(x,t) \;=\; \sup_{y \in B}\, \langle y, \nabla u(x,t)\rangle, \tag{1.58}$$
$$\partial_t u(x,t) \;=\; \inf_{y \in B}\, \langle y, \nabla u(x,t)\rangle. \tag{1.59}$$

[6]Henceforth, we frequently use the simplified notation $u(t)$ instead of $u(.,t)$

with f as initial condition [11, 213].

By choosing e.g. $B := \{y \in \mathbb{R}^2,\ |y| \leq 1\}$ one obtains

$$\begin{aligned} \partial_t u &= |\nabla u|, && (1.60)\\ \partial_t u &= -|\nabla u|. && (1.61) \end{aligned}$$

The solution $u(t)$ is the dilation (resp. erosion) of f with a disc of radius t and centre 0 as structuring element.

Connection to curve evolution

Morphological PDEs such as (1.60) or (1.61) are closely related to shape and curve evolutions.[7] To illustrate this, let us consider a smooth Jordan curve $C : [0, 2\pi] \times [0, \infty) \to \mathbb{R}^2$,

$$C(p,t) = \begin{pmatrix} x_1(p,t) \\ x_2(p,t) \end{pmatrix} \tag{1.62}$$

where p is the parametrization and t is the evolution parameter. We assume that C evolves in outer normal direction n with speed β, which may be a function of its curvature $\kappa := \frac{\det(C_p, C_{pp})}{|C_p|^3}$:

$$\begin{aligned} \partial_t C &= \beta(\kappa) \cdot n, && (1.63)\\ C(p,0) &= C_0(p). && (1.64) \end{aligned}$$

One can embed the curve $C(p,t)$ in an image $u(x,t)$ in such a way that C is just a level curve of u. The corresponding evolution for u is given by [188, 214, 15]

$$\partial_t u = \beta(\mathrm{curv}(u)) \cdot |\nabla u|. \tag{1.65}$$

where the curvature of u is

$$\mathrm{curv}(u) := \mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right). \tag{1.66}$$

Sethian [226] calls (1.65) the *Eulerian formulation* of the curve evolution (1.63), because it is written in terms of a fixed coordinate system.

We observe that (1.60) and (1.61) correspond to the simple cases $\beta = \pm 1$. Hence, they describe the curve evolutions

$$\partial_t C = \pm\, n. \tag{1.67}$$

This equation moves level sets in normal direction with constant speed. Such a process is also named *grassfire flow* or *prairie flow*. It is closely related to the Huygens principle for wave propagation [22].

[7] Besides this application, curve evolution approaches play an important role in numerous other image processing problems ranging from shape-from-shading to skeletons, see [43] for an overview.

1.4.4 Theoretical results

Equations such as (1.67) may develop singularities and intersections even for smooth initial data. Hence, concepts of jump conditions, entropy solutions, and shocks have to be applied to this shape evolution [138].

A suitable framework for the image evolution equation (1.65) is provided by the theory of *viscosity solutions* [68]. The advantage of this analysis is that it allows us to treat shapes with singularities such as corners, where the classical solution concept does not apply, but a unique weak solution in the viscosity sense still exists.

It can be shown [62, 83, 68], that for an initial value

$$f \in \mathrm{BUC}(\mathbb{R}^2) := \{\varphi \in L^\infty(\mathbb{R}^2) \mid \varphi \text{ is uniformly continuous on } \mathbb{R}^2\} \qquad (1.68)$$

the equations (1.58),(1.59) possess a unique global viscosity solution $u(x,t)$ which fulfils the maximum–minimum principle

$$\inf_{\mathbb{R}^2} f \;\le\; u(x,t) \;\le\; \sup_{\mathbb{R}^2} f \quad \text{on} \quad \mathbb{R}^2 \times [0,\infty). \qquad (1.69)$$

Moreover, it is L^∞-stable: for two different initial images f, g the corresponding solutions $u(t)$, $v(t)$ satisfy

$$\|u(t)-v(t)\|_{L^\infty(\mathbb{R}^2)} \;\le\; \|f-g\|_{L^\infty(\mathbb{R}^2)}. \qquad (1.70)$$

1.4.5 Scale-space properties

Brockett and Maragos [41] pointed out that the convexity of B is sufficient to ensure the semigroup property of the corresponding dilations and erosions. This establishes an important architectural scale-space property.

Similar results have been found by van den Boomgaard and Smeulders [35, 36]. Moreover, they conjecture a causality property where singularities play a similar role as zero-crossings in the Gaussian scale-space.

Jackway et al. [131, 132] prove a causality theorem for the dilation–erosion scale-space, which is also based on local extrema instead of zero-crossings. They establish that under erosion the number of local minima is decreasing, while dilation reduces the number of local maxima.

A complete scale-space interpretation is due to Alvarez, Guichard, Lions and Morel [11]: They prove that under three architectural assumptions (semigroup property, locality and regularity), one smoothing axiom (monotony) and additional invariance requirements (grey-level shift invariance, Euclidean invariance, morphological invariance), a two-dimensional scale-space equation has the following form:

$$\partial_t u = |\nabla u| \, F(t, \mathrm{curv}(u)) \qquad (1.71)$$

Clearly, dilation and erosion belong to the class (1.71), thus being good candidates for morphological scale-spaces. Indeed, in [11] it is shown that the converse is true as well: all axioms that lead to (1.71) are fulfilled.[8]

1.4.6 Numerical aspects

Morphological schemes for dilation or erosion which are based on *curve* evolution turn out be be difficult to handle: they require prohibitive small time steps, and suffer from the problem of coping with singularities and topological changes [188, 22, 213].

For this reason it is useful to discretize the corresponding *image* evolution equations. The *PSC schemes* [9] of Osher and Sethian [188] are based on the idea to derive numerical methods for such equations from techniques for hyperbolic conservation laws.

To illustrate this with a simple example, let us restrict ourselves to the one-dimensional dilation equation $\partial_t u = |\partial_x u|$. A first-order upwind PSC scheme for this process is given by

$$\frac{u_i^{n+1} - u_i^n}{\tau} = \sqrt{\left(\min\left(\frac{u_i^n - u_{i-1}^n}{h}, 0\right)\right)^2 + \left(\max\left(\frac{u_{i+1}^n - u_i^n}{h}, 0\right)\right)^2}, \qquad (1.72)$$

where h is the pixel size, τ is the time step size, and u_i^n denotes a discrete approximation of $u(ih, n\tau)$.

Such PDE-based morphological schemes possess two advantages over classical set-theoretic schemes for dilation/erosion [22, 213]:

(a) They give excellent results for non-digitally scalable structuring elements whose shapes cannot be represented correctly on a discrete grid, for instance discs or ellipses.

(b) The time t plays the role of a continuous scale parameter. Therefore, the size of a structuring element need not to be multiples of the pixel size, and it is possible to get results with sub-pixel accuracy.

However, they reveal also two disadvantages:

(a) They are slower than set-theoretic morphological schemes.

(b) Dissipative effects such as blurring of discontinuities occur.

To address the first problem, speed-up techniques for shape evolution have been proposed [6], which use only points close to the curve at every time step. Blurring of discontinuities can be minimized by applying shock-capturing techniques such as high-order ENO schemes [237, 234].

[8] Euclidean invariance is only satisfied for a disc centered in 0 as structuring element.

[9] PSC means "propagation of surfaces under curvature".

1.5 Curvature-based morphological processes

Besides providing a useful reinterpretation of classic continuous-scale morphology, the PDE approach has led to the discovery of new morphological operators. These processes are curvature-based, and – although they cannot be written in conservation form – they reveal interesting relations to diffusion processes. Two important representatives of this class are curvature motion and the so-called fundamental equation in image processing. In this subsection we shall discuss these PDEs, possible generalizations, numerical aspects, and applications.

1.5.1 Mean-curvature filtering

In order to motivate our first curvature-based morphological PDE, let us recall that the linear diffusion equation (1.9) can be rewritten as

$$\partial_t u = \partial_{\eta\eta} u + \partial_{\xi\xi} u, \tag{1.73}$$

where the gauge coordinates η and ξ point in the direction parallel and perpendicular to ∇u, respectively. The first term on the right-hand side of (1.73) describes smoothing along the flowlines, while the second one smoothes along isophotes. When we want to smooth the image anisotropically along its isophotes, we can neglect the first term and end up with the problem

$$\begin{aligned} \partial_t u &= \partial_{\xi\xi} u, && (1.74)\\ u(x,0) &= f(x). && (1.75) \end{aligned}$$

By straightforward calculations one verifies that (1.74) can also be written as

$$\begin{aligned} \partial_t u &= \frac{u_{x_2}^2 u_{x_1x_1} - 2u_{x_1}u_{x_2}u_{x_1x_2} + u_{x_1}^2 u_{x_2x_2}}{u_{x_1}^2 + u_{x_2}^2} && (1.76)\\ &= \Delta u - \frac{1}{|\nabla u|^2}\langle \nabla u, \mathrm{Hess}(u)\nabla u\rangle && (1.77)\\ &= |\nabla u|\,\mathrm{curv}(u). && (1.78) \end{aligned}$$

Processes of this type are of importance for flame propagation and crystal growth, see [188] and the references therein. Since $\mathrm{curv}(u) = \mathrm{div}\left(\frac{\nabla u}{|\nabla u|}\right)$ is the curvature of u (mean curvature for dimensions ≥ 3), equation (1.78) is named *(mean-) curvature motion (MCM).* The corresponding curve evolution

$$\partial_t C(p,t) = \kappa(p,t) \cdot n(p,t) \tag{1.79}$$

shows that (1.74) propagates isophotes in inner normal direction with a velocity that is given by their curvature $\kappa = \frac{\det(C_p, C_{pp})}{|C_p|^3}$. It is called *geometric heat equation* or *Euclidean shortening flow.* The subsequent discussions shall clarify these names.

Intrinsic heat flow

Interestingly, there exists a further connection between linear diffusion and motion by curvature. Let $v(p,t)$ denote the *Euclidean arc-length* of $C(p,t)$, i.e.

$$v(p,t) := \int_0^p |C_\rho(\rho,t)|\, d\rho, \tag{1.80}$$

where $C_\rho := \partial_\rho C$. The Euclidean arc-length is characterized by $|C_v| = 1$. It is invariant under *Euclidean transformations*, i.e. mappings

$$x \;\rightarrow\; Rx + b \tag{1.81}$$

where $R \in \mathbb{R}^{2\times 2}$ denotes a rotation matrix and $b \in \mathbb{R}^2$ is a translation vector. Since it is well-known from differential geometry (see e.g. [50], p. 14) that

$$\kappa(p,t) \cdot n(p,t) = \partial_{vv} C(p,t), \tag{1.82}$$

we recognize that curvature motion can be regarded as Euclidean invariant diffusion of isophotes:

$$\partial_t C(p,t) = \partial_{vv} C(p,t). \tag{1.83}$$

This *geometric heat equation* is intrinsic, as it is independent of the curve parametrization. However, the reader should be aware of the fact that – although this equation looks like a linear one-dimensional heat equation – it is in fact nonlinear, since the arc-length v is again a function of the curve.

Theoretical results

For the evolution of a smooth curve under its curvature, it has been shown in [123, 95, 110] that a smooth solution exists for some finite time interval $[0, T)$. A convex curve remains convex, a nonconvex one becomes convex and, for $t \rightarrow T$, the curve shrinks to a *circular point*, i.e. a point with a circle as limiting shape. Moreover, since under all flows $C_t = C_{qq}$ the Euclidean arc-length parametrization $q(p) := v(p)$ is the fastest way to shrink the Euclidean perimeter $\oint |C_p|\, dp$, equation (1.83) is called *Euclidean shortening flow* [94]. The time for shrinking a circle of radius σ to a point is given by

$$T = \tfrac{1}{2}\,\sigma^2. \tag{1.84}$$

In analogy to the dilation/erosion case it can be shown that, for an initial image $f \in \mathrm{BUC}(\mathbb{R}^2)$, equation (1.74) has a unique viscosity solution which is L^∞-stable and satisfies a maximum–minimum principle [11].

Scale-space interpretation

A shape scale-space interpretation for curve evolution under Euclidean heat flow is studied by Kimia and Siddiqi [137]. It is based on results of Evans and Spruck [83]. They establish the semigroup property as architectural quality, and smoothing properties follow from the fact that the total curvature decreases. Moreover, the number of extrema and inflection points of the curvature is nonincreasing.

As an image evolution, MCM belongs to the class of morphological scale-spaces which satisfy the general axioms of Alvarez, Guichard, Lions and Morel [11], that we have mentioned in 1.4.5.

When studying the evolution of isophotes under MCM, it can be shown that, if one isophote is enclosed by another, this ordering is preserved [83, 137]. Such a *shape inclusion principle* implies in connection with (1.84) that it takes the time $T = \frac{1}{2}\sigma^2$ to remove all isophotes within a circle of radius σ. This shows that the relation between temporal and spatial scale for MCM is the same as for linear diffusion filtering (cf. (1.14)).

Moreover, two level sets cannot move closer to one another than they were initially [83, 137]. Hence, contrast cannot be enhanced. This property is characteristic for all scale-spaces of the Alvarez–Guichard–Lions–Morel axiomatic and distinguishes them from nonlinear diffusion filters.

1.5.2 Affine invariant filtering

Motivation

Although Euclidean invariant smoothing methods are sufficient in many applications, there exist certain problems which also require invariance with respect to affine transformations. A *(full) affine transformation* is a mapping

$$x \;\to\; Ax + b \tag{1.85}$$

where $b \in \mathbb{R}^2$ denotes a translation vector and the matrix $A \in \mathbb{R}^{2\times 2}$ is invertible. Affine transformations arise as shape distortions of planar objects when being observed from a large distance under different angles.

Affine invariant intrinsic diffusion

In analogy to the Euclidean invariant heat flow, Sapiro and Tannenbaum [214, 215] constructed an affine invariant flow by replacing the Euclidean arc-length $v(p,t)$ in (1.83) by an "arc-length" $s(p,t)$ that is invariant with respect to affine transformations with $\det(A) = 1$ (and relative invariant to full affine transformations).

Such an affine arc-length was proposed by Blaschke [33, pp. 12–15] in 1923. It is characterized by $\det(C_s, C_{ss}) = 1$, and it can be calculated as

$$s(p,t) \;:=\; \int_0^p \Big(\det \Big(C_\rho(\rho,t),\, C_{\rho\rho}(\rho,t) \Big) \Big)^{\frac{1}{3}} \, d\rho. \tag{1.86}$$

By virtue of

$$\partial_{ss} C(p,t) = \Big(\kappa(p,t) \Big)^{\frac{1}{3}} \cdot n(p,t) \tag{1.87}$$

we obtain the *affine invariant heat flow*

$$\begin{aligned} \partial_t C(p,t) &= \Big(\kappa(p,t) \Big)^{\frac{1}{3}} \cdot n(p,t), && (1.88) \\ C(p,0) &= C_0(p). && (1.89) \end{aligned}$$

Affine invariant image evolution

When regarding the curve $C(p,t)$ as a level-line of an image $u(x,t)$, we end up with the evolution equation

$$\begin{aligned} \partial_t u &= |\nabla u| \Big(\mathrm{curv}(u) \Big)^{\frac{1}{3}} && (1.90) \\ &= \Big(u_{x_2}^2 u_{x_1 x_1} - 2 u_{x_1} u_{x_2} u_{x_1 x_2} + u_{x_1}^2 u_{x_2 x_2} \Big)^{\frac{1}{3}} && (1.91) \\ &= |\nabla u|^{\frac{2}{3}} \, u_{\xi\xi}^{\frac{1}{3}}, && (1.92) \end{aligned}$$

where ξ is the direction perpendicular to ∇u.

Besides the name affine invariant heat flow, this equation is also called *affine shortening flow*, *affine morphological scale-space (AMSS)*, and *fundamental equation in image processing*.

This image evolution equation has been discovered independently and simultaneously to the curve evolution approach of Sapiro and Tannenbaum by Alvarez, Guichard, Lions and Morel [9, 11] via an axiomatic scale-space approach. After having mentioned some theoretical results, we shall briefly sketch this reasoning below.

Theoretical results

The curve evolution properties of affine invariant heat flow can be shown to be the same as in the Euclidean invariant case, with three exceptions [215]:

(a) Closed curves shrink to points with an ellipse as limiting shape (*elliptical points*).

(b) The name *affine shortening flow* reflects the fact that, under all flows $C_t = C_{qq}$, the affine arc-length parametrization $q(p) := s(p)$ is the fastest way to shrink the affine perimeter

$$L(t) \;:=\; \oint \Big(\det \big(C_p(p,t),\, C_{pp}(p,t) \big) \Big)^{\frac{1}{3}} \, dp. \tag{1.93}$$

(c) The time for shrinking a circle of radius σ to a point is

$$T = \tfrac{3}{4}\, \sigma^{\frac{4}{3}}. \tag{1.94}$$

For the image evolution equation (1.90) we have the same results as for MCM and dilation/erosion concerning well-posedness of a viscosity solution which satisfies a maximum–minimum principle [11].

Scale-space properties

Alvarez, Guichard, Lions and Morel [9, 11] proved that (1.90) is unique (up to temporal rescalings) when imposing on the scale-space axioms for (1.71) an additional *affine invariance axiom*:

For every invertible $A \in \mathbb{R}^{2\times 2}$ *and for all* $t \geq 0$, *there exists a rescaled time* $t'(t, A) \geq 0$, *such that*

$$T_t(Af) = A(T_{t'} f) \quad \forall\, f \in \mathrm{BUC}(\mathbb{R}^2). \tag{1.95}$$

For this reason they call the AMSS equation also *fundamental equation in image analysis*. Simplifications of this axiomatic and related axioms for shape scale-spaces can be found in [15, 16].

The scale-space reasoning of Sapiro and Tannenbaum investigates properties of the curve evolution, see [214] and the references therein. Based on results of [139, 21] they point out that the Euclidean absolute curvature decreases as well as the number of extrema and inflection points of curvature. Moreover, a shape inclusion principle holds.

1.5.3 Generalizations

In order to analyse planar shapes in a way that does not depend on their location in $\mathbb{R}^3$, one requires a multiscale analysis which is invariant under a general projective mapping

$$(x_1, x_2)^\top \;\to\; \left(\frac{a_{11}x_1 + a_{21}x_2 + a_{31}}{a_{13}x_1 + a_{23}x_2 + a_{33}},\, \frac{a_{12}x_1 + a_{22}x_2 + a_{32}}{a_{13}x_1 + a_{23}x_2 + a_{33}} \right)^\top \tag{1.96}$$

with $A = (a_{ij}) \in \mathbb{R}^{3\times 3}$ and $\det A = 1$. Research in this direction has been carried out by Faugeras [84, 85], Bruckstein and Shaked [45], Olver et al. [184, 185], and Dibos [72]. It turns out that intrinsic heat-equation-like formulations for the projective group are more complicated than the Euclidean and affine invariant ones, and that there is some evidence that they do not reveal the same smoothing scale-space properties [184, 185]. A general study of intrinsic heat flows which are invariant under subgroups of the projective group can be found in [184, 185].

The Euclidean and affine invariant curve evolution may also be modified in order to obtain area- or length-preserving equations [94, 216, 219]. By adopting a Nordström-like modification (cf. (1.45)) to curve evolution under Euclidean shortening flow, one can avoid the drawback that corners get rounded and curves shrink to points [87]. Since these methods work well for curve evolution, it would be interesting to study whether they are practicable and useful for the evolution of images.

Generalizations of MCM or AMMS for 3D images are investigated in [54, 175, 185]. The crucial problem in this case is to combine the principal curvatures in such a way that the resulting process simplifies arbitrary surfaces by deforming them (e.g. into spheres) without creating singularities.

Affine scale-space axiomatics for image sequences (movies) have been established in [10, 11, 169], while generalizations of the Alvarez–Guichard–Lions–Morel axiomatic to scale-spaces for colour images are a topic of current research [59].

1.5.4 Numerical aspects

Due to the equivalence between curve evolution and morphological image processing PDEs, we have two main classes of numerical methods: curve or shape evolution methods, and approximation methods for the Eulerian formulation. A comparison of different methods of both classes can be found in [64].

When applying curve or shape evolution schemes to images, one separates the image into a (finite) number of level curves and evolves each curve separately. Afterwards, the results have to be superimposed to obtain the evolved image.

Curve evolution schemes are investigated by Mokhtarian and Mackworth [170], Bruckstein et al. [44], and Cohignac et al. [64]. In [44] discrete analogues of MCM and AMSS for the evolution of planar polygons are introduced. In complete analogy to the behaviour of the continuous equations, convergence to polygones, whose corners belong to circles and ellipses, respectively, is established. Convergent set-theoretic morphological schemes for MCM and AMSS have been proposed by Catté et al. [57, 56].

Although curve evolution schemes have an excellent behaviour with respect to morphological invariance (and affine invariance in the case of AMSS), they have mainly been used for the evolution of a single curve. Their application to images is less efficient because of the high computational cost to treat each level curve separately [15].

This high effort for evolving images can be circumvented by approximating directly the image evolution equations.

Most discretizations of morphological image evolution PDEs are based on the PSC schemes of Osher and Sethian [188]. In the case of MCM or AMSS, this leads to an explicit finite difference method which approximates the spatial derivatives by central differences. Such a scheme is used by Sapiro and Tannenbaum [214], while Cohignac et al. [64], and Alvarez and Morales [14] modify the approximation of the spatial derivatives in order to get better rotational invariance and stability properties. Niessen et al. [177, 178] approximate the spatial derivatives by Gaussian derivatives which are calculated in the Fourier domain.

Concerning stability one observes that the explicit PSC-like approaches do not preserve a discrete maximum–minimum principle and require small time steps to be experimentally stable. For the AMSS an additional constraint appears: the behaviour of this equation is highly nonlocal, since the affine invariance implies that circles are equivalent to ellipses of any arbitrary eccentricity. If one wants to have a good numerical approximation of affine invariance one has to decrease the time step size significantly below the step size of experimental stability [15].

One way to achieve unconditional L^∞-stability for MCM is to replace $u_{\xi\xi}$ by a linear combination of fixed smoothing directions and to apply a semi-implicit finite difference scheme (cf. [64, 12]). However, such approximations renounce consistency with the original equation as well as rotational invariance (round shapes evolve into polygonal structures).

1.5.5 Applications

The special invariances of AMSS are useful for shape recognition tasks [11, 86], for corner detection [14], and for texture discrimination [158, 15]. MCM and AMSS have also been applied to denoising [177, 178] and blob detection [100] in medical images.

If one aims to use these equations for image restoration one usually modifies them by multiplying them with a term that reduces smoothing at edges with large gradients [12, 217, 218]. Another modification results from omitting the factor $|\nabla u|$ in the mean curvature motion (1.78), see e.g. [81, 82]. This corresponds to nonlinear diffusion filters and a restoration method by total variation minimization [207] which we shall describe in 1.6.2. Related models to MCM also appear in the context of active contour models [52, 161, 160, 162, 259, 53, 55]. For image restoration tasks, MCM or AMSS are frequently combined with other processes such as linear diffusion [12], shock filtering ([13], cf. 1.6.1) or global PDEs for histogramme enhancement [212].

All these preceding modifications are at the expense of renouncing the morphological invariance of the genuine operators (and also affine invariance in the case of [217, 218], unless an "affine invariant gradient" is used). If one wants to remain within the morphological framework one can combine different morphological processes, for instance MCM and dilation/erosion. This so-called *entropy scale-space* is useful for analysing components of shape [138, 140, 233, 238, 269].

It is interesting to note that in 1965 Gabor (the inventor of optical holography and the so-called Gabor functions) proposed a deblurring algorithm based on solving MCM backwards in time [93, 154], a long-time forgotten method.

1.6 Total variation methods

Inspired by observations from fluid dynamics where the *total variation (TV)*

$$TV(u) := \int_\Omega |\nabla u| \, dx \tag{1.97}$$

plays an important role for shock calculations, one may ask if it is possible to apply related ideas to image processing. This would be useful to restore discontinuities such as edges.

Below we shall focus on two important TV-based image restoration techniques which have been pioneered by Osher and Rudin: TV-preserving methods and techniques which are TV-minimizing subject to certain constraints.[10]

1.6.1 TV-preserving methods

In 1990, Osher and Rudin have proposed to restore blurred images by *shock filtering* [186]. These filters calculate the restored image as the steady-state solution of the problem

$$\partial_t u = -|\nabla u| \, F(\mathcal{L}(u)), \tag{1.98}$$

$$u(x,0) = f(x). \tag{1.99}$$

Here, $\operatorname{sgn}(F(u)) = \operatorname{sgn}(u)$, and $\mathcal{L}(u)$ is a second-order elliptic operator whose zero-crossings correspond to edges, e.g. the Laplacian $\mathcal{L}(u) = \Delta u$ or the second-order directional derivative $\mathcal{L}(u) = u_{\eta\eta}$ with $\eta \,\|\, \nabla u$.

By means of our knowledge from morphological processes, we recognize that this filter aims to produce a flow field that is directed from the interior of a region

[10] Another image enhancement method that is close in spirit is due to Eidelman, Grossmann and Friedman [80]. It maps the image grey values to gas dynamical parameters and solves the compressible Euler equations using shock-capturing total variation diminishing (TVD) techniques based on Godunov's method.

towards its edges where it develops shocks. Thus, the goal is to obtain a piecewise constant steady-state solution with discontinuities only at the edges of the initial image.

It has been shown that a one-dimensional version of this filter preserves the total variation and satisfies a maximum–minimum principle, both in the continuous and discrete case. For the two-dimensional case not many theoretical results are available except for a discrete maximum–minimum principle.

Another TV-preserving deblurring method of Osher and Rudin solves the linear diffusion equation backwards in time under the constraint that the total variation remains constant [187].

From a practical point of view, TV-preserving methods suffer from the problem that fluctuations due to noise do also create shocks. For this reason, Alvarez and Mazorra [13] replace the operator $\mathcal{L}(u) = u_{\eta\eta}$ in (1.98) by a regularized version $\mathcal{L}(K_\sigma * u) = K_\sigma * u_{\eta\eta}$ and supplement the resulting equation with a noise-eliminating mean-curvature process. They prove that their semi-implicit finite-difference scheme has a unique solution which satisfies a maximum–minimum principle.

1.6.2 TV-minimizing methods

Total variation is good for quantifying the simplicity of an image since it measures oscillations without unduly punishing discontinuities. For this reason, blocky images (consisting only of a few almost piecewise constant segments) reveal very small total variation.

In order to recover noisy blocky images, Rudin, Osher and Fatemi [207] have proposed to minimize the total variation under constraints which reflect assumptions about noise.[11]

To fix ideas, let us study an example. Given an image f with additive noise of zero mean and known variance σ^2, we seek a restoration u satisfying

$$\min_u \int_\Omega |\nabla u| \, dx \tag{1.100}$$

subject to

$$\frac{1}{2} \int_\Omega (u-f)^2 \, dx = \sigma^2, \tag{1.101}$$

$$\int_\Omega u \, dx = \int_\Omega f \, dx. \tag{1.102}$$

[11] Related ideas have also been developed by Geman and Reynolds [97].

In order to solve this constraint variational problem, PDE methods can be applied: A solution of (1.100)–(1.102) verifies necessarily the Euler equation

$$\operatorname{div}\left(\frac{\nabla u}{|\nabla u|}\right) - \mu - \lambda(u-f) = 0 \tag{1.103}$$

with homogeneous Neumann boundary conditions. The (unknown) Lagrange multipliers μ and λ have to be determined in such a way that the constraints are fulfilled. Interestingly, (1.103) looks similar to the steady state equation of the diffusion–reaction processes of Nordström [182, 183] and Schnörr [220], but – in contrast to TV approaches – these methods are not intended to satisfy the noise constraint exactly [208]. Moreover, the divergence term in (1.103) is identical with the curvature, which relates this technique to MCM.

In [207] a gradient descent method is proposed to solve (1.103). It uses an explicit finite difference scheme with central and one-sided spatial differences and adapts the Lagrange multiplier by means of the gradient projection method of Rosen. General existence and uniqueness results for constrained nonlinear PDEs have been obtained by Lions et al. [156].

Besides their PDE interpretation, TV-minimizing image processing methods have gained much interest within the optimization community [1, 60, 61, 73, 75, 129, 130, 147, 242]. To overcome the problem that the total variation integral contains the nondifferentiable argument $|\nabla u|$, one applies regularization strategies or techniques from nonsmooth optimization. Much research is done in order to find efficient numerical methods for which convergence can be established.

The constrained TV-minimization idea can also be adapted to other constraints such as blur, noise with blur, or other types of noise [156, 206, 208, 74, 147, 243]. Lions et al. [156] and Dobson and Santosa [73] have shown the existence of a BV(Ω)-solution for problems of this type, whereas uniqueness seems to be doubtful [73]. Moreover, one can minimize the $L^1(\Omega)$-norm of higher than first-order derivatives [206].

Total variation methods have been applied to restoring images of military relevance [207, 208, 156, 147, 243], to improving material from criminal and civil investigations as court evidence [206], and to enhancing pictures from confocal microscopy [242] and electrical impedance tomography [73].

1.7 Conclusions and scope of the thesis

Now that we have acquired a survey on the main ideas behind PDEs in image processing, we are in a position to draw consequences for the objectives of the present work.

We have seen that both linear diffusion and morphological scale-spaces reveal well-posedness results and can be very well axiomatically justified. On the other hand, for some applications, they possess the undesirable property that they do not permit contrast enhancement and that they may blur and delocalize structures.

Pure restoration methods such as diffusion–reaction equations or TV-based techniques do allow contrast enhancement and lead to stable structures but suffer from theoretical or practical problems, for instance well-posedness questions or the search for efficient minimizers of nonconvex or nondifferentiable functionals. Moreover, most image-enhancing PDE methods focus on edge detection and segmentation problems. Other interesting image restoration topics have found less attention.

For both scale-space and restoration methods many questions concerning their discrete realizations are still open: discrete scale-space results are frequently missing, minimization algorithms may get trapped in a poor local minimum, or the use of explicit schemes causes restrictive step size limitations.

Nonlinear diffusion filtering seems to combine many of the abovementioned shortcomings: Neither stability results nor a satisfactory continuous, semidiscrete or discrete scale-space interpretation are available. Edge-enhancing diffusivities offer a lot of opportunities to run into trouble, and explicit algorithms are typically applied to edge detection and segmentation tasks.

The goal of the present work is to convince the reader that this impression should be revised:

We shall see that anisotropic nonlinear diffusion processes share most advantages of the scale-space and the image enhancement world. A scale-space interpretation is presented which does not exclude contrast enhancement, and well-posedness results are established. Both scale-space and well-posedness properties carry over from the continuous to the semidiscrete and discrete setting. The latter one comprises also (semi-)implicit techniques for which unconditional stability in the L^∞-norm is proved. Finally, it is demonstrated that everything can be considered within a general framework which includes linear and isotropic nonlinear diffusion filters. Anisotropic models are presented which permit applications beyond segmentation and edge enhancement tasks, for instance enhancement of coherent flow-like structures in textures.

This thesis is organized as follows:

Chapter 2 presents a general model for the continuous setting where the diffusion tensor depends on the structure tensor, a generalization of the Gaussian-smoothed gradient allowing a more sophisticated description of local image structure. For this model class we discuss existence, uniqueness, stability, and an extremum principle. Scale-space properties are investigated with respect to invariances and information-reducing qualities resulting from associated Lyapunov functionals.

Chapter 3 establishes conditions under which comparable well-posedness and scale-space results can be proved for the semidiscrete framework. This case is of special interest since it involves the spatial discretization which is characteristic for digital images but it keeps the scale-space idea of using a continuous scale parameter. It leads to nonlinear systems of ordinary differential equations. We shall investigate under which conditions it is possible to get consistent approximations of the continuous anisotropic filter class which satisfy the abovementioned requirements.

In practice, however, scale-spaces are always approximated with a finite number of scales. This corresponds to the fully discrete case which is treated in Chapter 4. The investigated discrete filter class comes down to solving linear systems of equations which may arise from semi-implicit time discretizations of the semidiscrete filters. We shall see that many numerical schemes share typical features with their semidiscrete counterparts, for instance well-posedness results, extremum principles, Lyapunov functionals, and convergence to a constant steady state.

In Chapter 5 specific models are proposed which are tailored towards two objectives: smoothing with edge enhancement or multiscale enhancement of coherent structures. Their qualities are illustrated using images arising from computer aided quality control and medical applications, but also fingerprint images and impressionistic paintings shall be processed. The results are compared to related methods.

Finally, Chapter 6 concludes the thesis by giving a summary of the mathematical results and their relevance for image processing.

Chapter 2

Continuous diffusion filtering

This chapter presents a general continuous model for anisotropic diffusion filters, analyses its theoretical properties and gives a scale-space interpretation. To this end, we adapt the diffusion process to the structure tensor, a well-known tool for analysing local orientation. Under fairly weak assumptions on the class of filters, it is possible to establish well-posedness and regularity results and to prove a maximum–minimum principle. Since the proof does not require any monotony assumption it applies also to contrast-enhancing diffusion processes. After sketching invariances of the resulting scale-space, we focus on analysing its smoothing properties. We shall see that, besides the extremum principle, a large class of associated Lyapunov functionals plays an important role in this context [249].

2.1 Basic filter structure

Let us consider a rectangular image domain $\Omega := (0, a_1) \times (0, a_2)$ with boundary $\Gamma := \partial\Omega$ and let an image be represented by a mapping $f \in L^\infty(\Omega)$. The class of anisotropic diffusion filters we are concerned with is represented by the initial boundary value problem

$$\begin{aligned} \partial_t u = \operatorname{div}(D\, \nabla u) \qquad & \text{on} \qquad \Omega \times (0,\infty), && (2.1)\\ u(x,0) = f(x) \qquad & \text{on} \qquad \Omega, && (2.2)\\ \langle D\nabla u,\, n\rangle = 0 \qquad & \text{on} \qquad \Gamma \times (0,\infty). && (2.3) \end{aligned}$$

Hereby, n denotes the outer normal and $\langle .,. \rangle$ the Euclidean scalar product on $\mathbb{R}^2$. In order to adapt the diffusion tensor $D \in \mathbb{R}^{2\times 2}$ to the local image structure, one would usually let it depend on the edge estimator ∇u_σ (cf. 1.2.3), where

$$u_\sigma(x,t) \;:=\; (K_\sigma * \tilde{u}(.,t))\,(x) \qquad (\sigma > 0) \tag{2.4}$$

and $\tilde{u}$ denotes an extension of u from Ω to $\mathbb{R}^2$, which may be obtained by mirroring at Γ (cf. [58]).

However, we shall choose a more general structure descriptor which comprises the edge detector ∇u_σ, but also allows to extract more information. This will be presented next.

2.2 The structure tensor

In order to identify features such as corners or to measure local coherence of structures, we need methods which take into account how the (smoothed) gradient changes within the vicinity of any investigated point.

The *structure tensor* (also called *scatter matrix* or *(windowed second) moment tensor*) is an important representative of this class. Matrices of this type are useful for analysing flow-like textures [203, 204, 25], corners and T-junctions [181, 179], shape cues [151, pp. 349–382] and spatio–temporal image sequences [133, pp. 147–153]. Related approaches may also be found in [135, 30, 31]. Let us focus on some aspects which are of importance in our case.

To this end, we reconsider the vector-valued structure descriptor ∇u_σ within a matrix framework. The matrix J_0 resulting from the tensor product

$$J_0(\nabla u_\sigma) := \nabla u_\sigma \otimes \nabla u_\sigma := \nabla u_\sigma \nabla u_\sigma^T \tag{2.5}$$

has an orthonormal basis of eigenvectors v_1, v_2 with $v_1 \parallel \nabla u_\sigma$ and $v_2 \perp \nabla u_\sigma$. The corresponding eigenvalues $|\nabla u_\sigma|^2$ and 0 give just the contrast (the squared gradient) in the eigendirections. By convolving $J_0(\nabla u_\sigma)$ componentwise with a Gaussian K_ρ we obtain the structure tensor

$$J_\rho(\nabla u_\sigma) := K_\rho * (\nabla u_\sigma \otimes \nabla u_\sigma) \qquad (\rho \ge 0). \tag{2.6}$$

It is not hard to verify that the symmetric matrix $J_\rho = \begin{pmatrix} j_{11} & j_{12} \\ j_{12} & j_{22} \end{pmatrix}$ is positive semidefinite and possesses orthonormal eigenvectors v_1, v_2 with

$$v_1 \parallel \begin{pmatrix} 2j_{12} \\ j_{22} - j_{11} + \sqrt{(j_{11}-j_{22})^2 + 4j_{12}^2} \end{pmatrix}. \tag{2.7}$$

The corresponding eigenvalues μ_1 and μ_2 are given by

$$\mu_{1,2} = \frac{1}{2}\left(j_{11}+j_{22} \pm \sqrt{(j_{11}-j_{22})^2 + 4j_{12}^2} \right), \tag{2.8}$$

where the + sign belongs to μ_1. As they integrate the variation of the grey values within a neighbourhood of size $O(\rho)$, they describe the average contrast in the eigendirections. Thus, the *integration scale* ρ should reflect the characteristic window size over which the orientation is to be analysed. Presmoothing in order to obtain ∇u_σ makes the structure tensor insensitive to noise and irrelevant details of scales smaller than $O(\sigma)$. The parameter σ is called *local scale*.

By virtue of $\mu_1 \geq \mu_2 \geq 0$, we observe that v_1 is the orientation with the highest grey value fluctuations, and v_2 gives the preferred local orientation, the *coherence direction*. Furthermore, μ_1 and μ_2 can be used as descriptors of local structure: Constant areas are characterized by $\mu_1 = \mu_2 = 0$, straight edges give $\mu_1 \gg \mu_2 = 0$, corners reveal $\mu_1 \geq \mu_2 \gg 0$, and the expression

$$(\mu_1 - \mu_2)^2 \;=\; (j_{11} - j_{22})^2 + 4j_{12}^2 \tag{2.9}$$

becomes large for anisotropic structures. It is a measure of the *local coherence*.

2.3 Theoretical results

In order to discuss well-posedness results, let us first recall some useful notations. Let $\mathrm{H}^1(\Omega)$ be the Sobolev space of functions $u(x) \in \mathrm{L}^2(\Omega)$ with all distributional derivatives of first order being in $\mathrm{L}^2(\Omega)$. We equip $\mathrm{H}^1(\Omega)$ with the norm

$$\|u\|_{\mathrm{H}^1(\Omega)} \;:=\; \left(\|u\|^2_{\mathrm{L}^2(\Omega)} + \sum_{i=1}^{2} \|\partial_{x_i} u\|^2_{\mathrm{L}^2(\Omega)} \right)^{1/2} \tag{2.10}$$

and identify it with its dual space. Let $\mathrm{L}^2(0,T;\mathrm{H}^1(\Omega))$ be the space of functions u, strongly measurable on $[0,T]$ with range in $\mathrm{H}^1(\Omega)$ (for the Lebesgue measure dt on $[0,T]$) such that

$$\|u\|_{\mathrm{L}^2(0,T;\mathrm{H}^1(\Omega))} \;:=\; \left(\int_0^T \|u(t)\|^2_{\mathrm{H}^1(\Omega)}\, dt \right)^{1/2} \;<\; \infty. \tag{2.11}$$

In a similar way, $\mathrm{C}([0,T];\mathrm{L}^2(\Omega))$ is defined as the space of continuous functions $u : [0,T] \to \mathrm{L}^2(\Omega)$ supplemented with the norm

$$\|u\|_{\mathrm{C}([0,T];\mathrm{L}^2(\Omega))} \;:=\; \max_{[0,T]} \|u(t)\|_{\mathrm{L}^2(\Omega)}. \tag{2.12}$$

As usual, we denote by $\mathrm{C}^p(X,Y)$ the set of C^p-mappings from X to Y.

Let us now give a precise formulation of the problem we are concerned with. We need the following prerequisites:

$$\left.\begin{array}{l}
\text{Assume that } f \in L^\infty(\Omega),\ \rho \geq 0, \text{ and } \sigma, T > 0. \\
\text{Let } a := \operatorname*{ess\,inf}_{\Omega} f,\ b := \operatorname*{ess\,sup}_{\Omega} f, \text{ and consider the problem} \\
\qquad \begin{array}{rcl}
\partial_t u = \operatorname{div}\left(D(J_\rho(\nabla u_\sigma))\ \nabla u\right) & \text{on} & \Omega \times (0,T], \\
u(x,0) = f(x) & \text{on} & \Omega, \\
\langle D(J_\rho(\nabla u_\sigma))\nabla u,\ n\rangle = 0 & \text{on} & \Gamma \times (0,T],
\end{array} \\
\text{where the diffusion tensor } D = (d_{ij}) \text{ satisfies the following properties:} \\
\text{(C1) Smoothness: } D \in C^\infty(\mathbb{R}^{2\times 2}, \mathbb{R}^{2\times 2}). \\
\text{(C2) Symmetry: } d_{12}(J) = d_{21}(J) \text{ for all symmetric matrices } J \in \mathbb{R}^{2\times 2}. \\
\text{(C3) Uniform positive definiteness: For all } w \in L^\infty(\Omega, \mathbb{R}^2) \text{ with } |w(x)| \leq K \text{ on } \bar{\Omega}, \\
\qquad \text{there exists a positive lower bound } \nu(K) \text{ for the eigenvalues of } D(J_\rho(w)).
\end{array}\right\} \quad (P_c)$$

Under these assumptions the following theorem, which generalizes and extends results from [58, 249], can be proved.

Theorem 1 (Well-posedness,[1] regularity, extremum principle)
The problem (P_c) has a unique solution $u(x,t)$ in the distributional sense which satisfies

$$u \ \in \ C([0,T]; L^2(\Omega)) \ \cap \ L^2(0,T; H^1(\Omega)), \tag{2.13}$$

$$\partial_t u \ \in \ L^2(0,T; H^1(\Omega)). \tag{2.14}$$

Moreover, $u \in C^\infty(\bar{\Omega} \times (0,T])$. *This solution depends continuously on* f *with respect to* $\|\,.\,\|_{L^2(\Omega)}$, *and it fulfils the extremum principle*

$$a \leq u(x,t) \leq b \quad \text{on} \quad \Omega \times (0,T]. \tag{2.15}$$

[1] For a complete well-posedness proof one also has to establish stability with respect to perturbations of the diffusion equation. This problem will not be addressed here.

Proof:

(a) **Existence, uniqueness and regularity**

Existence, uniqueness and regularity are straightforward anisotropic extensions of the proof for the isotropic case studied by Catté, Lions, Morel and Coll [58]. Therefore, we just sketch the basic ideas of this proof.

Existence can be proved using Schauder's fixed point theorem. One considers the solution $U(w)$ of a distributional linear version of (P_c) where D depends on some function w instead of u. Then one shows that U is a weakly continuous mapping from a nonempty, convex and weakly compact subset W_0 of $W(0,T) := \left\{w \in \mathrm{L}^2(0,T;\mathrm{H}^1(\Omega)),\ \frac{dw}{dt} \in \mathrm{L}^2(0,T;\mathrm{H}^1(\Omega))\right\}$ into itself. Since $W(0,T)$ is contained in $\mathrm{L}^2(0,T;\mathrm{L}^2(\Omega))$, with compact inclusion, U reveals a fixed point $u \in W_0$, i.e. $u = U(u)$.

Smoothness follows from classical bootstrap arguments and the general theory of parabolic equations [144]. Since $u(t) \in \mathrm{H}^1(\Omega)$ for all $t > 0$, one deduces that $u(t) \in \mathrm{H}^2(\Omega)$ for all $t > 0$. By iterating, one can establish that u is a strong solution of (P_c) and $u \in \mathrm{C}^\infty((0,T] \times \bar{\Omega})$.

The basic idea of the uniqueness proof consists of using energy estimates for the difference of two solutions, such that the Gronwall–Bellman inequality can be applied. Then, uniqueness follows from the fact that both solutions start with the same initial values.

Finally an iterative linear scheme is investigated, whose solution is shown to converge in $\mathrm{C}([0,T];\mathrm{L}^2(\Omega))$ to the strong solution of (P_c).

(b) **Extremum principle**

In order to prove a maximum–minimum principle, we utilize Stampacchia's truncation method (cf. [40], p. 211).

We restrict ourselves to proving only the maximum principle. The minimum principle follows from the maximum principle when being applied to the initial datum $-f$.
Let $G \in \mathrm{C}^1(\mathbb{R})$ be a function with $G(s) = 0$ on $(-\infty, 0]$ and $0 < G'(s) \le C$ on $(0,\infty)$ for some constant C. Now, we define

$$\begin{aligned} H(s) &:= \int_0^s G(\sigma)\,d\sigma, \qquad s \in \mathbb{R}, \\ \varphi(t) &:= \int_\Omega H(u(x,t)-b)\,dx, \qquad t \in [0,T]. \end{aligned}$$

By the Cauchy–Schwarz inequality, we have

$$\int_\Omega |G(u(x,t)-b)\ \partial_t u(x,t)|\ dx \ \le\ C \cdot \|u(t)-b\|_{\mathrm{L}^2(\Omega)} \cdot \|\partial_t u(t)\|_{\mathrm{L}^2(\Omega)}$$

and by virtue of (2.13), (2.14) we know that the right-hand side of this estimate exists. Therefore, φ is differentiable for $t > 0$, and we get

$$\begin{aligned}
\frac{d\varphi}{dt} &= \int_\Omega G(u-b)\, \partial_t u \, dx \\
&= \int_\Omega G(u-b)\, \mathrm{div}\,(D(J_\rho(\nabla u_\sigma))\, \nabla u)\, dx \\
&= \int_\Gamma G(u-b)\, \underbrace{\langle D(J_\rho(\nabla u_\sigma))\, \nabla u, n\rangle}_{=0}\, dS \\
&\quad - \int_\Omega \underbrace{G'(u-b)}_{\geq 0}\, \underbrace{\langle \nabla u, D(J_\rho(\nabla u_\sigma))\, \nabla u\rangle}_{\geq 0}\, dx \\
&\leq 0.
\end{aligned} \tag{2.16}$$

By means of $H(s) \leq \frac{C}{2}s^2$, we have

$$0 \;\leq\; \varphi(t) \;\leq\; \int_\Omega H(u(x,t)-f(x))\, dx \;\leq\; \frac{C}{2}\, \|u(t)-f\|^2_{\mathrm{L}^2(\Omega)}. \tag{2.17}$$

Since $u \in \mathrm{C}([0,T]; \mathrm{L}^2(\Omega))$, the right-hand side of (2.17) tends to $0 = \varphi(0)$ for $t \to 0^+$ which proves the continuity of $\varphi(t)$ in 0. Now from

$$\varphi \in \mathrm{C}[0,T], \qquad \varphi(0) = 0, \qquad \varphi \geq 0 \quad \text{on} \quad [0,T]$$

and (2.16), it follows that

$$\varphi \equiv 0 \quad \text{on} \quad [0,T].$$

Hence, for all $t \in [0,T]$, we obtain $u(x,t) - b \leq 0$ almost everywhere (a.e.) on Ω. Due to the smoothness of u for $t > 0$, we finally end up with the assertion

$$u(x,t) \leq b \quad \text{on} \quad \bar{\Omega} \times (0,T].$$

(c) **Continuous dependence on the initial image**
Let $f, h \in \mathrm{L}^\infty(\Omega)$ be two initial values and u, w the corresponding solutions. In the same way as in the uniqueness proof in [58], one shows that there exists some constant $c > 0$ such that

$$\frac{d}{dt}\left(\|u(t)-w(t)\|^2_{\mathrm{L}^2(\Omega)}\right) \;\leq\; c \cdot \|\nabla u(t)\|^2_{\mathrm{L}^2(\Omega)} \cdot \|u(t)-w(t)\|^2_{\mathrm{L}^2(\Omega)}.$$

Applying the Gronwall–Bellman lemma [39, pp. 156–137] yields

$$\|u(t)-w(t)\|^2_{\mathrm{L}^2(\Omega)} \;\leq\; \|f-h\|^2_{\mathrm{L}^2(\Omega)} \cdot \exp\left(c \cdot \int_0^t \|\nabla u(s)\|^2_{\mathrm{L}^2(\Omega)}\, ds\right).$$

By means of the extremum principle we know that u is bounded on $\bar{\Omega}\times[0,T]$. Thus, ∇u_σ is also bounded, and prerequisite (ii) implies that there exists some constant $\nu = \nu(\sigma, \|f\|_{\mathrm{L}^\infty(\Omega)}) > 0$, such that

$$\begin{aligned}
& \int_0^t \|\nabla u(s)\|^2_{\mathrm{L}^2(\Omega)}\, ds \\
\leq\ & \int_0^T \|\nabla u(s)\|^2_{\mathrm{L}^2(\Omega)}\, ds \\
\leq\ & \frac{1}{\nu}\int_0^T \left|\int_\Omega \langle \nabla u(x,s), D(J_\rho(\nabla u_\sigma(x,s)))\nabla u(x,s)\rangle\, dx\right| ds \\
=\ & \frac{1}{\nu}\int_0^T \left|\int_\Omega u(x,s)\cdot \operatorname{div}\Big(D(J_\rho(\nabla u_\sigma(x,s)))\nabla u(x,s)\Big)\, dx\right| ds \\
\leq\ & \frac{1}{\nu}\int_0^T \|u(s)\|_{\mathrm{L}^2(\Omega)}\, \|\partial_t u(s)\|_{\mathrm{L}^2(\Omega)}\, ds \\
\leq\ & \frac{1}{\nu}\, \|u\|_{\mathrm{L}^2(0,T;\mathrm{H}^1(\Omega))}\, \|\partial_t u\|_{\mathrm{L}^2(0,T;\mathrm{H}^1(\Omega))}.
\end{aligned}$$

By virtue of (2.13), (2.14), we know that the right-hand of this estimate exists. Now, let $\epsilon > 0$ and choose

$$\delta \ := \ \epsilon \cdot \exp\left(\frac{-c}{2\nu}\, \|u\|_{\mathrm{L}^2(0,T;\mathrm{H}^1(\Omega)}\, \|\partial_t u\|_{\mathrm{L}^2(0,T;\mathrm{H}^1(\Omega))}\right).$$

Then for $\|f-h\|_{\mathrm{L}^2(\Omega)} < \delta$, the preceding results imply

$$\|u(t)-w(t)\|_{\mathrm{L}^2(\Omega)} < \epsilon \qquad \forall\, t \in [0,T],$$

which proves the continuous dependence on the initial data. □

Remarks:

(a) We observe a strong smoothing effect which is characteristic for diffusion processes: under fairly weak assumptions on the initial image ($f \in \mathrm{L}^\infty(\Omega)$) we obtain an infinitely often differentiable solution for arbitrary small positive times. More restrictive requirements – for instance $f \in \mathrm{BUC}(\mathbb{R}^2)$ in order to apply the theory of viscosity solutions – are not necessary in our case.

(b) Moreover, our proof does not require any monotony assumption. This has the advantage that contrast-enhancing processes are permitted as well. Chapter 5 will illustrate this by providing examples where contrast is enhanced.

(c) The continuous dependence of the solution on the initial image has significant practical impact as it ensures stability with respect to perturbations of the original image. This is of importance when considering stereo image pairs, spatio-temporal image sequences or slices from medical CT or MRT sequences, since we know that similar images remain similar after filtering.[2]

(d) The extremum principle offers the practical advantage that, if we start for instance with an image within the range $[0, 255]$, we will never obtain results with grey value such as 257. It is also closely related to smoothing scale-space properties, as we shall see in 2.4.2.

(e) The well-posedness results are essentially based on the fact that the regularization by convolution with a Gaussian allows to estimate $\|\nabla u_\sigma\|_{L^\infty(\Omega)}$ by $\|u\|_{L^\infty(\Omega)}$. This property is responsible for the uniform positive definiteness of the diffusion tensor.

2.4 Scale-space properties

Let us now investigate scale-space properties of the class (P_c) and juxtapose the results to other scale-spaces. To this end, we shall not focus on further investigations of architectural requirements like recursivity, regularity and locality, as these qualities do not distinguish nonlinear diffusion scale-spaces from other ones. We start with briefly discussing invariances. Afterwards, we turn to a more crucial task, namely the question in which sense our evolution equation – which may allow contrast enhancement – can still be considered as a smoothing, information-reducing image transformation.

2.4.1 Invariances

Let $u(x, t)$ be the unique solution of (P_c) and define the scale-space operator T_t by

$$T_t f := u(t), \qquad t \geq 0, \tag{2.18}$$

where $u(t) := u(., t)$.
The properties we discuss now illustrate that an invariance of T_t with respect to some image transformation P is characterized by the fact that T_t and P commute. Much of the terminology used below is borrowed from [11].

[2] This does not contradict contrast enhancement: In the case of two similar images, where one leads to contrast enhancement and the other not, the regularization damps the enhancement process in such a way that both images do not differ much after filtering.

Grey level shift invariance

Since the diffusion tensor is only a function of $J_\rho(\nabla u_\sigma)$, but not of u, we may shift the grey level range by an arbitrary constant C, and the filtered images will also be shifted by the same constant. Moreover, a constant function is not affected by diffusion filtering. Therefore, we have

$$\begin{aligned} T_t(0) &= 0, && (2.19)\\ T_t(f+C) &= T_t(f) + C \qquad \forall\, t \geq 0. && (2.20) \end{aligned}$$

Reverse contrast invariance

From $D(J_\rho(-\nabla u_\sigma)) = D(J_\rho(\nabla u_\sigma))$, it follows that

$$T_t(-f) = -T_t(f) \qquad \forall\, t \geq 0. \tag{2.21}$$

This property is not fulfilled by classical morphological scale-space equations like dilation and erosion. When reversing the contrast, the role of dilation and erosion has to be exchanged as well.

Average grey level invariance

Average grey level invariance is a further property in which diffusion scale-spaces differ from morphological scale-spaces. In general, the evolution PDEs of the latter ones are not of divergence form and do not preserve the mean grey value.

Proposition 1 (Conservation of average grey value).
The average grey level

$$\mu := \frac{1}{|\Omega|} \int_\Omega f(x)\, dx \tag{2.22}$$

is not affected by nonlinear diffusion filtering:

$$\frac{1}{|\Omega|} \int_\Omega T_t f\, dx = \mu \qquad \forall\, t > 0. \tag{2.23}$$

Proof:
Define $I(t) := \int_\Omega u(x,t)\, dx$ for all $t \geq 0$. Then the Cauchy–Schwarz inequality implies

$$|I(t) - I(0)| = \left| \int_\Omega \Big(u(x,t) - f(x)\Big)\, dx \right| \leq |\Omega|^{1/2}\, \|u(t) - f\|_{L^2(\Omega)}.$$

Since $u \in \mathrm{C}([0,T];\mathrm{L}^2(\Omega))$, the preceding inequality gives the continuity of $I(t)$ in 0.
For $t > 0$, Theorem 1, the divergence theorem and the boundary conditions yield

$$\frac{dI}{dt} \;=\; \int\limits_{\Omega} \partial_t u\, dx \;=\; \int\limits_{\Gamma} \langle D(J_\rho(\nabla u_\sigma))\nabla u,\, n\rangle\, dS \;=\; 0.$$

Hence, $I(t)$ must be constant for all $t \geq 0$. □

Average grey level invariance may be described by commuting operators, when introducing an averaging operator $M : \mathrm{L}^1(\Omega) \to \mathrm{L}^1(\Omega)$ which maps f to a constant image with the same mean grey level:

$$(Mf)(y) := \frac{1}{|\Omega|} \int\limits_{\Omega} f(x)\, dx \qquad \forall\, y \in \Omega. \tag{2.24}$$

Then Proposition 1 and grey level shift invariance imply that the order of M and T_t is exchangable:

$$M(T_t f) = T_t(Mf) \qquad \forall\, t \geq 0. \tag{2.25}$$

When studying diffusion filtering as a pure initial value problem in the domain $\mathbb{R}^2$, it also makes sense to investigate Euclidean transformations of an image. This leads us to translation and isometry invariance.

Translation invariance

Define a translation τ_h by $(\tau_h f)(x) := f(x+h)$. Then diffusion filtering fulfils

$$T_t(\tau_h f) = \tau_h(T_t f) \qquad \forall\, t \geq 0. \tag{2.26}$$

This is a consequence of the fact that the diffusion tensor depends on $J_\rho(\nabla u_\sigma)$ solely, but not explicitly on x.

Isometry invariance

Let $R \in \mathbb{R}^{2\times 2}$ be an orthogonal transformation. If we apply R to f by defining $Rf(x) := f(Rx)$, then the eigenvalues of the diffusion tensor are unaltered and any eigenvector v is transformed into Rv. Thus, it makes no difference whether the orthogonal transformation is applied before or after diffusion filtering:

$$T_t(Rf) = R(T_t f) \qquad \forall\, t \geq 0. \tag{2.27}$$

2.4.2 Information-reducing properties

Nonenhancement of local extrema

Many smoothing scale-space properties are closely related to extremum principles: Hummel [124] for instance shows that, under certain conditions, the maximum principle for parabolic operators is equivalent to the property that the corresponding scale-space never creates additional level-crossings for $t > 0$. Another important smoothing axiom is the requirement that local extrema must not be amplified when increasing the scale parameter. This condition has first been used by Lindeberg [150] in the context of linear diffusion filtering. However, it is also satisfied by nonlinear diffusion scale-spaces, as we shall see now.[3]

Theorem 2 (Nonenhancement of local extrema).
Let u be the unique solution of (P_c) and consider some $\theta > 0$. Suppose that $\xi \in \Omega$ is a local extremum of $u(.,\theta)$. Then,

$$\partial_t u(\xi,\theta) \;\le\; 0, \qquad \text{if } \xi \text{ is a local maximum,} \tag{2.28}$$

$$\partial_t u(\xi,\theta) \;\ge\; 0, \qquad \text{if } \xi \text{ is a local minimum.} \tag{2.29}$$

Proof:
Let $D(J_\rho(\nabla u_\sigma)) =: (d_{ij}(J_\rho(\nabla u_\sigma)))$. Then we have

$$\partial_t u \;=\; \sum_{i=1}^{2}\sum_{j=1}^{2}\Big(\partial_{x_i} d_{ij}(J_\rho(\nabla u_\sigma))\Big)\,\partial_{x_j} u \;+\; \sum_{i=1}^{2}\sum_{j=1}^{2} d_{ij}(J_\rho(\nabla u_\sigma))\,\partial_{x_i x_j} u. \tag{2.30}$$

Since $\nabla u(\xi,\theta) = 0$ and $\partial_{x_i} d_{ij}(J_\rho(\nabla u_\sigma(\xi,\theta)))$ is bounded, the first term of the right-hand side of (2.30) vanishes in (ξ,θ).

We know that the diffusion tensor $D := D(J_\rho(\nabla u_\sigma(\xi,\theta)))$ is positive semidefinite. Hence, there exists an orthogonal transformation $S \in \mathbb{R}^{2\times 2}$ such that

$$S^T D S \;=\; \mathrm{diag}(\lambda_1,\lambda_2) \;=:\; \Lambda$$

with λ_1, λ_2 being the nonnegative eigenvalues of D.

Now, let us assume that (ξ,θ) is a local maximum. Then the Hessian $H := \mathrm{Hess}(u(\xi,\theta))$ and $B := (b_{ij}) := S^T H S$ are negative semidefinite. Therefore, we have

$$b_{ii} \le 0 \quad (i = 1, 2)$$

[3] As in the linear diffusion case, nonenhancement of local extrema generally does not imply that their number is nonincreasing, cf. 1.1.5.

and by the invariance of the trace with respect to orthogonal transformations it follows that

$$\begin{aligned}\partial_t u(\xi,\theta) &= \operatorname{trace}(DH)\\ &= \operatorname{trace}(S^T DS\, S^T HS)\\ &= \operatorname{trace}(\Lambda B)\\ &= \sum_{i=1}^{2} \lambda_i\, b_{ii}\\ &\le 0.\end{aligned}$$

If ξ is a local minimum of $u(x,\theta)$, one proceeds in the same way utilizing the positive semidefiniteness of the Hessian. □

Nonenhancement of local extrema distinguishes anisotropic diffusion from classical contrast enhancing methods such as high-frequency emphasis [105, pp. 182–183], which do violate this principle. Although possibly behaving like backward diffusion across edges, nonlinear diffusion is always in the forward region at extrema. This ensures its stability.

Lyapunov functionals and behaviour for $t \to \infty$

Since scale-spaces are intended to subsequently simplify an image, it is desirable that, for $t \to \infty$, we obtain the simplest possible image representation, namely a constant image with the same average grey value as the original one. The following theorem states that anisotropic diffusion filtering always leads to a constant steady-state. This is due to the class of Lyapunov functionals associated with the diffusion process.

Theorem 3 (Lyapunov functionals and behaviour for $t \to \infty$).
Suppose that u is the solution of (P_c) and let a, b, μ and M be defined as in (P_c), (2.22) and (2.24), respectively. Then the following properties are valid:

(a) *(Lyapunov functionals)*
For all $r \in \mathrm{C}^2[a,b]$ with $r'' \ge 0$ on $[a,b]$, the function

$$V(t) := \Phi(u(t)) := \int_\Omega r(u(x,t))\, dx \tag{2.31}$$

is a Lyapunov functional:

(i) $\Phi(u(t)) \ge \Phi(Mf)$ *for all* $t \ge 0$.

(ii) $V \in \mathrm{C}[0,\infty) \cap \mathrm{C}^1(0,\infty)$ *and* $V'(t) \le 0$ *for all* $t > 0$.

Moreover, if $r'' > 0$ on $[a,b]$, then $V(t) = \Phi(u(t))$ is a strict Lyapunov functional:

(iii) $\quad \Phi(u(t)) = \Phi(Mf) \quad \Longleftrightarrow \quad \begin{cases} u(t) = Mf & \text{on } \bar{\Omega} & (\text{if } t > 0) \\ u(t) = Mf & \text{a.e. on } \Omega & (\text{if } t = 0) \end{cases}$

(iv) *If $t > 0$, then $V'(t) = 0$ if and only if $u(t) = Mf$ on $\bar{\Omega}$.*

(v) $\quad V(0) = V(T) \quad \text{for } T > 0 \quad \Longleftrightarrow \quad \begin{cases} f = Mf & \text{a.e. on } \Omega \text{ and} \\ u(t) = Mf & \text{on } \bar{\Omega} \times (0,T] \end{cases}$

(b) (Convergence)

(i) $\lim_{t\to\infty} \|u(t) - Mf\|_{L^p(\Omega)} = 0 \quad$ *for* $\quad p \in [1,\infty)$.

(ii) *In the 1D case, the convergence $\lim_{t\to\infty} u(x,t) = \mu$ is uniform on $\bar{\Omega}$.*

Proof:

(a) (i) Since $r \in C^2[a,b]$ with $r'' \geq 0$ on $[a,b]$, we know that r is convex on $[a,b]$. Using the average grey level invariance and Jensen's inequality we obtain, for all $t \geq 0$,

$$\begin{aligned} \Phi(Mf) &= \int_\Omega r\left(\frac{1}{|\Omega|}\int_\Omega u(x,t)\,dx\right) dy \\ &\leq \int_\Omega \left(\frac{1}{|\Omega|}\int_\Omega r(u(x,t))\,dx\right) dy \\ &= \int_\Omega r(u(x,t))\,dx \\ &= \Phi(u(t)). \end{aligned} \tag{2.32}$$

(ii) Let us start by proving the continuity of $V(t)$ in 0. Thanks to the maximum–minimum principle, we may choose a constant

$$L := \max_{s\in[a,b]} |r'(s)|$$

such that for all $t > 0$, the Lipschitz condition

$$|r(u(x,t)) - r(f(x))| \;\leq\; L\,|u(x,t) - f(x)|$$

is verified a.e. on Ω. From this and the Cauchy–Schwarz inequality, we get

$$\begin{aligned} |V(t) - V(0)| &\leq |\Omega|^{1/2}\,\|r(u(t)) - r(f)\|_{L^2(\Omega)} \\ &\leq |\Omega|^{1/2}\,L\,\|u(t) - f\|_{L^2(\Omega)}, \end{aligned}$$

and by virtue of $u \in \mathrm{C}([0,T];\mathrm{L}^2(\Omega))$, the limit $t \to 0^+$ gives the announced continuity in 0.
By Theorem 1 and the boundedness of r' on $[a,b]$, we know that V is differentiable for $t>0$ and $V'(t) = \int_\Omega r'(u)\, u_t\, dx$. Thus, the divergence theorem yields

$$\begin{aligned} V'(t) &= \int_\Omega r'(u) \operatorname{div}\left(D(J_\rho(\nabla u_\sigma))\nabla u\right) dx \\ &= \int_\Gamma r'(u) \underbrace{\langle D(J_\rho(\nabla u_\sigma))\nabla u, n\rangle}_{=0} dS \\ &- \int_\Omega \underbrace{r''(u)}_{\geq 0} \underbrace{\langle \nabla u, D(J_\rho(\nabla u_\sigma))\nabla u\rangle}_{\geq 0} dx \\ &\leq 0. \end{aligned}$$

(iii) Let $\Phi(u(t)) = \Phi(Mf)$.
If $t>0$, then $u(t)$ is continuous in $\bar\Omega$. Let us now show that equality in the estimate (2.32) implies that $u(t) = \text{const.}$ on $\bar\Omega$. To this end, assume that u is not constant on $\bar\Omega$. Then, by the continuity of u, there exists a partition $\Omega = \Omega_1 \cup \Omega_2$ with $|\Omega_1|, |\Omega_2| \in (0, |\Omega|)$ and

$$\alpha := \frac{1}{|\Omega_1|}\int_{\Omega_1} u\, dx \neq \frac{1}{|\Omega_2|}\int_{\Omega_2} u\, dx =: \beta.$$

From $r''>0$ on $[a,b]$ it follows that r is strictly convex on $[a,b]$ and

$$\begin{aligned} r\left(\frac{1}{|\Omega|}\int_\Omega u\, dx\right) &= r\left(\frac{|\Omega_1|}{|\Omega|}\alpha + \frac{|\Omega_2|}{|\Omega|}\beta\right) \\ &< \frac{|\Omega_1|}{|\Omega|} r(\alpha) + \frac{|\Omega_1|}{|\Omega|} r(\beta) \\ &\leq \frac{1}{|\Omega|}\int_{\Omega_1} r(u)\, dx + \frac{1}{|\Omega|}\int_{\Omega_2} r(u)\, dx \\ &= \frac{1}{|\Omega|}\int_\Omega r(u)\, dx. \end{aligned}$$

If we utilize this result in the estimate (2.32) we observe that, for $t>0$, $\Phi(u(t)) = \Phi(Mf)$ implies that $u(t) = \text{const.}$ on $\bar\Omega$. Thanks to the average grey value invariance we finally obtain $u(t) = Mf$ on $\bar\Omega$.

So let us turn to the case $t=0$. From (i) and (ii), we conclude that $\Phi(u(\theta)) = \Phi(Mf)$ for all $\theta > 0$. Thus, we have $u(\theta) = Mf$ for all

$\theta > 0$.

For $\theta > 0$, the Cauchy–Schwarz inequality gives

$$\int_\Omega |u(x,\theta) - \mu| \, dx \;\le\; |\Omega|^{1/2} \, \|u(\theta) - Mf\|_{\mathrm{L}^2(\Omega)} \;=\; 0.$$

Since $u \in \mathrm{C}([0,T]; \mathrm{L}^2(\Omega))$, the limit $\theta \to 0^+$ finally yields $u(0) = Mf$ a.e. on Ω.

Conversely, it is obvious that $u(t) = Mf$ (a.e.) on Ω implies $\Phi(u(t)) = \Phi(Mf)$.

(iv) Let $t > 0$ and $V'(t) = 0$. Then from

$$0 \;=\; V'(t) \;=\; -\int_\Omega \underbrace{r''(u(x,t))}_{>0} \langle \nabla u(x,t), D(J_\rho(\nabla u_\sigma(x,t)))\nabla u(x,t)\rangle \, dx$$

and the smoothness of u we obtain

$$\langle \nabla u, D(J_\rho(\nabla u_\sigma))\nabla u\rangle \;=\; 0 \quad \text{on } \bar{\Omega}.$$

By the uniform boundedness of D, there exists some constant $\nu > 0$, such that

$$\nu \, |\nabla u|^2 \;\le\; \langle \nabla u, D(J_\rho(\nabla u_\sigma))\nabla u\rangle \quad \text{on } \bar{\Omega} \times (0,\infty).$$

Thus, we have $\nabla u(x,t) = 0$ a.e. on Ω. Due to the continuity of ∇u, this yields $u(x,t) = \text{const.}$ for all $x \in \Omega$, and the average grey level invariance finally gives $u(x,t) = \mu$ on Ω.

Conversely, let $u(x,t) = \mu$ on Ω. Then,

$$V'(t) \;=\; -\int_\Omega r''(u) \, \langle \nabla u, D(J_\rho(\nabla u_\sigma))\nabla u\rangle \, dx \;=\; 0.$$

(v) Suppose that $V(T) = V(0)$. Since V is decreasing, we have

$$V(t) = \text{const.} \qquad \text{on } [0,T].$$

Let $\epsilon > 0$. Then for any $t \in [\epsilon, T]$, we have $V'(t) = 0$, and part (iv) implies that $u(t) = Mf$ on Ω. Now, the Cauchy–Schwarz inequality gives

$$\int_\Omega |f - Mf| \, dx \;\le\; |\Omega|^{1/2} \, \|f - u(t)\|_{\mathrm{L}^2(\Omega)}.$$

As $u \in \mathrm{C}([0,T]; \mathrm{L}^2(\Omega))$, the limit $t \to 0^+$ yields $f = Mf$ a.e. on Ω.

Conversely, if $u(t) = Mf$ (a.e.) on Ω holds for all $t \in [0,T]$, it is evident that $V(0) = V(T)$.

(b) (i) By the grey level shift invariance we know that $e := u - Mf$ satisfies the diffusion equation as well. We multiply this equation by e, integrate, and use the divergence theorem to obtain

$$\int_\Omega e e_t \, dx \;=\; -\int_\Omega \langle \nabla e, D(J_\rho(\nabla e_\sigma)) \nabla e \rangle \, dx.$$

Since ∇e_σ is bounded, we find some $\nu > 0$ such that

$$\frac{1}{2}\frac{d}{dt}\left(\|e\|^2_{\mathrm{L}^2(\Omega)}\right) \;\leq\; -\nu\, \|\nabla e\|^2_{\mathrm{L}^2(\Omega)}.$$

For $t > 0$, there exists some x_0 with $e(x_0) = 0$. Therefore, the Poincaré inequality (cf. [7, p. 122]) tells us that

$$\|e\|^2_{\mathrm{L}^2(\Omega)} \;\leq\; C_0\, \|\nabla e\|^2_{\mathrm{L}^2(\Omega)}$$

with some constant $C_0 = C_0(\Omega) > 0$. This yields

$$\frac{d}{dt}\|e\|^2_{\mathrm{L}^2(\Omega)} \;\leq\; -2\nu\, C_0\, \|e\|^2_{\mathrm{L}^2(\Omega)}$$

and hence the exponential decay of $\|e\|_{\mathrm{L}^2(\Omega)}$ to 0.
By the maximum principle, we know that $\|e(t)\|_{\mathrm{L}^\infty(\Omega)}$ is bounded by $\|f - Mf\|_{\mathrm{L}^\infty(\Omega)}$. Thus, for $q \in \mathbb{N}$, $q \geq 2$, we get

$$\|e(t)\|^q_{L^q(\Omega)} \;\leq\; \|f - Mf\|^{q-2}_{\mathrm{L}^\infty(\Omega)} \cdot \|e(t)\|^2_{\mathrm{L}^2(\Omega)} \;\to\; 0,$$

and Hölder's inequality gives, for $1 \leq p < q < \infty$,

$$\|e(t)\|_{L^p(\Omega)} \;\leq\; |\Omega|^{(1/p)-(1/q)} \cdot \|e(t)\|_{L^q(\Omega)} \;\to\; 0.$$

This proves the assertion.

(ii) To prove uniform convergence in the one-dimensional setting, we can generalize and adapt methods from [127] to our case.
Let $\Omega = (0, a)$. From part (a) we know that $V(t) := \int_0^a u^2(x,t)\, dx$ is nonincreasing and bounded from below. Thus, the sequence $(V(i))_{i \in \mathbb{N}}$ converges.
Since $V \in \mathrm{C}[0,\infty) \cap \mathrm{C}^1(0,\infty)$ the mean value theorem implies

$$\exists\, t_i \in (i,\, i{+}1): \qquad V'(t_i) \;=\; V(i+1) - V(i).$$

Thus, $(t_i)_{i \in \mathbb{N}} \to \infty$ and from the convergence of $(V(i))_{i \in \mathbb{N}}$ it follows that

$$V'(t_i) \to 0. \tag{2.33}$$

Thanks to the uniform positive definiteness of D there exists some $\nu > 0$ such that, for $t > 0$,

$$\begin{aligned} V'(t) &= -2 \int_0^a u_x^2 \, D(J_\rho(\partial_x u_\sigma)) \, dx \\ &\leq -2\nu \int_0^a u_x^2 \, dx \\ &\leq 0. \end{aligned} \tag{2.34}$$

Equations (2.33) and (2.34) yield

$$\|u_x(t_i)\|_{\mathrm{L}^2(\Omega)} \to 0.$$

Hence, $u(t_i)$ is a bounded sequence in $\mathrm{H}^1(0,a)$. By virtue of the Rellich–Kondrachov theorem [5, p. 144] we know that the embedding from $\mathrm{H}^1(0,a)$ into $\mathrm{C}^{0,\alpha}[0,a]$, the space of Hölder-continuous functions on $[0,a]$ [5, pp. 9–12], is compact for $\alpha \in (0, \frac{1}{2})$. Therefore, there exists a subsequence $(t_{i_j}) \to \infty$ and some $\bar{u}$ with

$$u(t_{i_j}) \to \bar{u} \quad \text{in} \quad \mathrm{C}^{0,\alpha}[0,a].$$

This also gives $u(t_{i_j}) \to \bar{u}$ in $\mathrm{L}^2(0,a)$. Since we already know from (b)(i) that $u(t_{i_j}) \to Mf$ in $\mathrm{L}^2(0,a)$, it follows that $\bar{u} = Mf$. Hence,

$$\lim_{j \to \infty} \|u(t_{i_j}) - Mf\|_{\mathrm{L}^\infty(\Omega)} = 0. \tag{2.35}$$

Part (a) tells us that $\|u(t) - Mf\|^p_{\mathrm{L}^p(\Omega)}$ is a Lyapunov function for $p \geq 2$. Thus,

$$\|u(t) - Mf\|_{\mathrm{L}^\infty(\Omega)} = \lim_{p \to \infty} \|u(t) - Mf\|_{\mathrm{L}^p(\Omega)}$$

is also nonincreasing. Therefore, $\lim\limits_{t \to \infty} \|u(t) - Mf\|_{\mathrm{L}^\infty(\Omega)}$ exists and from (2.35) we conclude that

$$\lim_{t \to \infty} \|u(t) - Mf\|_{\mathrm{L}^\infty(\Omega)} = 0.$$

The smoothness of u establishes finally that

$$\lim_{t \to \infty} u(x,t) = \mu$$

uniform on $\bar{\Omega}$. $\square$

Since our class (P_c) does not forbid contrast enhancement it admits processes where forward diffusion has to compete with backward diffusion. Theorem 3 is of importance as it states that the regularization by convolving with K_σ tames the backward diffusion in such a way that forward diffusion wins in the long run. Moreover, the competition evolves in a certain direction all the time: although backward diffusion may be locally superior, the global result – denoted by the Lyapunov functional – becomes permanently better for forward diffusion. So let us have a closer look at what might be the meaning of this global result in the context of image processing.

Considering the Lyapunov functions associated with $r(s) := |s|^p$, $r(s) := (s-\mu)^{2n}$ and $r(s) := s\,\ln s$, respectively, the preceding theorem gives the following corollary.

Corollary 1 (Special Lyapunov functionals).
Let u be the solution of (P_c) and a and μ be defined as in (P_c) and (2.22). Then the following functions are decreasing for $t \in [0,\infty)$:

(a) $\|u(t)\|_{L^p(\Omega)}$ *for all* $p \geq 2$.

(b) $M_{2n}[u(t)] := \dfrac{1}{|\Omega|}\displaystyle\int_\Omega (u(x,t)-\mu)^{2n}\,dx$ *for all* $n \in \mathbb{N}$.

(c) $H[u(t)] := \displaystyle\int_\Omega u(x,t)\,\ln(u(x,t))\,dx,$ *if* $a > 0$.

Corollary 1 offers multiple possibilities of how to interpret nonlinear anisotropic diffusion filtering as a smoothing transformation.

As a special case of (a) it follows that the energy $\|u(t)\|^2_{L^2(\Omega)}$ is reduced by diffusion.

Part (b) gives a probabilistic interpretation of anisotropic diffusion filtering. Consider the intensity in an image f as a random variable Z_f with distribution $F_f(z)$, i.e. $F_f(z)$ is the probability that an arbitrary grey value Z_f of f does not exceed z. By the average grey level invariance, μ is equal to the expected value

$$EZ_{u(t)} := \int_{\mathbb{R}} z\,dF_{u(t)}(z), \tag{2.36}$$

and it follows that $M_{2n}[u(t)]$ is just the even central moment

$$\int_{\mathbb{R}} \left(z - EZ_{u(t)}\right)^{2n} dF_{u(t)}(z). \tag{2.37}$$

The second central moment (the variance) characterizes the spread of the intensity about its mean. It is a common tool for constructing measures for the relative smoothness of the intensity distribution. The fourth moment is frequently used to describe the relative flatness of the grey value distribution. Higher moments are more difficult to interprete, although they do provide important information for tasks like texture discrimination [105, pp. 414–415]. All decreasing even moments demonstrate that the image becomes smoother during diffusion filtering. Hence, local effects such as edge enhancement, which object to increase central moments, are overcompensated by smoothing in other areas.

If we choose another probabilistic model of images, then part (c) characterizes the information-theoretical side of our scale-space. Provided the initial image f is strictly positive on Ω, we may regard it also as a two-dimensional density.[4] Then,

$$S[u(t)] \;:=\; -\int_{\Omega} u(x,t)\, \ln(u(x,t))\, dx \tag{2.38}$$

is called the *entropy* of $u(t)$, a measure of uncertainty and missing information [46]. Since anisotropic diffusion filters increase the entropy the corresponding scale-space embeds the genuine image f into a family of subsequently likelier versions of it which contain less information.[5] Moreover, for $t \to \infty$, the process reaches the state with the lowest possible information, namely a constant image.

From all the previous considerations, we observe that, in spite of possible contrast-enhancing properties, anisotropic diffusion does really simplify the original image in a steady way.

Let us finally point out another interpretation of the Lyapunov functionals. In a classic scale-space representation, the time t plays the role of the scale parameter. By increasing t, one transforms the image from a local to a more global representation. We have seen in Chapter 1 that, for linear diffusion scale-spaces and morphological scale-spaces, it is possible to associate with the evolution time a corresponding spatial scale.

In the nonlinear diffusion case, however, the situation is more complicated. Since the smoothing is nonuniform, one can only define an average measure for the globality of the representation. This can be achieved by taking some Lyapunov function $\Phi(u(t))$ and investigating the expression

$$\Psi(u(t)) \;:=\; \frac{\Phi(f) - \Phi(u(t))}{\Phi(f) - \Phi(Mf)}. \tag{2.39}$$

We observe that $\Psi(t)$ increases from 0 to 1. It gives the average globality of $u(t)$ and its value can be used to measure the distance of $u(t)$ from the initial state f and the final state Mf. Prescribing a certain value for Ψ provides us with an a-posteriori criterion for the stopping time of the nonlinear diffusion process.

[4] Without loss of generality we omit the normalization.

[5] This might make anisotropic diffusion attractive for image compression.

Chapter 3

Semidiscrete diffusion filtering

The goal of this chapter is to study a semidiscrete framework for diffusion scale-spaces where the image is sampled on a finite grid and the scale parameter is continuous. This leads to a system of nonlinear ordinary differential equations (ODEs). We shall investigate conditions under which one can establish similar properties as in the continuous setting concerning well-posedness, extremum principles, average grey level invariance, Lyapunov functionals, and convergence to a constant steady-state. Afterwards we shall discuss whether it is possible to obtain such filters from spatial discretizations of the continuous models that have been investigated in Chapter 2. We will see that there exists a finite stencil on which a difference approximation of the spatial derivatives fits the semidiscrete scale-space framework.

3.1 The general model

A discrete image can be regarded as a vector $f \in \mathbb{R}^N$, $N \geq 2$, whose components f_j, $j = 1,...,N$ represent the grey values at the pixels. We denote the index set $\{1, ..., N\}$ by J. In order to specify our requirements for our semidiscrete filter class we first recall a useful definition of irreducible matrices [241, pp. 18–20].

Definition 1 (Irreducibility). *A matrix $A = (a_{ij}) \in \mathbb{R}^{N \times N}$ is called irreducible if for any $i, j \in J$ there exist $k_0,...,k_r \in J$ with $k_0 = i$ and $k_r = j$ such that $a_{k_p k_{p+1}} \neq 0$ for $p = 0,...,r-1$.*

The semidiscrete problem class (P_s) we are concerned with is defined in the following way:

$$\left.\begin{minipage}{0.8\textwidth}
Let $f \in \mathbb{R}^N$. Find a function $u \in \mathrm{C}^1([0,\infty),\mathbb{R}^N)$ which satisfies the initial value problem
$$\begin{aligned} \frac{du}{dt} &= A(u)\,u, \\ u(0) &= f, \end{aligned}$$
where $A = (a_{ij})$ has the following properties:

(S1) Lipschitz-continuity of $A \in \mathrm{C}(\mathbb{R}^N, \mathbb{R}^{N\times N})$ for every bounded subset of $\mathbb{R}^N$,

(S2) symmetry: $a_{ij}(u) = a_{ji}(u) \quad \forall\, i,j \in J, \quad \forall\, u \in \mathbb{R}^N$,

(S3) vanishing row sums: $\sum_{j\in J} a_{ij}(u) = 0 \quad \forall\, i \in J, \quad \forall\, u \in \mathbb{R}^N$,

(S4) nonnegative off-diagonals: $a_{ij}(u) \geq 0 \quad \forall\, i \neq j, \quad \forall\, u \in \mathbb{R}^N$,

(S5) irreducibility for all $u \in \mathbb{R}^N$.
\end{minipage}\right\}\quad (P_s)$$

Not all of these requirements are necessary for every theoretical result below. (S1) is needed for well-posedness, the proof of a maximum–minimum principle involves (S3) and (S4), while average grey value invariance uses (S2) and (S3). The existence of Lyapunov functionals can be established by means of (S2)–(S4), and strict Lyapunov functionals and the convergence to a constant steady state require (S5) in addition to (S2)–(S4).

This indicates that these properties reveal some interesting parallels to the continuous setting from Chapter 2: In both cases we need smoothness assumptions to ensure well-posedness; (S2) and (S3) correspond to the specific structure of the divergence expression with a symmetric diffusion tensor D, while (S4) and (S5) play a similar role as the nonnegativity of the eigenvalues of D and its uniform positive definiteness, respectively.

3.2 Theoretical results

Before we can establish scale-space results, it is of importance to ensure the existence of a unique solution. This is done in the theorem below which also states the continuous dependence of the solution and a maximum–minimum principle.

Theorem 4 (Well-posedness, extremum principle).
For every $T > 0$ the problem (P_s) has a unique solution $u(t) \in \mathrm{C}^1([0,T], \mathbb{R}^N)$. This solution depends continuously on the initial value and the right-hand side of the ODE system, and it satisfies the extremum principle

$$a \;\le\; u_i(t) \;\le\; b \qquad \forall\, i \in J, \quad \forall\, t \in [0,T], \tag{3.1}$$

where

$$a := \min_{j \in J} f_j, \tag{3.2}$$
$$b := \max_{j \in J} f_j. \tag{3.3}$$

Proof:

(a) **Local existence and uniqueness**

Local existence and uniqueness are proved by showing that our problem satisfies the requirements of the Picard–Lindelöf theorem [254, p. 59].

Let $t_0 := 0$ and $\beta > 0$. Evidently, $\phi(t,u) := \psi(u) := A(u)\, u$ is continuous on

$$B_0 := [0,T] \times \left\{ u \in \mathbb{R}^N \;\middle|\; \|u\|_\infty \le \|f\|_\infty + \beta \right\},$$

since it is a composition of continuous functions. Moreover, by the compactness of B_0 there exists some $c > 0$ with

$$\|\phi(t,u)\|_\infty \le c \qquad \forall\, (t,u) \in B_0.$$

In order to prove existence and uniqueness of a solution of (P_s) in

$$R_0 := \left\{ (t,u) \;\middle|\; t \in [t_0, t_0 + \min(\tfrac{\beta}{c}, T)],\; \|u - f\|_\infty \le \beta \right\} \subset B_0$$

we have to show that $\phi(t,u)$ satisfies a global Lipschitz condition on R_0 with respect to u. However, this follows directly from the fact that A is Lipschitz-continuous on $\{u \in \mathbb{R}^N \mid \|u - f\|_\infty \le \beta\}$.

(b) **Maximum–minimum principle**

We prove only the maximum principle, since the proof for the minimum principle is analogous.

Assume that the problem (P_s) has a unique solution on $[0,\theta]$. First we show that the derivative of the largest component of $u(t)$ is nonpositive for every

$t \in [0, \theta]$. Let $u_k(\vartheta) := \max_{j \in J} u_j(\vartheta)$ for some arbitrary $\vartheta \in [0, \theta]$. If we keep this k fixed we obtain, for $t = \vartheta$,

$$\begin{aligned} \frac{du_k}{dt} &= \sum_{j \in J} a_{kj}(u)\, u_j \\ &= a_{kk}(u)\, u_k + \sum_{j \in J \setminus \{k\}} \underbrace{a_{kj}(u)}_{\geq 0} \underbrace{u_j}_{\leq u_k} \\ &\leq u_k \cdot \sum_{j \in J} a_{kj}(u) \\ &\overset{(S4)}{=} 0. \end{aligned} \tag{3.4}$$

Let us now prove that this implies a maximum principle (cf. [126]).
Let $\varepsilon > 0$ and set

$$u_\varepsilon(t) := u(t) - \begin{pmatrix} \varepsilon t \\ \vdots \\ \varepsilon t \end{pmatrix}.$$

Moreover, let $P := \{p \in J \mid u_{\varepsilon p}(0) = \max_{j \in J} u_{\varepsilon j}(0)\}$. Then, by (3.4),

$$\left(\frac{du_{\varepsilon p}}{dt}\right)(0) \;=\; \underbrace{\left(\frac{du_p}{dt}\right)(0)}_{\leq 0} - \varepsilon \;<\; 0 \qquad \forall\, p \in P. \tag{3.5}$$

By means of

$$\max_{i \in J \setminus P} u_{\varepsilon i}(0) \;<\; \max_{j \in J} u_{\varepsilon j}(0),$$

and the continuity of u there exists some $t_1 \in (0, \theta)$ such that

$$\max_{i \in J \setminus P} u_{\varepsilon i}(t) \;<\; \max_{j \in J} u_{\varepsilon j}(0) \qquad \forall\, t \in [0, t_1). \tag{3.6}$$

Next, let us consider some $p \in P$. Due to (3.5) and the smoothness of u we may find a $\vartheta_p \in (0, \theta)$ with

$$\left(\frac{du_{\varepsilon p}}{dt}\right)(t) \;<\; 0 \qquad \forall\, t \in [0, \vartheta_p).$$

Thus, we have

$$u_{\varepsilon p}(t) \;<\; u_{\varepsilon p}(0) \qquad \forall\, t \in (0, \vartheta_p)$$

and, for $t_2 := \min_{p \in P} \vartheta_p$, it follows that

$$\max_{p \in P} u_{\varepsilon p}(t) \;<\; \max_{j \in J} u_{\varepsilon j}(0) \qquad \forall\, t \in (0, t_2). \tag{3.7}$$

Hence, for $t_0 := \min(t_1, t_2)$, the estimates (3.6) and (3.7) give

$$\max_{j\in J} u_{\varepsilon j}(t) \; < \; \max_{j\in J} u_{\varepsilon j}(0) \qquad \forall\, t \in (0, t_0). \tag{3.8}$$

Now we want to prove that this estimate extends to the case $t \in (0, \theta)$. To this end, assume the opposite is true. Then, by virtue of the intermediate value theorem, there exists some t_3 which is the smallest time in $(0, \theta)$ such that

$$\max_{j\in J} u_{\varepsilon j}(t_3) \; = \; \max_{j\in J} u_{\varepsilon j}(0).$$

Let $u_{\varepsilon k} := \max\limits_{j\in J} u_{\varepsilon j}(t_3)$. Then the minimality of t_3 yields

$$u_{\varepsilon k}(t) < u_{\varepsilon k}(t_3) \qquad \forall\, t \in (0, t_3), \tag{3.9}$$

and inequality (3.4) gives

$$\left(\frac{du_{\varepsilon k}}{dt}\right)(t_3) \; = \; \underbrace{\left(\frac{du_k}{dt}\right)(t_3)}_{\leq 0} - \varepsilon \; < \; 0.$$

Due to the continuity of $\frac{du}{dt}$ there exists some $t_4 \in (0, t_3)$ with

$$\left(\frac{du_{\varepsilon k}}{dt}\right)(t) \; < \; 0 \qquad \forall\, t \in (t_4, t_3]. \tag{3.10}$$

The mean value theorem, however, implies that we find a $t_5 \in (t_4, t_3)$ with

$$\left(\frac{du_{\varepsilon k}}{dt}\right)(t_5) \; = \; \frac{u_{\varepsilon k}(t_3) - u_{\varepsilon k}(t_4)}{t_3 - t_4} \; \overset{(3.9)}{>} \; 0,$$

which contradicts (3.10). Hence, (3.8) must be valid on the entire interval $(0, \theta)$.

Together with $u = \lim\limits_{\varepsilon\to 0} u_\varepsilon$ and the continuity of u this yields the announced maximum principle

$$\max_{j\in J} u_j(t) \; \leq \; \max_{j\in J} u_j(0) \qquad \forall\, t \in [0, \theta].$$

(c) **Global existence and uniqueness**

Global existence and uniqueness follow from local existence and uniqueness when being combined with the extremum principle.

Using the notations and results from (a), we know that the problem (P_s) has a unique solution $u(t)$ for $t \in [t_0, t_0 + \min(\frac{\beta}{c}, T)]$.

Now let $t_1 := t_0 + \min(\frac{\beta}{c}, T)$, $g := u(t_1)$, and consider the problem

$$\begin{aligned} \frac{du}{dt} &= A(u)\,u, \\ u(t_1) &= g. \end{aligned}$$

Clearly, $\phi(t,u) = A(u)u$ is continuous on

$$B_1 := [0,T] \times \left\{ u \in \mathbb{R}^N \;\middle|\; \|u\|_\infty \le \|g\|_\infty + \beta \right\},$$

and by the extremum principle we know that $B_1 \subset B_0$. Hence,

$$\|\phi(t,u)\|_\infty \le c \qquad \forall\,(t,u) \in B_1.$$

with the same c as in (a). Using the same considerations as in (a) one shows that ϕ is Lipschitz-continuous on

$$R_1 := \left\{ (t,u) \;\middle|\; t \in [t_1, t_1 + \min(\tfrac{\beta}{c}, T)],\ \|u - g\|_\infty \le \beta \right\}.$$

Hence, the considered problem has a unique solution on $[t_1, t_1 + \min(\frac{\beta}{c}, T)]$. Therefore, (P_s) reveals a unique solution on $[0, \min(\frac{2\beta}{c}, T)]$, and, by iterating this reasoning, the existence of a unique solution can be extended to the entire interval $[0,T]$. As a consequence, the extremum principle is valid on $[0,T]$ as well.

(d) **Continuous dependence**

Let $u(t)$ be the solution of

$$\begin{aligned} \frac{du}{dt} &= \phi(t,u), \\ u(0) &= f \end{aligned}$$

for $t \in [0,T]$ and $\phi(u,t) = \psi(u) = A(u)u$. In order to show that $u(t)$ depends continuously on the initial data and the right-hand side of the ODE system, it is sufficient to prove that $\phi(t,u)$ is continuous, and that there exists some $\alpha > 0$ such that $\phi(t,u)$ satisfies a global Lipschitz condition on

$$S_\alpha := \left\{ (t,v) \;\middle|\; t \in [0,T],\ \|v - u\|_\infty \le \alpha \right\}.$$

with respect to its second argument. In this case the results in [245, p. 93] ensure that for every $\varepsilon > 0$ there exists a $\delta > 0$ such that the solution $\tilde{u}$ of the perturbed problem

$$\begin{aligned} \frac{d\tilde{u}}{dt} &= \tilde{\phi}(t,\tilde{u}), \\ \tilde{u}(0) &= \tilde{f} \end{aligned}$$

with continuous $\tilde{\phi}$ and

$$\begin{aligned} \|\tilde{f} - f\|_\infty \;&<\; \delta, \\ \|\tilde{\phi}(t,v) - \phi(t,v)\|_\infty \;&<\; \delta \quad \text{for} \quad \|v-u\|_\infty \;<\; \alpha \end{aligned}$$

exists in $[0,T]$ and satisfies the inequality

$$\|\tilde{u}(t) - u(t)\|_\infty < \varepsilon.$$

Similar to the the local existence and uniqueness proof, the global Lipschitz condition on S_α follows direcly from the fact that A is Lipschitz-continuous on $\{v \in \mathbb{R}^N \mid \|v-u\|_\infty \le \alpha\}$. □

3.3 Scale-space properties

It is evident that properties such as grey level shift invariance or reverse contrast invariance are automatically satisfied by every consistent semidiscrete approximation of the continuous filter class (P_c). On the other hand, translation invariance only makes sense for translations in grid direction with multiples of the grid size, and isometry invariance is satisfied for consistent schemes up to an discretization error. So let us focus on average grey level invariance now.

Proposition 2 (Conservation of average grey value).
The average grey level

$$\mu \;:=\; \frac{1}{N} \sum_{j \in J} f_j \tag{3.11}$$

is not affected by the semidiscrete diffusion filter:

$$\frac{1}{N} \sum_{j \in J} u_j(t) = \mu \qquad \forall\, t \ge 0. \tag{3.12}$$

Proof:
By virtue of (S2) and (S3) we have $\sum_{j \in J} a_{jk}(u) = 0$ for all $k \in J$. Thus, for $t \ge 0$,

$$\sum_{j \in J} \frac{du_j}{dt} \;=\; \sum_{j \in J} \sum_{k \in J} a_{jk}(u)\, u_k \;=\; \sum_{k \in J} \Big(\sum_{j \in J} a_{jk}(u) \Big) u_k \;=\; 0,$$

which shows that $\sum_{j \in J} u_j(t)$ is constant on $[0,\infty)$ and concludes the proof. □

This property is in complete accordance with the result for the continuous filter class.

Similar to the continuous setting, it is possible to find a large class of Lyapunov functionals which establish smoothing scale-space properties and ensure that the image tends to a constant steady-state with the same average grey level as the initial image.

Theorem 5 (Lyapunov functionals and behaviour for $t \to \infty$).
Let $u(t)$ be the solution of (P_s), let a, b, and μ be defined as in (3.2), (3.3), and (3.11), respectively, and let $c := (\mu, \mu, ..., \mu)^\top \in \mathbb{R}^N$.
Then the following properties are valid:

(a) (Lyapunov functionals)
For all $r \in \mathrm{C}^1[a,b]$ with increasing r' on $[a,b]$, the function

$$V(t) := \Phi(u(t)) := \sum_{i \in J} r(u_i(t))$$

is a Lyapunov functional:

(i) $\Phi(u) \geq \Phi(c)$ *for all* $t \geq 0$.

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

Moreover, if r' is strictly increasing on $[a,b]$, then $V(t) = \Phi(u(t))$ is a strict Lyapunov functional:

(iii) $\Phi(u) = \Phi(c) \quad \Longleftrightarrow \quad u = c$

(iv) $V'(t) = 0 \quad \Longleftrightarrow \quad u = c$

(b) (Convergence)
$\lim_{t \to \infty} u(t) = c.$

Proof:

(a) (i) Since r' is increasing on $[a,b]$ we know that r is convex on $[a,b]$. Average grey level invariance and this convexity yield, for all $t \geq 0$,

$$\begin{aligned} \Phi(c) &= \sum_{i=1}^N r\left(\sum_{j=1}^N \frac{1}{N} u_j\right) \\ &\leq \sum_{i=1}^N \left(\frac{1}{N} \sum_{j=1}^N r(u_j)\right) \\ &= \sum_{j=1}^N r(u_j) \\ &= \Phi(u). \end{aligned} \tag{3.13}$$

(ii) Since $u \in \mathrm{C}^1([0,\infty), \mathbb{R}^N)$ and $r \in \mathrm{C}^1[a,b]$, it follows that $V \in \mathrm{C}^1[0,\infty)$. Using the prerequisites (S2) and (S3) we get

$$\begin{aligned} V'(t) &= \sum_{i=1}^{N} \frac{du_i}{dt}\, r'(u_i) \\ &\overset{(S3)}{=} \sum_{i=1}^{N}\sum_{j=1}^{N} a_{ij}(u)\ (u_j - u_i)\ r'(u_i) \\ &= \sum_{i=1}^{N} \left(\sum_{j=i+1}^{N} + \sum_{j=1}^{i-1} \right) a_{ij}(u)\ (u_j - u_i)\ r'(u_i) \\ &= \sum_{i=1}^{N}\sum_{k=1}^{N-i} a_{i,i+k}(u)\ (u_{i+k} - u_i)\ r'(u_i) \\ &+ \sum_{i=1}^{N}\sum_{k=1}^{N-i} a_{i+k,i}(u)\ (u_i - u_{i+k})\ r'(u_{i+k}) \\ &\overset{(S2)}{=} \sum_{i=1}^{N}\sum_{k=1}^{N-i} a_{i,i+k}(u)\ (u_{i+k} - u_i)\ \Big(r'(u_i) - r'(u_{i+k})\Big). \qquad (3.14) \end{aligned}$$

Since r' is increasing, we always have

$$(u_{i+k} - u_i)\ \Big(r'(u_i) - r'(u_{i+k})\Big)\ \le\ 0.$$

With this and (S4), equation (3.14) implies that $V'(t) \le 0$ for $t \ge 0$.

(iii) Let us first prove that equality in the estimate (3.13) implies that all components of u are equal.

To this end, suppose that $u_{i_0} := \min_i u_i < \max_j u_j =: u_{j_0}$ and let

$$\eta\ :=\ \sum_{\substack{j=1 \\ j \ne j_0}}^{N} \frac{\frac{1}{N}\, u_j}{1 - \frac{1}{N}}.$$

Then, $\eta < u_{j_0}$. Since r' is strictly increasing on $[a,b]$, we know that r is strictly convex. Hence, we get

$$\begin{aligned} r\left(\sum_{i=1}^{N} \frac{1}{N}\, u_j \right) &= r\left(\frac{1}{N}\, u_{j_0}\ +\ \left(1 - \frac{1}{N}\right)\ \eta \right) \\ &< \frac{1}{N}\, r(u_{j_0})\ +\ \left(1 - \frac{1}{N}\right)\ r(\eta) \\ &\le \frac{1}{N}\, r(u_{j_0})\ +\ \sum_{\substack{j=1 \\ j \ne j_0}}^{N} \frac{1}{N}\, r(u_j) \\ &= \sum_{j=1}^{N} \frac{1}{N}\, r(u_j). \end{aligned}$$

This shows that equality in (3.13) implies that $u_1 = ... = u_N$. By virtue of the grey level shift invariance we conclude that $u = c$.

Conversely, it is trivial that $\Phi(u) = \Phi(c)$ for $u = c$.

(iv) Let $V'(t) = 0$. From (3.14) we have

$$0 \;=\; V'(t) \;=\; \sum_{i=1}^{N}\sum_{k=1}^{N-i} \underbrace{a_{i,i+k}(u)\,(u_{i+k}-u_i)\,\Big(r'(u_i)-r'(u_{i+k})\Big)}_{\leq 0},$$

and by virtue of the symmetry of $A(u)$ it follows that

$$a_{ij}(u)\,(u_j-u_i)\,\Big(r'(u_i)-r'(u_j)\Big) \;=\; 0 \qquad \forall\, i,j \in J. \tag{3.15}$$

Now consider two arbitrary $i_0, j_0 \in J$. The irreducibility of $A(u)$ implies that there exist $k_0,...,k_r \in J$ with $k_0 = i_0$, $k_r = j_0$, and

$$a_{k_p k_{p+1}}(u) \neq 0, \qquad p = 0, ..., r-1.$$

As r' is strictly increasing we have, for $p = 0,...,r-1$,

$$(u_{k_p}-u_{k_{p+1}})\,\Big(r'(u_{k_{p+1}})-r'(u_{k_p})\Big) \;=\; 0 \quad\Longleftrightarrow\quad u_{k_p} = u_{k_{p+1}}.$$

From this and (3.15) we get

$$u_{i_0} = u_{k_0} = u_{k_1} = ... = u_{k_r} = u_{j_0}.$$

Since i_0 and j_0 are arbitrary, we obtain $u_i = \text{const.}$ for all $i \in J$, and the average grey level invariance gives $u = c$. This proves the first implication.

Conversely, let $u_i = \text{const.}$ for all $i \in J$. Then from the representation (3.14) we immediately conclude that $V'(t) = 0$.

(b) The convergence proof is based on classical Lyapunov reasonings, see e.g. [116] for an introduction to these techniques.

Consider the Lyapunov functional $V(t) := \Phi(u(t)) := |u(t) - c|^2$, which results from the choice $r(s) := (s–\mu)^2$. Since $V(t)$ is decreasing and bounded from below by 0, we know that $\lim_{t\to\infty} V(t) =: \eta$ exists and $\eta \geq 0$.

Now assume that $\eta > 0$.

Since $|u(t)-c|$ is bounded from above by $\alpha := |f-c|$ we have

$$|u(t)-c| \leq \alpha \qquad \forall\, t \geq 0. \tag{3.16}$$

By virtue of $\Phi(x) = |x-c|^2$ we know that, for $\beta \in (0, \sqrt{\eta})$,

$$\Phi(x) < \eta \qquad \forall\, x \in \mathbb{R}^N, \quad |x-c| < \beta.$$

Let w.l.o.g. $\beta < \alpha$. Since $\Phi(u(t)) \geq \eta$ we conclude that

$$|u(t)-c| \geq \beta \qquad \forall\, t \geq 0. \tag{3.17}$$

So from (3.16) and (3.17) we have

$$u(t) \in \{x \in \mathbb{R}^N \mid \beta \leq |x-c| \leq \alpha\} =: S \qquad \forall\, t \geq 0.$$

By (a)(ii),(iv), the compactness of S, and $\beta > 0$ there exists some $M > 0$ such that

$$V'(t) \leq -M \qquad \forall\, t \geq 0.$$

Therefore, it follows

$$V(t) \;=\; V(0) + \int_0^t V'(\theta)\, d\theta \;\leq\; V(0) - tM$$

which implies $\lim\limits_{t\to\infty} V(t) = -\infty$ and, thus, contradicts (a)(i).

Hence the assumption $\eta > \rho$ is wrong and we must have $\eta = \rho$.

According to (a)(iii) this yields $\lim\limits_{t\to\infty} u(t) = c$. □

As in the continuous case, we can consider the Lyapunov functions associated with $r(s) := |s|^p$, $r(s) := (s-\mu)^{2n}$ and $r(s) := s \ln s$, respectively, and obtain the following corollary.

Corollary 2 (Special Lyapunov functionals).
Let u be the solution of (P_s) and a and μ be defined as in (3.2) and (3.11). Then the following functions are decreasing for $t \in [0, \infty)$:

(a) $\quad \|u(t)\|_p \quad$ *for all* $p \geq 2$.

(b) $\quad M_{2n}[u(t)] \;:=\; \frac{1}{N} \sum\limits_{j=1}^{N} (u_j(t) - \mu)^{2n} \quad$ *for all* $n \in \mathbb{N}$.

(c) $\quad H[u(t)] \;:=\; \sum\limits_{j=1}^{N} u_j(t)\, \ln(u_j(t)), \quad$ *if* $a > 0$.

Since all p-norms ($p \geq 2$) and all central moments are decreasing, while the discrete entropy

$$S[u(t)] \;:=\; -\sum_{j=1}^{N} u_j(t)\, \ln(u_j(t)) \tag{3.18}$$

is increasing with respect to t, we observe that the semidiscrete setting reveals smoothing scale-space properties which are closely related to the continuous case.

3.4 Relation to continuous models

Let us now investigate whether it is possible to use spatial discretizations of the continuous filter class (P_c) in order to find semidiscrete diffusion models satisfying (S1)–(S5). First we shall verify that this is easily done for isotropic models. In the anisotropic case, however, the mixed derivative terms make it more difficult to ensure nonnegative off-diagonal elements. We shall present a constructive existence proof that for a sufficiently large stencil it is always possible to find such a nonnegative discretization. This concept is illustrated by investigating the situation on a (3×3)-stencil in detail.

3.4.1 Isotropic case

Let the rectangle $\Omega = (0,a_1)\times(0,a_2)$ be discretized by a grid of $N = n_1 \cdot n_2$ pixels such that a pixel (i,j) with $1 \le i \le n_1$ and $1 \le j \le n_2$ represents the location (x_i, y_j) where

$$x_i = (i - \tfrac{1}{2})\, h_1, \tag{3.19}$$

$$y_j = (j - \tfrac{1}{2})\, h_2, \tag{3.20}$$

and the grid sizes h_1, h_2 are given by $h_1 := a_1/n_1$ and $h_2 := a_2/n_2$, respectively.

These pixels can be numbered by means of an arbitrary bijection

$$p: \quad \{1,...,n_1\} \times \{1,...,n_2\} \quad \to \quad \{1,...,N\}. \tag{3.21}$$

Thus, pixel (i,j) is represented by a single index $p(i,j)$.

Let us now show that a standard FD space discretization of an isotropic variant of (P_c) leads to a semidiscrete filter satisfying the requirements (S1)–(S5). To this end, we replace the diffusion tensor $D(J_\rho(\nabla u_\sigma))$ by some scalar-valued function $g(J_\rho(\nabla u_\sigma))$. The structure tensor requires the calculations of convolutions with ∇K_σ and K_ρ, respectively. In the spatially discrete case this comes down to specific vector–matrix multiplications. For this reason, we may approximate the structure tensor by some matrix $H(u) = (h_{ij}(u))$ where $H \in C^\infty(\mathbb{R}^N, \mathbb{R}^{2\times 2})$.

Let us consider some pixel $k = p(i,j)$. We define four expressions $g_{kE}(u)$, $g_{kW}(u)$, $g_{kN}(u)$, $g_{kS}(u)$, which characterize the interactions of pixel (i,j) with its four nearest neighbours and take into account the homogeneous Neumann boundary condition:

$$g_{kE}(u) := \begin{cases} \frac{g((H(u))_{p(i,j)}) + g((H(u))_{p(i+1,j)})}{2} & (i \in \{1,...,n_1-1\}) \\ 0 & (\text{else}), \end{cases} \tag{3.22}$$

$$g_{kW}(u) := \begin{cases} \frac{g((H(u))_{p(i,j)}) + g((H(u))_{p(i-1,j)})}{2} & (i \in \{2,...,n_1\}) \\ 0 & (\text{else}), \end{cases} \tag{3.23}$$

$$g_{kN}(u) \;:=\; \begin{cases} \frac{g((H(u))_{p(i,j)})+g((H(u))_{p(i,j+1)})}{2} & (j \in \{1,...,n_2-1\}) \\ 0 & (\text{else}), \end{cases} \tag{3.24}$$

$$g_{kS}(u) \;:=\; \begin{cases} \frac{g((H(u))_{p(i,j)})+g((H(u))_{p(i,j-1)})}{2} & (j \in \{2,...,n_2\}) \\ 0 & (\text{else}). \end{cases} \tag{3.25}$$

When defining $A(u) = (a_{kl}(u))$ by means of

$$a_{kl}(u) \;:=\; \begin{cases} \frac{g_{kE}(u)}{h_1^2} & (l = p(i+1,j)), \\ \frac{g_{kW}(u)}{h_1^2} & (l = p(i-1,j)), \\ \frac{g_{kN}(u)}{h_2^2} & (l = p(i,j+1)), \\ \frac{g_{kS}(u)}{h_2^2} & (l = p(i,j-1)), \\ -\frac{g_{kE}(u)+g_{kW}(u)}{h_1^2} - \frac{g_{kN}(u)+g_{kS}(u)}{h_2^2} & (l = p(i,j)), \\ 0 & (\text{else}), \end{cases} \tag{3.26}$$

a semidiscrete version of the isotropic diffusion filter can be written as

$$\frac{du}{dt} \;=\; A(u)\, u. \tag{3.27}$$

Let us now verify that (S1)–(S5) are fulfilled.

Since $H \in \mathrm{C}^\infty(\mathbb{R}^N, \mathbb{R}^{2\times 2})$ and $g \in \mathrm{C}^\infty(\mathbb{R}^{2\times 2})$, we have $A \in \mathrm{C}^\infty(\mathbb{R}^N, \mathbb{R}^{N\times N})$. This proves (S1).

The symmetry of A follows directly from (3.26): For instance, if $l = p(i+1,j)$, we obtain

$$\begin{aligned} a_{kl}(u) \;&=\; \frac{g_{kE}(u)}{h_1^2} \\ &=\; \begin{cases} \frac{g((H(u))_{p(i,j)})+g((H(u))_{p(i+1,j)})}{2h_1^2} & (i \in \{1,...,n_1-1\}), \\ 0 & (\text{else}) \end{cases} \\ &=\; \begin{cases} \frac{g((H(u))_{p(i+1,j)})+g((H(u))_{p((i+1)-1,j)})}{2h_1^2} & (i+1 \in \{2,...,n_1\}), \\ 0 & (\text{else}) \end{cases} \\ &=\; \frac{g_{lW}(u)}{h_1^2} \;=\; a_{lk}(u). \end{aligned}$$

By the construction of A it is also evident that all row sums vanish, i.e. (S3) is satisfied. Moreover, since g is positive, it follows that $a_{kl} \geq 0$ for all $k \neq l$ and, thus, (S4) holds.

In order to show that A is irreducible, let us consider two arbitrary pixels $s_1 = p(i,j)$ and $s_2 = p(n,m)$. Then we have to find k_0,...,$k_r \in J$ with $k_0 = s_1$ and

$k_r = s_2$ such that $a_{k_q k_{q+1}} \neq 0$ for $q = 0,...,r-1$. To this end, let $r := |n-i| + |m-j|$ and

$$k_q \;:=\; \begin{cases} p(i + q \cdot \operatorname{sgn}(n-i), j) & (q = 0, ..., |n-i|) \\ p(n, j + (q-|n-i|) \cdot \operatorname{sgn}(m-j)) & (q = |n-i|+1, ..., r) \end{cases} \tag{3.28}$$

Then, using (3.26) and the positivity of g we obtain $k_0 = s_1$, $k_r = s_2$, and $a_{k_q k_{q+1}} > 0$ for $q = 0,..., r-1$. This establishes (S5).

Remarks:

(a) We observe that (S1)–(S5) are properties which are valid for all arbitrary pixel numberings.

(b) The filter class (P_c) is not the only family which leads to semidiscrete filters satisfying (S1)–(S5). Interestingly, a semidiscrete version of the Perona–Malik filter (which is claimed to be ill-posed in the continuous setting) on a fixed grid also satisfies (S1)–(S5) and, thus, reveals all the beforementioned well-posedness and scale-space properties. This is due to the fact that the extremum principle limits the modulus of discrete gradient approximations. Hence, the spatial discretization implicitly causes a regularization.

3.4.2 Anisotropic case

If one wishes to transfer the results from the isotropic case to the general anisotropic setting the main difficulty arises from the fact that, due to the mixed derivative expressions, it is not obvious how to ensure (S4), the nonnegativity of all off-diagonal elements of $A(u)$. The theorem below states that this is always possible for a sufficiently large stencil.

Theorem 6 (Existence of a nonnegative discretization).
Let $D \in \mathbb{R}^{2\times 2}$ be symmetric positive definite with a spectral condition number κ. Then there exists some $m(\kappa) \in \mathbb{N}$ such that $\operatorname{div}(D\,\nabla u)$ *reveals a second-order nonnegative FD discretization on a $(2m+1)\times(2m+1)$-stencil.*

Proof:
Let us consider some $m \in \mathbb{N}$ and the corresponding $(2m+1)\times(2m+1)$-stencil. The "boundary pixels" of this stencil define $4m$ principal orientations $\beta_i \in (-\frac{\pi}{2}, \frac{\pi}{2}]$, $i = -2m+1, ..., 2m$ according to

$$\beta_i \;:=\; \begin{cases} \arctan\left(\frac{ih_2}{mh_1}\right) & (|i| \le m), \\ \operatorname{arccot}\left(\frac{(2m-i)h_1}{mh_2}\right) & (m < i \le 2m), \\ \operatorname{arccot}\left(\frac{(i-2m)h_1}{mh_2}\right) & (-2m+1 \le i < -m). \end{cases}$$

Now let $J_m := \{1, ..., 2m-1\}$ and define a partition of $(-\frac{\pi}{2}, \frac{\pi}{2}]$ into $4m-2$ subintervals I_i, $|i| \in J_m$:

$$(-\tfrac{\pi}{2}, \tfrac{\pi}{2}] = \bigcup_{i=-2m+1}^{-1} \underbrace{(\theta_i, \theta_{i+1}]}_{=:I_i} \cup \bigcup_{i=1}^{2m-1} \underbrace{(\theta_{i-1}, \theta_i]}_{=:I_i},$$

where

$$\theta_i \;:=\; \begin{cases} 0 & (i = 0), \\ \frac{1}{2} \arctan\left(\frac{2}{\cot\beta_i - \tan\beta_{i+1}}\right) & (i \in \{1, ..., 2m-2\},\ \beta_i + \beta_{i+1} < \frac{\pi}{2}), \\ \frac{\pi}{4} & (i \in \{1, ..., 2m-2\},\ \beta_i + \beta_{i+1} = \frac{\pi}{2}), \\ \frac{\pi}{2} + \frac{1}{2} \arctan\left(\frac{2}{\cot\beta_i - \tan\beta_{i+1}}\right) & (i \in \{1, ..., 2m-2\},\ \beta_i + \beta_{i+1} > \frac{\pi}{2}), \\ \frac{\pi}{2} & (i = 2m-1), \end{cases}$$

and

$$\theta_i \;:=\; -\theta_{-i} \qquad (i \in \{-2m+1, ..., -1\}).$$

It is not hard to verify that $\beta_i \in I_i$ for $|i| \in J_m$.

Let $\lambda_1 \geq \lambda_2 > 0$ be the eigenvalues of D with corresponding eigenvectors $(\cos\psi, \sin\psi)^\top$ and $(-\sin\psi, \cos\psi)^\top$, where $\psi \in (-\frac{\pi}{2}, \frac{\pi}{2}]$. Now we show that for a suitable m there exists a stencil direction β_k, $|k| \in J_m$ such that the splitting

$$\operatorname{div}(D\,\nabla u) \;=\; \partial_{e_{\beta_0}}\Big(\alpha_0\, \partial_{e_{\beta_0}} u\Big) + \partial_{e_{\beta_k}}\Big(\alpha_k\, \partial_{e_{\beta_k}} u\Big) + \partial_{e_{\beta_{2m}}}\Big(\alpha_{2m}\, \partial_{e_{\beta_{2m}}} u\Big) \tag{3.29}$$

with $e_{\beta_i} := (\cos\beta_i, \sin\beta_i)^\top$ reveals nonnegative "directional diffusivities" α_0, α_k, α_{2m} along the stencil orientations β_0, β_k, β_{2m}. This can be done by proving the following properties:

(a) Let $\psi \in I_k$ and $D = \binom{a\ b}{b\ c}$. Then a nonnegative splitting of type (3.29) is possible if

$$\min\Big(a - b\,\cot\beta_k,\ c - b\,\tan\beta_k\Big) \;\geq\; 0. \tag{3.30}$$

(b) Inequality (3.30) is satisfied for

$$\frac{\lambda_1}{\lambda_2} \;\leq\; \min\Big(\cot(\rho_k - \beta_k)\,\tan\rho_k,\ \cot(\beta_k - \eta_k)\,\cot\eta_k\Big) \;=:\; \kappa_{k,m} \tag{3.31}$$

with

$$\rho_k \;:=\; \begin{cases} \theta_k & (|k| \in \{1, ..., 2m-2\}), \\ \frac{1}{2}(\theta_k + \beta_k) & (|k| = 2m-1), \end{cases}$$

$$\eta_k \;:=\; \begin{cases} \frac{1}{2}\beta_k & (|k| = 1), \\ \theta_{k-1} & (|k| \in \{2, ..., 2m-1\}). \end{cases}$$

(c) $\lim\limits_{m\to\infty}\Big(\min\limits_{|i|\in J_m}\kappa_{i,m}\Big) \;=\; \infty.$

Once these assertions are proved a nonnegative second-order discretization of (3.29) arises in a natural way, as we shall see at the end of this chapter. So let us now verify (a)–(c).

(a) In order to use subsequent indices, let $\varphi_0 := 0$, $\varphi_1 := \beta_k$ where $\psi \in I_k$, and $\varphi_2 := \frac{\pi}{2}$. Furthermore, let $\gamma_0 := \alpha_0$, $\gamma_1 := \alpha_k$, and $\gamma_2 := \alpha_{2m}$. Then (3.29) requires that

$$
\begin{aligned}
\operatorname{div}\left(\begin{pmatrix} a & b \\ b & c \end{pmatrix}\nabla u\right) \;&=\; \sum_{i=0}^{2}\frac{\partial}{\partial e_{\varphi_i}}\left(\gamma_i\frac{\partial u}{\partial e_{\varphi_i}}\right)\\
=\;& \frac{\partial}{\partial x}\sum_{i=0}^{2}\cos\varphi_i\Big(\gamma_i(u_x\cos\varphi_i+u_y\sin\varphi_i)\Big)\\
+\;& \frac{\partial}{\partial y}\sum_{i=0}^{2}\sin\varphi_i\Big(\gamma_i(u_x\cos\varphi_i+u_y\sin\varphi_i)\Big)\\
=\;& \operatorname{div}\left(\begin{pmatrix} \sum_{i=0}^{2}\gamma_i\cos^2\varphi_i & \sum_{i=0}^{2}\gamma_i\sin\varphi_i\cos\varphi_i \\ \sum_{i=0}^{2}\gamma_i\sin\varphi_i\cos\varphi_i & \sum_{i=0}^{2}\gamma_i\sin^2\varphi_i \end{pmatrix}\nabla u\right).
\end{aligned}
$$

By comparing the coefficients and using the definition of φ_0, φ_1 and φ_2 we obtain the linear system

$$
\begin{pmatrix} 1 & \cos^2\beta_k & 0 \\ 0 & \sin\beta_k\cos\beta_k & 0 \\ 0 & \sin^2\beta_k & 1 \end{pmatrix}\begin{pmatrix}\gamma_0\\ \gamma_1\\ \gamma_2\end{pmatrix} = \begin{pmatrix} a\\ b\\ c\end{pmatrix}
$$

which has the unique solution

$$\gamma_0 \;=\; a - b\,\cot\beta_k, \tag{3.32}$$

$$\gamma_1 \;=\; \frac{b}{\sin\beta_k\cos\beta_k}, \tag{3.33}$$

$$\gamma_2 \;=\; c - b\,\tan\beta_k. \tag{3.34}$$

From the structure of the eigenvalues and eigenvectors of D it is easily seen that

$$b = (\lambda_1 - \lambda_2)\,\sin\psi\cos\psi.$$

Now, $\lambda_1 - \lambda_2 \geq 0$, and since $\psi, \beta_k \in I_k$ we conclude that ψ and β_k belong to the same quadrant. Thus, γ_1 is always nonnegative. In order to satisfy the nonnegativity of γ_0 and γ_2 we need that

$$\min\Big(a - b\,\cot\beta_k,\; c - b\,\tan\beta_k\Big) \;\geq\; 0.$$

(b) Let $\frac{\lambda_1}{\lambda_2} \le \kappa_{k,m}$ and consider the case $0 < \beta_k < \frac{\pi}{2}$. By defining

$$\begin{aligned} B(\varphi) &:= \cos^2\varphi - \sin\varphi\cos\varphi\cot\beta_k, \\ C(\varphi) &:= \sin^2\varphi + \sin\varphi\cos\varphi\cot\beta_k \end{aligned}$$

we get

$$\frac{\lambda_1}{\lambda_2} \le \cot(\rho_k - \beta_k)\,\tan\rho_k = -\frac{C(\rho_k)}{B(\rho_k)} = \min_{\varphi\in(\beta_k,\theta_k)}\left(-\frac{C(\varphi)}{B(\varphi)}\right).$$

Since $B(\varphi) < 0$ on $(\beta_k, \frac{\pi}{2})$ we have

$$\lambda_1 B(\varphi) + \lambda_2 C(\varphi) \ge 0 \qquad \forall\, \varphi \in (\beta_k, \theta_k). \tag{3.35}$$

Because of

$$\begin{aligned} &B(\varphi) \ge 0 \qquad \forall\, \varphi \in [-\tfrac{\pi}{2}, \beta_k], \\ &C(\varphi) \ge 0 \qquad \forall\, \varphi \in [0, \tfrac{\pi}{2}], \end{aligned}$$

and the continuity of $B(\varphi)$ and $C(\varphi)$ we may extend (3.35) to the entire interval $I_k = (\theta_{k-1}, \theta_k]$. In particular, since $\psi \in I_k$, we have

$$\begin{aligned} 0 &\le \lambda_1 B(\psi) + \lambda_2 C(\psi) \\ &= (\lambda_1\cos^2\psi + \lambda_2\sin^2\psi) - (\lambda_1 - \lambda_2)\,\sin\psi\cos\psi\cot\beta_k. \end{aligned}$$

By the representation

$$\begin{aligned} \begin{pmatrix} a & b \\ b & c \end{pmatrix} &= \begin{pmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} \cos\psi & \sin\psi \\ -\sin\psi & \cos\psi \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1\cos^2\psi + \lambda_2\sin^2\psi & (\lambda_1-\lambda_2)\,\sin\psi\cos\psi \\ (\lambda_1-\lambda_2)\,\sin\psi\cos\psi & \lambda_1\sin^2\psi + \lambda_2\cos^2\psi \end{pmatrix} \end{aligned}$$

we recognize that this is just the desired condition

$$a - b\,\cot\beta_k \ge 0. \tag{3.36}$$

For the case $-\frac{\pi}{2} < \beta_k < 0$ a similar reasoning can be applied leading also to (3.36).

In an analogous way one verifies that

$$\frac{\lambda_1}{\lambda_2} \le \cot(\beta_k - \eta_k)\,\cot\eta_k \qquad \Longrightarrow \qquad c - b\,\tan\beta_k \ge 0.$$

(c) Let us first consider the case $1 \le i \le 2m-2$. Then, $\rho_i = \theta_i$, and the definition of θ_i implies that

$$\cot\rho_i - \tan\rho_i \;=\; \cot\beta_i - \tan\beta_{i+1}.$$

Solving for $\cot\rho_i$ and $\tan\rho_i$, respectively, yields

$$\begin{aligned}
\cot\rho_i &= \frac{1}{2}\left(\cot\beta_i - \tan\beta_{i+1} + \sqrt{(\cot\beta_i - \tan\beta_{i+1})^2 + 4}\right),\\
\tan\rho_i &= \frac{1}{2}\left(-\cot\beta_i + \tan\beta_{i+1} + \sqrt{(\cot\beta_i - \tan\beta_{i+1})^2 + 4}\right).
\end{aligned}$$

By means of these results we obtain

$$\cot(\rho_i - \beta_i)\,\tan\rho_i \;=\; \frac{\cot\beta_i + \tan\rho_i}{\cot\beta_i - \cot\rho_i} \;=\; 1 + \frac{2}{\sqrt{\frac{(\cot\beta_i + \tan\beta_{i+1})^2}{(\cot\beta_i - \tan\beta_{i+1})^2 + 4}} - 1}.$$

Let us now assume that $1 \le i \le m-1$. Then we have

$$\begin{aligned}
\tan\beta_{i+1} &= \frac{(i+1)\,h_2}{mh_1},\\
\cot\beta_i &= \frac{mh_1}{ih_2}.
\end{aligned}$$

This gives

$$\frac{(\cot\beta_i + \tan\beta_{i+1})^2}{(\cot\beta_i - \tan\beta_{i+1})^2 + 4} \;=\; \frac{1}{1 - \frac{4m^2 i}{\left(m^2\frac{h_1}{h_2} + i(i+1)\frac{h_2}{h_1}\right)^2}} \;=:\; \frac{1}{1 - g_m(i)} \;=:\; f_m(i).$$

For $m > \frac{1}{2}\frac{h_2}{h_1}$ the function $g_m(x)$ is bounded and attains its global maximum in

$$x_m \;:=\; -\tfrac{1}{6} + \tfrac{1}{6}\sqrt{1 + 12\,m^2\,\tfrac{h_1^2}{h_2^2}}.$$

Thus, for $1 \le i \le m-1$,

$$g_m(i) \;\le\; g_m(x_m) \;\to\; 0^+ \quad \text{for} \quad m \to \infty,$$

which yields

$$f_m(i) \;\le\; \frac{1}{1 - g_m(x_m)} \;\to\; 1^+ \quad \text{for} \quad m \to \infty.$$

This gives

$$\min_{1 \le i \le m-1}\Big(\cot(\rho_i - \beta_i)\,\tan\rho_i\Big) \;\ge\; 1 + \frac{2}{\sqrt{f_m(x_m)} - 1} \;\to\; \infty \quad \text{for} \quad m \to \infty.$$

For $m \le i \le 2m-2$ similar calculations show that by means of

$$\begin{aligned} \tan\beta_{i+1} &= \frac{mh_2}{(2m-i-1)\,h_1}, \\ \cot\beta_i &= \frac{(2m-i)\,h_1}{mh_2} \end{aligned}$$

one obtains

$$\min_{m\le i\le 2m-2}\Big(\cot(\rho_i-\beta_i)\,\tan\rho_i\Big) \to \infty \quad \text{for} \quad m\to\infty.$$

For $i = 2m-1$ we have

$$\begin{aligned} \cot\Big(\rho_{2m-1}-\beta_{2m-1}\Big)\,\tan\rho_{2m-1} &= \cot\left(\frac{\pi}{4}-\frac{\beta_{2m-1}}{2}\right)\tan\left(\frac{\pi}{4}+\frac{\beta_{2m-1}}{2}\right) \\ &= \tan^2\left(\frac{\pi}{4}+\frac{\beta_{2m-1}}{2}\right) \\ &\to \infty \quad \text{for} \quad m\to\infty. \end{aligned}$$

It is not hard to verify that for $-2m+1 \le i \le -1$ the preceding results carry over. Hence,

$$\lim_{m\to\infty}\left(\min_{|i|\in J_m}\Big(\cot(\rho_i-\beta_i)\,\tan\rho_i\Big)\right) = \infty. \tag{3.37}$$

Now, in a similar way as above, one establishes that

$$\lim_{m\to\infty}\left(\min_{|i|\in J_m}\Big(\cot(\beta_i-\eta_i)\,\cot\eta_i\Big)\right) = \infty. \tag{3.38}$$

From (3.37) and (3.38) we finally end up with the assertion

$$\lim_{m\to\infty}\Big(\min_{|i|\in J_m}\kappa_{i,m}\Big) = \infty.$$

□

Remarks:

(a) We observe that the preceding existence proof is constructive. Moreover, only three directions are sufficient to guarantee a nonnegative directional splitting. Thus, unless m is very small, most of the stencil coefficients can be set to zero.

(b) Especially for large m, a $(2m+1)\times(2m+1)$-stencil reveals much more directions than those $4m$ that are induced by the $8m$ "boundary pixels". Therefore, even if we use only 3 directions, we may expect to find stricter estimates than those given in the proof. These estimates might be improved further by admitting more than 3 directions.

(c) For a specified diffusion tensor function D it is possible to give a-priori estimates for the required stencil size: using the extremum principle it is not hard to show that

$$|\nabla u_\sigma(x,t)| \;=\; |(\nabla K_\sigma * u)(x,t)| \;\le\; \frac{4\,\|f\|_{L^\infty(\Omega)}}{\sqrt{2\pi}\,\sigma} \quad \text{on} \quad \bar{\Omega}\times(0,\infty),$$

where the notations from Chapter 2 have been used. Thanks to the uniform positive definiteness of D there exists an upper limit for the spectral condition number of D. This condition limit can be used to fix a suitable stencil size.

(d) The existence of a nonnegative directional splitting distinguishes the filter class (P_c) from morphological anisotropic equations such as mean curvature motion. In this case it has been proved that it is impossible to find a nonnegative directional splitting on a finite stencil [12].

Let us now illustrate the ideas in the proof of Theorem 6 by applying them to a practical example: We want to find a nonnegative spatial discretization of $\operatorname{div}(D\nabla u)$ on a (3×3)-stencil, where

$$D = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

and a, b and c may be functions of $J_\rho(\nabla u_\sigma)$.

Since $m=1$ we have a partition of $(-\frac{\pi}{2},\frac{\pi}{2}]$ into $4m-2=2$ subintervals:

$$(-\tfrac{\pi}{2},\tfrac{\pi}{2}] \;=\; (-\tfrac{\pi}{2},0] \cup (0,\tfrac{\pi}{2}] \;=:\; I_{-1}\cup I_1.$$

I_{-1} and I_1 belong to the grid angles

$$\beta_{-1} \;=\; \arctan\left(-\frac{h_2}{h_1}\right),$$
$$\beta_1 \;=\; \arctan\left(\frac{h_2}{h_1}\right) \;=:\; \beta.$$

First we focus on the case $\psi\in I_1$ where $(\cos\psi,\sin\psi)$ denotes the eigenvector to the larger eigenvalue λ_1 of D. With the notations from the proof of Theorem 6 we obtain

$$\begin{aligned}
\theta_1 &= \frac{\pi}{2}, \\
\rho_1 &= \frac{\theta_1 + \beta_1}{2} \;=\; \frac{\pi}{4} + \frac{\beta}{2}, \\
\eta_1 &= \frac{\beta}{2}.
\end{aligned}$$

Therefore, we get

$$\begin{aligned}
\cot(\rho_1 - \beta_1)\,\tan\rho_1 &= \cot\left(\frac{\pi}{4} - \frac{\beta}{2}\right)\tan\left(\frac{\pi}{4} + \frac{\beta}{2}\right) \;=\; \frac{1+\sin\beta}{1-\sin\beta}, \\
\cot(\beta_1 - \eta_1)\,\cot\eta_1 &= \cot^2\left(\frac{\beta}{2}\right) \;=\; \frac{1+\cos\beta}{1-\cos\beta},
\end{aligned}$$

which restricts the upper condition number for a nonnegative discretization with $\psi \in I_1$ to

$$\kappa_{1,1} \;:=\; \min\left(\frac{1+\sin\beta}{1-\sin\beta}, \frac{1+\cos\beta}{1-\cos\beta}\right). \tag{3.39}$$

Thanks to the symmetry we obtain the same condition restriction for $\psi \in I_{-1}$. These admissable condition numbers attain their maximal value for $h_1 = h_2$. In this case $\beta = \frac{\pi}{4}$ gives

$$\kappa_{1,1} \;=\; \kappa_{-1,1} \;=\; \frac{2+\sqrt{2}}{2-\sqrt{2}} \;\approx\; 5.8284. \tag{3.40}$$

By virtue of (3.32)–(3.34) we obtain as expressions for the directional diffusivities

$$\begin{aligned}
\alpha_{-1} &= \Big(|b| - b\Big) \cdot \frac{h_1^2 + h_2^2}{2h_1h_2}, \\
\alpha_0 &= a - |b| \cdot \frac{h_1}{h_2}, \\
\alpha_1 &= \Big(|b| + b\Big) \cdot \frac{h_1^2 + h_2^2}{2h_1h_2}, \\
\alpha_2 &= c - |b| \cdot \frac{h_2}{h_1}.
\end{aligned}$$

This induces in a natural way the following second-order discretization for $\operatorname{div}(D\nabla u)$:

$\frac{\lvert b_{i-1,j+1}\rvert - b_{i-1,j+1}}{4h_1h_2} + \frac{\lvert b_{i,j}\rvert - b_{i,j}}{4h_1h_2}$	$\frac{c_{i,j+1}+c_{i,j}}{2h_2^2} - \frac{\lvert b_{i,j+1}\rvert + \lvert b_{i,j}\rvert}{2h_1h_2}$	$\frac{\lvert b_{i+1,j+1}\rvert + b_{i+1,j+1}}{4h_1h_2} + \frac{\lvert b_{i,j}\rvert + b_{i,j}}{4h_1h_2}$
$\frac{a_{i-1,j}+a_{i,j}}{2h_1^2} - \frac{\lvert b_{i-1,j}\rvert + \lvert b_{i,j}\rvert}{2h_1h_2}$	$-\frac{a_{i-1,j}+2a_{i,j}+a_{i+1,j}}{2h_1^2} - \frac{\lvert b_{i-1,j+1}\rvert - b_{i-1,j+1} + \lvert b_{i+1,j+1}\rvert + b_{i+1,j+1}}{4h_1h_2} - \frac{\lvert b_{i-1,j-1}\rvert + b_{i-1,j-1} + \lvert b_{i+1,j-1}\rvert - b_{i+1,j-1}}{4h_1h_2} + \frac{\lvert b_{i-1,j}\rvert + \lvert b_{i+1,j}\rvert + \lvert b_{i,j-1}\rvert + \lvert b_{i,j+1}\rvert + 2\lvert b_{i,j}\rvert}{2h_1h_2} - \frac{c_{i,j-1}+2c_{i,j}+c_{i,j+1}}{2h_2^2}$	$\frac{a_{i+1,j}+a_{i,j}}{2h_1^2} - \frac{\lvert b_{i+1,j}\rvert + \lvert b_{i,j}\rvert}{2h_1h_2}$
$\frac{\lvert b_{i-1,j-1}\rvert + b_{i-1,j-1}}{4h_1h_2} + \frac{\lvert b_{i,j}\rvert + b_{i,j}}{4h_1h_2}$	$\frac{c_{i,j-1}+c_{i,j}}{2h_2^2} - \frac{\lvert b_{i,j-1}\rvert + \lvert b_{i,j}\rvert}{2h_1h_2}$	$\frac{\lvert b_{i+1,j-1}\rvert - b_{i+1,j-1}}{4h_1h_2} + \frac{\lvert b_{i,j}\rvert - b_{i,j}}{4h_1h_2}$

All nonvanishing entries of the p-th row of $A(u)$ are represented in this stencil, where $p(i,j)$ is the index of some inner pixel (i,j). Thus, for instance, the upper left stencil entry gives the element $(p(i,j),\, p(i{-}1,j{+}1))$ of $A(u)$. The other notations should be clear from the context as well, e.g. $b_{i,j}$ denotes a finite difference approximation of $b(J_\rho(\nabla u_\sigma))$ at some grid point (x_i, y_j).

It should be noted that related nonnegative difference approximations on a $(3{\times}3)$-stencil have been proposed in [111, 76, 112] for the simpler operator

$$a(x,y)\,\partial_{xx}u + 2b(x,y)\,\partial_{xy}u + c(x,y)\,\partial_{yy}u.$$

Our approach extends these results to

$$\partial_x\Big(a(x,y)\,\partial_x u\Big) + \partial_x\Big(b(x,y)\,\partial_y u\Big) + \partial_y\Big(b(x,y)\,\partial_x u\Big) + \partial_y\Big(c(x,y)\,\partial_y u\Big)$$

and clarifies the relation between the condition number of $\binom{a\ b}{b\ c}$ and the nonnegativity of the difference operator.

Chapter 4

Discrete diffusion filtering

This chapter presents a discrete class of diffusion processes for which one can establish similar properties as in the semidiscrete case concerning existence, uniqueness, continuous dependence of the solution on the initial image, maximum-minimum principle, average grey level invariance, Lyapunov functionals and convergence to a constant steady state. We shall see that this class comprises α-semi-implicit discretizations of the semidiscrete filter class (P_s).

4.1 The general model

As in Chapter 3 we regard a discrete image as a vector $f \in \mathbb{R}^N$, $N \geq 2$, and denote the index set $\{1, ..., N\}$ by J. Our discrete filter class (P_d) is as follows:

$$\left.\begin{array}{l}
\text{Let } f \in \mathbb{R}^N. \text{ Calculate a sequence } (u^{(k)})_{k\in\mathbb{N}_0} \text{ of processed versions} \\
\text{of } f \text{ by means of} \\
\qquad\qquad \begin{array}{rcl} u^{(0)} &=& f, \\ u^{(k+1)} &=& Q(u^{(k)})\, u^{(k)}, \qquad \forall\, k \in \mathbb{N}_0, \end{array} \\
\text{where } Q = (q_{ij}) \text{ has the following properties:} \\
\text{(D1)} \quad \text{continuity in its argument:} \qquad Q \in \mathrm{C}(\mathbb{R}^N, \mathbb{R}^{N\times N}), \\
\text{(D2)} \quad \text{symmetry:} \qquad q_{ij}(v) = q_{ji}(v) \quad \forall\, i,j \in J, \quad \forall\, v \in \mathbb{R}^N, \\
\text{(D3)} \quad \text{unit row sum:} \qquad \sum_{j\in J} q_{ij}(v) = 1 \quad \forall\, i \in J, \quad \forall\, v \in \mathbb{R}^N, \\
\text{(D4)} \quad \text{nonnegativity:} \qquad q_{ij}(v) \geq 0 \quad \forall\, i,j \in J, \quad \forall\, v \in \mathbb{R}^N, \\
\text{(D5)} \quad \text{irreducibility for all } v \in \mathbb{R}^N, \\
\text{(D6)} \quad \text{positive diagonal:} \qquad q_{ii}(v) > 0 \quad \forall\, i \in J, \quad \forall\, v \in \mathbb{R}^N.
\end{array}\right\} \quad (P_d)$$

Remarks:

(a) Although the basic idea behind scale-spaces is to have a continuous scale parameter, it is evident that fully discrete results are of importance since in practise scale-space evolutions are evaluated exclusively at a finite number of scales.

(b) The requirements (D1)–(D5) have a similar meaning as their semidiscrete counterparts (S1)–(S5). Indeed, (D1) immediately gives well-posedness results, while the proof of the extremum principle requires (D3) and (D4), and average grey value invariance is based on (D2) and (D3). The existence of Lyapunov functionals is a consequence of (D2)–(D4), strict Lyapunov functionals need (D5) and (D6) in addition to (D2)–(D4), and the convergence to a constant steady state utilizes (D2)–(D5).

(c) Nonnegative matrices $Q = (q_{ij}) \in \mathbb{R}^{N\times N}$ satisfying $\sum_{j\in J} q_{ij} = 1$ for all $i \in J$ are also called *stochastic matrices*. Moreover, if Q is stochastic and $\sum_{i\in J} q_{ij} = 1$ for all $j \in J$, then Q is *doubly stochastic*. This indicates that our discrete diffusion processes are related to the theory of Markov chains [224, 91, 136].

4.2 Theoretical results

It is obvious that for a fixed filter belonging to the class (P_d) every initial image $f \in \mathbb{R}^N$ generates a unique sequence $(u^{(k)})_{k\in\mathbb{N}_0}$. Moreover, by means of (D1) we know that, for every finite k, $u^{(k)}$ depends continuously on f. Therefore, let us now prove a maximum–minimum principle.

Proposition 3 (Extremum principle).
Let $f \in \mathbb{R}^N$ and let $(u^{(k)})_{k\in\mathbb{N}_0}$ be the sequence of filtered images according to (P_d). Then,

$$a \;\le\; u_i^{(k)} \;\le\; b \qquad \forall\, i \in J, \quad \forall\, k \in \mathbb{N}_0, \tag{4.1}$$

where

$$a := \min_{j\in J} f_j, \tag{4.2}$$

$$b := \max_{j\in J} f_j. \tag{4.3}$$

Proof:
The maximum–minimum principle follows directly from the fact that, for all $i \in J$ and $k \in \mathbb{N}_0$, the following inequalities hold:

(i) $$u_i^{(k+1)} \;=\; \sum_{j\in J} q_{ij}(u^{(k)})u_j^{(k)} \;\overset{(D4)}{\leq}\; \max_{m\in J} u_m^{(k)} \sum_{j\in J} q_{ij}(u^{(k)}) \;\overset{(D3)}{=}\; \max_{m\in J} u_m^{(k)}.$$

(ii) $$u_i^{(k+1)} \;=\; \sum_{j\in J} q_{ij}(u^{(k)})u_j^{(k)} \;\overset{(D4)}{\geq}\; \min_{m\in J} u_m^{(k)} \sum_{j\in J} q_{ij}(u^{(k)}) \;\overset{(D3)}{=}\; \min_{m\in J} u_m^{(k)}.$$

□

4.3 Scale-space properties

All statements from Chapter 3 with respect to invariances are valid in the discrete framework as well. Below we focus on proving average grey level invariance.

Proposition 4 (Conservation of average grey value).
The average grey level

$$\mu \;:=\; \frac{1}{N}\sum_{j\in J} f_j \tag{4.4}$$

is not affected by the discrete diffusion filter:

$$\frac{1}{N}\sum_{j\in J} u_j^{(k)} = \mu \qquad \forall\, k \in \mathbb{N}_0. \tag{4.5}$$

Proof:
By virtue of (D2) and (D3) we have $\sum_{i\in J} q_{ij}(u^{(k)}) = 1$ for all $j \in J$ and $k \in \mathbb{N}_0$. This so-called *redistribution property* [107] ensures that, for all $k \in \mathbb{N}_0$,

$$\sum_{i\in J} u_i^{(k+1)} \;=\; \sum_{i\in J}\sum_{j\in J} q_{ij}(u^{(k)})u_j^{(k)} \;=\; \sum_{j\in J}\Big(\sum_{i\in J} q_{ij}(u^{(k)})\Big)u_j^{(k)} \;=\; \sum_{j\in J} u_j^{(k)},$$

which proves the proposition. □

As one might expect, the class (P_d) allows an interpretation as a transformation which is smoothing in terms of Lyapunov functions. These functions ensure that $u^{(k)}$ converges to a constant image as $k \to \infty$. However, we need less regularity than in the semidiscrete case: The convex function r, which generates the Lyapunov functionals, needs only to be continuous, but no more differentiable.

Theorem 7 (Lyapunov functionals and behaviour for $n \to \infty$).
Assume that $(u^{(k)})_{k\in\mathbb{N}_0}$ satisfies the requirements of (P_d), let a, b, and μ be defined as in (4.2), (4.3), and (4.4), respectively, and let $c := (\mu, \mu, ..., \mu)^\top \in \mathbb{R}^N$.
Then the following properties are fulfilled:

(a) *(Lyapunov functionals)*
For all convex $r \in \mathrm{C}[a,b]$ the sequence

$$V^{(k)} := \Phi(u^{(k)}) := \sum_{i\in J} r(u_i^{(k)}), \qquad k \in \mathbb{N}_0$$

is a Lyapunov sequence:

(i) $\Phi(u^{(k)}) \geq \Phi(c) \quad \forall\, k \in \mathbb{N}_0$

(ii) $V^{(k+1)} - V^{(k)} \leq 0 \quad \forall\, k \in \mathbb{N}_0$

Moreover, if r is strictly convex, then $V^{(k)} = \Phi(u^{(k)})$ is a strict Lyapunov sequence:

(iii) $\Phi(u^{(k)}) = \Phi(c) \quad \Longleftrightarrow \quad u^{(k)} = c$

(iv) $V^{(k+1)} - V^{(k)} = 0 \quad \Longleftrightarrow \quad u^{(k)} = c$

(b) *(Convergence)*
$\lim_{k\to\infty} u^{(k)} = c.$

Proof:

(a) (i) Average grey level invariance and the convexity of r give

$$\begin{aligned} \Phi(c) &= \sum_{i=1}^{N} r\left(\sum_{j=1}^{N} \frac{1}{N} u_j^{(k)}\right) \\ &\leq \sum_{i=1}^{N} \left(\frac{1}{N} \sum_{j=1}^{N} r(u_j^{(k)})\right) \\ &= \sum_{j=1}^{N} r(u_j^{(k)}) \\ &= \Phi(u^{(k)}). \end{aligned} \tag{4.6}$$

(ii) For $i,j \in J$ we define

$$a_{ij}(u^{(k)}) := \begin{cases} q_{ij}(u^{(k)}) - 1 & (i=j) \\ q_{ij}(u^{(k)}) & (i \neq j). \end{cases} \tag{4.7}$$

Using the convexity of r, the preceding definition, and the prerequisites (D2) and (D3) we obtain

$$V^{(k+1)} - V^{(k)} = \sum_{i=1}^{N} \left(r\left(\sum_{j=1}^{N} q_{ij}(u^{(k)})\, u_j^{(k)}\right) - r(u_i^{(k)})\right)$$

$$
\begin{aligned}
&\overset{\text{conv.}}{\leq} && \sum_{i=1}^{N} \left(\sum_{j=1}^{N} q_{ij}(u^{(k)})\, r(u_j^{(k)}) - r(u_i^{(k)}) \right) \\
&\overset{(4.7)}{=} && \sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij}(u^{(k)})\ r(u_j^{(k)}) \\
&\overset{(D3)}{=} && \sum_{i=1}^{N} \sum_{j=1}^{N} a_{ij}(u^{(k)})\ \left(r(u_j^{(k)}) - r(u_i^{(k)}) \right) \\
&= && \sum_{i=1}^{N} \sum_{m=1}^{N-i} a_{i+m,i}(u^{(k)})\ \left(r(u_i^{(k)}) - r(u_{i+m}^{(k)}) \right) \\
&+ && \sum_{i=1}^{N} \sum_{m=1}^{N-i} a_{i,i+m}(u^{(k)})\ \left(r(u_{i+m}^{(k)}) - r(u_i^{(k)}) \right) \\
&\overset{(D2)}{=} && 0. \qquad (4.8)
\end{aligned}
$$

(iii) This part of the proof can be shown in exactly the same manner as in the semidiscrete case (Chapter 3, Theorem 5): Equality in the estimate (4.6) holds due to the strict convexity of r if and only if $u^{(k)} = c$.

(iv) In order to verify the first implication, let us start with a proof that $V^{(k+1)} = V^{(k)}$ implies $u_1^{(k)} = ... = u_N^{(k)}$. To this end, assume that $u^{(k)}$ is not constant:

$$u_{i_0}^{(k)} \ := \ \min_{i \in J} u_i^{(k)} \ < \ \max_{j \in J} u_j^{(k)} \ =: \ u_{j_0}^{(k)}.$$

Then, by the irreducibility of $Q(u^{(k)})$, we find $l_0, ..., l_r \in J$ with $l_0 = i_0$, $l_r = j_0$ and $q_{l_p l_{p+1}} \neq 0$ for $p = 0, ..., r-1$. Hence, there exists some $p_0 \in \{0, ..., r-1\}$ such that $n := l_{p_0}$, $m := l_{p_0+1}$, $q_{nm}(u^{(k)}) \neq 0$, and $u_m^{(k)} \neq u_n^{(k)}$. Moreover, the nonnegativity of $Q(u^{(k)})$ gives $q_{nm}(u^{(k)}) > 0$, and by (D6) we have $q_{nn}(u^{(k)}) > 0$. Together with the strict convexity of r these properties lead to

$$
\begin{aligned}
& r\left(\sum_{j=1}^{N} q_{nj}(u^{(k)})\, u_j^{(k)} \right) \\
= \quad & r\left(\sum_{\substack{j=1 \\ j \neq n,m}}^{N} q_{nj}(u^{(k)})\, u_j^{(k)} + q_{nn}(u^{(k)})\, u_n^{(k)} + q_{nm}(u^{(k)})\, u_m^{(k)} \right) \\
< \quad & \sum_{\substack{j=1 \\ j \neq n,m}}^{N} q_{nj}(u^{(k)})\, r(u_j^{(k)}) + q_{nn}(u^{(k)})\, r(u_n^{(k)}) + q_{nm}(u^{(k)})\, r(u_m^{(k)}) \\
= \quad & \sum_{j=1}^{N} q_{nj}(u^{(k)})\, r(u_j^{(k)}).
\end{aligned}
$$

If we combine this with the results in (4.8), we obtain

$$\begin{aligned}
V^{(k+1)} - V^{(k)} &= \sum_{\substack{i=1\\ i\neq n}}^{N} \left(r\left(\sum_{j=1}^{N} q_{ij}(u^{(k)})\, u_j^{(k)} \right) - r(u_i^{(k)}) \right) \\
&+ \; r\left(\sum_{j=1}^{N} q_{nj}(u^{(k)})\, u_j^{(k)} \right) - r(u_n^{(k)}) \\
&< \; \sum_{i=1}^{N} \left(\sum_{j=1}^{N} q_{ij}(u^{(k)})\, r(u_j^{(k)}) - r(u_i^{(k)}) \right) \\
&\overset{(4.8)}{=} \; 0.
\end{aligned}$$

This establishes that $V^{(k+1)} = V^{(k)}$ implies $u_1^{(k)} = ... = u_N^{(k)}$. Then, by virtue of the grey value invariance, we conclude that $u^{(k)} = c$.

Conversely, let $u^{(k)} = c$. By means of prerequisite (D3) we obtain

$$V^{(k+1)} - V^{(k)} \;=\; \sum_{i=1}^{N} r\left(\sum_{j=1}^{N} q_{ij}(u^{(k)})\mu \right) - \sum_{i=1}^{N} r(\mu) \;=\; 0.$$

(c) In order to prove convergence to a constant steady state, we can argue exactly in the same way as in the semidiscrete case if we replace Lyapunov functionals by Lyapunov sequences and integrals by sums. See Chapter 3, Theorem 5 for more details. □

In analogy to the semidiscrete case the preceding theorem comprises many Lyapunov functionals which demonstrate the information-reducing qualities of our filter class. Choosing the convex functions $r(s) := |s|^p$, $r(s) := (s-\mu)^{2n}$ and $r(s) := s \ln s$, we immediately obtain the following corollary.

Corollary 3 (Special Lyapunov functionals).
Let $(u^{(k)})_{k\in\mathbb{N}_0}$ be a diffusion sequence according to (P_d), and let a and μ be defined as in (4.2) and (4.4). Then the following functions are decreasing in k:

(a) $\|u^{(k)}\|_p$ *for all* $p \geq 1$.

(b) $M_{2n}[u^{(k)}] \;:=\; \frac{1}{N} \sum_{j=1}^{N} (u_j^{(k)} - \mu)^{2n}$ *for all* $n \in \mathbb{N}$.

(c) $H[u^{(k)}] \;:=\; \sum_{j=1}^{N} u_j^{(k)} \ln(u_j^{(k)}),$ *if* $a > 0$.

An interpretation of these results in terms of decreasing energy, decreasing central moments and increasing entropy is evident.

4.4 Relation to semidiscrete models

Let us now investigate in which sense our discrete filter class covers in a natural way time discretizations of semidiscrete filters. To this end, we regard $u^{(k)}$ as an approximation of the solution u of (P_s) at time $t = k\tau$, where τ denotes the time step size. We consider an FD scheme with two time levels where the operator A – which depends nonlinearly on u – is evaluated in an explicit way, while the linear remainder is discretized in an α-implicit manner. Such schemes are called *α-semi-implicit*. They reveal the advantage that the linear implicit part ensures good stabililty properties, while the explicit evaluation of the nonlinear terms avoids the necessity to solve nonlinear systems of equations. The theorem below states that this class of schemes is covered by the discrete framework, for which we have established scale-space results.

Theorem 8 (Scale-space interpretation for α-semi-implicit schemes). *Let $\alpha \in [0,1]$, $\tau > 0$, and let $A = (a_{ij}) : \mathbb{R}^N \to \mathbb{R}^{N\times N}$ satisfy the requirements (S1)–(S5) of section 3.1. Then the α-semi-implicit scheme*

$$\frac{u^{(k+1)} - u^{(k)}}{\tau} \;=\; A(u^{(k)}) \left(\alpha u^{(k+1)} + (1-\alpha) u^{(k)}\right) \tag{4.9}$$

fulfils the prerequisites (D1)–(D6) for discrete diffusion models provided that

$$\tau \;\le\; \frac{1}{(1-\alpha)\, \max\limits_{i\in J} |a_{ii}(u^{(k)})|} \tag{4.10}$$

for $\alpha \in (0,1)$. In the explicit case $(\alpha=0)$ the properties (D1)–(D6) hold for

$$\tau \;<\; \frac{1}{\max\limits_{i\in J} |a_{ii}(u^{(k)})|}, \tag{4.11}$$

and the semi-implicit case $(\alpha=1)$ satisfies (D1)–(D6) unconditionally.

Proof: Let

$$\begin{aligned} B(u^{(k)}) \;&:=\; (b_{ij}(u^{(k)})) \;&&:=\; I - \alpha\tau A(u^{(k)}), \\ C(u^{(k)}) \;&:=\; (c_{ij}(u^{(k)})) \;&&:=\; I + (1-\alpha)\tau A(u^{(k)}), \end{aligned}$$

where $I \in \mathbb{R}^N$ denotes the unit matrix. Since (4.9) can be written as

$$B(u^{(k)})\, u^{(k+1)} \;=\; C(u^{(k)})\, u^{(k)}$$

we first have to show that $B(u^{(k)})$ is invertible for all $u^{(k)} \in \mathbb{R}^N$. Henceforth, the argument $u^{(k)}$ is suppressed frequently since the considerations below are valid for all $u^{(k)} \in \mathbb{R}^N$.

If $\alpha=0$, then $B=I$ and hence invertible. Now assume that $\alpha > 0$. Then B is strictly diagonally dominant, since

$$b_{ii} \;=\; 1-\alpha\tau a_{ii} \;\overset{(S3)}{=}\; 1+\alpha\tau\sum_{\substack{j\in J\\ j\neq i}} a_{ij} \;>\; \alpha\tau\sum_{\substack{j\in J\\ j\neq i}} a_{ij} \;\overset{(S4)}{=}\; \sum_{\substack{j\in J\\ j\neq i}} |b_{ij}| \qquad \forall\, i\in J.$$

This also shows that $b_{ii} > 0$ for all $i \in J$, and by the structure of the off-diagonal elements of B we observe that the irreducibility of A implies the irreducibility of B. Thanks to the fact that B is irreducibly diagonally dominant, $b_{ij} \le 0$ for all $i \neq j$, and $b_{ii} > 0$ for all $i \in J$, we know from [241, p. 85] that $B^{-1} =: H =: (h_{ij})$ exists and $h_{ij} > 0$ for all $i, j \in J$. Thus, $Q := (q_{ij}) := B^{-1}C$ exists and by (S1) it follows that $Q \in \mathrm{C}(\mathbb{R}^N, \mathbb{R}^{N\times N})$. This proves (D1).

The requirement (D2) is not hard to satisfy: Since B^{-1} and C are symmetric and reveal the same set of eigenvectors – namely those of A – it follows that $Q = B^{-1}C$ is symmetric as well.

Let us now verify (D3). By means of (S3) we obtain

$$\sum_{j\in J} b_{ij} = 1 = \sum_{j\in J} c_{ij} \qquad \forall\, i \in J. \tag{4.12}$$

Let $v := (1, ..., 1)^\top \in \mathbb{R}^N$. Then (4.12) is equivalent to

$$Bv = v = Cv, \tag{4.13}$$

and the invertibility of B gives

$$v = B^{-1}v = Hv. \tag{4.14}$$

Therefore, from

$$Qv \;=\; HCv \;\overset{(4.13)}{=}\; Hv \;\overset{(4.14)}{=}\; v$$

we conclude that $\sum_{j\in J} q_{ij} = 1$ for all $i \in J$. This proves (D3).

In order to show that (D4) is fulfilled, we first check the nonnegativity of C. For $i \neq j$ we have

$$c_{ij} \;=\; (1-\alpha)\tau a_{ij} \;\overset{(S4)}{\geq}\; 0.$$

The diagonal entries yield

$$c_{ii} \;=\; 1+(1-\alpha)\tau a_{ii}.$$

If $\alpha = 1$ we have $c_{ii} = 1$ for all $i \in J$. For $0 \le \alpha < 1$, however, nonnegativity of C is not automatically guaranteed: Using (S3)–(S5) we obtain

$$a_{ii} \;\overset{(S3)}{=}\; -\sum_{\substack{j\in J\\ j\neq i}} a_{ij} \;\overset{(S4),(S5)}{<}\; 0 \qquad \forall\, i\in J. \tag{4.15}$$

Hence, $C(u^{(k)})$ is nonnegative if

$$\tau \;\le\; \frac{1}{(1-\alpha)\,\max\limits_{i\in J}|a_{ii}(u^{(k)})|} \;=:\; \tau_\alpha(u^{(k)}).$$

Since H is nonnegative, we know that the nonnegativity of C implies the non-negativity of $Q = HC$.

Now we want to prove (D5). If $\alpha = 1$, then $C = I$, and by the positivity of H we have $q_{ij} > 0$ for all $i, j \in J$. Thus, Q is irreducible.

Next let us consider the case $0 < \alpha < 1$ and $\tau \le \tau_\alpha(u^{(k)})$. Then we know that C is nonnegative. Using this information, the positivity of H, the symmetry of C, and (4.12) we obtain

$$q_{ij} \;=\; \sum_{k\in J} h_{ik}c_{kj} \;\ge\; \underbrace{\min_{k\in J} h_{ik}}_{>0} \cdot \underbrace{\sum_{k\in J} c_{kj}}_{=1} \;>\; 0 \qquad \forall\, i,j \in J,$$

which establishes the irreducibility of Q.

Finally, for $\alpha = 0$, we have $Q = C$. For $i, j \in J$ with $i \neq j$ we know that $a_{ij}(u^{(k)}) > 0$ implies $c_{ij}(u^{(k)}) > 0$. Now for

$$\tau \;<\; \frac{1}{\max\limits_{i\in J}|a_{ii}(u^{(k)})|}$$

it follows that $c_{ii}(u^{(k)}) > 0$ for all $i \in J$ and, thus, the irreducibility of $A(u^{(k)})$ carries over to $Q(u^{(k)})$.

In all the abovementioned cases the time step size restrictions for ensuring irreducibility imply that all diagonal elements of $Q(u^{(k)})$ are positive. This establishes (D6). □

Remarks:

(a) We have seen that the discrete filter class (P_d) – although at first glance looking like a pure explicit discretization – covers the α-semi-implicit case as well. Explicit two-level schemes are comprised by the choice $\alpha = 0$. Equation (4.11) shows that they reveal the most prohibitive time step size restrictions.

(b) The conditions (4.10) and (4.11) can be satisfied by means of an a-priori estimate. Since the semi-implicit scheme fulfils (D1)–(D6) we know by Theorem 3 that the solution obeys an extremum principle. This means that $u^{(k)}$ belongs to the compact set $\{v \in \mathbb{R}^N \mid \|v\|_\infty \le \|f\|_\infty\}$ for all $k \in \mathbb{N}_0$. By $A \in \mathrm{C}(\mathbb{R}^N, \mathbb{R}^{N\times N})$ it follows that

$$K_f := \max\left\{|a_{ii}(v)| \;\middle|\; i \in J,\; v \in \mathbb{R}^N,\; \|v\|_\infty \le \|f\|_\infty\right\}$$

exists, and (4.15) shows that $K_f > 0$. Thus, choosing

$$\tau \;\le\; \frac{1}{(1-\alpha)\,K_f}$$

ensures that (4.10) is always satisfied, and

$$\tau \;<\; \frac{1}{K_f}$$

guarantees that (4.11) holds.

(c) If $\alpha > 0$, a large linear system of equations has to be solved. The system matrix, however, is symmetric, diagonally dominant, and sparse.[1] Hence, one can apply standard iterative techniques. The author's experiments indicate that a conjugate gradient algorithm with SSOR preconditioning is highly recommendable. More details about this method can be found in [168, pp. 154–161].

(d) For $\alpha = 1$ we obtain semi-implicit schemes which do not suffer from time step size restrictions. In spite of the fact that the nonlinearity is discretized in an explicit way they are absolutely stable in the maximum norm, and they inherit the scale-space properties from the semidiscrete setting regardless of the step size. Compared to explicit schemes, this advantage usually overcompensates for the additional effort of resolving a linear system.

(e) By the explicit discretization of the nonlinear operator A it follows that all schemes in the preceding theorem are of first order in time. This should not give rise to concern, since in image processing one is in general more interested in maintaining qualitative properties such as maximum principles or invariances rather than having an accurate approximation of the continuous equation. However, if one insists in second-order schemes, one may for instance use the predictor–corrector approach by Douglas and Jones [77]:

$$\begin{aligned}
\frac{u^{(k+1/2)} - u^{(k)}}{\tau/2} &= A(u^{(k)})\,u^{(k+1/2)},\\
\frac{u^{(k+1)} - u^{(k)}}{\tau} &= A(u^{(k+1/2)})\left(\tfrac{1}{2}u^{(k+1)} + \tfrac{1}{2}u^{(k)}\right).
\end{aligned}$$

This scheme satisfies the properties (D1)–(D6) for $\tau \le 2/K_f$.

[1] If the system matrix results from an FD discretization on a $(p\times p)$-stencil it contains at most p^2 nonvanishing entries per row.

(f) One may also wish to apply splitting techniques [76, 164, 168, 211, 265]. In this case it should be kept in mind that in the general nonlinear framework the split operators do not commute. Especially for large time steps this may lead to visible bias towards certain preferred directions.

(g) The assumptions (S1)–(S5) are sufficient conditions for the α-semi-implicit scheme to fulfil (D1)–(D6), but they are not necessary. Nonnegativity of $Q(u^{(k)})$ may also be achieved using spatial discretizations where $A(u^{(k)})$ violates nonnegativity (see [38] for examples).

Chapter 5

Examples and applications

The scale-space theory that we have discussed in Chapters 2–4 covers methods such as linear or nonlinear isotopic diffusion filtering, for which many interesting applications have already been mentioned in Chapter 1. Therefore, the goal of the present chapter is to show that a generalization to anisotropic models with diffusion tensors depending on the structure tensor offer novel interesting properties and application fields. Thus, we focus mainly on these anisotropic techniques and juxtapose the results to other methods. In order to demonstrate the flexibility of anisotropic diffusion filtering, we shall pursue two different objectives:

- smoothing with simultaneous edge-enhancement,
- smoothing with enhancement of coherent flow-like textures.

All calculations for diffusion filtering are performed using semi-implicit FD schemes with time steps $\Delta t \in [2, 5]$, and the regularizing convolution with a Gaussian is implemented via an implicit FD scheme for linear diffusion. In order to compare anisotropic diffusion to other methods, morphological scale-spaces have been discretized as well. For MCM and AMSS this is achieved by means of (explicit) PSC schemes (cf. 1.5.4) with $\Delta t := 0.1$ and $\Delta t := 0.01$, respectively. On a small workstation such as an HP 712/80 (92 MIPS, 27.6 Megaflops) it takes about 3 CPU seconds to calculate one anisotropic diffusion step for a 256×256 image, and MCM or AMSS require approximately 0.4 seconds.

5.1 Edge-enhancing diffusion

5.1.1 Filter design

In accordance with the notations in 2.2, let μ_1, μ_2 with $\mu_1 \geq \mu_2$ be the eigenvalues of the structure tensor J_ρ, and v_1, v_2 the corresponding orthonormal eigenvectors.

Since the diffusion tensor should reflect the local image structure it ought to be chosen in such a way that it reveals the same set of eigenvectors v_1, v_2 als J_ρ. The choice of the corresponding eigenvalues λ_1, λ_2 depends on the desired goal of the filter.

If one wants to smooth preferably within each region and aims to inhibit diffusion across edges, then one can reduce the diffusivity λ_1 perpendicular to edges the more the higher the contrast μ_1 is, see 1.2.3 and [247, 250]. This behaviour may be accomplished by the following choice ($m \in \mathbb{N}$, $C_m > 0$, $\lambda > 0$):

$$\lambda_1(\mu_1) \ := \ g(\mu_1), \tag{5.1}$$
$$\lambda_2 \ := \ 1 \tag{5.2}$$

with

$$g(s) \ := \ \begin{cases} 1 & (s \leq 0) \\ 1 - \exp\left(\frac{-C_m}{(s/\lambda)^m}\right) & (s > 0). \end{cases} \tag{5.3}$$

This exponentially decreasing function is chosen in order to fulfil the smoothness requirement stated in (P_c), cf. 2.3. Since ∇u_σ remains bounded on $\Omega \times [0, \infty)$ and $\mu_1 = |\nabla u_\sigma|^2$, we know that the uniform positive definiteness of D is automatically satisfied by this filter.

The constant C_m is calculated in such a way that the flux $sg(s)$ is increasing for $s \in [0, \lambda]$ and decreasing for $s \in (\lambda, \infty)$. Thus, the preceding filter strategy can be regarded as an anisotropic regularization of the Perona–Malik model.

The choice $m := 4$ (which implies $C_4 = 3.31488$) gives visually good results and is used exclusively in the examples below. Since in this section we are only interested in edge-enhancing diffusion we may set the integration scale ρ of the structure tensor equal to 0. Applications which require nonvanishing integration scales shall be studied in section 5.2.

5.1.2 Applications

Figure 5.1 illustrates that anisotropic diffusion filtering is still capable of possessing the contrast-enhancing properties of the Perona–Malik filter (provided the regularization parameter σ is not too large).[1] It depicts the temporal evolution of a Gaussian-like function and its isolines. It can be observed that two regions with almost constant grey value evolve which are separated by a fairly steep edge. Edge enhancement is caused by the fact that, due to the rapidly decreasing diffusivity, smoothing within each region is strongly preferred to diffusion between

[1]Except for Figs. 5.1, 5.3 and 5.4, where contrast enhancement is to be demonstrated, all images in the present work are depicted in such a way that the lowest value is black and the highest one appears white. They reveal a range within the interval $[0, 255]$ and all pixels have unit length in both directions.

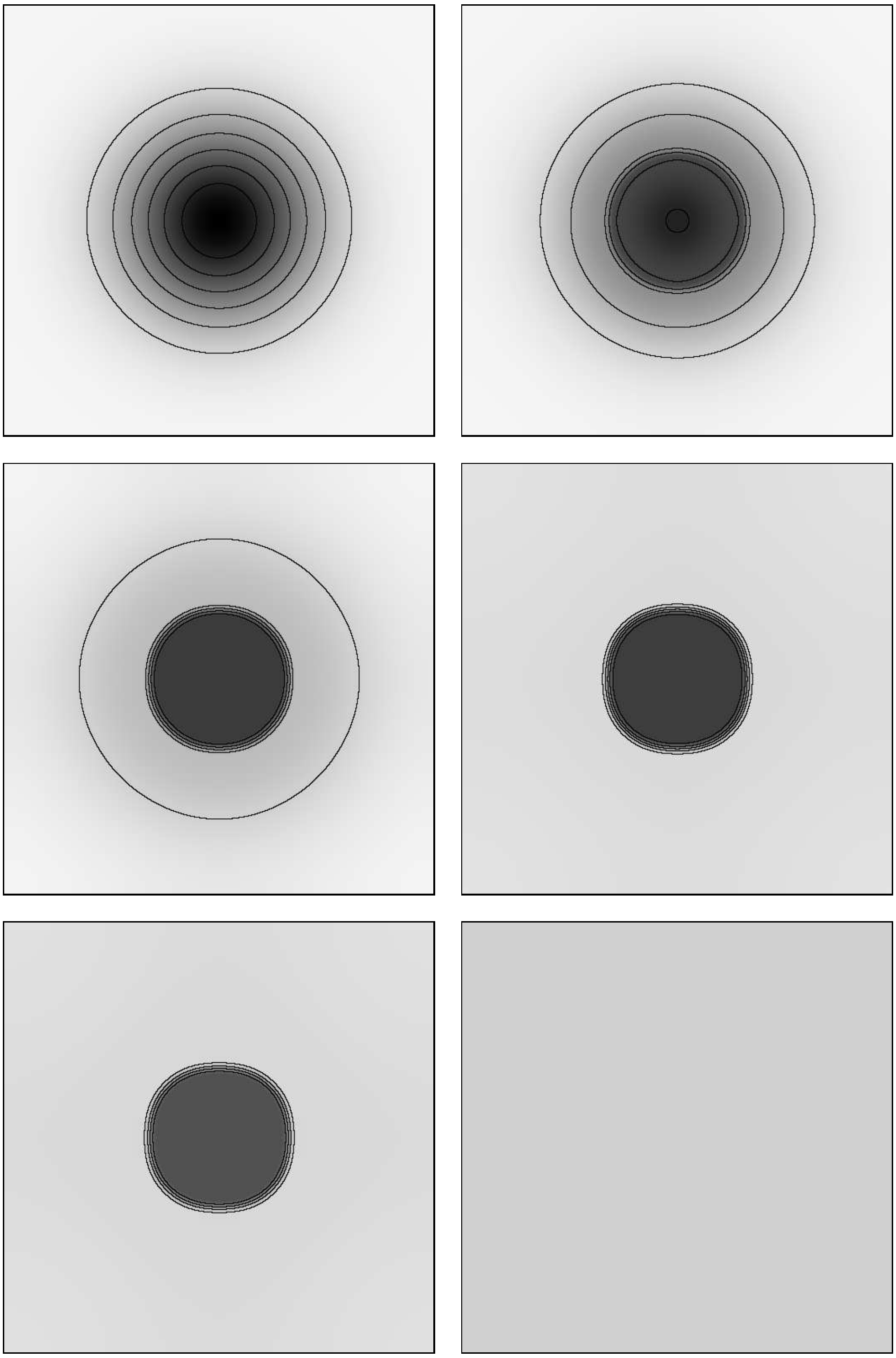

Figure 5.1: Anisotropic diffusion filtering of a Gaussian-type function, $\Omega = (0, 256)^2$. From top left to bottom right: $t = 0$, 125, 625, 3125, 15625, 78125.

the two adjacent regions. The edge location remains stable over a very long time interval. This indicates that, in practise, the determination of a suitable stopping time is not a critical problem. After the process of contrast enhancement is concluded, the steepness of edges decreases very slowly until the gradient reaches a value where no backward diffusion is possible anymore. Then the image converges quickly towards a constant image.

Let us now compare the denoising properties of different diffusion filters. Figure 5.2(a) consists of a triangle and a rectangle with 70 % of all pixels being completely degraded by noise.[2] In Fig. 5.2(b) we observe that linear diffusion filtering is capable of removing all noise, but we have to pay a price: the image becomes completely blurred. Besides the fact that edges get smoothed so that they are harder to identify, the correspondence problem appears: edges become dislocated. Thus, once they are identified at a coarse scale, they have to be traced back in order to find their true location, a numerically very difficult problem.

Fig. 5.2(c) shows the effect when applying the isotropic nonlinear diffusion equation [58]

$$\partial_t u = \operatorname{div}\left(g(|\nabla u_\sigma|)\nabla u\right) \tag{5.4}$$

with g as in (5.3). Since edges are hardly affected by this process, nonlinear isotropic diffusion does not reveal the correspondence problems which are characteristic for linear filtering. On the other hand, the drastically reduced diffusivity at edges is also responsible for the drawback that noise at edges is preserved.

Figure 5.2(d) demonstrates that nonlinear anisotropic filtering shares the advantages of both beforementioned methods. It combines the good noise eliminating properties of linear diffusion with the stable edge structure of nonlinear isotropic filtering. Due to the permitted smoothing along edges, however, corners get more rounded than in the nonlinear isotropic case.

The scale-space behaviour of the different diffusion filters is juxtaposed in Fig. 5.3, where an MRT slice of a human head is processed [249].

We observe that linear diffusion does not only blur all structures in an equal amount but also dislocates them more and more with increasing scale.

Isotropic nonlinear diffusion gives very well localized edges, so no correspondence problem appears anymore. One the other hand, small or thin structures reveal an extremely extended lifetime across the scales if they can be distinguished from their neighbourhood by a sufficiently large grey value difference. As a consequence, they may be present for too long, as can be seen from the last

[2] This image is taken from the Software package MegaWave which was developed at the CEREMADE (University Paris IX). Test images of this type have been used to study the behaviour of filters such as [12, 14, 15, 65, 66].

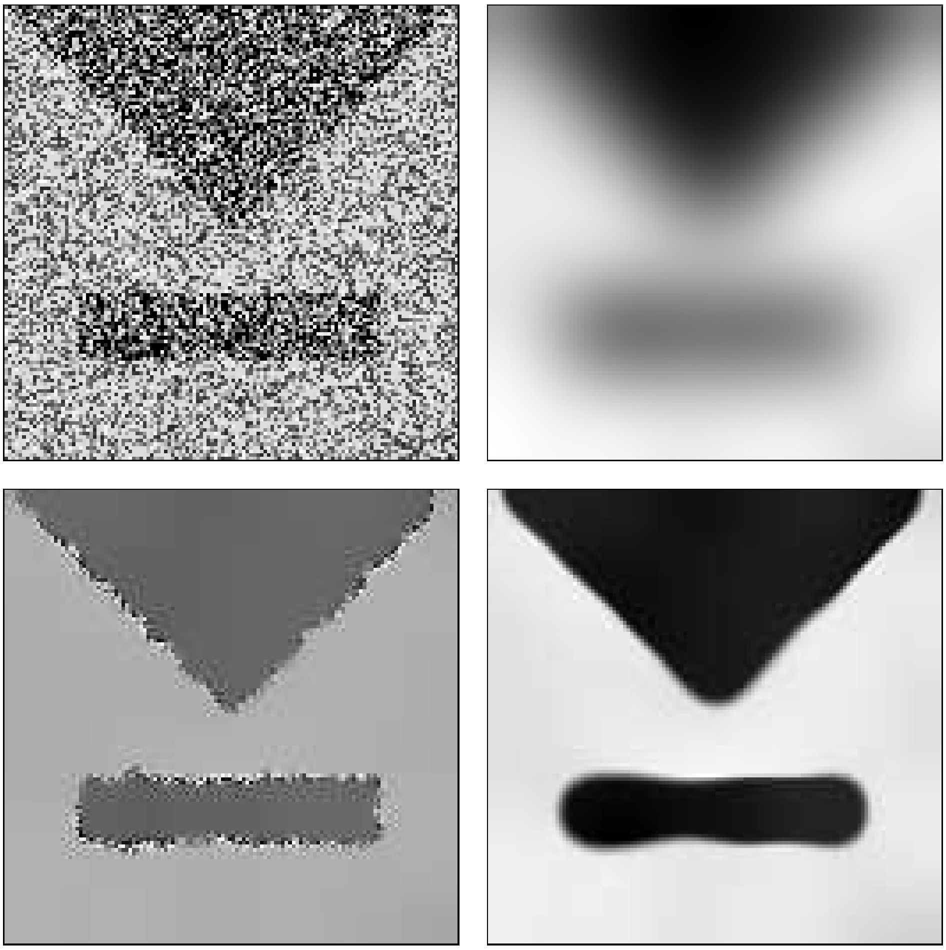

Figure 5.2: Restoration properties of diffusion filters. (a) Top Left: Test image, $\Omega = (0,128)^2$. (b) Top Right: Linear diffusion, $t = 80$. (c) Bottom Left: Nonlinear isotropic diffusion, $\lambda = 3.5$, $\sigma = 3$, $t = 80$. (d) Bottom Right: Nonlinear anisotropic diffusion, $\lambda = 3.5$, $\sigma = 3$, $t = 80$.

image in the middle column. The whole nonlinear isotropic diffusion scale-space evolution is very slow.

Anisotropic nonlinear diffusion has a faster scale-space behaviour since diffusion along edges is still permitted. As in (5.2(d)), this causes a stronger rounding of structures, which can be seen at the nose. As a positive consequence of this slight shrinking effect, small or elongated and thin structures are better eliminated than in the isotropic case. Thus, we recognize that most of the depicted "segments" coincide with the semantically correct objects that one would expect at these scales. Finally the image turns to a silhouette of the head, before it converges to a constant image.

As can be seen for instance at the chin, both nonlinear diffusion processes are capable of enhancing edges. The tendency to produce piecewise almost constant regions indicates that these scale-spaces are ideal prepocessing tools for segmentation. Unlike diffusion–reaction models aiming to yield *one* segmentation-like result for $t \to \infty$ (cf. 1.3), the temporal evolution of these models generates a complete *hierarchical family* of segmentation-like images. The contrast-enhancing quality distinguishes nonlinear diffusion filters from other scale-spaces. It should be noted that contrast enhancement is a local phenomenon which cannot be replaced by simple global rescalings of the grey value range. Therefore, it is generally not possible to obtain similar segmentation-like results by just rescaling the results from a scale-space which is only contrast-reducing.

The contrast and noise parameters λ and σ give the user the liberty to adapt nonlinear diffusion scale-spaces to the desired purpose in order to reward interesting features with a longer lifetime. Suitable values for them should result in a natural way from the specific problem. In this sense, the time t is rather a parameter of importance, with respect to the specified task, than a descriptor of spatial scale. The common view that the evolution paramater t of scale-spaces should be related to the spatial scale reflects the assumption that a scale-space analysis should be uncommitted. Nonlinear diffusion filtering renounces this requirement by allowing to incorporate a-priori information (e.g. about the contrast of semantically important structures) into the evolution process. The basic idea of scale-spaces, however, is maintained: to provide a family of subsequently simplified versions of the original image, which gives a hierarchy of structures and allows to pick up the relevant information from a certain scale.

Besides these specific features of nonlinear diffusion scale-spaces it should be mentioned that, due to the homogeneous Neumann boundary condition and the divergence form, both linear and nonlinear diffusion filters preserve the average grey level of the image.

This is not true for the morphological filters which have been applied in Fig. 5.4. We observe that the images become increasingly darker as the time proceeds.

On the other hand, edges are not blurred as strongly as in the linear diffusion case.

The curvature motion (1.74) at the left column of Fig. 5.4 is a pure anisotropic technique. Since MCM shrinks level lines with a velocity that is proportional to their curvature, low-curved object boundaries are less affected by this process, while high-curved structures (e.g. the nose) exhibit roundings at an earlier stage. After some time, however, the head looks almost like a ball. This is in accordance with the theory which predicts convergence of all closed level lines to a circular point.

A similar behaviour can be observed for the affine invariant morphological scale-space (1.92). Since it takes the time $T = \frac{3}{4}s^{\frac{4}{3}}$ to remove all isolines within a circle of radius s – in contrast to $T = \frac{1}{2}s^2$ for MCM – we see that, for a comparable elimination of small structure, the shrinking effect of large structures is stronger for AMSS than for MCM. Thus, although edges are easier to detect than in the linear diffusion case, the correspondence problem remains. Nevertheless, the advantage of having affine invariance may counterbalance the correspondence problem in certain applications. Since the AMSS involves no additional parameters and offers more invariances than other scale-spaces, it is ideal for uncommitted image analysis and shape recognition.

One possibility to avoid the correspondence problem of morphological scale-spaces is to damp the curve evolution at high-contrast edges, which is at the expense of withdrawing morphology. Processes of this type are studied in [12, 217, 218]. As a simple prototype for this idea, let us investigate the modified MCM

$$\partial_t u \;=\; \frac{|\nabla u|\,\mathrm{curv}(u)}{1+(|\nabla u_\sigma|/\lambda)^2}. \tag{5.5}$$

Its corresponding evolution is depicted in the third column of Fig. 5.4. We observe that structures remain much better localized than in the original MCM. On the other hand, the experiments give evidence that this process is probably not contrast-enhancing, see e.g. the chin. As a consequence, the results appear less segmentation-like than those for nonlinear diffusion filtering.

Let us now study two applications of nonlinear diffusion filtering in computer aided quality control (CAQ): the grading of fabrics and wood surfaces (see also [248]).

The quality of a fabric is determined by two criteria, namely clouds and stripes. Clouds result from isotropic inhomogeneities of the density distribution, whereas stripes are an anisotropic phenomenon caused by adjacent fibers pointing in the same direction. Anisotropic diffusion filters are capable of visualizing both quality-relevant features simultaneously (Fig. 5.5). For a suitable parameter choice, they perform isotropic smoothing at clouds and diffuse in an anisotropic

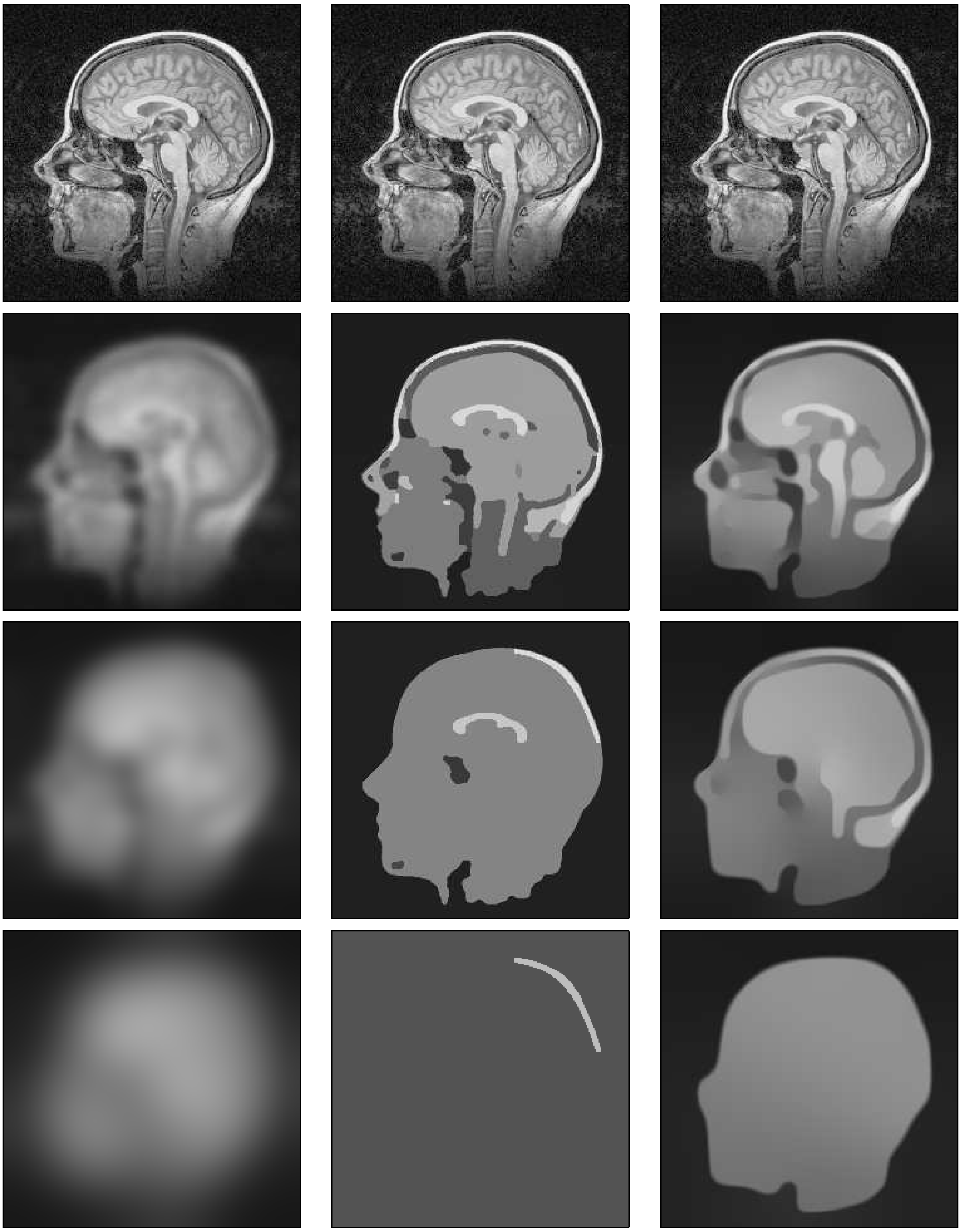

Figure 5.3: Diffusion scale-spaces. Top: Original image, $\Omega = (0, 236)^2$. Left Column: Linear diffusion, top to bottom: $t = 0, 12.5, 50, 200$. Middle Column: Isotropic nonlinear diffusion ($\lambda = 3$, $\sigma = 1$), $t = 0, 25000, 500000, 7000000$. Right Column: Anisotropic nonlinear diffusion ($\lambda = 3$, $\sigma = 1$), $t = 0$, 250, 875, 3000.

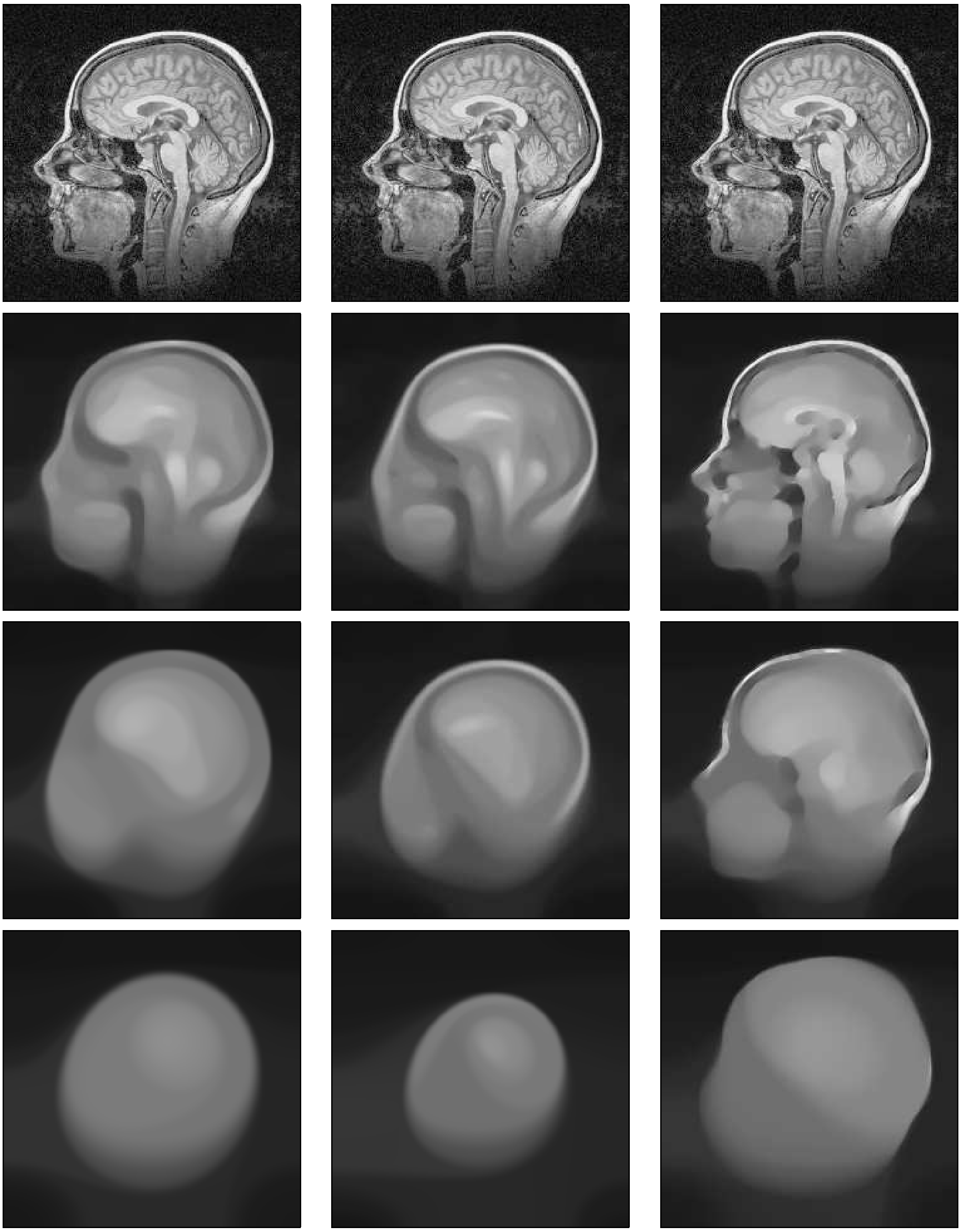

Figure 5.4: Morphological scale-spaces and related methods. TOP: Original image, $\Omega = (0, 236)^2$. LEFT COLUMN: Mean curvature motion, $t = 0$, 70, 275, 1275. MIDDLE COLUMN: Affine morphological scale-space, $t = 0$, 20, 50, 140. RIGHT COLUMN: Modified mean curvature motion ($\lambda = 3$, $\sigma = 1$), $t = 0$, 100, 350, 1500.

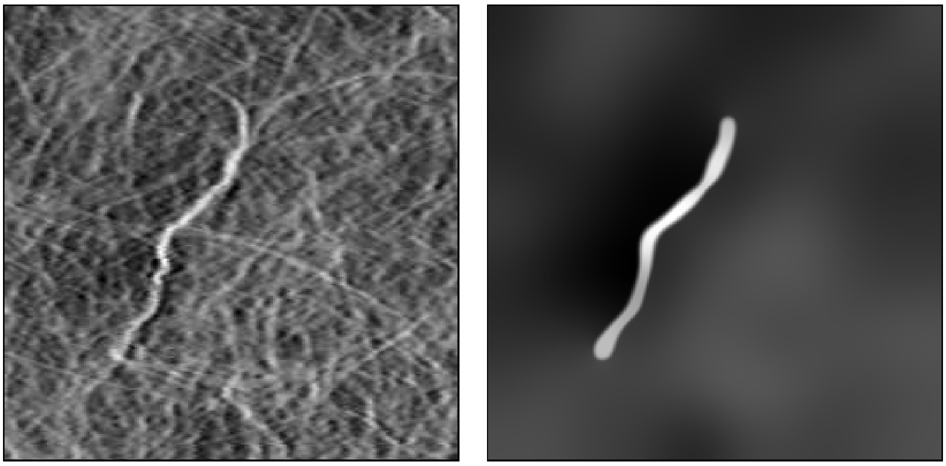

Figure 5.5: Preprocessing of a fabric image. (a) LEFT: Fabric, $\Omega = (0, 257)^2$. (b) RIGHT: Anisotropic diffusion, $\lambda = 4$, $\sigma = 2$, $t = 240$.

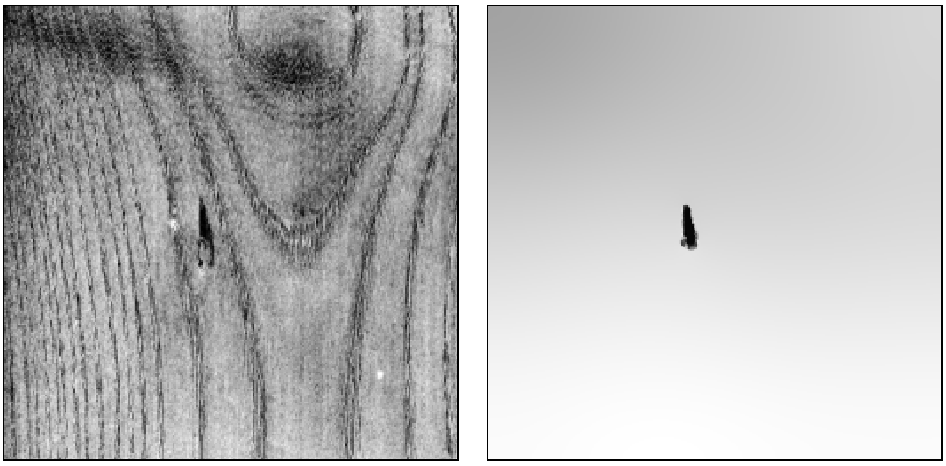

Figure 5.6: Defect detection in wood. (a) LEFT: Wood surface, $\Omega = (0, 256)^2$. (b) RIGHT: Isotropic nonlinear diffusion, $\lambda = 4$, $\sigma = 2$, $t = 2000$.

way along fibres in order to enhance them. However, if one wants to visualize both features separately, one can use a fast pyramid algorithm based on linear diffusion filtering for the clouds [251], whereas stripes can be enhanced by a special nonlinear diffusion filter which is designed for closing interrupted lines and which shall be discussed in Section 5.2.

For furniture production it is of importance to classify the quality of wood surfaces. If one aims to automize this evaluation, one has to process the image in such a way that quality relavant features become better visible und unimportant structures disappear. Fig. 5.6(a) depicts a wood surface possessing one defect. To visualize this defect, equation (5.4) can be applied with good success (Fig. 5.6(b)). In [248] it is demonstrated how a modified anisotropic diffusion process yields even more accurate results with less roundings at the corners.

Fig. 5.7(a) gives an example for possible medical applications of nonlinear diffusion filtering as a preprocessing tool for segmentation (see also [247, 250] for another example). It depicts an MRT slice[3] of the human head. For detecting Alzheimer's disease one is interested in determining the ratio between the ventricle areas (the two white longitudinal objects in the centre) and the entire head area.

In order to make the diagnosis more objective and reliable, it is intended to extract these features by a segmentation algorithm. Figure 5.7(c) shows a segmentation according to the following simplification of the Mumford–Shah functional (1.47):

$$E_f(u,K) = \int_\Omega (u-f)^2\,dx \;+\; \alpha|K|. \tag{5.6}$$

It has been obtained by a MegaWave programme using a hierarchical region-growing algorithm due to Koepfler et al. [142]. As is seen in Fig. 5.7(d), one gets a better segmentation when processing the original image slightly by means of nonlinear diffusion filtering (Fig. 5.7(b)) prior to segmenting it. In order to extend these promising results to the 3D setting, an anisotropic diffusion algorithm for 3D images has been devoloped which allows to filter simultaneously an entire MRT or CT set [202].

[3] kindly provided by Aldo von Wangenheim, German Research Center for Artificial Intelligence, Kaiserslautern.

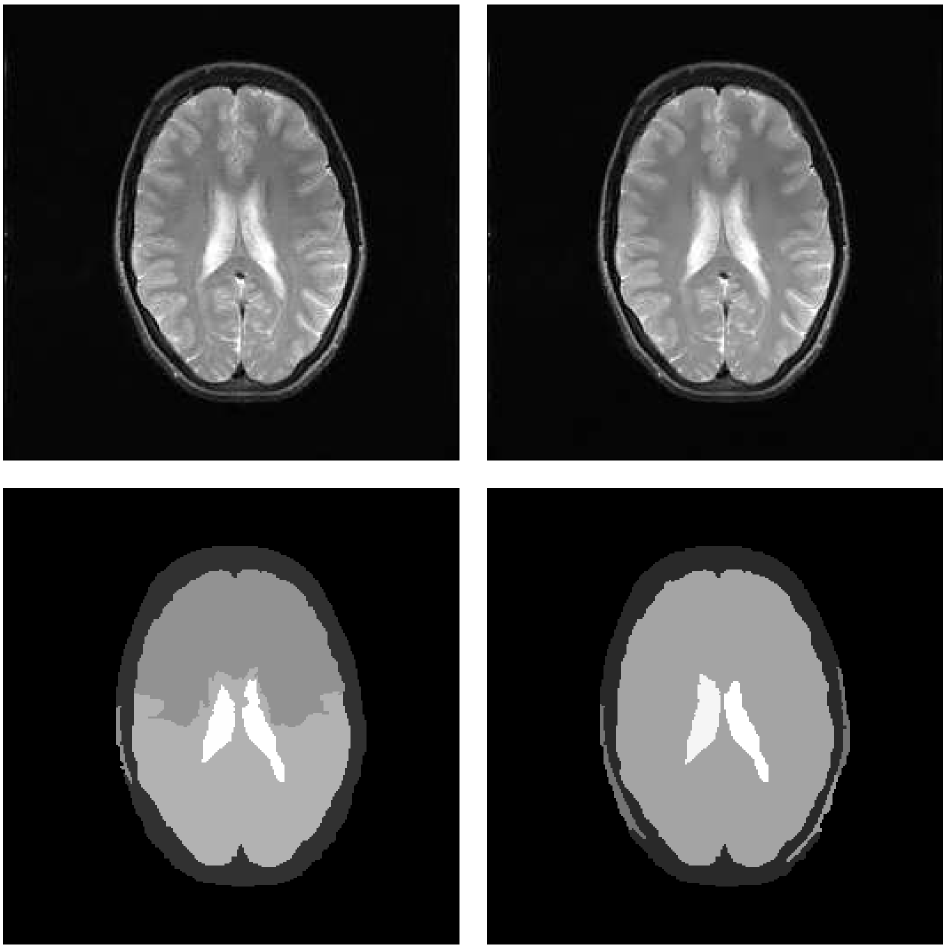

Figure 5.7: Preprocessing of an MRT image. (a) TOP LEFT: Head, $\Omega = (0, 256)^2$. (b) TOP RIGHT: Diffusion-filtered, $\lambda = 5$, $\sigma = 0.1$, $t = 2.5$. (c) BOTTOM LEFT: Segmented original image, $\alpha = 8192$. (d) BOTTOM RIGHT: Segmented filtered image, $\alpha = 8192$.

5.2 Coherence-enhancing diffusion

5.2.1 Filter design

In this section we shall investigate how the structure tensor information can be used to design anisotropic diffusion scale-spaces which enhance the coherence of flow-like textures [252]. This requires a nonvanishing integration scale ρ.

Let again μ_1, μ_2 with $\mu_1 \geq \mu_2$ be the eigenvalues of J_ρ, and v_1, v_2 the corresponding orthonormal eigenvectors. As in 5.1 the diffusion tensor $D(J_\rho(\nabla u_\sigma))$ ought to possess the same set of eigenvectors as $J_\rho(\nabla u_\sigma)$.

If one wants to enhance coherent structures, one should smooth preferably along the coherence direction v_2 with a diffusivity λ_2 which increases with respect to the coherence $(\mu_1 - \mu_2)^2$. This may be achieved by the following choice for the eigenvalues of the diffusion tensor ($C > 0$, $m \in \mathbb{N}$):

$$\begin{aligned} \lambda_1 &:= \alpha, \\ \lambda_2 &:= \begin{cases} \alpha & \text{if } \mu_1 = \mu_2, \\ \alpha + (1-\alpha)\exp\left(\frac{-C}{(\mu_1-\mu_2)^{2m}}\right) & \text{else,} \end{cases} \end{aligned}$$

where the exponential function was chosen to ensure the smoothness of D and the small positive parameter $\alpha \in (0,1)$ keeps $D(J_\rho(\nabla u_\sigma))$ uniformly positive definite.[4]

All examples below are calculated using $C := 1$, $m := 1$, and $\alpha := 0.001$.

5.2.2 Applications

Figure 5.8 illustrates the advantages of local orientation analysis by means of the structure tensor. In order to detect the local orientation of the fingerprint depicted in Fig. 5.8(a), the gradient orientation of a slightly smoothed image has been calculated (Fig. 5.8(b)). Horizontally oriented structures appear black, while vertical structures are represented in white. We observe very high fluctuations in the local orientation. When applying a larger smoothing kernel it is clear that adjacent gradients having the same orientation but opposite direction cancel out. Therefore, the results in (c) are much worse than in (b). The structure tensor, however, averages the gradient orientation instead of its direction. This is the reason for the reliable estimates of local orientation that can be obtained with this method (Fig. 5.8(d)).

To illustrate how the result of anisotopic PDE methods depends on the direction in which they smooth, let us recall the example of mean curvature motion

[4]Evidently, filters of this type are not regularizations of the Perona–Malik process: the limit $\sigma \to 0$, $\rho \to 0$ leads to a linear diffusion process with constant diffusivity α.

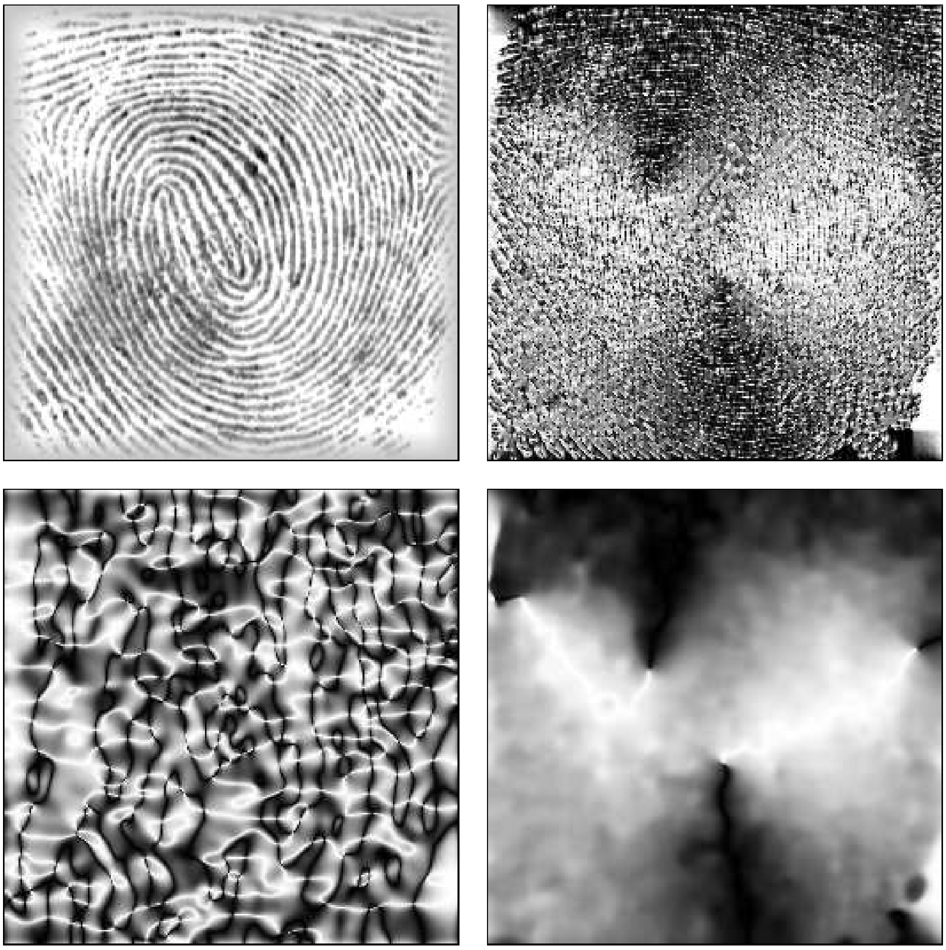

Figure 5.8: Local orientation in a fingerprint image. (a) Top Left: Original fingerprint, $\Omega = (0, 256)^2$. (b) Top Right: Orientation of smoothed gradient, $\sigma = 0.5$. (c) Bottom Left: Orientation of smoothed gradient, $\sigma = 5$. (d) Bottom Right: Structure tensor orientation, $\sigma = 0.5$, $\rho = 4$.

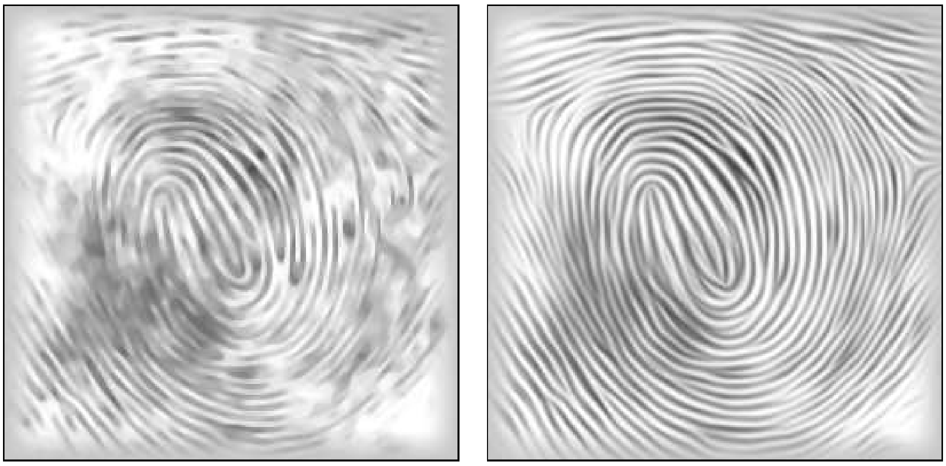

Figure 5.9: Anisotropic equations applied to the fingerprint image. (a) LEFT: Mean-curvature motion, $t = 5$. (b) RIGHT: Coherence-enhancing anisotropic diffusion, $\sigma = 0.5$, $\rho = 4$, $t = 20$.

(cf. 1.5.1):

$$\partial_t u \; = \; u_{\xi\xi} \; = \; |\nabla u| \, \mathrm{curv}(u) \tag{5.7}$$

with ξ being the direction perpendicular to ∇u. Since MCM smoothes by propagating level lines in inner normal direction we recognize that its smoothing direction depends exclusively on ∇u. Thus, although this method is in a complete anisotropic spirit, we should not expect it to be capable of closing interrupted line-like structures. The results in Fig. 5.9(a) confirm this impression.

The proposed anisotropic diffusion filter, however, rotates the diffusive flux towards the coherence orientation v_2 and is therefore well-suited for closing interrupted lines in coherent flow-like textures, see Fig. 5.9(b). Due to its reduced diffusivity at noncoherent structures, the locations of the semantically important singularities in the fingerprint remain the same. This is an important prerequisite that any image processing method has to satisfy if it is to be applied to fingerprint analysis.

Figure 5.10 depicts the scale-space behaviour of coherence-enhancing anisotropic diffusion applied to the fabric image from Fig. 5.5. The temporal behaviour of this diffusion filter seems to be appropriate for visualizing coherent fibre agglomerations (stripes) at different scales, a difficult problem for the automatic grading of nonwovens.

Let us now investigate the impact of coherence-enhancing diffusion on images, which are not typical texture images, but still reveal a flow-like character. To this end, we shall process impressionistic paintings by Vincent van Gogh.

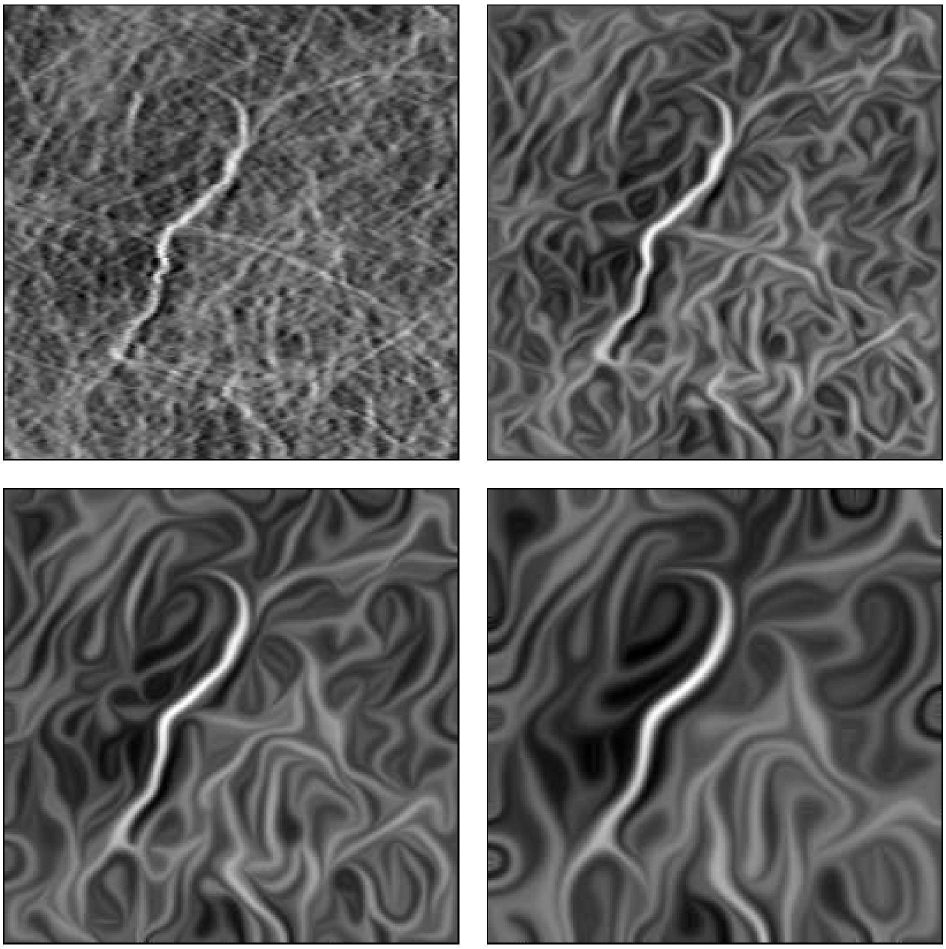

Figure 5.10: Scale-space behaviour of coherence-enhancing diffusion ($\sigma = 0.5$, $\rho = 2$). (a) Top Left: Original fabric image, $\Omega = (0, 257)^2$. (b) Top Right: $t = 20$. (c) Bottom Left: $t = 120$. (d) Bottom Right: $t = 640$.

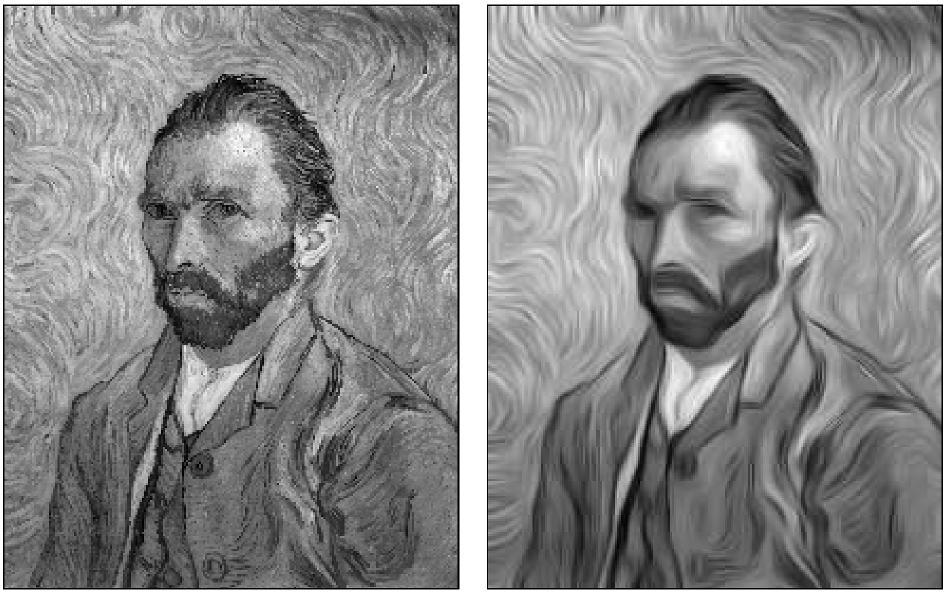

Figure 5.11: Image restoration using coherence-enhancing anisotropic diffusion. (a) LEFT: "Selfportrait" by van Gogh (Saint-Rémy, 1889; Paris, Museé d'Orsay), $\Omega = (0, 215) \times (0, 275)$. (b) RIGHT: Filtered, $\sigma = 0.5$, $\rho = 4$, $t = 6$.

Fig. 5.11 shows the restoration properties of cohererence-enhancing anisotropic diffusion when being applied to a selfportrait of the artist [103]. We observe that the diffusion filter can close interrupted lines and enhance the flow-like character which is typical for van Gogh paintings.

The next painting we are concerned with is called "Lane under Cypresses Below the Starry Sky" [104]. It is depicted in Fig. 5.12. In order to demonstrate the influence of the integration scale ρ, all filter parameters are fixed except for ρ. In Fig. 5.12(b) we observe that a value for ρ which is too small does not lead to the visually dominant coherence orientation and, thus, the filtered structures reveal a lot of undesired fluctuations. Increasing the value for ρ improves the image significantly (Fig. 5.13(c)). Interestingly, a further increasing of ρ does hardly alter this result (Fig. 5.13(d)), which indicates that this van Gogh painting possesses a uniform "texture scale" reflecting the characteristic painting style of the artist.

In a last example the temporal evolution of flow-like images is illustrated by virtue of the "Starry Night" painting in Fig. 5.13 [102, 253]. Due to the established scale-space properties, the image becomes gradually simpler with respect to many

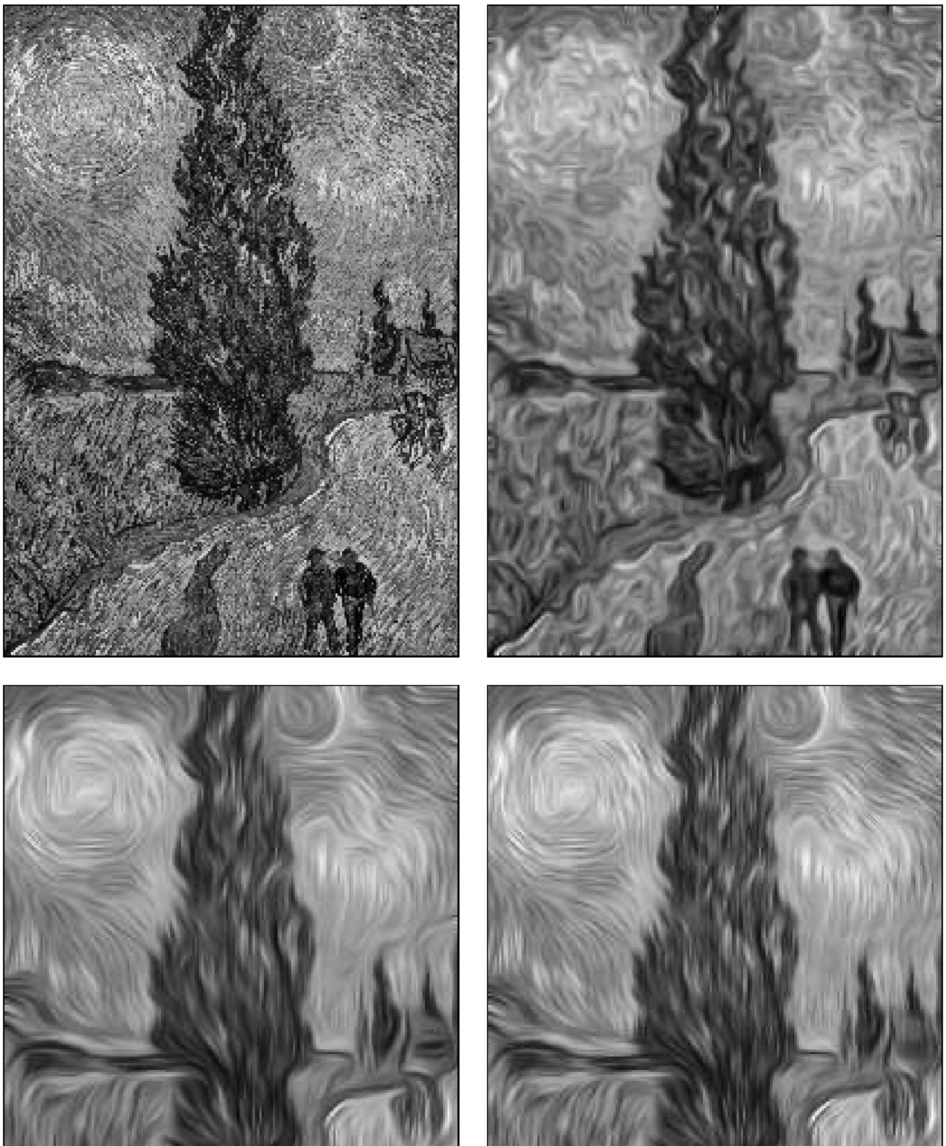

Figure 5.12: Impact of the integration scale on coherence-enhancing anisotropic diffusion ($\sigma = 0.5$, $t = 8$). (a) Top Left: "Lane under Cypresses below the Starry Sky" by van Gogh (Auvers-sur-Oise, 1890; Otterlo, Rijksmuseum Kröller–Müller), $\Omega = (0, 203) \times (0, 290)$. (b) Top Right: Filtered with $\rho = 1$. (c) Bottom Left: $\rho = 4$. (d) Bottom Right: $\rho = 6$.

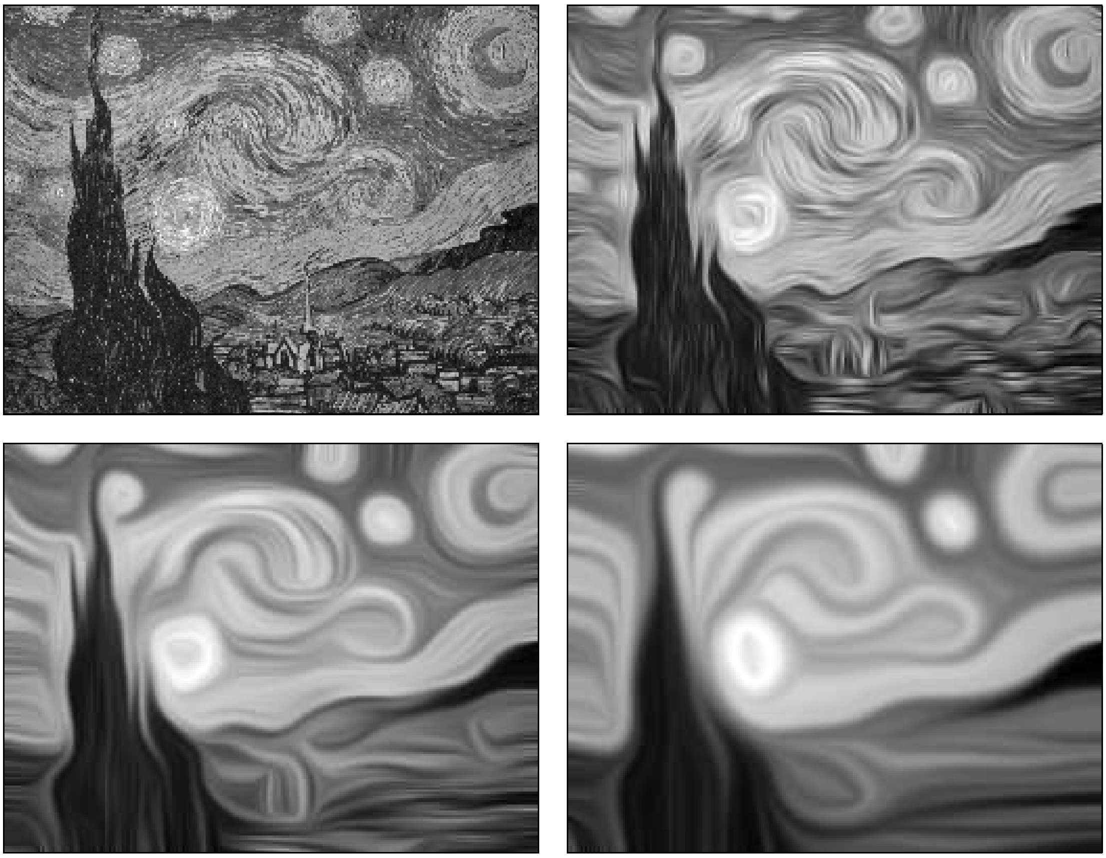

Figure 5.13: Scale-space properties of coherence-enhancing anisotropic diffusion ($\sigma = 0.5$, $\rho = 4$). (a) Top Left: "Starry Night" by van Gogh (Saint-Rémy, 1889; New York, The Museum of Modern Art), $\Omega = (0, 255) \times (0, 199)$. (b) Top Right: $t = 8$. (c) Bottom Left: $t = 64$. (d) Bottom Right: $t = 512$.

aspects, before it finally will tend to its simplest representation, a constant image with the same average grey value as the original one. The flow-like character, however, is maintained for a very long time.[5]

[5]Results for AMSS filtering of this image can be found in [178].

Chapter 6

Conclusions

In this work a scale-space framework has been presented which does not require any monotony assumption (comparison principle). We have seen that, besides the fact that many global smoothing scale-space properties are maintained, new possibilities with respect to image restoration appear.

Rather than deducing a unique equation from first principles we have analysed well-posedness and scale-space properties of a general family of regularized anisotropic diffusion filters. Existence and uniqueness results, continuous dependence of the solution on the initial image, maximum–minimum principles, invariances, Lyapunov functionals, and convergence to a constant steady state have been established.

The large class of Lyapunov functionals permits to regard these filters in numerous ways as simplifying, information-reducing transformations. These global smoothing properties do not contradict seemingly opposite local effects such as edge enhancement. For this reason it is possible to design scale-spaces with restoration properties giving segmentation-like results.

Prerequisites have been stated under which one can prove well-posedness and scale-space results in the continuous, semidiscrete and discrete setting. Each of these frameworks stands on its own and does not require the others. On the other hand, the prerequisites in all three settings reveal many similarities and, as a consequence, representatives of the semidiscrete class can be obtained by suitable spatial discretizations of the continuous class, while representatives of the discrete class may arise from time discretizations of semidiscrete filters.

The degree of freedom within the proposed class of filters can be used to tailor the filters towards specific restoration tasks. Therefore, these scale-spaces do not need to be uncommitted; they give the user the liberty to incorporate a-priori knowledge, for instance concerning size and contrast of especially interesting features.

The analysed class comprises linear diffusion filtering and the nonlinear isotropic model of Catté, Lions, Morel, Coll [58] and Whitaker and Pizer [255], but also novel approaches have been proposed: The use of diffusion tensors instead of scalar-valued diffusivities puts us in a position to design real anisotropic diffusion processes which may reveal advantages at noisy edges. Last but not least, the fact that these filters are steered by the structur tensor instead of the regularized gradient allows to adapt them to more sophisticated tasks such as the enhancement of coherent flow-like structures.

In view of these results, anisotropic diffusion deserves to be regarded as much more than an ad-hoc strategy for transforming a degraded image into a more pleasant looking one. It is a flexible and mathematically sound class of methods which ties the advantages of two worlds: scale-space analysis and image restoration.

Bibliography

[1] R. Acar, C.R. Vogel *Analysis of bounded variation penalty methods for ill-posed problems*, Inverse Problems, Vol. 10, 1217–1229, 1994.

[2] A. Ackah–Miezan, A. Gagalowicz, *Discrete models for energy-minimizing segmentation*, Proc. Fourth Int. Conf. on Computer Vision (ICCV '93, Berlin, May 11–14, 1993), IEEE Computer Society Press, Los Alamitos, 200–207, 1993.

[3] S.T. Acton, A.C. Bovik, M.M. Crawford, *Anisotropic diffusion pyramids for image segmentation*, Proc. First IEEE Int. Conf. Image Processing (ICIP–94, Austin, Nov. 13–16, 1994), IEEE Computer Society Press, Los Alamitos, Vol. 3, 478–482, 1994.

[4] S.T. Acton, M.M. Crawford, *A mean field solution to anisotropic edge detection of remotely sensed data*, Proc. 12th Int. Geoscience and Remote Sensing Symposium (IGARSS '92, Houston, May 26–29, 1992), Vol. 2, 845–847, 1992.

[5] R.A. Adams, *Sobolev spaces*, Academic Press, New York, 1975.

[6] D. Adalsteinsson, J.A. Sethian, *A fast level set method for propagating interfaces*, J. Comput. Phys., Vol. 118, 269–277, 1995.

[7] H.W. Alt, *Lineare Funktionalanalysis*, Springer, Berlin, 1992.

[8] L. Alvarez, F. Guichard, P.-L. Lions, J.-M. Morel, *Axiomes et équations fondamentales du traitement d'images. (Analyse multiéchelle et E.D.P.)*, C. R. Acad. Sci. Paris, t. 315, Série I, 135–138, 1992.

[9] L. Alvarez, F. Guichard, P.-L. Lions, J.-M. Morel, *Axiomatisation et nouveaux opérateurs de la morphologie mathématique*, C. R. Acad. Sci. Paris, t. 315, Série I, 265–268, 1992.

[10] L. Alvarez, F. Guichard, P.-L. Lions, J.-M. Morel, *Équations fondamentales de l'analyse multiéchelle des films*, C. R. Acad. Sci. Paris, t. 315, Série I, 1145–1148, 1992.

[11] L. Alvarez, F. Guichard, P.-L. Lions, J.-M. Morel, *Axioms and fundamental equations in image processing*, Arch. Rational Mech. Anal., Vol. 123, 199–257, 1993.

[12] L. Alvarez, P.-L. Lions, J.-M. Morel, *Image selective smoothing and edge detection by nonlinear diffusion. II*, SIAM J. Numer. Anal., Vol. 29, 845–866, 1992.

[13] L. Alvarez, L. Mazorra, *Signal and image restoration using shock filters and anisotropic diffusion*, SIAM J. Numer. Anal., Vol. 31, 590–605, 1994.

[14] L. Alvarez, F. Morales, *Affine morphological multiscale analysis of corners and multiple junctions*, Report No. 9402, Departamento de Informática y Sistemas, Universidad de Las Palmas, Campus Universitario de Tafira, 35017 Las Palmas, Spain, 1994; to appear in Int. J. Comput. Vision.

[15] L. Alvarez, J.-M. Morel, *Formalization and computational aspects of image analysis*, Acta Numerica, 1–59, 1994.

[16] L. Alvarez, J.-M. Morel, *Morphological approach to multiscale analysis: From principles to equations*, B.M. ter Haar Romeny (Ed.), Geometry-driven diffusion in computer vision, Kluwer, Dordrecht, 229–254, 1994.

[17] L. Ambrosio, *A compactness theorem for a special class of functions of bounded variations*, Boll. Un. Math. Ital., Vol. 3-B, 857–881, 1990.

[18] L. Ambrosio, V. Tortorelli, *On the approximation of functionals depending on jumps by elliptic functionals*, Comm. Pure Appl. Math., Vol. 43, 999–1036, 1990.

[19] L. Ambrosio, V. Tortorelli, *Approximation of functionals depending on jumps by elliptic functionals via* Γ*-convergence*, Boll. Un. Math. Ital., Vol. 7, 105–123, 1992.

[20] S. Angenent, *Parabolic equations on curves and surfaces. Part I. Curves with p-integrable curvature*, Annals of Mathematics, Vol. 132, 451–483, 1990.

[21] S. Angenent, *Parabolic equations on curves and surfaces. Part II. Intersections, blowup, and generalized solutions*, Annals of Mathematics, Vol. 133, 171–215, 1991.

[22] A.B. Arehart, L. Vincent, B.B. Kimia, *Mathematical morphology: The Hamilton–Jacobi connection*, Proc. Fourth Int. Conf. on Computer Vision

(ICCV '93, Berlin, May 11–14, 1993), IEEE Computer Society Press, Los Alamitos, 215–219, 1993.

[23] I. Bajla, M. Marušiak, M. Šrámek, *Anisotropic filtering of MRI data based upon image gradient histogram*, D. Chetverikov, W.G. Kropatsch (Eds.), Computer analysis of images and patterns, Lecture Notes in Comp. Science, Vol. 719, Springer, Berlin, 90–97, 1993.

[24] J. Babaud, A.P. Witkin, M. Baudin, R.O. Duda, *Uniqueness of the Gaussian kernel for scale space filtering*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 8, 26–33, 1986.

[25] C. Ballester Nicolau, *An affine invariant model for image segmentation: Mathematical analysis and applications*, Ph.D. thesis, University of Illes Baleares, Palma de Mallorca, Spain, 1995.

[26] C. Ballester, V. Caselles, M. González, *Affine invariant segmentation by variational method*, Technical Report, Dept. of Math. and Comp. Science, Univ. of Illes Baleares, 07071 Palma de Mallorca, Spain, 1994; to appear in SIAM J. Appl. Math.

[27] G.I. Barenblatt, M. Bertsch, R. Dal Passo, M. Ughi, *A degenerate pseudoparabolic regularization of a nonlinear forward–backward heat equation arising in the theory of heat and mass exchange in stably stratified turbulent shear flow*, SIAM J. Math. Anal., Vol. 24, 1414–1439, 1993.

[28] B. Benhamouda, *Parameter adaptation for nonlinear diffusion in image processing*, master thesis, Dept. of Mathematics, University of Kaiserslautern, P.O. Box 3049, 67653 Kaiserslautern, Germany, 1994.

[29] F. Bergholm, *Edge focusing*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 9, 726–741, 1987.

[30] J. Bigün, G.H. Granlund, *Optimal orientation detection of linear symmetry*, Proc. First Int. Conf. on Computer Vision (ICCV '87, London, June 8–11, 1987), IEEE Computer Society Press, Washington, 433–438, 1987.

[31] J. Bigün, G.H. Granlund, J. Wiklund, *Multidimensional orientation estimation with applications to texture analysis and optical flow*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 13, 775–790, 1991.

[32] A. Blake, A. Zisserman, *Visual reconstruction*, MIT Press, Cambridge (Mass.), 1987.

[33] W. Blaschke, *Vorlesungen über Differentialgeometrie II*, Springer, Berlin, 1923.

[34] R. van den Boomgaard, *Mathematical morphology: Extensions towards computer vision*, Ph.D. thesis, University of Amsterdam, 1992.

[35] R. van den Boomgaard, A.W.M. Smeulders, *Towards a morphological scale-space theory*, Y.L. O, A. Toet, D. Foster, H.J.A.M. Heijmans, P. Meer (Eds.), Shape in picture, Springer, 631–640, 1994.

[36] R. van den Boomgaard, A. Smeulders, *The morphological structure of images: The differential equations of morphological scale-space*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 16, 1101–1113, 1994.

[37] A. Bonnet, *On the Mumford–Shah functional for image segmentation*, oral presentation, Third Int. Congress on Industrial and Applied Mathematics (ICIAM '95, Hamburg, July 3–7, 1995).

[38] J.H. Bramble, B.E. Hubbard, *New monotone type approximations for elliptic problems*, Math. Comp., Vol. 18, 349–367, 1964.

[39] H. Brezis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North Holland, Amsterdam, 1973.

[40] H. Brezis, *Analyse fonctionelle*, Masson, Paris, 1992.

[41] R.W. Brockett, P. Maragos, *Evolution equations for continuous-scale morphology*, Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP–92, San Francisco, March 23–26, 1992), Vol. 3, 125–128, 1992.

[42] R.W. Brockett, P. Maragos, *Evolution equations for continuous-scale morphological filtering*, IEEE Trans. Signal Processing, Vol. 42, 3377–3386, 1994.

[43] A.M. Bruckstein, *Analyzing and synthesizing images by evolving curves*, Proc. First IEEE Int. Conf. Image Processing (ICIP–94, Austin, Nov. 13–16, 1994), IEEE Computer Society Press, Los Alamitos, Vol. 1, 11–15, 1994.

[44] A.M. Bruckstein, G. Sapiro, D. Shaked, *Evolution of planar polygons*, CIS Report No. 9202 (2nd revision), Computer Science Dept., Technion, Haifa 32000, Israel, 1993.

[45] A.M. Bruckstein, D. Shaked, *On projective invariant smoothing and evolution of planar curves and polygons*, CIS Report No. 9328, Computer Science Dept., Technion, Haifa 32000, Israel, 1993.

[46] B. Buck, V. Macaulay (Eds.), *Maximum entropy in action*, Clarendon, Oxford, 1991.

[47] P.J. Burt, E.H. Adelson, *The Laplacian pyramid as a compact image code*, IEEE Trans. Comm., Vol. 31, 532–540, 1983.

[48] L.D. Cai, *Some notes on repeated averaging smoothing*, J. Kittler (Ed.), Pattern recognition, Lecture Notes in Comp. Science, Vol. 301, Springer, Berlin, 597–605, 1988.

[49] L.D. Cai, *A "small leakage" model for diffusion smoothing of image data*, Proc. 11th Int. Joint Conf. on Artificial Intelligence (IJCAI '89, Detroit, Aug. 20–25, 1989), Vol. 2, 1585–1590, 1989.

[50] M.P. do Carmo, *Differentialgeometrie von Kurven und Flächen*, Vieweg, Braunschweig, 1983.

[51] J. Canny, *A computational approach to edge detection*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 8, 679–698, 1986.

[52] V. Caselles, F. Catté, T. Coll, F. Dibos, *A geometric model for active contours in image processing*, Numer. Math., Vol. 66, 1–31, 1993.

[53] V. Caselles, B. Coll, *Snakes in movement*, Technical Report, Dept. of Math. and Comp. Science, Univ. of Illes Baleares, 07071 Palma de Mallorca, Spain, 1993.

[54] V. Caselles, C. Sbert, *What is the best causal scale-space for 3-D images?*, Technical Report, Dept. of Math. and Comp. Science, Univ. of Illes Baleares, 07071 Palma de Mallorca, Spain, 1994; submitted to SIAM J. Appl. Math.

[55] V. Caselles, R. Kimmel, G. Sapiro, *Geodesic snakes*, Technical Report, Dept. of Math. and Comp. Science, Univ. of Illes Baleares, 07071 Palma de Mallorca, Spain, 1994.

[56] F. Catté, *Convergence of iterated affine and morphological filters by nonlinear semi-group theory*, Preprint No. 9528, CEREMADE, Université Paris IX – Dauphine, Place du Maréchal de Lattre de Tassigny, 75775 Paris Cedex 16, France, 1995.

[57] F. Catté, F. Dibos, G. Koepfler, *A morphological scheme for mean curvature motion and applications to anisotropic diffusion and motion of level sets*, Proc. First IEEE Int. Conf. Image Processing (ICIP–94, Austin, Nov. 13–16, 1994), IEEE Computer Society Press, Los Alamitos, Vol. 1, 26–30, 1994.

[58] F. Catté, P.-L. Lions, J.-M. Morel, T. Coll, *Image selective smoothing and edge detection by nonlinear diffusion*, SIAM J. Numer. Anal., Vol. 29, 182–193, 1992.

[59] A. Chambolle, *Partial differential equations and image processing*, Proc. First IEEE Int. Conf. Image Processing (ICIP–94, Austin, Nov. 13–16, 1994), IEEE Computer Society Press, Los Alamitos, Vol. 1, 16–20, 1994.

[60] R.H. Chan, T.F. Chan, C.K. Wong, *Cosine transform based preconditioners for total variation minimization problems in image processing*, CAM Report 95-23, Dept. of Mathematics, Univ. of Califormia, Los Angeles, CA 90024, U.S.A., 1995.

[61] T.F. Chan, H.M. Zhou, R.H. Chan, *Continuation method for total variation minimizing problems in image processing*, CAM Report 95-18, Dept. of Mathematics, Univ. of Califormia, Los Angeles, CA 90024, U.S.A, 1995; to appear in Advanced Signal Processing Algorithms, SPIE Vol. 2563, 1995.

[62] Y.-G. Chen, Y. Giga, S. Goto, *Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations*, J. Differential Geometry, Vol. 33, 749–786, 1991.

[63] M.A. Cohen, S. Grossberg, *Neural dynamics of brightness perception: Features, boundaries, diffusion and resonance*, Perception and Psychophysics, Vol. 36, 428–456, 1984.

[64] T. Cohignac, F. Eve, F. Guichard, C. Lopez, J.-M. Morel, *Numerical analysis of the fundamental equation of image processing*, Preprint No. 9254, CEREMADE, Université Paris IX – Dauphine, Place du Maréchal de Lattre de Tassigny, 75775 Paris Cedex 16, France, 1992; to appear in Int. J. Comput. Vision.

[65] G.-H. Cottet, *Diffusion approximation on neural networks and applications for image processing*, F. Hodnett (Ed.), Proc. Sixth European Conf. on Mathematics in Industry, Teubner, Stuttgart, 3–9, 1992.

[66] G.-H. Cottet, L. Germain, *Image processing through reaction combined with nonlinear diffusion*, Math. Comp., Vol. 61, 659–673, 1993.

[67] G.-H. Cottet, *Neural networks: Continuous approach and applications to image processing*, Technical Report No. 113, LMC – IMAG, Université Joseph Fourier, B.P. 53, 38041 Grenoble Cedex 9, France, 1994; submitted to Proc. 2nd Europ. Conf. on Mathematics Applied to Biology and Medicine (Lyon, Dec. 15–18, 1993).

[68] M.G. Crandall, I. Ishii, P.-L. Lions, *User's guide to viscosity solutions of second order partial differential equations*, Bull. Am. Math. Soc., Vol. 27, 1–67, 1992.

[69] J. Damon, *Local Morse theory for solutions to the heat equation and Gaussian blurring*, J. Differential Equations, Vol. 115, 368–401, 1995.

[70] E. De Giorgi, M. Carriero, A. Laeci, *Existence theorem for a minimum problem with a free discontinuity set*, Arch. Rat. Mech. Anal., Vol. 108, 195–218, 1989.

[71] D. Dettwiler, *Stabile Bildglättung und Extraktion von Strukturprimitiven mittels nichtlinearer Diffusion*, Semesterarbeit, Communication Technology Laboratory, Image Science Division, ETH–Zentrum, CH–8092 Zürich, Switzerland, 1993.

[72] F. Dibos, *Analyse multiéchelle invariante par transformations projectives*, C. R. Acad. Sci. Paris, t. 320, Série I, 917–921, 1995.

[73] D.C. Dobson, F. Santosa, *An image enhancement technique for electrical impedance tomography*, Inverse Problems, Vol. 10, 317–334, 1994.

[74] D.C. Dobson, F. Santosa, *Recovery of blocky images from noisy and blurred data*, Technical Report, Dept. of Mathematics, Texas A&M University, College Station, TX 77843–3368, U.S.A.; to appear in SIAM J. Appl. Math.

[75] D.C. Dobson, C.R. Vogel, *Convergence of an iterative method for total variation denoising*, Preprint, Dept. of Mathematical Sciences, Montana State University, Bozeman, MT 59717-0240, U.S.A., 1995; submitted to SIAM J. Numer. Anal.

[76] J. Douglas, J.E. Gunn, *A general formulation of alternating direction methods. Part I. Parabolic and hyperbolic problems*, Numer. Math., Vol. 6, 428–453, 1964.

[77] J. Douglas, B.F. Jones, *On predictor–corrector methods for non-linear parabolic differential equations*, J. Soc. Ind. Appl. Math., Vol. 11, 195–204, 1963.

[78] S.K. Dzhu Magazieva, *Numerical study of a partial differential equation*, U.S.S.R. Comput. Maths. Math. Phys., Vol. 23, No. 4, 45–49, 1983.

[79] D. Eberly, *A differential geometric approach to anisotropic diffusion*, B.M. ter Haar Romeny (Ed.), Geometry-driven diffusion in computer vision, Kluwer, Dordrecht, 371–391, 1994.

[80] S. Eidelman, W. Grossmann, A. Friedman, *Nonlinear signal processing using integration of fluid dynamics equations*, A.G. Tescher (Ed.): Applications of digital image processing XIV, SPIE Vol. 1567, 439–450, 1991.

[81] A.I. El-Fallah, G.E. Ford, *Nonlinear adaptive image filtering based on inhomogeneous diffusion and differential geometry*, S.A. Rajala, R.L. Stevenson (Eds.), Image and video processing II, SPIE Vol. 2182, 49–63, 1994.

[82] A.I. El-Fallah, G.E. Ford, *The evolution of mean curvature in image filtering*, Proc. First IEEE Int. Conf. Image Processing (ICIP–94, Austin, Nov. 13–16, 1994), IEEE Computer Society Press, Los Alamitos, Vol. 1, 298–302, 1994.

[83] L.C. Evans, J. Spruck, *Motion of level sets by mean curvature I.*, J. Differential Geometry, Vol. 33, 635–681, 1991.

[84] O. Faugeras, *Sur l'évolution de courbes simples du plan projectif réel*, C. R. Acad. Sci. Paris, t. 317, Série I, 565–570, 1993.

[85] O. Faugeras, *Cartan's moving frame method and its application to the geometry and evolution of curves in the Euclidean, affine and projective planes*, J.L. Mundy, A. Zisserman, D. Forsyth (Eds.), Applications of invariance in computer vision, Lecture Notes in Comp. Science, Vol. 825, Springer, Berlin, 11–46, 1994.

[86] O. Faugeras, R. Keriven, *Scale-spaces and affine curvature*, W. Chenke, R. Mohr (Eds.), Proc. Workshop on Geometrical Modelling and Invariants for Computer Vision (Xi'an, April 27–30, 1995), Xidian University Press, Xi'an, 1995.

[87] P. Fiddelaers, E.J. Pauwels, L.J. Van Gool, *Geometry-driven curve evolution*, J.-O. Eklundh (Ed.), Computer vision – ECCV '94, Lecture Notes in Comp. Science, Vol. 800, Springer, Berlin, 427–432, 1994.

[88] L.M.J. Florack, *The syntactical structure of scalar images*, Ph.D. thesis, Utrecht University, The Netherlands, 1993.

[89] L.M.J. Florack, B.M. ter Haar Romeny, J.J. Koenderink, M.A. Viergever, *Linear scale-space*, J. Math. Imag. Vision, Vol. 4, 325–351, 1994.

[90] L.M.J. Florack, A.H. Salden, B.M. ter Haar Romeny, J.J. Koenderink, M.A. Viergever, *Nonlinear scale-space*, B.M. ter Haar Romeny (Ed.), Geometry-driven diffusion in computer vision, Kluwer, Dordrecht, 339–370, 1994.

[91] F.J. Fritz, B. Huppert, W. Willems, *Stochastische Matrizen*, Springer, Berlin, 1979.

[92] J. Fröhlich, J. Weickert, *Image processing using a wavelet algorithm for nonlinear diffusion*, Report No. 104, Laboratory of Technomathematics, University of Kaiserslautern, P.O. Box 3049, 67653 Kaiserslautern, Germany, 1994.

[93] D. Gabor, *Information theory in electron microscopy*, Laboratory Investigation, Vol. 14, 801–807, 1965.

[94] M. Gage, *On area-preserving evolution equation for plane curves*, Contemp. Math., Vol. 23, 51–62, 1986.

[95] M. Gage, R.S. Hamilton, *The heat equation shrinking convex plane curves*, J. Differential Geometry, Vol. 23, 69–96, 1986.

[96] D. Geiger, A. Yuille, *A common framework for image segmentation*, Int. J. Comput. Vision, Vol. 6, 227–243, 1991.

[97] D. Geman, G. Reynolds, *Constrained restoration and the recovery of discontinuities*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 14, 367–383, 1992.

[98] S. Geman, D. Geman, *Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 6, 721–741, 1984.

[99] G. Gerig, O. Kübler, R. Kikinis, F.A. Jolesz, *Nonlinear anisotropic filtering of MRI data*, IEEE Trans. Medical Imaging, Vol. 11, 221–232, 1992.

[100] G. Gerig, G. Székely, G. Israel, M. Berger, *Detection and characterization of unsharp blobs by curve evolution*, Y. Bizais et al. (Eds.), Information processing in medical imaging (IPMI '95), Series on Computational Imaging and Vision, Kluwer, Dordrecht, 165–176, 1995.

[101] D. Giuiliani, *Directed diffusion equations and edge detection*, master thesis, Dept. of Mathematics, University of Kaiserslautern, P.O. Box 3049, 67653 Kaiserslautern, Germany, 1991.

[102] V. van Gogh, *Starry night*, Saint-Rémy, 1889. New York, The Museum of Modern Art.

[103] V. van Gogh, *Selfportrait*, Saint-Rémy, 1889. Paris, Museé d'Orsay.

[104] V. van Gogh, *Lane under cypresses below the starry sky*, Auvers-sur-Oise, 1890. Otterlo, Rijksmuseum Kröller–Müller.

[105] R.C. Gonzalez, P. Wintz, *Digital image processing*, Addison–Wesley, Reading, 1987.

[106] R. Gorenflo, *Differenzenschemata monotoner Art für lineare parabolische Randwertaufgaben*, ZAMM, Vol. 51, 595–610, 1971.

[107] R. Gorenflo, *Conservative difference schemes for diffusion problems*, Preprint No. 39, FB Mathematik, FU Berlin, Arnimallee 2–6, W-1000 Berlin 33, Germany, 1977.

[108] A.R. Gourlay, *Implicit convolution*, Image Vision Comput., Vol. 3, 15–23, 1985.

[109] R. Graham, *Snow removal: A noise-stripping process for TV signals*, IRE Trans. Information Theory, Vol. 8, 129–144, 1962.

[110] M. Grayson, *The heat equation shrinks embedded plane curves to round points*, J. Differential Geometry, Vol. 26, 285–314, 1987.

[111] D. Greenspan, *Partial differential approximations with non-negative coefficients*, J. Franklin Inst., Vol. 275, 481–490, 1963.

[112] D. Greenspan, P.C. Jain, *On non-negative difference analogues of elliptic differential equations*, J. Franklin Inst., Vol. 279, 360–365, 1965.

[113] B.M. ter Haar Romeny (Ed.), *Geometry-driven diffusion in computer vision*, Kluwer, Dordrecht, 1994.

[114] B.M. ter Haar Romeny, L.M.J. Florack, J.J. Koenderink, M.A. Viergever, *Scale space: Its natural operators and differential invariants*, A.C.F. Colchester, D.J. Hawkes (Eds.), Information processing in medical imaging, Lecture Notes in Comp. Science, Vol. 511, Springer, Berlin, 239–255, 1991.

[115] B.M. ter Haar Romeny, L.M.J. Florack, M. de Swart, J. Wilting, M.A. Viergever, *Deblurring Gaussian blur*, F.L. Bookstein, J.S. Duncan, N. Lange, D.C. Wilson (Eds.), Mathematical methods in medical imaging III, SPIE Vol. 2299, 139–148, 1994.

[116] W. Hahn, *Theorie und Anwendung der direkten Methode von Ljapunov*, Springer, Berlin, 1959.

[117] H. Hamy, P. Prez, *Processus de diffusion en vision computationelle: Analyse critique de la notion d'échelle*, Report No. 9518, CREA, Ecole Polytechnique, 1, rue Descartes, 75005 Paris, France, 1995.

[118] R.M. Haralick, S.R. Sternberg, X. Zhuang, *Image analysis using mathematical morphology*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 9, 532–550, 1987.

[119] H.J.A.M. Heijmans, *Mathematical morphology: A modern approach in image processing based on algebra and geometry*, SIAM Review, Vol. 37, 1–36, 1995.

[120] G. Hellwig, *Partial differential equations*, Teubner, Stuttgart, 1977.

[121] K. Höllig, *Existence of infinitely many solutions for a forward–backward heat equation*, Trans. Amer. Math. Soc., Vol. 278, 299–316, 1983.

[122] K. Höllig, J.A. Nohel, *A diffusion equation with a nonmonotone constitutive function*, J.M. Ball (Ed.), Systems of nonlinear partial differential equations, Reidel, Dordrecht, 409–422, 1983.

[123] G. Huisken, *Flow by mean curvature of convex surfaces into spheres*, J. Differential Geometry, Vol. 20, 237–266, 1984.

[124] R.A. Hummel, *Representations based on zero-crossings in scale space*, Proc. IEEE Comp. Soc. Conf. Computer Vision and Pattern Recognition (CVPR '86, Miami Beach, June 22–26, 1986), IEEE Computer Society Press, Washington, 204–209, 1986.

[125] R. Hummel, R. Moniot, *Reconstructions from zero-crossings in scale space*, IEEE Trans. Acoustics, Speech and Signal Processing, Vol. 37, 2111–2130, 1989.

[126] R. Illner, *Global existence for two-velocity models of the Boltzmann equation*, Math. Meth. Appl. Sci., Vol. 1, 187–193, 1979.

[127] R. Illner, H. Neunzert, *Relative entropy maximization and directed diffusion equations*, Math. Meth. Appl. Sci., Vol. 16, 545–554, 1993.

[128] R. Illner, J. Tie, *On directed diffusion with measurable background*, Math. Meth. Appl. Sci., Vol. 16, 681–690, 1993.

[129] K. Ito, K. Kunisch, *Augmented Lagrangian methods for nonsmooth, convex optimization in Hilbert spaces*, Preprint No. 409/1994, Fachbereich Mathematik (3), Technische Universität Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany, 1994.

[130] K. Ito, K. Kunisch, *An active set strategy for image restoration based on the augmented Lagrangian formulation*, Preprint No. 410/1994, Fachbereich Mathematik (3), Technische Universität Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany, 1994.

[131] P. T. Jackway, *Morphological scale-space*, Proc. 11th Int. Conf. Pattern Recognition (ICPR 11, The Hague, Aug. 30 – Sept. 3, 1992), Vol. C, 252–255, 1992.

[132] P. Jackway, W.W. Boles, M. Deriche, *Morphological scale-space fingerprints and their use in object recognition in range images*, Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP–94, Adelaide, April 19–22, 1994), Vol. 5, 5–8, 1994.

[133] B. Jähne, *Spatio-temporal image processing*, Lecture Notes in Comp. Science, Vol. 751, Springer, Berlin, 1993.

[134] F. John, *Partial differential equations*, Springer, New York, 1978.

[135] M. Kass, A. Witkin, *Analyzing oriented patterns*, Computer Vision, Graphics, and Image Processing, Vol. 37, 362–385, 1987.

[136] J.G. Kemeny, J.L. Snell, *Finite Markov chains*, D. Van Nostrand, Princeton, 1960.

[137] B.B. Kimia, K. Siddiqi, *Geometric heat equation and non-linear diffusion of shapes and images*, Proc. IEEE Comp. Soc. Conf. Computer Vision and Pattern Recognition (CVPR '94, Seattle, June 21–23, 1994), IEEE Computer Society Press, Los Alamitos, 113–120, 1994.

[138] B.B. Kimia, A. Tannenbaum, S.W. Zucker, *Toward a computational theory of shape: An overview*, O. Faugeras (Ed.), Computer vision – ECCV '90, Lecture Notes in Comp. Science, Vol. 427, Springer, Berlin, 402–407, 1990.

[139] B.B. Kimia, A. Tannenbaum, S.W. Zucker, *On the evolution of curves via a function of curvature. I. The classical case*, J. Math. Anal. Appl., Vol. 163, 438–458, 1992.

[140] B.B. Kimia, A.R. Tannenbaum, S.W. Zucker, *Non-linear shape approximation via the entropy scale space*, B.C. Vemuri (Ed.), Geometric methods in computer vision II, SPIE Vol. 2031, 218–233, 1993.

[141] H.E. Knutsson, R. Wilson, G.H. Granlund, *Anisotropic nonstationary image estimation and its applications: Part I – Restoration of noisy images*, IEEE Trans. Comm., Vol. 31, 388–397, 1983.

[142] G. Koepfler, C. Lopez, J.-M. Morel, *A multiscale algorithm for image segmentation by variational method*, SIAM J. Numer. Anal., Vol. 31, 282–299, 1994.

[143] J.J. Koenderink, *The structure of images*, Biol. Cybern., Vol. 50, 363–370, 1984.

[144] O.A. Ladyženskaja, V.A. Solonnikov, N.N. Ural'ceva, *Linear and quasilinear equations of parabolic type*, American Mathematical Society, Providence, 1968.

[145] A. Leaci, S. Solimini, *Variational problems with a free discontinuity set*, B.M. ter Haar Romeny (Ed.), Geometry-driven diffusion in computer vision, Kluwer, Dordrecht, 147–154, 1994.

[146] A. Lev, S. Zucker, A. Rosenfeld, *Iterative enhancement of noisy images*, IEEE Trans. Systems, Man and Cybernetics, Vol. 7, 435–442, 1977.

[147] V. Li, F. Santosa, *An affine scaling algorithm for minimizing total variation in image enhancement*, Technical Report CTC94TR201, Cornell Theory Center, Cornell University, Ithaca, NY 14853–3801, U.S.A., 1994.

[148] X. Li, T. Chen, *Nonlinear diffusion with multiple edginess thresholds*, Pattern Recognition, Vol. 27, 1029–1037, 1994.

[149] L.M. Lifshitz, S.M. Pizer, *A multiresolution hierarchical approach to image segmentation based on intensity extrema*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 12, 529–540, 1990.

[150] T. Lindeberg, *Scale-space for discrete signals*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 12, 234–254, 1990.

[151] T. Lindeberg, *Scale-space theory in computer vision*, Kluwer, Boston, 1994.

[152] T. Lindeberg, J. Gårding, *Shape-adapted smoothing in estimation of 3-D depth cues from affine distortions of local 2-D brightness structure*, J.-O. Eklundh (Ed.), Computer vision – ECCV '94, Lecture Notes in Comp. Science, Vol. 800, Springer, Berlin, 389–400, 1994.

[153] T. Lindeberg, B. M. ter Haar Romeny, *Linear scale-space II: Early visual operations*, B.M. ter Haar Romeny (Ed.), Geometry-driven diffusion in computer vision, Kluwer, Dordrecht, 39–71, 1994.

[154] M. Lindenbaum, M. Fischer, A. Bruckstein, *On Gabor's contribution to image enhancement*, Pattern Recognition, Vol. 27, 1–8, 1994.

[155] P.-L. Lions, *Axiomatic derivation of image processing models*, Math. Mod. Meth. Appl. Sci., Vol. 4, 467–475, 1994.

[156] P.-L. Lions, S. Osher, L.I. Rudin, *Denoising and deblurring images using constrained nonlinear partial differential equations*, Preprint, Cognitech, Inc., 2800 – 28 St., Suite 101, Santa Monica, CA 90405, U.S.A., 1993; submitted to SIAM J. Numer. Anal.

[157] M.H. Loew, J. Rosenman, J. Chen, *A clinical tool for enhancement of portal images*, M.H. Loew (Ed.), Image processing, SPIE Vol. 2167, 543–550, 1994.

[158] C. Lopez, J.-M. Morel, *Axiomatization of shape analysis and applications to texture hyperdiscrimination*, Proc. IEEE Comp. Soc. Conf. Computer Vision and Pattern Recognition (CVPR '93, New York, June 15–17, 1993), IEEE Computer Society Press, Los Alamitos, 646–647, 1993.

[159] A. Lukschin, H. Neunzert, J. Struckmeier, *Interim report of the project DPH 6473/91 – Coupling of Navier-Stokes and Boltzmann regions*, Internal Report, Laboratory of Technomathematics, Kaiserslautern, 1993.

[160] R. Malladi, J.A. Sethian, *A unified framework for shape segmentation, representation and recognition*, Report PAM-614, Center for Pure and Applied Mathematics, University of California, Berkeley, CA 94720, U.S.A., 1994.

[161] R. Malladi, J.A. Sethian, B.C. Vemuri, *A topology independent shape modeling scheme*, B.C. Vemuri (Ed.), Geometric methods in computer vision II, SPIE Vol. 2031, 246–258, 1993.

[162] R. Malladi, J.A. Sethian, B.C. Vemuri, *Shape modeling with front propagation: A level set approach*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 17, 158–175, 1995.

[163] P. Maragos, *A representation theory for morphological image and signal processing*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 11, 586–599, 1989.

[164] G.I. Marchuk, *Splitting and alternating direction methods*, P.G. Ciarlet, J.-L. Lions (Eds.), Handbook of numerical analysis, Vol. I, 197–462, 1990.

[165] D. Marr, E. Hildreth, *Theory of edge detection*, Proc. R. Soc. London Ser. B, Vol. 207, 187–217, 1980.

[166] J.L. Marroquin, *Deterministic interactive particle models for image processing and computer graphics*, CVGIP: Graphical Models and Image Processing, Vol. 55, 408–417, 1993.

[167] G. Matheron, *Random sets and integral geometry*, Wiley, New York, 1975.

[168] A.R. Mitchell, D.F. Griffiths, *The finite difference method in partial differential equations*, Wiley, Chichester, 1980.

[169] L. Moisan, *Analyse multiéchelle de films pour la reconstruction du relief*, C. R. Acad. Sci. Paris, t. 320, Série I, 279–284, 1995.

[170] F. Mokhtarian, A.K. Mackworth, *A theory of multiscale, curvature-based shape representation for planar curves*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 14, 789–805, 1992.

[171] J.-M. Morel, S. Solimini, *Variational methods in image segmentation*, Birkhäuser, Boston, 1994.

[172] D. Mumford, J. Shah, *Boundary detection by minimizing functionals, I*, Proc. IEEE Comp. Soc. Conf. Computer Vision and Pattern Recognition (CVPR '85, San Francisco, June 19–23, 1985), IEEE Computer Society Press, Washington, 22–26, 1985.

[173] D. Mumford, J. Shah, *Optimal approximations by piecewise smooth functions and associated variational problems*, Comm. Pure Appl. Math., Vol. 42, 577–685, 1989.

[174] M. Nagao, T. Matsuyama, *Edge preserving smoothing*, Computer Graphics and Image Processing, Vol. 9, 394–407, 1979.

[175] P. Neskovic, B.B. Kimia, *Three-dimensional shape representation from curvature dependent surface evolution*, Proc. First IEEE Int. Conf. Image Processing (ICIP–94, Austin, Nov. 13–16, 1994), IEEE Computer Society Press, Los Alamitos, Vol. 1, 6–10, 1994.

[176] T.G. Newman, H. Dirilten, *A nonlinear transformation for digital picture processing*, IEEE Trans. Computers, Vol. 22, 869–873, 1973.

[177] W.J. Niessen, B.M. ter Haar Romeny, L.M.J. Florack, A.H. Salden, M.A. Viergever, *Nonlinear diffusion of scalar images using well-posed differential operators*, Proc. IEEE Comp. Soc. Conf. Computer Vision and Pattern Recognition (CVPR '94, Seattle, June 21–23, 1994), IEEE Computer Society Press, Los Alamitos, 92–97, 1994.

[178] W.J. Niessen, B.M. ter Haar Romeny, M.A. Viergever, *Numerical analysis of geometry-driven diffusion equations*, B.M. ter Haar Romeny (Ed.), Geometry-driven diffusion in computer vision, Kluwer, Dordrecht, 392–410, 1994.

[179] M. Nitzberg, D. Mumford, T. Shiota, *Filtering, segmentation and depth*, Lecture Notes in Comp. Science, Vol. 662, Springer, Berlin, 1993.

[180] M. Nitzberg, T. Shiota, *Nonlinear image filtering with edge and corner enhancement*, Technical Report No. 90–2, Division of Applied Sciences, Harvard University, Cambridge, MA 02138, U.S.A., 1990.

[181] M. Nitzberg, T. Shiota, *Nonlinear image filtering with edge and corner enhancement*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 14, 826–833, 1992.

[182] N. Nordström, *Biased anisotropic diffusion - a unified regularization and diffusion approach to edge detection*, O. Faugeras (Ed.), Computer vision – ECCV '90, Lecture Notes in Comp. Science, Vol. 427, Springer, Berlin, 18–27, 1990.

[183] N. Nordström, *Biased anisotropic diffusion - a unified regularization and diffusion approach to edge detection*, Image Vision Comput., Vol. 8, 318–327, 1990.

[184] P.J. Olver, G. Sapiro, A. Tannenbaum, *Classification and uniqueness of invariant geometric flows*, C. R. Acad. Sci. Paris, t. 319, Série I, 339–344, 1994.

[185] P. Olver, G. Sapiro, A. Tannenbaum, *Differential invariant signatures and flows in computer vision: A symmetry group approach*, B.M. ter Haar Romeny (Ed.), Geometry-driven diffusion in computer vision, Kluwer, Dordrecht, 255–306, 1994.

[186] S. Osher, L. Rudin, *Feature-oriented image enhancement using shock filters*, SIAM J. Numer. Anal., Vol. 27, 919–949, 1990.

[187] S. Osher, L. Rudin, *Shocks and other nonlinear filtering applied to image processing*, A.G. Tescher (Ed.): Applications of digital image processing XIV, SPIE Vol. 1567, 414–425, 1991.

[188] S. Osher, J.A. Sethian, *Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton–Jacobi formulations*, J. Comput. Phys., Vol. 79, 12–49, 1988.

[189] K. Ottenberg, *Model-based extraction of geometric structure from digital images*, Ph.D. thesis, Utrecht University, The Netherlands, 1993.

[190] P. Perona, J. Malik, *Scale space and edge detection using anisotropic diffusion*, Proc. IEEE Comp. Soc. Workshop on Computer Vision (Miami Beach, Nov. 30 – Dec. 2, 1987), IEEE Computer Society Press, Washington, 16–22, 1987.

[191] P. Perona, J. Malik, *A network for multiscale image segmentation*, Proc. IEEE Int. Symp. Circuits and Systems (ISCAS–88, Espoo, June 7–9, 1988), 2565–2568, 1988.

[192] P. Perona, J. Malik, *Scale space and edge detection using anisotropic diffusion*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 12, 629–639, 1990.

[193] P. Perona, T. Shiota, J. Malik, *Anisotropic diffusion*, B.M. ter Haar Romeny (Ed.), Geometry-driven diffusion in computer vision, Kluwer, Dordrecht, 72–92, 1994.

[194] P. Perona, M. Tartagni, *Diffusion networks for on-chip image contrast normalization*, Proc. First IEEE Int. Conf. Image Processing (ICIP–94, Austin, Nov. 13–16, 1994), IEEE Computer Society Press, Los Alamitos, Vol. 1, 1–5, 1994.

[195] I.G. Petrowski, *Vorlesungen über partielle Differentialgleichungen*, Teubner, Leipzig, 1955.

[196] E.S. Posmentier, *The generation of salinity finestructure by vertical diffusion*, J. Phys. Oceanogr., Vol. 7, 298–300, 1977.

[197] C.B. Price, P. Wambacq, A. Oosterlinck, *Applications of reaction–diffusion equations to image processing*, IEE Third Int. Conf. Image Processing and its Applications (Warwick, 18–20 July 1989), IEE, London, 49–53, 1989.

[198] C.B. Price, P. Wambacq, A. Oosterlinck, *Computing with reaction–diffusion systems: applications in image processing*, D. Roose et al. (Eds.), Continuation and bifurcations: Numerical techniques and applications, Kluwer, 379–387, 1990.

[199] M. Proesmans, E.J. Pauwels, L.J. Van Gool, *Coupled geometry-driven diffusion equations for low-level vision*, B.M. ter Haar Romeny (Ed.), Geometry-driven diffusion in computer vision, Kluwer, Dordrecht, 191–228, 1994.

[200] M. Proesmans, E.J. Pauwels, L.J. Van Gool, T. Moons, A. Oosterlinck, *Image enhancement using non-linear diffusion*, Proc. IEEE Comp. Soc. Conf. Computer Vision and Pattern Recognition (CVPR '93, New York, June 15–17, 1993), IEEE Computer Society Press, Los Alamitos, 680–681, 1993.

[201] M. Proesmans, L. Van Gool, E. Pauwels, A. Oosterlinck, *Determination of optical flow and its discontinuities using non-linear diffusion*, J.-O. Eklundh (Ed.), Computer vision – ECCV '94, Lecture Notes in Comp. Science, Vol. 801, Springer, Berlin, 295–304, 1994.

[202] I. Rambaux, P. Garçon, *Nonlinear anisotropic diffusion filtering of 3D images*, project work, Département Génie Mathématique, INSA de Rouen and Laboratory of Technomathematics, University of Kaiserslautern, 1994.

[203] A.R. Rao, B.G. Schunck, *Computing oriented texture fields*, Proc. IEEE Comp. Soc. Conf. Computer Vision and Pattern Recognition (CVPR '89, San Diego, June 4–8, 1989), IEEE Computer Society Press, Washington, 61–68, 1989.

[204] A.R. Rao, B.G. Schunck, *Computing oriented texture fields*, CVGIP: Graphical Models and Image Processing, Vol. 53, 157–185, 1991.

[205] T. Richardson, S. Mitter, *Approximation, computation, and distortion in the variational formulation*, B.M. ter Haar Romeny (Ed.), Geometry-driven diffusion in computer vision, Kluwer, Dordrecht, 169–190, 1994.

[206] L.I. Rudin, S. Osher, *Total variation based image restoration with free local constraints*, Proc. First IEEE Int. Conf. Image Processing (ICIP–94, Austin, Nov. 13–16, 1994), IEEE Computer Society Press, Los Alamitos, Vol. 1, 31–35, 1994.

[207] L.I. Rudin, S. Osher, E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Physica D, Vol. 60, 259–268, 1992.

[208] L.I. Rudin, S. Osher, C. Fu, *Total variation based restoration of noisy, blurred images*, Preprint, Cognitech, Inc., 2800 – 28 St., Suite 101, Santa Monica, CA 90405, U.S.A., 1992; SIAM J. Numer. Anal., in press.

[209] P. Saint-Marc, J.S. Chen, G. Medioni, *Adaptive smoothing: a general tool for early vision*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 13, 514–529, 1990.

[210] A.H. Salden, *The nonlinear rescaling process*, Report 3DCV 92-28, 3D Computer Vision, AZU, Room E.02.222, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands, 1992.

[211] A.A. Samarskij, *Theorie der Differenzenverfahren*, Geest & Portig, Leipzig, 1984.

[212] G. Sapiro, V. Caselles, *Histogram modification via partial differential equations*, Technical Report HPL-95-66, Hewlett–Packard Laboratories, 1501 Page Mill Road, Palo Alto, CA 94304, U.S.A., 1995; to appear in Comm. Appl. Nonlinear Analysis.

[213] G. Sapiro, R. Kimmel, D. Shaked, B.B. Kimia, A.M. Bruckstein, *Implementing continuous-scale morphology via curve evolution*, Pattern Recognition, Vol. 26, 1363–1372, 1993.

[214] G. Sapiro, A. Tannenbaum, *Affine invariant scale-space*, Int. J. Comput. Vision, Vol. 11, 25–44, 1993.

[215] G. Sapiro, A. Tannenbaum, *On affine plane curve evolution*, J. Funct. Anal., Vol. 119, 79–120, 1994.

[216] G. Sapiro, A. Tannenbaum, *Area and length preserving geometric invariant scale-spaces*, J.-O. Eklundh (Ed.), Computer vision – ECCV '94, Lecture Notes in Comp. Science, Vol. 801, Springer, Berlin, 449–458, 1994.

[217] G. Sapiro, A. Tannenbaum, *Edge preserving geometric smoothing of MRI data*, Technical Report, Dept. of Electical Engineering, University of Minnesota, Minneapolis, MN 55455, U.S.A., 1994; submitted to IEEE Trans. Medical Imaging.

[218] G. Sapiro, A. Tannenbaum, Y.-L. You, M. Kaveh, *Experiments on geometric image enhancement*, Proc. First IEEE Int. Conf. Image Processing (ICIP–94, Austin, Nov. 13–16, 1994), IEEE Computer Society Press, Los Alamitos, Vol. 2, 472–476, 1994.

[219] G. Sapiro, A. Tannenbaum, *Area and length preserving geometric invariant scale-spaces*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 17, 67–72, 1995.

[220] C. Schnörr, *Unique reconstruction of piecewise smooth images by minimizing strictly convex non-quadratic functionals*, J. Math. Imag. Vision, Vol. 4, 189–198, 1994.

[221] C. Schnörr, *Segmentation of visual motion by minimizing convex non-quadratic functionals*, Proc. 12th Int. Conf. Pattern Recognition (ICPR 12, Jerusalem, Oct. 9–13, 1994), Vol. A, 661–663, 1994.

[222] C. Schnörr, *Bewegungssegmentation von Bildfolgen durch die Minimierung konvexer nicht-quadratischer Funktionale*, W.G. Kropatsch, H. Bischof (Eds.), Tagungsband Mustererkennung 1994, Informatik Xpress 5, Wien, 178–185, 1994.

[223] C. Schnörr, R. Sprengel, B. Neumann, *A variational approach to the design of early vision algorithms*, W. Kropatsch, R. Klette, F. Solina (Eds.), Theoretical foundations of computer vision, Computing Supplement 11, Springer, Wien, 149–165, 1996.

[224] E. Seneta, *Non-negative matrices and Markov chains*, Springer, New York, 1981.

[225] J. Serra, *Image analysis and mathematical morphology*, Academic Press, London, 1982.

[226] J.A. Sethian, *Numerical algorithms for propagating interfaces: Hamilton–Jacobi equations and conservation laws*, J. Differential Geometry, Vol. 31, 131–161, 1990.

[227] J. Shah, *Segmentation by nonlinear diffusion*, Proc. IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition (CVPR '91, Maui, June 3–6, 1992), IEEE Computer Society Press, Los Alamitos, 202–207, 1991.

[228] J. Shah, *Segmentation by nonlinear diffusion, II*, Proc. IEEE Comp. Soc. Conf. Computer Vision and Pattern Recognition (CVPR '92, Champaign, June 15–18, 1992), IEEE Computer Society Press, Los Alamitos, 644–647, 1992.

[229] J. Shah, *Properties of energy-minimizing segmentations*, SIAM J. Control and Optimization, Vol. 30, 99–111, 1992.

[230] J. Shah, *A nonlinear diffusion model for discontinuous disparity and half-occlusions in stereo*, Proc. IEEE Comp. Soc. Conf. Computer Vision and Pattern Recognition (CVPR '93, New York, June 15–17, 1993), IEEE Computer Society Press, Los Alamitos, 34–40, 1993.

[231] J. Shah, *Piecewise smooth approximations of functions*, Calc. Var., Vol. 2, 315–328, 1994.

[232] J. Shah, *Use of elliptic approximations in computer vision*, to appear in R. Serapioni, F. Tomarelli (Eds.), Variational methods for discontinuous structures, Birkhäuser, Boston.

[233] K. Siddiqi, B.B. Kimia, *Parts of visual form: computational aspects*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 17, 239–251, 1995.

[234] K. Siddiqi, B.B. Kimia, C.-W. Shu, *Geometric shock-capturing ENO schemes for subpixel interpolation, computation and curve evolution*, Technical Report LEMS-142, Division of Engineering, Brown University, Providence, R.I. 02912, U.S.A., 1995.

[235] W. Snyder, Y.-S. Han, G. Bilbro, R. Whitaker, S. Pizer, *Image relaxation: Restoration and feature extraction*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 17, 620–624, 1995.

[236] E. Steen, B. Olstad, *Scale-space and boundary detection in ultrasonic imaging using nonlinear signal-adaptive anisotropic diffusion*, M.H. Loew (Ed.), Image processing, SPIE Vol. 2167, 116–127, 1994.

[237] P. Stoll, C.-W. Shu, B.B. Kimia, *Schock-capturing numerical methods for viscosity solutions of certain PDEs in computer vision: The Godunov, Osher–Sethian and ENO schemes*, Technical Report LEMS-132, Division of Engineering, Brown University, Providence, R.I. 02912, U.S.A., 1994.

[238] H. Tek, B.B. Kimia, *Shock-based reaction–diffusion bubbles for image segmentation*, N. Ayache (Ed.), Computer vision, virtual reality and robotics in medicine, Lecture Notes in Comp. Science, Vol. 905, Springer, Berlin, 434–438, 1995.

[239] V. Torre, T.A. Poggio, *On edge detection*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 8, 148–163, 1986.

[240] G. Turk, *Generating textures on arbitrary surfaces using reaction–diffusion*, Computer Graphics, Vol. 25, No. 4, 289–298, 1991.

[241] R.A. Varga, *Matrix iterative analysis*, Prentice Hall, Englewood Cliffs, 1962.

[242] C.R. Vogel, M.E. Oman, *Iterative methods for total variation denoising*, Preprint, Dept. of Mathematical Sciences, Montana State University, Bozeman, MT 59717-0240, U.S.A., 1994; to appear in SIAM J. Sci. Comput.

[243] C.R. Vogel, M.E. Oman, *Fast numerical methods for total variation minimization in image reconstruction*, Preprint, Dept. of Mathematical Sciences, Montana State University, Bozeman, MT 59717-0240, U.S.A., 1995; to appear in SPIE Proc. Vol. 2563 "Advanced Signal Processing Algorithms", 1995.

[244] W. Walter, *Differential and integral inequalities*, Springer, Berlin, 1970.

[245] W. Walter, *Gewöhnliche Differentialgleichungen*, Springer, Berlin, 1976.

[246] J. Weickert, *Zwischenbericht zum Projekt "Nichtlineare Diffusionsfilter"*, Bericht über die wissenschaftliche Tätigkeit Januar 1991 – Dezember 1991, Center for Applied Mathematics, Darmstadt – Kaiserslautern, 133–142, 1992.

[247] J. Weickert, *Abschlußbericht zum Projekt "Nichtlineare Diffusionsfilter"*, Abschlußbericht und Bericht über die wissenschaftliche Tätigkeit Januar 1992 – Dezember 1993, Center for Applied Mathematics, Darmstadt – Kaiserslautern, 191–209, 1994.

[248] J. Weickert, *Anisotropic diffusion filters for image processing based quality control*, A. Fasano, M. Primicerio (Eds.), Proc. Seventh European Conf. on Mathematics in Industry, Teubner, Stuttgart, 355–362, 1994.

[249] J. Weickert, *Scale-space properties of nonlinear diffusion filtering with a diffusion tensor*, Report No. 110, Laboratory of Technomathematics, University of Kaiserslautern, P.O. Box 3049, 67653 Kaiserslautern, Germany, 1994.

[250] J. Weickert, *Theoretical foundations of anisotropic diffusion in image processing*, W. Kropatsch, R. Klette, F. Solina (Eds.), Theoretical foundations of computer vision, Computing Supplement 11, Springer, Wien, 221–236, 1996.

[251] J. Weickert, *A model for the cloudiness of fabrics*, H. Neunzert (Ed.), Proc. Eighth European Conf. on Mathematics in Industry, Teubner, Stuttgart, 1995, in press.

[252] J. Weickert, *Multiscale texture enhancement*, V. Hlaváč, R. Šára (Eds.), Computer analysis of images and patterns, Lecture Notes in Comp. Science, Vol. 970, Springer, Berlin, 230–237, 1995.

[253] J. Weickert, *Foundations and applications of nonlinear anisotropic diffusion filtering*, submitted to Proc. Third Int. Congress on Industrial and Applied Mathematics (ICIAM '95, Hamburg, July 3–7, 1995), Zeitschrift für Angewandte Mathematik und Mechanik (ZAMM), 1996.

[254] H. Werner, H. Arndt, *Gewöhnliche Differentialgleichungen*, Springer, Berlin, 1986.

[255] R.T. Whitaker, S.M. Pizer, *A multi-scale approach to nonuniform diffusion*, CVGIP: Image Understanding, Vol. 57, 99–110, 1993.

[256] R.T. Whitaker, *Geometry-limited diffusion*, Ph.D. thesis, Dept. of Computer Science, University of North Carolina, Chapel Hill, NC 27514, U.S.A., 1993.

[257] R.T. Whitaker, *Geometry limited diffusion in the characterization of geometric patches in images*, CVGIP: Image Understanding, Vol. 57, 111–120, 1993.

[258] R.T. Whitaker, *Characterizing first and second-order patches using geometry-limited diffusion*, H.H. Barrett, A.F. Gmitro (Eds.), Information processing in medical imaging, Lecture Notes in Comp. Science, Vol. 687, Springer, Berlin, 149–167, 1993.

[259] R.T. Whitaker, D.T. Chen, *Embedded active surfaces for volume visualization*, M.H. Loew (Ed.), Image processing, SPIE Vol. 2167, 340–352, 1994.

[260] R. Whitaker, G. Gerig, *Vector-valued diffusion*, B.M. ter Haar Romeny (Ed.), Geometry-driven diffusion in computer vision, Kluwer, Dordrecht, 93–134, 1994.

[261] R.T. Whitaker, S.M. Pizer, *Geometry-based image segmentation using anisotropic diffusion*, Y.L. O, A. Toet, D. Foster, H.J.A.M. Heijmans, P. Meer (Eds.), Shape in picture, Springer, 641–650, 1994.

[262] R. Wilson, H.E. Knutsson, G.H. Granlund, *Anisotropic nonstationary image estimation and its applications: Part II – Predictive image coding*, IEEE Trans. Comm., Vol. 31, 398–406, 1983.

[263] A.P. Witkin, *Scale-space filtering*, Proc. Eighth Int. Joint Conf. on Artificial Intelligence (IJCAI '83, Karlsruhe, Aug. 8–12, 1983), Vol. 2, 1019–1022, 1983.

[264] A. Witkin, M. Kass, *Reaction–diffusion textures*, Computer Graphics, Vol. 25, No. 4, 299–308, 1991.

[265] N.N. Yanenko, *The method of fractional steps: the solution of problems of mathematical physics in several variables*, Springer, New York, 1971.

[266] G.Z. Yang, P. Burger, D.N. Firmin, S.R. Underwood, *Structure adaptive anisotropic filtering for magnetic resonance image enhancement*, V. Hlaváč, R. Šára (Eds.), Computer analysis of images and patterns, Lecture Notes in Comp. Science, Vol. 970, Springer, Berlin, 384–391, 1995.

[267] Y.-L. You, M. Kaveh, W.-Y. Xu, A. Tannenbaum, *Analysis and design of anisotropic diffusion for image processing*, Proc. First IEEE Int. Conf. Image Processing (ICIP–94, Austin, Nov. 13–16, 1994), IEEE Computer Society Press, Los Alamitos, Vol. 2, 496–501, 1994.

[268] A.L. Yuille, T.A. Poggio, *Scaling theorems for zero crossings*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 8, 15–25, 1986.

[269] S.W. Zucker, A. Dobbins, L. Iverson, B. Kimia, A. Tannenbaum, *From curve detection to shape description: an outline*, A. Basu, X. Li (Eds.), Computer vision: systems, theory and applications, World Scientific, Singapore, 25–39, 1993.

Wissenschaftlicher Werdegang

15. 03. 1965	geboren in Ludwigshafen am Rhein
1971 – 1975	Grundschule Hettenleidelheim
1975 – 1984	Staatliches Leininger–Gymnasium Grünstadt
06/84	Abitur
10/84 – 01/91	Studium der Mathematik und Physik an der Universität Kaiserslautern
09/86	Zwischenprüfung in Mathematik für das Lehramt an Gymnasien
04/87	Zwischenprüfung in Physik für das Lehramt an Gymnasien und Vordiplom in Physik mit Nebenfach Informatik
08/87	Vordiplom in Technomathematik mit Nebenfach Technische Physik
01/91	Diplom in Technomathematik mit Nebenfach Technische Physik
02/91 – 07/93	Wissenschaftlicher Mitarbeiter am Zentrum für Praktische Mathematik Darmstadt–Kaiserslautern im Projekt "Nichtlineare Diffusionsfilter"
08/93 – 01/96	Wissenschaftlicher Mitarbeiter in der Arbeitsgruppe Technomathematik der Universität Kaiserslautern im Projekt "Entwicklung neuer mathematischer Werkzeuge in der Bildverarbeitung zur automatischen Qualitätskontrolle von Oberflächen"