International Journal of Computer Vision, Vol. 31, No. 2/3, pp. 111–127, April 1999.

Coherence-Enhancing Diffusion Filtering

Joachim Weickert[1]

Image Sciences Institute, Utrecht University,
Heidelberglaan 100, Room E01.334,
NL-3584 CX Utrecht, The Netherlands.
E-mail: joachim@cv.ruu.nl

Abstract

The completion of interrupted lines or the enhancement of flow-like structures is a challenging task in computer vision, human vision, and image processing. We address this problem by presenting a multiscale method in which a nonlinear diffusion filter is steered by the so-called interest operator (second-moment matrix, structure tensor). An m-dimensional formulation of this method is analysed with respect to its well-posedness and scale-space properties. An efficient scheme is presented which uses a stabilization by a semi-implicit additive operator splitting (AOS), and the scale-space behaviour of this method is illustrated by applying it to both 2-D and 3-D images.

Keywords: image enhancement, scale-space, texture, nonlinear diffusion

1 Introduction

Oriented flow-like structures arise in many computer vision and image processing problems: Within the field of texture analysis they appear for instance in the automatic grading of fabrics or wood surfaces, but they are also of importance for fingerprint analysis in forensic applications. They are present in many scientific imaging problems ranging from fluid dynamics to meteorology, and last but not least in medical imaging, for instance in the analysis of trabecular structures in bones. Interestingly, related tasks such as gap completion or the completion of interrupted lines also play a role in human vision tasks such as perceptual grouping. Moreover, the success of paintings by Munch or van Gogh suggests that emphasizing flow-like structures may create effects which fascinate many people.

Analysing flow-like patterns is an active research topic for certainly more than one decade; see for instance Kass and Witkin [28]. Besides methods such as Gabor filters or steerable filters, many of these approaches are equivalent to the so-called structure tensor (interest operator, second moment matrix) [17, 44]. In the meantime this field has even entered textbooks [21], and it can be regarded as well-established.

Much less attention has been payed to the question how to *enhance* flow-like patterns. Poor quality of fingerprint or trabecular bone images is not unusual. In those cases it would be desirable to have a tool which improves the quality of flow-like structures without destroying for instance semantically important singularities like the minutiae in fingerprints.

For problems like the grading of fabrics or applications to fluid dynamics it is also useful to have a multiscale simplification of the original image by embedding it into a scale-space in order to obtain a subsequently coarser, more global impression of the main flow-like structures. Of course, such a scale-space should take into account the coherence of the structures by smoothing mainly along their preferred orientation instead of perpendicular to it.

Since flow-like structures can also be present in higher dimensional data sets, e.g. 3-D images of trabecular bones, it should be possible to generalize such a method to arbitrary dimensions.

The preceding problems will be addressed in this paper by presenting an m-dimensional scale-space for the enhancement of coherent structures. The underlying concept can be motivated from ideas of Perona

[1] Current address: Department of Computer Science, University of Copenhagen, Universitetsparken 1, DK–2100 Copenhagen, Denmark.

and Malik for improving edge detection by creating a feedback loop of an edge detector and a nonlinear diffusion process [42]. In a similar way we embed a classical method for describing flow-like structures – namely the structure tensor approach – into a nonlinear diffusion process. This turns a method for *analysing* coherent pattern into a technique for *enhancing* and *simplifying* them. In contrast to most nonlinear diffusion filters, however, we use an approach where the process is steered by a diffusion tensor instead of a scalar-valued diffusivity. This enables true anisotropic behaviour by adapting the diffusion process not only to the location, but allowing also different smoothing in different directions. We shall see that this filter belongs to a class of nonlinear diffusion methods for which many well-posedness and scale-space properties can be proved. This is also in contrast to the Perona–Malik filter, whose continuous formulation is only locally well-posed [29]. Details of the numerical implementation will be discussed as well. In particular, an efficient novel stabilization technique based on a semi-implicit additive operator splitting (AOS) is presented. It extends previous encouraging experiments with AOS schemes for nonlinear diffusion filtering with a scalar-valued diffusivity to the fully anisotropic case with a diffusion tensor. All theoretical and numerical discussions hold in the m-dimensional case. As examples, results from 2-D and 3-D implementations are presented.
The outline of the papers is as follows: Section 2 surveys the underlying structure tensor method for describing coherence in images. This method is used in Section 3 for constructing a nonlinear diffusion process which smoothes along coherent flow-like structures. This process is called coherence-enhancing diffusion (CED). Section 4 gives a detailed analysis of the theoretical properties of a more general class of diffusion filters comprising CED. In Section 5 numerical and algorithmical questions are addressed; in particular, a more efficient alternative to the explicit scheme is presented. Section 6 applies CED to 2-D and 3-D images from different application areas. The paper is concluded with a summary in Section 7. Some results in this paper have been presented earlier at conferences [50, 56].

Related work. Since the work of Perona and Malik [42] numerous nonlinear diffusion filters have been proposed; see e.g. [23, 53] for an overview. Nevertheless, most of them use a (spatially varying) scalar diffusivity, not a diffusion tensor. Thus, they act inhomogeneously (nonuniformly) on the image, but – in our terminology – they remain isotropic.

True anisotropic diffusion filtering is studied in the reaction-diffusion model of Cottet and Germain [15]. Its diffusion tensor uses the eigenvectors $v_1 \parallel \nabla u_\sigma$, $v_2 \perp \nabla u_\sigma$ and its eigenvalues are given by

$$\begin{aligned} \lambda_1 &:= 0, \\ \lambda_2 &:= \frac{\eta |\nabla u_\sigma|^2}{1 + (|\nabla u_\sigma|/\sigma)^2} \qquad (\eta > 0). \end{aligned}$$

This choice is similar to our method it that sense that it diffuses mainly along strongly anisotropic structures. However, there are two important differences: Firstly, we observe that this diffusion tensor cannot be treated within a scale-space framework using uniformly positive definite diffusion tensors. Moreover, the Cottet–Germain model uses an additional reaction term which leads to nontrivial steady-states and qualifies it as a pure restoration method without scale-space ambitions. Secondly, the eigendirections of D are adapted to ∇u_σ, not to the eigendirections of the structure tensor. We shall see that the introduction of an integration scale in the structure tensor is an essential model feature in order to obtain reliable orientation estimates for flow-like structures. Recently Cottet and El Ayyadi have proposed a modified image restoration model which replaces the spatial regularization by a temporal regularization [14].

Nitzberg and Shiota [39] pioneered shape-adapted Gaussian smoothing, where the shape of an anisotropic Gaussian convolution kernel is a function of the structure tensor. Later on similar proposals have been made in [33, 61] and supplemented with scale-selection methods [1]. It should be noted that a common feature of all the preceding shape-adapted Gaussian smoothing methods is the combination of isotropic smoothing inside a region with anisotropic smoothing along edges. Space-variant Gaussian blurring is, however, not equivalent to an inhomogeneous diffusion process, and it does not preserve the average grey value. Anisotropic diffusion filters with isotropic diffusion within regions and anisotropic diffusion along edges can be found in [51] and the references therein. These methods are different from the present approach: CED is basically a pure 1-D diffusion, where a minimal amount of isotropic smoothing is added only for regularization purposes.

Other anisotropic partial differential equations for smoothing images rely on morphological methods such as the *mean-curvature motion (geometric heat equation, Euclidean shortening flow)* [30, 4]

$$\partial_t u \;=\; u_{\xi\xi} \;=\; |\nabla u|\,\mathrm{curv}(u) \tag{1}$$

with ξ being the direction perpendicular to ∇u.
Since mean-curvature motion propagates isophotes in inner normal direction with curvature-dependent speed, we should not expect such completely local methods to be capable of closing interrupted line-like structures. The same holds for the affine invariant version of this process [3, 46]. All these methods do not take into account semilocal information as can be gained from the structure tensor. Recent results by Carmona and Zhong [12] confirm the importance of semilocal estimates of the smoothing direction also for equations of mean-curvature type, if one is interested in specific goals such as enhancement of line-like structures.

Three-dimensional nonlinear diffusion filters have been investigated first by Gerig *et al.* [18] in the isotropic case, and by Rambaux and Garçon [43] in the edge-enhancing anisotropic case. A generalization of coherence-enhancing anisotropic diffusion to higher dimensions was first proposed in [56]. A recent three-dimensional PDE-based filter by Krissian *et al.* [32] and a 3-D reaction-diffusion process by Payot *et al.* [41] may be related to these anisotropic diffusion techniques.

The discussion above shows that the distinctive features of our approach is a semi-local analysis by means of the structure tensor combined with 1-D diffusion along one of its eigenvectors.

2 The structure tensor

First we review a reliable tool for analysing coherent flow-like structures. Consider an m-dimensional image domain $\Omega := (0,a_1)\times\cdots\times(0,a_m)$, and let an image $u(x)$ be represented by a bounded mapping $u:\Omega\to\mathbb{R}$.
A very simple structure descriptor is given by ∇u_σ, the gradient of a Gaussian-smoothed version of u:

$$K_\sigma(x) \;:=\; \frac{1}{(2\pi\sigma^2)^{m/2}}\cdot\exp\left(-\frac{|x|^2}{2\sigma^2}\right), \tag{2}$$

$$u_\sigma(x,t) \;:=\; (K_\sigma * u(.,t))\,(x) \qquad (\sigma>0). \tag{3}$$

The standard deviation σ denotes the *noise scale*, since it makes the edge detector ignorant of details smaller than $O(\sigma)$. Convolution on a finite domain Ω can be defined as convolution between K_σ and $\tilde{u}$ on $\mathbb{R}^m$, where $\tilde{u}$ denotes an extension of u by mirroring. This results in applying no-flux boundary conditions.
Although ∇u_σ is useful for detecting edges, it is unsuited for finding parallel structures, as we can see from Figure 1. The left image shows an original fingerprint. Figure 1(b),(c) illustrates the gradient orientation using grey values: vertical gradients are depicted in black, and horizontal ones in white. We observe that for small σ high fluctuations remain, while larger σ lead to entirely useless results. This is due to the fact that for larger σ neighbouring gradients with the same orientation, but opposite sign cancel one another. Gradient smoothing averages directions instead of orientations[2]. To make the structure descriptor invariant under sign changes, we may replace ∇u_σ by its tensor product

$$J_0(\nabla u_\sigma) := \nabla u_\sigma \otimes \nabla u_\sigma := \nabla u_\sigma \nabla u_\sigma^T. \tag{4}$$

This matrix is symmetric and positive semidefinite, and its eigenvectors are parallel and orthogonal to ∇u_σ, respectively. The corresponding eigenvalues $|\nabla u_\sigma|^2$ and 0 describe just the contrast in the eigendirections. Now that we have replaced directions by orientations, we can average the orientations by applying a componentwise convolution with a Gaussian K_ρ:

$$J_\rho(\nabla u_\sigma) := K_\rho * (\nabla u_\sigma \otimes \nabla u_\sigma) \qquad (\rho \ge 0). \tag{5}$$

This matrix is named *structure tensor*, *interest operator* or *second-moment matrix*. It is useful for many different tasks, for instance for analysing flow-like textures [44], corners and T-junctions [17, 39], shape cues [33] and spatio–temporal image sequences [27]. Equivalent approaches have been discovered

[2] In our terminology, gradients with opposite sign share the same orientation, but point in opposite directions.

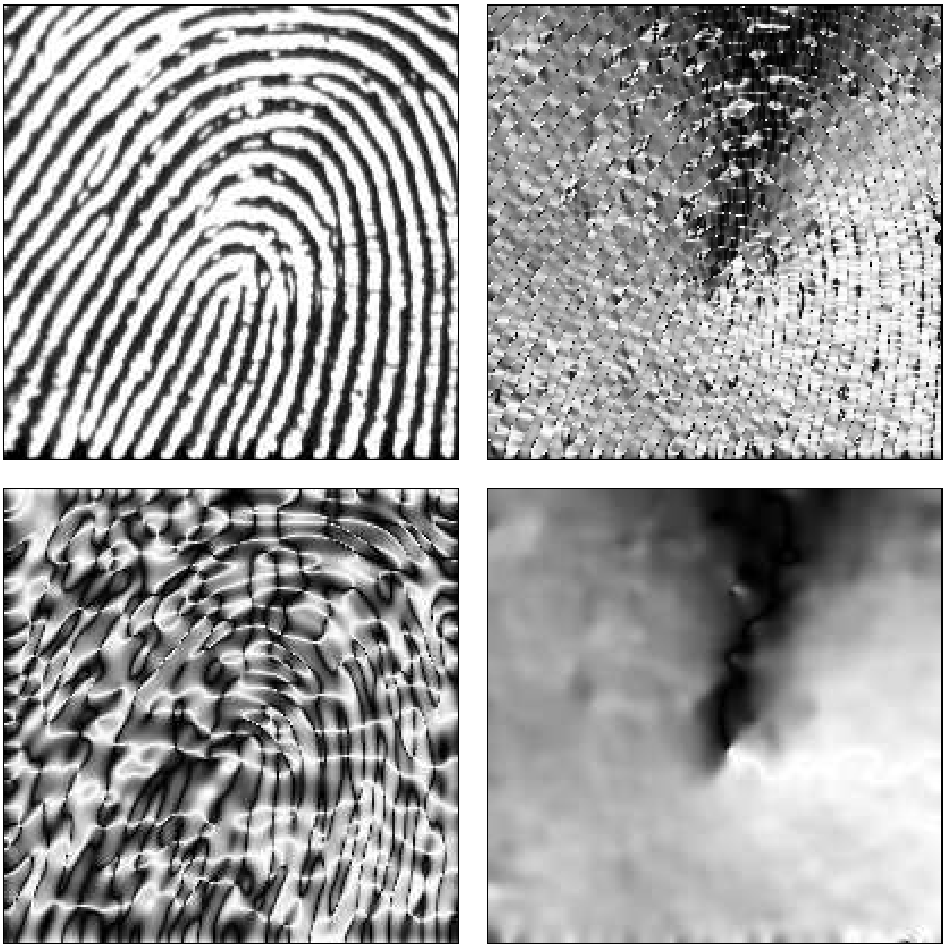

Figure 1: Local orientation in a fingerprint image. FROM TOP LEFT TO BOTTOM RIGHT: (a) Original fingerprint, $\Omega = (0, 200)^2$. (b) Orientation of smoothed gradient, $\sigma = 0.5$. (c) Orientation of smoothed gradient, $\sigma = 5$. (d) Structure tensor orientation, $\sigma = 0.5$, $\rho = 4$

independently in [7, 28]. A book by Jähne [27] gives a nice overview of these methods and clarifies their mutual relations.
It is not hard to verify that the symmetric matrix J_ρ is positive semidefinite. Let its eigenvalues $\mu_1,...,\mu_m$ be ordered such that

$$\mu_1 \ \geq \ \mu_2 \ \geq \ \cdots \ \geq \ \mu_m \tag{6}$$

and let $\{w_1, ..., w_m\}$ denote the corresponding orthonormal set of eigenvectors. Since the eigenvalues integrate the variation of the grey values within a neighbourhood of size $O(\rho)$, they describe the average contrast in the eigendirections $w_1,...,w_m$. Thus, the *integration scale* ρ should reflect the characteristic size of the texture. Usually, it is large in comparison to the noise scale σ. The eigenvector w_m corresponds to the smallest eigenvalue μ_m. It is the orientation with the lowest fluctuations, the so-called *coherence orientation*[3].
Figure 1(d) depicts the coherence orientation for the fingerprint image. We observe that it is exactly the desired average orientation of the lines. It should be noted how well the singularity corresponds to the singularity in the original fingerprint image.
The eigenvalues of J_ρ provide useful information on the coherence of a structure, i.e. the actual amount

[3]For the case that we have p identical minimal eigenvalues, one may regard every orientation within the span of $w_{m-p+1},...,w_m$ as a coherence orientation. For natural images, this situation happens almost never and can be neglected.

of anisotropy. As a measure for the coherence, one can define

$$\kappa := \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} (\mu_i - \mu_j)^2. \tag{7}$$

It becomes large for strongly differing eigenvalues, and it tends to zero for isotropic structures.

3 Coherence-enhancing anisotropic diffusion in m dimensions

Now that we know how to analyse coherent structures, we draw our attention to the question of how to enhance them. This can be done by embedding the structure tensor analysis into a nonlinear diffusion filter.

The principle of nonlinear diffusion filtering is as follows. One calculates a processed version $u(x,t)$ of $f(x)$ with a scale parameter $t \geq 0$ as the solution of a diffusion equation with f as initial condition and reflecting boundary conditions:

$$\partial_t u = \operatorname{div}(D\, \nabla u) \quad \text{on} \quad \Omega \times (0, \infty), \tag{8}$$

$$u(x,0) = f(x) \quad \text{on} \quad \Omega, \tag{9}$$

$$\langle D\nabla u,\, n\rangle = 0 \quad \text{on} \quad \partial\Omega \times (0, \infty). \tag{10}$$

Hereby, n denotes the outer normal and $\langle .,. \rangle$ the usual inner product. If one wants to adapt the diffusion process to the image itself one should choose the symmetric positive definite diffusion tensor $D = (d_{ij}) \in \mathbb{R}^{m \times m}$ as a function of the local image structure. In the case of enhancing flow-like patterns, it is natural to adapt the diffusion tensor D to the structure tensor $J_\rho(\nabla u_\sigma)$.

How should this function look like? For enhancing coherence in images with flow-like structures, we need a smoothing process which acts mainly along the flow direction w_m (with the notations from Section 2) and the smoothing should increase with the strength of its orientation given by the coherence κ. This may be achieved in the following way:

We construct D such that it has the same eigenvectors as J_ρ and its eigenvalues are given by

$$\lambda_i \; := \; \alpha \tag{11}$$

for $i = 1,...,m-1$, and by

$$\lambda_m \; := \; \begin{cases} \alpha & \text{if } \kappa = 0, \\ \alpha + (1-\alpha) \exp\left(\frac{-C}{\kappa}\right) & \text{else.} \end{cases} \tag{12}$$

$C > 0$ serves as a threshold parameter: For $\kappa \gg C$ we get $\lambda_m \approx 1$, and $\kappa \ll C$ leads to $\lambda_m \approx \alpha$. The exponential function and the small positive parameter $\alpha \in (0,1)$ were introduced mainly for two theoretical reasons: First, this guarantees that the smoothness of the structure tensor carries over to the diffusion tensor. The second reason is that the process never stops: Even if the structure becomes isotropic ($\kappa \to 0$), there remains some small linear diffusion with diffusivity $\alpha > 0$. Thus, α serves as a regularization parameter which keeps the diffusion tensor uniformly positive definite. In the next section we shall see that these are useful requirements in order to establish many theoretical properties for CED.

4 A general well-posedness and scale-space framework

Coherence-enhancing anisotropic diffusion filtering can be regarded as a special case within a more general framework for nonlinear diffusion filtering, for which many well-posedness and scale-space properties can be found. This framework shall be analysed now.

We study a diffusion filter with the following properties:

Assume that $f : \Omega \to \mathbb{R}$ is bounded, $\rho \geq 0$, and $\sigma > 0$.
Let $a := \inf_{\Omega} f$, $b := \sup_{\Omega} f$, and consider the problem

$$\begin{aligned} \partial_t u = \operatorname{div}\,(D(J_\rho(\nabla u_\sigma))\, \nabla u) \quad & \text{on} \quad \Omega \times (0,\infty), \\ u(x,0) = f(x) \quad & \text{on} \quad \Omega, \\ \langle D(J_\rho(\nabla u_\sigma))\nabla u,\, n\rangle = 0 \quad & \text{on} \quad \partial\Omega \times (0,\infty), \end{aligned}$$

where the diffusion tensor $D = (d_{ij})$ satisfies the following properties:

(C1) Smoothness:
$D \in \mathrm{C}^\infty(\mathbb{R}^{m\times m}, \mathbb{R}^{m\times m})$.

(C2) Symmetry:
$d_{ij}(J) = d_{ji}(J)$ for all symmetric matrices $J \in \mathbb{R}^{m\times m}$.

(C3) Uniform positive definiteness:
If $w : \Omega \to \mathbb{R}^2$ satisfies $|w(x)| \leq K$ on $\bar{\Omega}$, there exists a positive lower bound $\nu(K)$ for the eigenvalues of $D(J_\rho(w))$.

(P_c)

Evidently, coherence-enhancing anisotropic diffusion satisfies the preceding requirements. Under the assumptions (P_c) the following theorem, which generalizes and extends results from [13, 51], can be proved.

Theorem 1 (Well-posedness properties, smoothness, extremum principle)
The problem (P_c) has a unique solution $u(x,t)$ in the distributional sense, which is smooth for $t > 0$:

$$u \in \mathrm{C}^\infty(\bar{\Omega} \times (0,\infty)).$$

This solution depends continuously on f with respect to $\|\,.\,\|_{\mathrm{L}^2(\Omega)}$, and it fulfils the extremum principle

$$a \leq u(x,t) \leq b \quad \textit{on} \quad \Omega \times (0,\infty). \tag{13}$$

Proof: See Appendix A.1.

Remarks:

(a) We observe a strong smoothing effect which is characteristic for many diffusion processes: boundedness of the initial image is sufficient to obtain an infinitely often differentiable solution for arbitrary small positive times. Additional requirements – for instance that f should be uniformly continuous in order to apply the theory of viscosity solutions – are not necessary in our case.

(b) 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 MRI sequences, since we know that similar images remain similar after filtering.

(c) 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.

(d) Coherence-enhancing anisotropic diffusion in not the only filter which is covered by the theorem above. The well-posedness theory (as well as the subsequent scale-space framework) comprises the regularized Perona–Malik filter by Catté, Lions, Morel and Coll [13] as well as the edge-enhancing anisotropic diffusion filter described in [51]. It is also possible to establish a semidiscrete and fully discrete well-posedness and scale-space theory which proceeds in the same way; see [54] for more details.

4.1 Scale-space properties

Scale-space representations embed an image f into a family $\{T_t f \,|\, t \geq 0\}$ of gradually smoother, simplified versions of it. Long before this concept became popular in the western world by the work of Witkin [59] and Koenderink [31], it had been discovered and axiomatically justified in Japan by Iijima [26, 55]. Most axiomatic scale-space representations are devoted to the linear diffusion scale-space (Gaussian scale-space), but one can also create nonlinear scale-spaces which can be classified in a unique way [3, 8, 40]. All these approaches are uncommitted, since their unique classification does not give us the liberty to incorporate a-priori knowledge.
Anisotropic nonlinear diffusion filtering, however, does offer the possibility to incorporate knowledge into its evolution. So let us now study some of its scale-space properties.
It is evident that the filter class (P_c) satisfies typical scale-space properties such as the semi-group property, invariance under Euclidean image transformations, grey-level shifts, or contrast reversion. Moreover, since it can be written in divergence form and uses reflecting boundary conditions, it is also not hard to verify that the average grey level

$$\mu \;:=\; \frac{1}{|\Omega|}\int_{\Omega} f(x)\,dx \tag{14}$$

is not affected by nonlinear diffusion filtering [54, pp. 63–64]:

$$\frac{1}{|\Omega|}\int_{\Omega} T_t f\,dx \;=\; \mu \qquad \forall\, t > 0. \tag{15}$$

Average grey level invariance is a 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. A constant average grey level is useful for scale-space based segmentation algorithms such as the hyperstack [37, 49]. It is also a desirable quality for applications in medical imaging where grey values measure physical qualities of the depicted object.

4.1.1 Causality in terms of nonenhancement of local extrema

Let us now turn to the question in which sense an evolution equation of type (P_c) can be considered as a smoothing, information-reducing image transformation.
Koenderink [31] required that a scale-space evolution should not create new level curves when increasing the scale parameter. If this is satisfied, iso-intensity linking through the scales is possible and a structure at a coarse scale can (in principle) be traced back to the original image *(causality)*. For this reason, he required that at spatial extrema with nonvanishing determinant of the Hessian isophotes in scale-space are upwards convex. He showed that this constraint can be written as

$$\operatorname{sign}(\partial_t u) = \operatorname{sign}(\Delta u). \tag{16}$$

A sufficient condition for the causality equation (16) to hold is requiring that local extrema with positive or negative definite Hessians are not enhanced: an extremum in ξ at scale θ satisfies $\partial_t u > 0$ if ξ is a minimum, and $\partial_t u < 0$ if ξ is a maximum. This implication is easily seen: In the first case, for instance, the eigenvalues $\eta_1,...,\eta_m$ of $\operatorname{Hess}(u)$ are positive. Thus,

$$\Delta u = \operatorname{trace}(\operatorname{Hess}(u)) = \sum_{i=1}^{m} \eta_i > 0, \tag{17}$$

which gives just the causality requirement (16).
Nonenhancement of local extrema has first been used by Babaud *et al.* [5] in the context of linear diffusion filtering. However, it is also satisfied by nonlinear diffusion scale-spaces, as we shall see now.[4]

[4] As in the linear diffusion case, nonenhancement of local extrema generally does not imply that their number is nonincreasing for dimensions $m \geq 2$; cf. [45].

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 $u(.,\theta)$ with nonvanishing Hessian. Then,

$$\partial_t u(\xi,\theta) < 0, \qquad \text{if } \xi \text{ is a local maximum,} \tag{18}$$
$$\partial_t u(\xi,\theta) > 0, \qquad \text{if } \xi \text{ is a local minimum.} \tag{19}$$

Proof: See Appendix A.2.

Nonenhancement of local extrema is just one possibility to end up with Koenderink's causality requirement. Another way to establish causality is via the extremum principle (13) following Hummel's reasoning; see [25] for more details.

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

Interestingly, causality in terms of nonenhancement of local extrema can be regarded as a first representative of a much larger class of smoothing properties of nonlinear diffusion filtering. Other representatives result from studying the Lyapunov functionals of these processes. Lyapunov functionals are energy-like expressions which decrease during the evolution and which have a lower bound. They can be useful for finding attractors of a process for $t \to \infty$. The theorem below establishes a class of Lyapunov functionals and shows that processes of type (P_c) converge to a flat steady-state which is given by the average grey value of the original image. This is a desirable property for scale-spaces, since such a steady-state can be regarded as the coarsest, most global representation of the original image.

Theorem 3 (Lyapunov functionals and behaviour for $t \to \infty$).
Suppose that u is the solution of (P_c) and let a, b, μ be defined as in (P_c) and (14). Then the following properties are valid:

(a) (Lyapunov functionals)
For all $r \in \mathrm{C}^2[a,b]$ which are convex on $[a,b]$ the function

$$V(t) := \Phi(u(t)) := \int_\Omega r(u(x,t))\,dx \tag{20}$$

is a Lyapunov functional:

(i) It is bounded from below by the value of Φ for a constant image with the same average grey level:

$$\Phi(u(t)) \geq \Phi(Mf) \text{ for all } t \geq 0, \text{ where } (Mf)(x) := \mu. \tag{21}$$

(ii) $V(t)$ is continuous in 0, and decreasing for all $t > 0$.

(b) (Convergence)
$u(x,t)$ converges to a constant image with the same average grey value:

$$\lim_{t\to\infty} \|u(t) - Mf\|_{\mathrm{L}^p(\Omega)} = 0 \qquad \forall\, p \in [1,\infty). \tag{22}$$

Proof: See Appendix A.3.

What are interesting representatives of this family of Lyapunov functionals? 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 (14). Then the following functions are decreasing for $t \in [0,\infty)$:

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

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

$$\text{(c)} \quad H[u(t)] := \int_\Omega u(x,t)\, \ln(u(x,t))\, dx, \qquad \text{if } a > 0.$$

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. Using Parseval's equality we know that a decreasing energy is also equivalent to a decreasing sum of the squared Fourier coefficients.

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

$$EZ_{u(t)} := \int_{\mathbb{R}} z\, dF_{u(t)}(z), \tag{23}$$

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

$$\int_{\mathbb{R}} \left(z - EZ_{u(t)}\right)^{2n} dF_{u(t)}(z). \tag{24}$$

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 interpret, although they do provide important information for tasks like texture discrimination [22, pp. 414–415]. All decreasing even moments demonstrate that the image becomes smoother during diffusion filtering. Hence, local enhancement effects, 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.[5] Then,

$$S[u(t)] := -\int_\Omega u(x,t)\, \ln(u(x,t))\, dx \tag{25}$$

is called the *entropy* of $u(t)$, a measure of uncertainty and missing information [11]. 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. Moreover, for $t \to \infty$, the process reaches the state with the lowest possible information, namely a constant image. This information-reducing property indicates that anisotropic diffusion might be generally useful in the context of image compression. In particular, it helps to explain the success of nonlinear diffusion filtering as a preprocessing step for subsampling as observed in [16]. The interpretation of the entropy in terms of Lyapunov functionals carries also over to generalized entropies, see [48] for more details.

From all the previous considerations, we recognize that anisotropic diffusion does really simplify the original image in a steady way.

5 Numerical aspects

To approximate CED numerically, we replace the derivatives by finite differences. Since continuous CED has the structure

$$\partial_t u = \sum_{i,j=1}^{m} \partial_{x_i}(d_{ij}\partial_{x_j} u), \tag{26}$$

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

its simplest discretization is given by the finite difference scheme

$$\frac{U^{k+1}-U^k}{\Delta t} = \sum_{i,j=1}^{m} L_{ij}^k U^k. \tag{27}$$

In this notation, U describes a vector containing the values at each pixel. The upper index denotes the time level and L_{ij} is a central difference approximation to the operator $\partial_{x_i}(d_{ij}\partial_{x_j})$.
Since (27) can be rewritten as

$$U^{k+1} = \Big(I + \Delta t \sum_{i,j=1}^{m} L_{ij}^k\Big) U^k \tag{28}$$

we observe that U^{k+1} can be calculated explicitly from U^k without any matrix inversions. For this reason it is called an *explicit* scheme.
Unfortunately, such explicit schemes require very small time steps Δt in order to be stable. Therefore, it is desirable to replace (28) by an implicit scheme which has the same first-order Taylor expansion in Δt, but better stability properties. One possibility is the *AOS-stabilized scheme*

$$U^{k+1} = \frac{1}{m}\sum_{l=1}^{m}\Big(I - m\Delta t L_{ll}^k\Big)^{-1}\Big(I + \Delta t \sum_{i=1}^{m}\sum_{j\neq i} L_{ij}^k\Big) U^k. \tag{29}$$

This method achieves a stabilization through the nonnegative matrices $(I - m\Delta t L_{ll}^k)^{-1}$. They describe a semi-implicit discretization of the diffusion caused by the l-th diagonal entry of the diffusion tensor. The typically used step size $\Delta t = 2$ is about one order of magnitude larger than the ones for which an explicit scheme of type (27) is stable in 2-D or 3-D. For the standard approximations with central derivatives within a (3×3)-stencil, the matrix inversions in (29) come down to solving diagonally dominant tridiagonal systems of linear equations. This can be performed in linear complexity with a modified Gaussian algorithm (Thomas algorithm). It should be noted that (29) uses an additive operator splitting (AOS) instead of the usual multiplicative operator splitting from the mathematical literature such as [34]. This guarantees that all axes are treated in the same way. AOS schemes have been introduced in [57] as an efficient and reliable method for isotropic nonlinear diffusion filtering, which perform well on parallel computer architectures [58]. Equation (29) shows one way how to extend them to anisotropic processes with a diffusion tensor.
The final CED algorithm is as follows:

1. Calculation of the structure tensor in each pixel. This requires one convolution with a Gaussian K_σ, and $\sum_{k=1}^{m} k = \frac{1}{2}m(m+1)$ convolutions with K_ρ. The convolutions were implemented in the spatial domain exploiting the separability and symmetry of the Gaussian.
2. Principal axis transformation of the structure tensor in each pixel. In our case this was done by a cyclic Jacobi algorithm [47].
3. Calculation of the diffusion tensor in each pixel. Here a significant speed-up is possible by creating a look-up table for the function (12) at the beginning of the programme run.
4. Calculation of
$$V^k := \Big(I + \Delta t \sum_{i=1}^{m}\sum_{j\neq i} L_{ij}^k\Big) U^k.$$
5. Calculation of
$$W_l^{k+1} := \Big(I - m\Delta t L_{ll}^k\Big)^{-1} V^k \qquad (l = 1, ..., m).$$
by means of the Thomas algorithm [57].
6. Calculation of
$$U^{k+1} := \frac{1}{m}\sum_{l=1}^{m} W_l^{k+1}.$$

This algorithm is computationally less complex than it looks at first glance: typical execution times for one iteration on a HP 9000/889 are 0.3 seconds for a 256×256 image and 9 seconds for a $64 \times 64 \times 64$ image.

Figure 2: Anisotropic equations applied to the fingerprint image. (a) TOP LEFT: Original image. (b) TOP RIGHT: Mean curvature motion, $t = 10$. (c) BOTTOM LEFT: Anisotropic diffusion similar to the Cottet–Germain model, $\alpha = 0$, $\sigma = 2$, $t = 20$. (d) BOTTOM RIGHT: Coherence-enhancing anisotropic diffusion, $\sigma = 0.5$, $\rho = 4$, $t = 20$.

6 Examples

Figure 2 illustrates the importance of choosing the appropriate smoothing direction in anisotropic processes. In Figure 2(b) mean curvature motion is applied to the fingerprint image. We observe that – although mean curvature motion is a completely anisotropic technique with many merits in computer vision – it cannot be used for closing interrupted line-like structures: topologically connected structures remain connected, but gaps become larger, since both ends move in opposite directions. MCM does not exploit any semilocal information.

Figure 2(c) shows the effect of coherence-enhancing anisotropic diffusion where α and the integration scale ρ are set to 0. This model is close in spirit of the Cottet–Germain filter [15]. For such a model it is difficult to find a useful noise scale σ for smoothing along flow-like structures: for $\sigma \to 0$, ∇u becomes eigenvector of D and the diffusion is halted, since the corresponding eigenvalue λ_1 is zero. If σ becomes larger, the effect illustrated in Figure 1(c) dominates: opposite gradients cancel one another leading to more or less random smoothing directions.

Coherence-enhancing diffusion filtering with a nonvanishing integration scale is depicted in Figure 2(d). We observe that diffusion along the coherence orientation w_2 is well-suited for closing interrupted lines in flow-like textures. Due to its reduced diffusivity at noncoherent structures, the location of the semantically important singularity in the fingerprint remains the same. This is an important prerequisite that any image processing method has to satisfy if it is to be applied to fingerprint

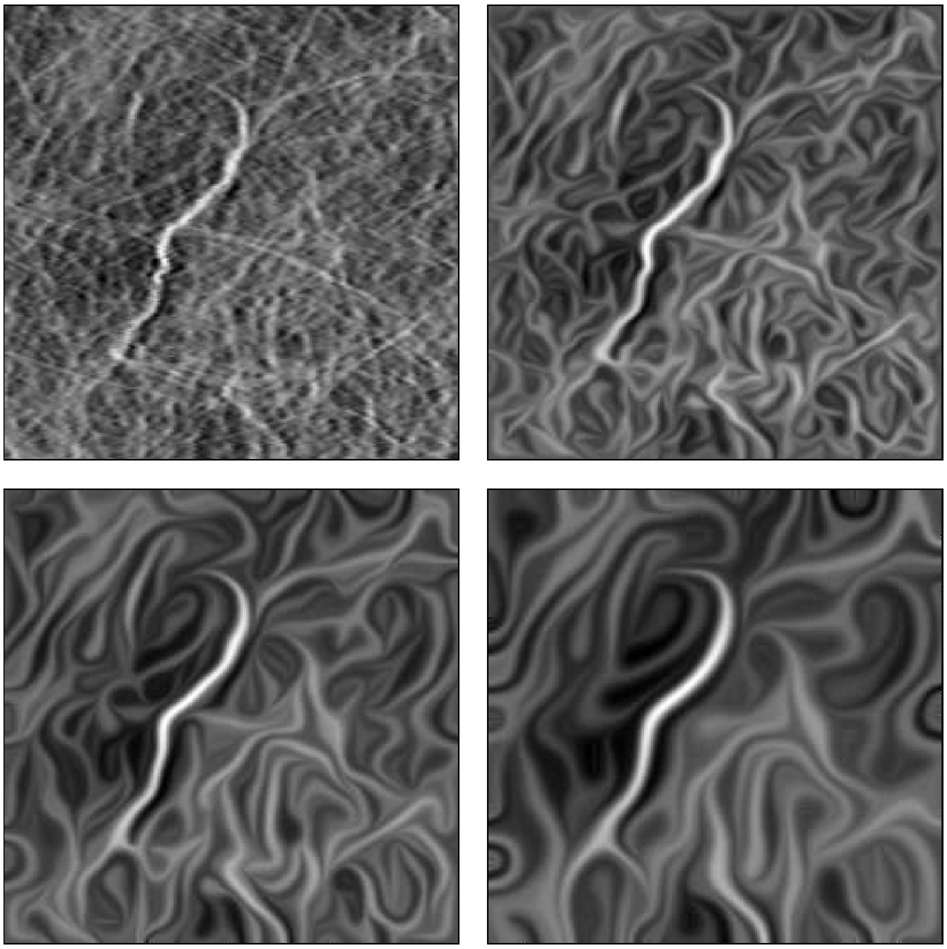

Figure 3: Scale-space behaviour of coherence-enhancing diffusion ($\sigma = 0.5$, $\rho = 2$). (a) TOP LEFT: Original fabric image, $\Omega = (0, 257)^2$. (b) TOP RIGHT: $t = 20$. (c) BOTTOM LEFT: $t = 120$. (d) BOTTOM RIGHT: $t = 640$.

analysis. In this and all subsequent CED images the parameters $C = 1$, and $\alpha = 0.001$ have been used.

Figure 3 depicts the scale-space behaviour of coherence-enhancing anisotropic diffusion applied to a fabric image arising in computer aided quality control. 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 expressionistic paintings by Vincent van Gogh.

Fig. 4 shows the restoration properties of coherence-enhancing anisotropic diffusion when being applied to a selfportrait of the artist [19]. 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 "Road with Cypress and Star" [20]. It is depicted in Fig. 5. In order to demonstrate the influence of the integration scale ρ, all filter parameters are fixed except for ρ. In Fig. 5(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(c)). Interestingly, a further increasing of ρ does hardly alter this result (Fig. 5(d)), which indicates that this van Gogh painting possesses a uniform "texture scale" reflecting the characteristic painting style of the artist.

Figure 6 illustrates the potential of CED for medical applications. It depicts a human bone. Its internal

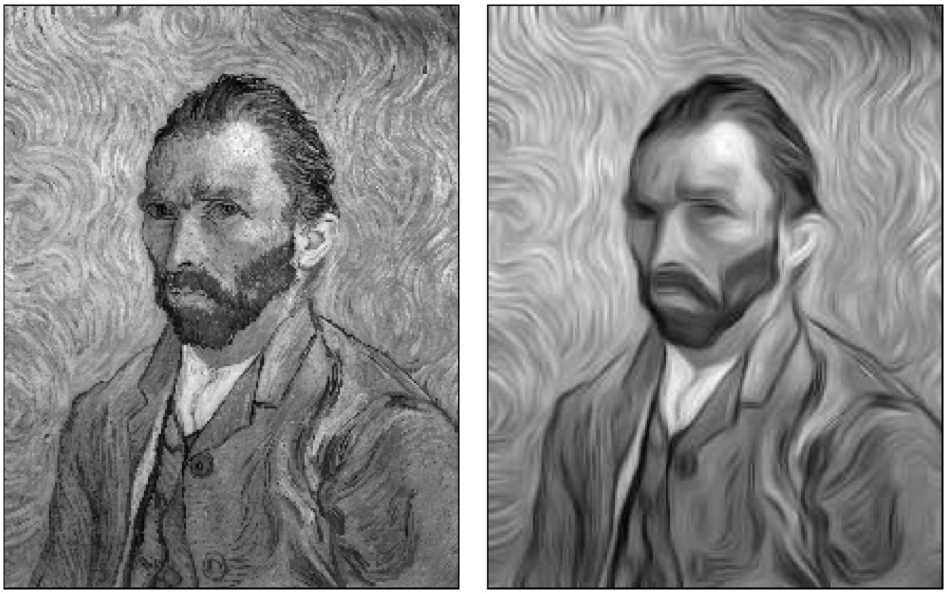

Figure 4: Image restoration using coherence-enhancing anisotropic diffusion. (a) LEFT: "Selfportrait" by van Gogh (Saint-Rémy, 1889; Paris, Museé d'Orsay), $\Omega = (0, 215) \times (0, 275)$. (b) RIGHT: Filtered, $\sigma = 0.5$, $\rho = 4$, $t = 6$.

structure has a distinctive texture through the presence of tiny elongated bony structural elements, the *trabeculae*. There is evidence that the trabecular formation is for a great deal determined by the external load [60, 35, 6]. For this reason the trabecular structure constitute an important clinical parameter in orthopedics. Examples are the control of recovery after surgical procedures, such as the placement or removal of metal implants, quantifying the rate of progression of rheumatism and osteoporosis, the determination of left-right deviations of symmetry in the load or establishing optimal load corrections for physiotherapy. The high resolution of contemporary CT using slipring technology and MR scanners now enables in vivo analysis of detailed trabecular structure.
While in vivo analysis of the trabecular bone by means of a structure tensor is investigated in [24, 36], the present paper points out ways to enhance the coherence information in medical images, in order to ease such an analysis.
From Figure 6(b),(c) we observe that CED is indeed capable of closing interrupted lines. Parallel flow-like structures are enhanced and a subsequent coherence analysis becomes much easier. Increasing the time t gives a coarser representation of the coherence. The entire evolution creates a task-driven scale-space which is dedicated to flow-like patterns.

In medical applications physicians often insist in only very small amounts of smoothing, in order to reduce some noise without affecting clinically relevant structures too much. Simplifications as in Figure 6(c) would be considered as misleading. Figure 7 shows the results for two iterations of three-dimensional CED filtering. The data set reveals a size of $256 \times 256 \times 128$, and the voxel dimensions are $0.25 \times 0.25 \times 0.5\,\mathrm{mm}^3$. It depicts a CT scan of a foot area by means of a Philips SR8000 slipring CT at 140 kV. Even for 3-D data sets, such a slight preprocessing using only a few iterations can easily be performed within the acquisition time. The price of a suitable hardware would be a small fraction of the price for a CT scanner.
It should be noted that in some 3-D applications it might also be desirable to smooth along two directions instead of one as in CED. Diffusion methods of this type have been studied by Rambaux and Garçon [43]. They are 3-D versions of the edge-enhancing anisotropic diffusion filter from [51].

7 Summary and conclusions

In the present paper we have treated the problem of enhancing flow-like patterns. For such tasks a reliable measurement of local orientation is needed. Our experiments demonstrate that the structure

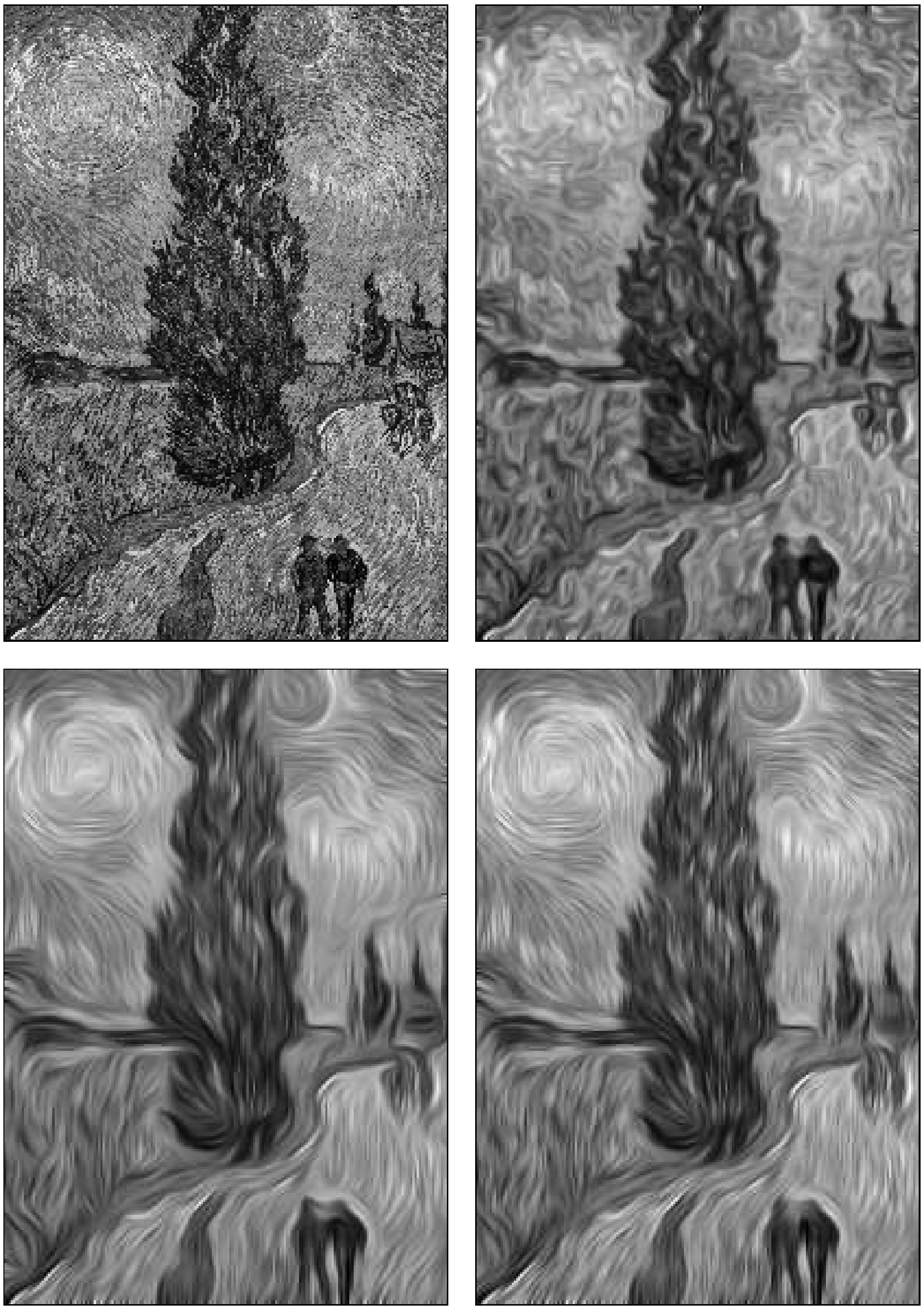

Figure 5: Impact of the integration scale on coherence-enhancing anisotropic diffusion ($\sigma = 0.5$, $t = 8$). (a) Top Left: "Road with Cypress and Star" by van Gogh (Auvers-sur-Oise, 1890; Otterlo, Rijksmuseum Kröller–Müller), $\Omega = (0, 203) \times (0, 290)$. (b) Top Right: Filtered with $\rho = 1$. (c) Bottom Left: $\rho = 4$. (d) Bottom Right: $\rho = 6$.

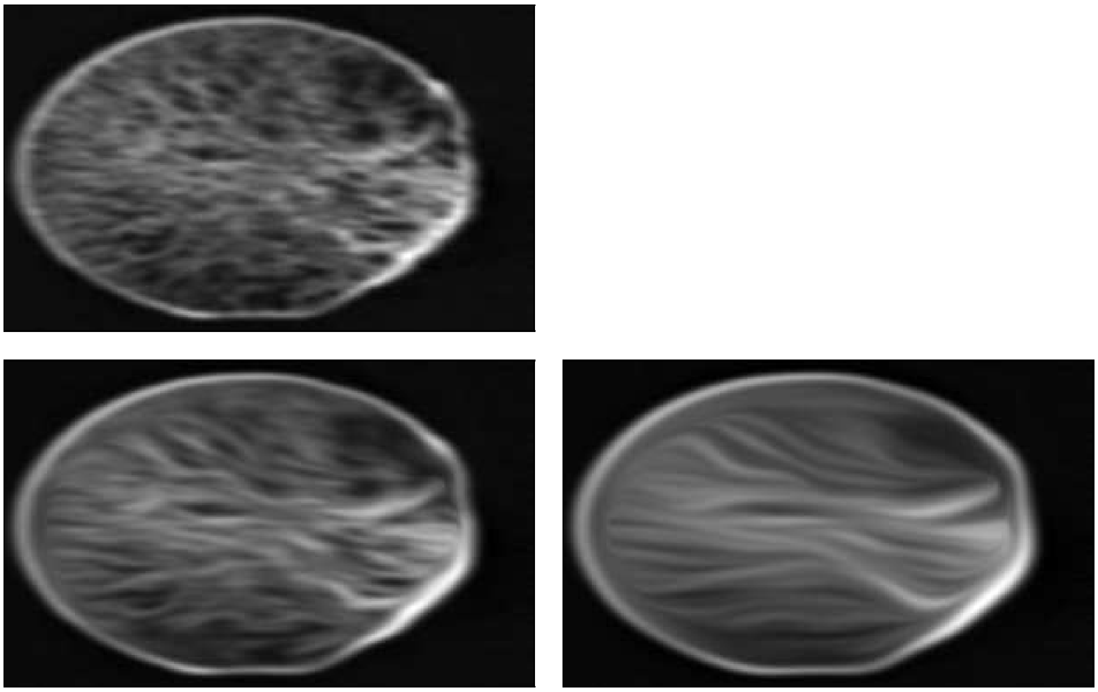

Figure 6: (a) TOP: High resolution slipring CT scan of a femural bone, showing the trabecular formation. Slice thickness 1 mm, field of view 6 cm. $\Omega = (0,300) \times (0,186)$. (b) BOTTOM LEFT: Filtered by coherence-enhancing anisotropic diffusion, $\sigma = 0.5$, $\rho = 6$, $t = 16$. (c) BOTTOM RIGHT: Dito with $t = 128$.

tensor satisfies this requirement. Unlike many other applications, we do not restrict its application to pure image *analysis*, we use it as a tool for *steering* a scale-space evolution.
To this end, we evolve the original image by means of a nonlinear anisotropic diffusion equation. Its diffusion tensor reflects the local image structure by using the same set of eigenvectors as the structure tensor. The eigenvalues are chosen in such a way that diffusion acts mainly along the direction with the highest coherence, and becomes stronger when the coherence increases. The resulting coherence-enhancing diffusion (CED) process gives a theoretically well-founded scale-space representation: proofs have been presented which show that its unique solution is stable under perturbations of the initial image, satisfies Koenderink's causality requirement, and creates a large family of Lyapunov functionals which ensure that the process is image simplifying and converges to a constant steady-state.
This theoretical framework is valid in any dimension. We have also presented an m-dimensional numerical scheme where a stabilization based on additive operator splitting (AOS) allows time steps which are one order of magnitude larger than in the explicit case.
Finally, the use of CED was illustrated by applying it to 2-D and 3-D data sets. The results indicate a variety of possible application areas ranging from computer aided quality control over fingerprint enhancement to medical imaging. This, however, gives only one aspect of the filter concept. Coherence-enhancing diffusion allows also generalizations to vector-valued images. Details in this direction can be found in [52].

Acknowledgments. The author thanks Robert Maas and Stiliyan Kalitzin for useful comments on a draft version of this paper, and Wiro Niessen for providing the 3-D trabecular bone data set.

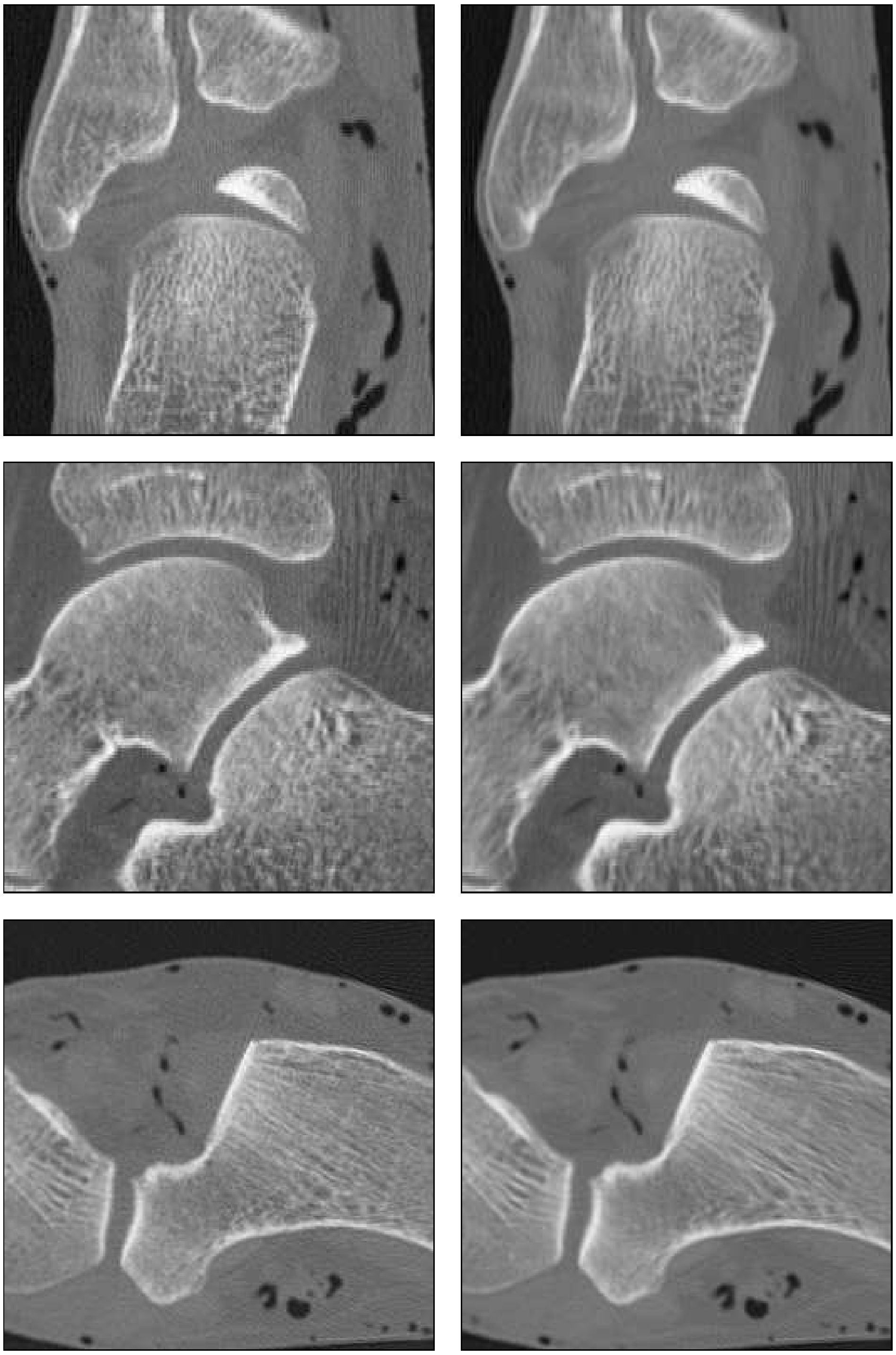

Figure 7: Three-dimensional CED applied to a CT data set of size $256 \times 256 \times 128$. LEFT COLUMN: 2-D sections depicting slice no. 180 in x, no. 128 in y, and no. 100 in z, respectively. RIGHT COLUMN: Filtered, $\sigma = 0.5$, $\rho = 6$, $t = 4$.

References

[1] A. Almansa, T. Lindeberg, *Enhancement of fingerprint images using shape-adapted scale-space operators*, J. Sporring, M. Nielsen, L. Florack, P. Johansen (Eds.), Gaussian scale-space theory, Kluwer, Dordrecht, 3–19, 1997.

[2] H.W. Alt, *Lineare Funktionalanalysis*, Springer, Berlin, 1992.

[3] 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.

[4] 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.

[5] 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.

[6] G.E. Bacon, P.J. Bacon, R.K. Griffiths, *The orientation of apatite crystals in bone*, J. Appl. Crystallography, Vol. 12, 99–103, 1979.

[7] 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.

[8] R. van den Boomgaard, *The morphological equivalent of the Gauss convolution*, Nieuw Archief Voor Wiskunde, Vierde Serie, Deel 10, 219–236, 1992.

[9] H. Brezis, *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North Holland, Amsterdam, 1973.

[10] H. Brezis, *Analyse fonctionelle*, Masson, Paris, 1992.

[11] B. Buck, V. Macaulay (Eds.), *Maximum entropy in action*, Clarendon, Oxford, 1991.

[12] R. Carmona, S. Zhong, *Adaptive smoothing respecting feature directions*, IEEE Trans. Image Proc., Vol. 7, 353–358, 1998.

[13] 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.

[14] G.-H. Cottet, M. El Ayyadi, *Nonlinear PDE operators with memory terms for image processing*, Proc. IEEE Int. Conf. Image Processing (ICIP–96, Lausanne, Sept. 16–19, 1996), Vol. 1, 481–483, 1996.

[15] G.-H. Cottet, L. Germain, *Image processing through reaction combined with nonlinear diffusion*, Math. Comp., Vol. 61, 659–673, 1993.

[16] G.E. Ford, R.R. Estes, H. Chen, *Scale-space analysis for image sampling and interpolation*, Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP–92, San Francisco, March 23–26, 1992), Vol. 3, 165–168, 1992.

[17] M.A. Förstner, E. Gülch, *A fast operator for detection and precise location of distinct points, corners and centres of circular features*, Proc. ISPRS Intercommission Conf. on Fast Processing of Photogrammetric Data (Interlaken, June 2–4, 1987), 281–305, 1987.

[18] G. Gerig, O. Kübler, R. Kikinis, F.A. Jolesz, *Nonlinear anisotropic filtering of MRI data*, IEEE Trans. Medical Imaging, Vol. 11, 221–232, 1992.

[19] V. van Gogh, *Selfportrait*, Saint-Rémy, 1889. Paris, Museé d'Orsay.

[20] V. van Gogh, *Road with cypress and star*, Auvers-sur-Oise, 1890. Otterlo, Rijksmuseum Kröller–Müller.

[21] G.H. Granlund, H. Knutsson, *Signal processing for computer vision*, Kluwer, Dordrecht, 1995.

[22] R.C. Gonzalez, P. Wintz, *Digital image processing*, Addison–Wesley, Reading, 1987.

[23] B.M. ter Haar Romeny (Ed.), *Geometry-driven diffusion in computer vision*, Kluwer, Dordrecht, 1994.

[24] B.M. ter Haar Romeny, W.J. Niessen, J. Weickert, P. van Roermund, W.J. van Enk, A. Lopez, R. Maas, *Orientation detection of trabecular bone*, Progress in Biophysics and Molecular Biology (Proc. 12th Int. Biophysics Congress, Amsterdam, Aug. 11–16, 1996), Vol. 65, Poster P-H5-43, 1996.

[25] 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.

[26] T. Iijima, *Basic theory on normalization of pattern (in case of typical one-dimensional pattern)*, Bulletin of the Electrotechnical Laboratory, Vol. 26, 368–388, 1962 (in Japanese).

[27] B. Jähne, *Spatio-temporal image processing*, Lecture Notes in Comp. Science, Vol. 751, Springer, Berlin, 1993.

[28] M. Kass, A. Witkin, *Analyzing oriented patterns*, Computer Vision, Graphics, and Image Processing, Vol. 37, 362–385, 1987.

[29] B. Kawohl, N. Kutev, *Maximum and comparison principles for anisotropic diffusion*, Preprint, Mathematical Institute, University of Cologne, 50923 Cologne, Germany, 1997.

[30] 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.

[31] J.J. Koenderink, *The structure of images*, Biol. Cybern., Vol. 50, 363–370, 1984.

[32] K. Krissian, G. Malandain, N. Ayache, *Directional anisotropic diffusion applied to segmentation of vessels in 3D images*, B. ter Haar Romeny, L. Florack, J. Koenderink, M. Viergever (Eds.), Scale-space theory in computer vision, Lecture Notes in Comp. Science, Vol. 1252, Springer, Berlin, 345–348, 1997.

[33] T. Lindeberg, J. Gårding, *Shape-adapted smoothing in estimation of 3-D depth cues from affine distortions of local 2-D brightness structure*, Image and Vision Computing, Vol. 15, 415–434, 1997.

[34] G.I. Marchuk, *Splitting and alternating direction methods*, P.G. Ciarlet, J.-L. Lions (Eds.), Handbook of numerical analysis, Vol. I, 197–462, 1990.

[35] H. Meyer, M. Culman, *Die Architektur der Spongiosa*, Arch. Anat. Physiol., Vol. 47, 615–628, 1867.

[36] W.J. Niessen, A.M. López, W.J. van Enk, P.M. van Roermund, B.M. ter Haar Romeny, M.A. Viergever, *In vivo analysis of trabecular bone architecture*, J.S. Duncan, G. Gindi (Eds.), Information processing in medical imaging, Lecture Notes in Comp. Science, Vol. 1230, Springer, Berlin, 435–440, 1997.

[37] W.J. Niessen, K.L. Vincken, J. Weickert, M.A. Viergever, *Nonlinear multiscale representations for image segmentation*, Computer Vision and Image Understanding, Vol. 66, 233–245, 1997.

[38] W.J. Niessen, K.L. Vincken, J. Weickert, M.A. Viergever, *Three-dimensional MR brain segmentation*, Proc. Sixth Int. Conf. on Computer Vision (ICCV '98, Bombay, Jan. 4–7, 1998), 53–58, 1998.

[39] M. Nitzberg, T. Shiota, *Nonlinear image filtering with edge and corner enhancement*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 14, 826–833, 1992.

[40] 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.

[41] E. Payot, R. Guillemaud, Y. Trousset, F. Preteux, *An adaptive and constrained model for 3D X-ray vascular reconstruction*, P. Grangeat, J.-L. Amans (Eds.), Three-dimensional image reconstruction in radiation and nuclear medicine, Kluwer, Dordrecht, 47–57, 1996.

[42] P. Perona, J. Malik, *Scale space and edge detection using anisotropic diffusion*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 12, 629–639, 1990.

[43] 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.

[44] A.R. Rao, B.G. Schunck, *Computing oriented texture fields*, CVGIP: Graphical Models and Image Processing, Vol. 53, 157–185, 1991.

[45] J.H. Rieger, *Generic evolution of edges on families of diffused greyvalue surfaces*, J. Math. Imag. Vision, Vol. 5, 207–217, 1995.

[46] G. Sapiro, A. Tannenbaum, *Affine invariant scale-space*, Int. J. Comput. Vision, Vol. 11, 25–44, 1993.

[47] H.R. Schwarz, *Numerische Mathematik*, Teubner, Stuttgart, 1988.

[48] J. Sporring, J. Weickert, *On generalized entropies and scale-space*, B. ter Haar Romeny, L. Florack, J. Koenderink, M. Viergever (Eds.), Scale-space theory in computer vision, Lecture Notes in Comp. Science, Vol. 1252, Springer, Berlin, 53–64, 1997.

[49] K.L. Vincken, A.S.E. Koster, M.A. Viergever, *Probabilistic multiscale image segmentation*, IEEE Trans. Pattern Anal. Mach. Intell., 109–120, 1997.

[50] 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.

[51] J. Weickert, *Theoretical foundations of anisotropic diffusion in image processing*, Computing, Suppl. 11, 221–236, 1996.

[52] J. Weickert, *Coherence-enhancing diffusion of colour images*, A. Sanfeliu, J.J. Villanueva, J. Vitrià (Eds.), Pattern Recognition and Image Analysis (VII NSPRIA, Barcelona, April 21–25, 1997), Vol. 1, 239–244, 1997. Extended version to appear in Image and Vision Computing.

[53] J. Weickert, *A review of nonlinear diffusion filtering*, B. ter Haar Romeny, L. Florack, J. Koenderink, M. Viergever (Eds.), Scale-space theory in computer vision, Lecture Notes in Comp. Science, Vol. 1252, Springer, Berlin, 3–28, 1997.

[54] J. Weickert, *Anisotropic diffusion in image processing*, Teubner-Verlag, Stuttgart, 1998.

[55] J. Weickert, S. Ishikawa, A. Imiya, *On the history of Gaussian scale-space axiomatics*, J. Sporring, M. Nielsen, L. Florack, P. Johansen (Eds.), Gaussian scale-space theory, Kluwer, Dordrecht, 45–59, 1997.

[56] J. Weickert, B.M. ter Haar Romeny, A. Lopez, W.J. van Enk, *Orientation analysis by coherence-enhancing diffusion*, Proc. Symp. Real World Computing (RWC '97, Tokyo, Jan. 29–31, 1997), 96–103, 1997.

[57] J. Weickert, B.M. ter Haar Romeny, M.A. Viergever, *Efficient and reliable schemes for nonlinear diffusion filtering*, IEEE Trans. Image Proc., Vol. 7, 398–410, 1998.

[58] J. Weickert, K.J. Zuiderveld, B.M. ter Haar Romeny, W.J. Niessen, *Parallel implementations of AOS schemes: A fast way of nonlinear diffusion filtering*, Proc. 1997 IEEE International Conference on Image Processing (ICIP–97, Santa Barbara, Oct. 26–29, 1997), Vol 3., 396–399, 1997.

[59] 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.

[60] J. Wolff, *Über die innere Architektur der Knochen*, Virchows Arch., Vol. 50, 389–453, 1870.

[61] G.Z. Yang, P. Burger, D.N. Firmin, S.R. Underwood, *Structure adaptive anisotropic filtering*, Image and Vision Computing, Vol. 14, 135–145, 1996.

A Appendix

A.1 Proof of Theorem 1

(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 [13], and the extremum principle follows from Stampacchia's truncation method (cf. [10], p. 211); see [54] for more details.

(b) **Continuous dependence on the initial image**
In order to discuss the continuous dependence on f, let us first define 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

$$\|u\|_{\mathrm{H}^1(\Omega)} := \left(\|u\|^2_{\mathrm{L}^2(\Omega)} + \sum_{i=1}^{m} \|\partial_{x_i} u\|^2_{\mathrm{L}^2(\Omega)} \right)^{1/2} \tag{30}$$

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

$$\|u\|_{\mathrm{L}^2(0,T;\mathrm{H}^1(\Omega))} := \left(\int_0^T \|u(t)\|^2_{\mathrm{H}^1(\Omega)}\, dt \right)^{1/2} < \infty. \tag{31}$$

The first part of the proof is similar to the uniqueness proof in [13]. Let $f, h \in \mathrm{L}^\infty(\Omega)$ be two initial values and u, w the corresponding solutions. Then one shows that

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|u(t)-w(t)\|^2_{\mathrm{L}^2(\Omega)} \;\le\; & \|\nabla u(t)-\nabla w(t)\|_{\mathrm{L}^2(\Omega)} \cdot (\|D(J_\rho(\nabla u_\sigma(t)))-D(J_\rho(\nabla w_\sigma(t)))\|_\infty)_{\mathrm{L}^\infty(\Omega)} \cdot \\ & \cdot \|\nabla u(t)\|_{\mathrm{L}^2(\Omega)} - \nu\, \|\nabla u(t)-\nabla w(t)\|^2_{\mathrm{L}^2(\Omega)} \end{aligned} \tag{32}$$

where $\|B\|_\infty := \max\limits_{1=1..m} \sum\limits_{j=1}^{m} |b_{ij}|$ for $B = (b_{ij}) \in \mathbb{R}^{m\times m}$, and ν denotes the lower positive bound for the eigenvalues of D.

In order to apply the Gronwall–Bellman lemma, we have to estimate

$$(\|D(J_\rho(\nabla u_\sigma(t)))-D(J_\rho(\nabla w_\sigma(t)))\|_\infty)_{\mathrm{L}^\infty(\Omega)}$$

by $\|u(t)-w(t)\|_{\mathrm{L}^2(\Omega)}$. By the boundedness of ∇u_σ and ∇w_σ, and the smoothness property (C1) we know that there exist a Lipschitz constant L such that

$$(\|D(J_\rho(\nabla u_\sigma(t)))-D(J_\rho(\nabla w_\sigma(t)))\|_\infty)_{\mathrm{L}^\infty(\Omega)} \le L \operatorname*{ess\,sup}_{x\in\Omega} \|\nabla u(t)-\nabla w(t)\|_2$$

By iteratively reflecting Ω at its boundaries we get a partitioning $\mathbb{R}^m = \dot{\bigcup\limits_i} \Omega_i$. Using Jensen's inequality, our definition of convolution on a finite domain Ω, and this partitioning we obtain

$$\begin{aligned} & (\|D(J_\rho(\nabla u_\sigma(t)))-D(J_\rho(\nabla w_\sigma(t)))\|_\infty)_{\mathrm{L}^\infty(\Omega)} \\ \le\; & L \cdot \operatorname*{ess\,sup}_{x\in\Omega} \left(\int_{\mathbb{R}^m} \|\nabla K_\sigma(y)\|_2^2\, ((\tilde{u}-\tilde{w})(x-y))^2\, dy \right)^{1/2} \\ \le\; & L \left(\sum_i \max_{z\in\bar{\Omega}_i} \|\nabla K_\sigma(z)\|_2^2 \right)^{1/2} \|u(t)-w(t)\|_{\mathrm{L}^2(\Omega)} \\ =\; & c \cdot \|u(t)-w(t)\|_{\mathrm{L}^2(\Omega)} \end{aligned} \tag{33}$$

with some finite constant c. Plugging (33) into (32) gives

$$\begin{aligned}\frac{1}{2}\,\frac{d}{dt}\,\|u(t)-w(t)\|^2_{\mathrm{L}^2(\Omega)} \;\;\le\;\; & \overbrace{\frac{c}{\sqrt{2\nu}}\|\nabla u(t)\|_{\mathrm{L}^2(\Omega)}\|u(t)-w(t)\|_{\mathrm{L}^2(\Omega)}}^{=:a}\cdot\overbrace{\sqrt{2\nu}\|\nabla u(t)-\nabla w(t)\|_{\mathrm{L}^2(\Omega)}}^{=:b} \\ & -\nu\,\|\nabla u(t)-\nabla w(t)\|^2_{\mathrm{L}^2(\Omega)}.\end{aligned}$$

To get at the right hand side a quadratic expression in $\|\nabla u(t)–\nabla w(t)\|_{\mathrm{L}^2(\Omega)}$, we use $|ab| \le \frac{a^2}{2}+\frac{b^2}{2}$ and end up with

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

Applying the Gronwall–Bellman lemma [9, pp. 156–157] yields

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

Next we show that $\int_0^t \|\nabla u(s)\|^2_{\mathrm{L}^2(\Omega)}\,ds$ is finite:

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

By virtue of the existence proof in [13], we know that the right-hand of this estimate exists. Now, let $\epsilon > 0$ and choose

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

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.

A.2 Proof of Theorem 2

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

$$\partial_t u \;=\; \sum_{i=1}^m\sum_{j=1}^m \left(\partial_{x_i} d_{ij}(J_\rho(\nabla u_\sigma))\right)\partial_{x_j} u \;+\; \sum_{i=1}^m\sum_{j=1}^m d_{ij}(J_\rho(\nabla u_\sigma))\,\partial_{x_i x_j} u. \tag{34}$$

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 (34) vanishes in (ξ,θ).

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

$$S^T DS \;=\; \mathrm{diag}(\lambda_1, ..., \lambda_m) \;=:\; \Lambda$$

with $\lambda_1,...,\lambda_m$ being the positive eigenvalues of D.
Now, let us assume that (ξ,θ) is a local maximum where the Hessian $H := \mathrm{Hess}(u(\xi,\theta))$ and, thus, $B := (b_{ij}) := S^T HS$ are negative definite. Then we have

$$b_{ii} < 0 \quad (i = 1, ..., m),$$

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

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

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

A.3 Proof of Theorem 3

(a) (i) Let $r \in \mathrm{C}^2[a,b]$ be convex on $[a,b]$. Using the average grey level invariance and Jensen's inequality we obtain, for all $t \geq 0$,

$$\begin{aligned}
\Phi(Mf) &= \int_\Omega r\left(\frac{1}{|\Omega|}\int_\Omega u(x,t)\,dx\right) dy \\
&\leq \int_\Omega \left(\frac{1}{|\Omega|}\int_\Omega r(u(x,t))\,dx\right) dy \\
&= \int_\Omega r(u(x,t))\,dx \\
&= \Phi(u(t)).
\end{aligned} \tag{35}$$

(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 almost everywhere on Ω. From this and the Cauchy–Schwarz inequality, we get

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

From [13] we know that $u \in \mathrm{C}([0,T];\mathrm{L}^2(\Omega))$. Thus, 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

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

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

$$\int_\Omega v v_t\, dx \;=\; -\int_\Omega \langle \nabla v, D(J_\rho(\nabla v_\sigma))\nabla v\rangle\, dx.$$

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

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

For $t > 0$, there exists some x_0 with $v(x_0) = 0$. Therefore, we may apply Poincaré's inequality (cf. [2, p. 122]), which tells us that

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

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

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

and hence the exponential decay of $\|v\|_{L^2(\Omega)}$ to 0.

By the maximum principle, we know that $\|v(t)\|_{L^\infty(\Omega)}$ is bounded by $\|f - Mf\|_{L^\infty(\Omega)}$. Thus, for $q \in \mathbb{N}$, $q \geq 2$, we get

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

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

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

This proves the assertion.