Computer Vision, Graphics, and Pattern Recognition Group
Department of Mathematics and Computer Science
University of Mannheim
D-68131 Mannheim, Germany

Reihe Informatik
8/2000

PDE–Based Preprocessing of Medical Images

Joachim Weickert and Christoph Schnörr

Technical Report 8/2000
Computer Science Series
February 2000

The technical reports of the CVGPR Group are listed under
`http://www.ti.uni-mannheim.de/~bmg/Publications-e.html`

PDE–Based Preprocessing of Medical Images

Joachim Weickert and Christoph Schnörr

Computer Vision, Graphics, and Pattern Recognition Group
Department of Mathematics and Computer Science
University of Mannheim, 68131 Mannheim, Germany

{Joachim.Weickert,Christoph.Schnoerr}@ti.uni-mannheim.de
http://www.ti.uni-mannheim.de/~bmg

Abstract

Medical imaging often requires a preprocessing step where filters are applied that remove noise while preserving semantically important structures such as edges. This may help to simplify subsequent tasks such as segmentation. One class of recent adaptive denoising methods consists of methods based on nonlinear partial differential equations (PDEs). In the present paper we survey our recent results on PDE-based preprocessing methods that may be applied to medical imaging problems. We focus on nonlinear diffusion filters and variational restoration methods. We explain the basic ideas, sketch some algorithmic aspects, illustrate the concepts by applying them to medical images such as mammograms, computerized tomography (CT), and magnetic resonance (MR) images. In particular we show the use of these filters as preprocessing steps for segmentation algorithms.

1 Introduction

A basic component of most computer–supported medical image analysis systems is a preprocessing stage for the enhancement of raw image data. This includes both noise reduction and elimination of spurious details in order to improve the result of a subsequent segmentation algorithm, for example, and enhancement of various features relevant for visual inspection in some diagnostic task.

The most basic operation for noise reduction or feature detection is to combine given image data linearly within some local neighbourhood. Here one applies a linear filter in order to compute for instance a local average, or to estimate a partial derivative for the detection of signal transitions. A fundamental drawback of a linear processing stage, however, is the fact that there is no feedback from filter outputs to the processing stage that may be used to control the spatial support of smoothing or to switch from local averaging to anisotropic smoothing in order to preserve signal transitions. Partial differential equations (PDEs) are the appropriate concept to model this and similar functionalities in a mathematical sound way. They lead to well–defined algorithms for the preprocessing of raw image data that are by far more powerful than linear processing stages. In particular, PDEs encode adaptive behaviour in a purely data–driven way that is flexible enough to cope with the rich image structure commonly found in medical images. As a result, an (interactive) user is typically left with just two well–defined global parameters that can be used to browse given image data.

Nonlinear PDE–based image processing has been introduced to the field of computer vision by Perona and Malik [18], and to the field of medical imaging by Gerig *et al.* [8]. In the last years there has been an intensive research reaching from the mathematical foundations and properties of PDE–based image processing to sound discretizations and stable numerical algorithms on parallel computer architectures. However, since these more advanced issues are not the objective of this article, we confine ourselves to a short presentation of the elementary mathematical definitions and the formalism underlying PDE–based image processing in the next section. Rather, the present article aims at surveying some of our PDE–based image processing applications with emphasis on medical image computing. Relevant references will be given below for the reader interested in a more detailed exposition of the underlying mathematical and computational issues as well as the different generalizations that have been used to compute the examples that will be discussed in the remainder of this article.

2 Basic mathematical formulations

Assume that a greyscale image is given by a bounded real-valued function $f(x)$ with $x = (x_1, x_2)^T \in \mathbb{R}^2$. Typically, large values of f represent bright structures, while low values correspond to dark features. This is illustrated in Figure 1(a). One may as well regard the function f as a surface in $\mathbb{R}^3$, as is depicted in Figure 1(b).

One important class of PDE techniques for adaptive image smoothing consists of diffusion filters. Here the grey values of the image $f(x)$ may be regarded as space-variant concentrations of some chemical substance. If we assume that the concentrations $f(x)$ represent the state at time $t = 0$, and that they are diffused over time, we may use the physical laws of diffusion phenomena to describe this evolution. Let us assume that $u(x, t)$ denotes the concentration at time $t \geq 0$ and that the initial condition $u(x, 0) = f(x)$ holds, then the evolution is governed by the so-called diffusion equation

$$u_t = \operatorname{div}(D \nabla u) \tag{1}$$

where the subscript denote partial derivatives, $\nabla u := (u_{x_1}, u_{x_2})^T$, and div is the divergence operator (i.e. $\operatorname{div} \binom{a}{b} = a_{x_1} + b_{x_2}$). Roughly speaking, this equation tells us that the temporal evolution of our image $u(x, t)$ is determined by its second order spatial derivatives. D is a positive definite 2×2 matrix that is called the diffusion tensor. It steers the diffusion process in such a way that the eigenvectors prescribe the diffusion directions and the corresponding eigenvalues determine the amount of diffusion along these directions. Diffusion filters differ from each other by the way this diffusion tensor is chosen.

Let us start by studying the simplest diffusion filter. It has been axiomatically derived almost four decades ago in Japan [11, 25]. If we have two identical eigenvalues (say $\lambda_1 = \lambda_2 = 1$), then the process is isotropic and the directions of the eigenvectors do not matter. The result of such an isotropic linear diffusion process can be seen in Figure 1(c),(d). Although it removes noise and small-scale details very well, it is of restricted use only: it cannot distinguish between noise and semantically important structures such as edges. Both are blurred in the same way.

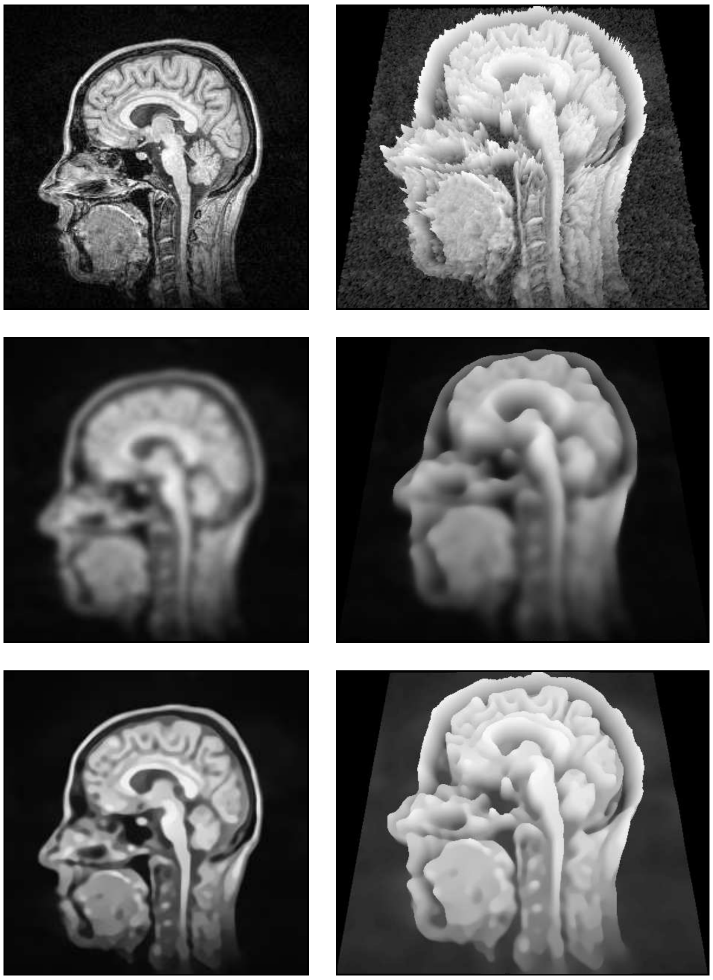

Figure 1: (a) Top left: slice of an MR image. (b) Top right: surface representation of (a). (c) Middle left: after linear diffusion filtering. (d) Middle right: surface representation of (c). (e) Bottom left: after anisotropic edge-enhancing diffusion filtering. (f) Bottom right: surface representation of (e).

As a remedy, nonlinear diffusion filters can be considered. As simple representative of this class would try to reduce smoothing at edges. How can this be achieved? We may identify edges as locations where $|\nabla u|$ is large, and reduce the diffusion process there by choosing eigenvalues that are decreasing in $|\nabla u|$, e.g.

$$\lambda_1(|\nabla u|^2) = \lambda_2(|\nabla u|^2) = \frac{1}{\sqrt{1 + |\nabla u|^2/c^2}}. \tag{2}$$

Such diffusivities are well-suited for denoising purposes, and the resulting diffusion equation has a unique solution that is stable under perturbations of the initial data and the parameters [22, 23].

Using faster decreasing diffusivities as is done e.g. in [18] even allows contrast enhancing behaviour. One the other hand, this may create problems such as unsolved existence and uniqueness questions and high sensitivity to noise. However, also in this case there exists a mathematically sound theory that shows that regularizations of these filters do not suffer from such theoretical and practical problems [6, 25], while still retaining contrast-enhancing properties. This framework can also be extended to the algorithmically important discretizations of these filters [25]. The basic idea behind these regularizations is to replace the edge detector ∇u by a Gaussian-smoothed version of it.

The whole concept can be improved further by introducing anisotropic behaviour into the diffusion process. Figure 1(e),(f) shows an example where an anisotropic diffusion filter has been specifically designed for the enhancement of edges [25]. It uses a diffusion tensor with eigenvectors parallel and orthogonal to the image edges. The eigenvalue that steers the diffusion across the edge is chosen such that it becomes very small when the edge contrast is high. In order to achieve good noise removal, smoothing parallel to the edge is permitted by keeping the corresponding eigenvalue to a fixed value. As an edge detector, a Gaussian-smoothed version of the evolving image gradient is used. In Figure 1(e),(f) it can be seen that this diffusion filter, which adapts itself in a nonlinear way to the evolving image, is well-suited for smoothing noise while simultaneously preserving important features such as edges. For further applications of nonlinear diffusion filtering to medical images we refer to [2, 3, 8, 10, 15, 20, 29] and the references therein.

Another important concept for PDE-based image restoration results from the consideration of variational methods [7, 16, 22]. Many methods of this type use two assumptions: 1. the restored image $u(x,t)$ should not deviate too much from the original image f(x) and 2. it should be piecewise smooth. These requirements are assembled in an energy which is minimized by the optimal restoration. A typical structure of such an energy is given by

$$E_f(u) := \int \Big((u-f)^2 + \; t\,\Psi(|\nabla u|^2) \Big)\, dx \tag{3}$$

where the so-called smoothness potential Ψ is an increasing function in its argument, e.g. $\Psi(s^2) = c^2\sqrt{1+s^2/c^2}$. One can guarantee that this energy has a unique minimum if $\Psi(s^2)$ is a convex function in s [22]. In nonconvex cases such as [16], this is not necessarily the case, and algorithms may get stuck in a local minimum.

The first summand of $E_f(u)$ encourages similarity between the restored image and the original one, while the second summand rewards smoothness. The smoothness weight $t > 0$ is called regularization parameter. From variational calculus it follows that the minimizer of $E_f(u)$ satisfies the Euler–Lagrange equation

$$\frac{u-f}{t} = \operatorname{div}\Big(\Psi'(|\nabla u|^2)\, \nabla u \Big). \tag{4}$$

where Ψ' is the derivative of Ψ. The left hand side of this equation may be regarded as an approximation to u_t. Hence, the variational method approximates a diffusion filter with diffusion tensor $\Psi'(|\nabla u|^2)I$ at time t. The eigenvalues of this tensor are given by

$$\Psi'(|\nabla u|^2) = \frac{1}{\sqrt{1+|\nabla u|^2/c^2}}. \tag{5}$$

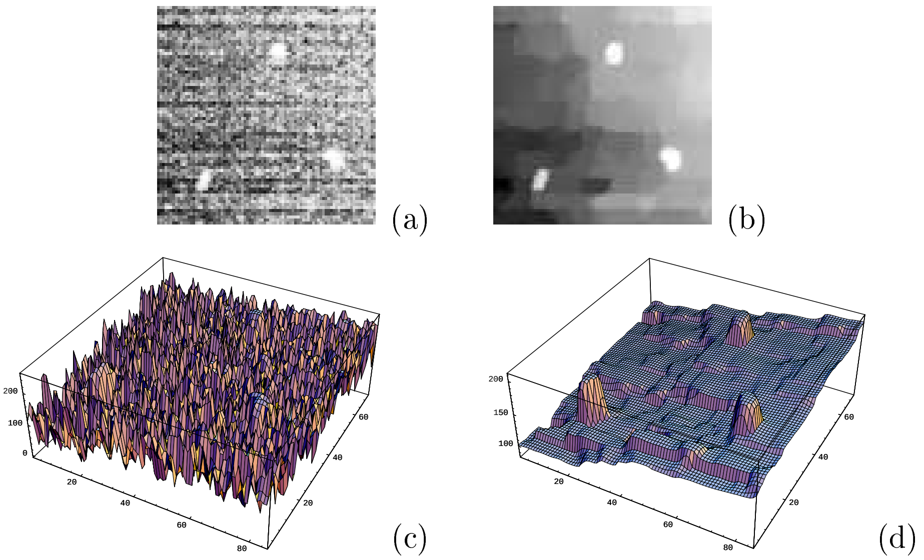

Figure 2: Denoising of a mammogram. (a): Section of the original data. (b): Restored image. (c),(d): Pseudo–3D plots of (a),(b). From [24].

which is just the diffusivity from (2). Hence, the diffusion is slowed down at edges where $|\nabla u|$ is large. This ensures that filters of this type are discontinuity preserving while simultaneously smoothing within the interior of regions. More details on the relations between diffusion filtering and variational methods are described in [21].

Variational methods are very useful for image restoration. Image restoration refers to denoising of severly perturbed image data or to the restoration of image data distorted by the imaging device (point spread function, physical effects). An application of such a variational restoration method is shown in Figure 2, where an approximation to the so–called total variation smoothness potential $\Psi(|\nabla u|^2) = |\nabla u|$ has been used [19]. It can be seen that this technique is well-suited for removing noise while retaining the diagnostically important microcalcifications. Consequently, such a preprocessing constitutes an important tool for the clinician. We remark that this result cannot be computed using traditional methods such as median filtering.

Both diffusion filters and variational methods reveal essentially two natural parameters: a smoothness parameter t and a contrast parameter c. Larger values for t correspond to a stronger image simplication. Locations with gradient magnitudes larger than c are regarded as edges where the smoothing process is inhibited, while locations with gradient magnitudes smaller than c are supposed to belong to the interior of a segment. Here smoothing is desired in order to simplify these structures. The choice of the parameters t and c has of course to depend on the image data and the desired application. There are, however, some heuristic guidelines that help to ease these parameter adaptations [27].

Other important classes of PDE-based image processing methods describe evolution processes that can be linked to mathematical morphology and level set methods. They propagate each level set of the image independently and are thus are invariant under

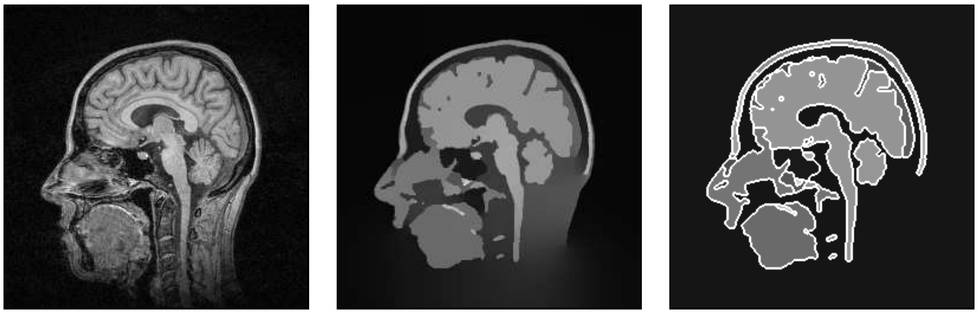

Figure 3: (a) Left: MR image from Figure 1(a). (b) Middle: Filtered with the isotropic nonlinear diffusion process of Catté *et al.* [6]. (c) Right: Watershed segmentation with region merging applied to (b). From [26].

monotone rescalings of the greyvalues (such as histogram equalizations or gamma corrections). For an axiomatic classification of these methods we refer to [1], and related curve evolutions are derived in [17]. It should be observed that the greyscale invariance of these methods implies that contrast does not carry any important information. Since this is not necessarily the case in medical imaging, we do not treat these methods any further in this paper. More details on recent PDE-based image processing methods in general can be found in [5, 9, 14, 26].

3 Nonlinear smoothing and segmentation

The prototypical PDEs described in Section 2 lead to adaptive algorithms that filter out noise and spurious details in homogeneous image regions, but locally adapt to significant signal transitions so as to preserve the relevant image structure. As a result, many traditional segmentation schemes like thresholding or employing watersheds become more robust and hence can be applied successfully in more applications. We illustrate this with two examples.

Figure 3 demonstrates the use of nonlinear diffusion filtering as a preprocessing tool for the watershed algorithm, a classical morphological segmentation method. We observe that this fully automatic segmentation is able to capture many semantically correct objects.

Figure 4 shows three–dimensional variational image restoration of CT data combined with thresholding. It should be mentioned that diffusion filters and variational methods generalize to arbitrary dimensional data sets in a straightforward manner. This example demonstrates that image structures can be discriminated from the background even when a substantial amount of noise is present.

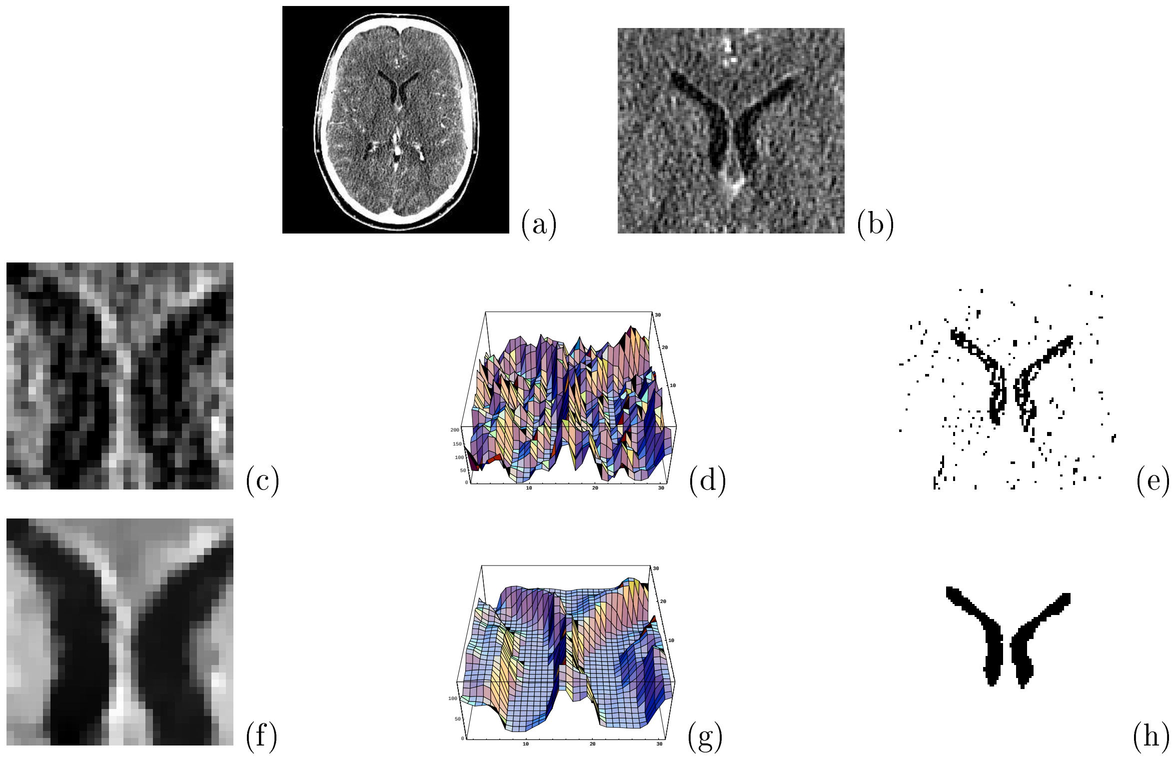

Figure 4: Three-dimensional variational restoration of a computer tomogram. (a): Slice of the 3D CT image data. (b),(c): Sections with an object of interest. (d): Pseudo–3D plot of the section depicted in (c). (e): It is not possible to discriminate object and background with thresholding: Either parts of the object are lost (threshold too high) or the result is contaminated with noise (threshold too low). (f)-(h): Variational restoration exploits homogeneous image structures in 3D and filters out noise while preserving signal transitions. As a result, thresholding succeeds in this case. From [24].

4 Algorithms

Diffusion filtering and variational image restoration are continuous concepts. Since we have to apply them to digital images, it is necessary to use discretizations for the partial differential equations.

For diffusion filtering, a direct way to achieve this is to use finite difference methods. They replace all derivatives by finite differences and proceed iteratively from time 0 to larger times. In its simplest case each such iteration consists of a convolution of the image with a small space- and time-dependent mask [27]. However, such so-called explicit schemes are only stable when small time steps are used. In order to proceed with larger time steps one can apply slightly more complicated schemes (linear implicit schemes) that require to solve a linear system of equations in each step. In recent years quite some efforts have been undertaken in order to find reliable and efficient numerical schemes for nonlinear diffusion filtering; see e.g. [20, 28]. Nonlinear diffusion filtering on current PCs or workstations can be achieved in the order of a second in 2D, and in the order of a minute for typical 3D data sets that arise in medical imaging.

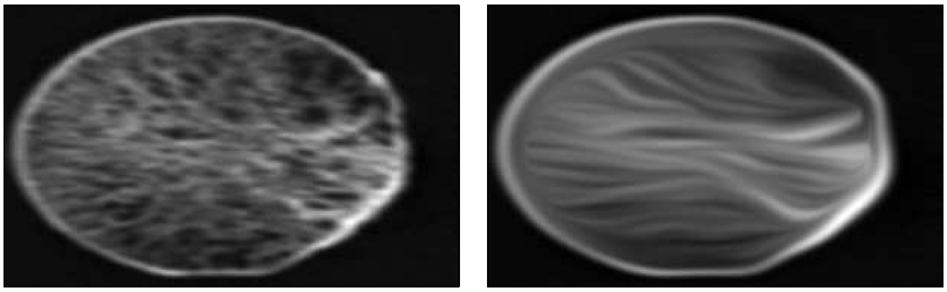

Figure 5: (a) Left: High resolution slipring CT scan of a femural bone. (b) Right: Filtered with coherence-enhancing anisotropic diffusion. From [25].

An alternative to finite difference methods are so-called finite element methods. They are closely linked to variational problems and are therefore considered a natural choice for variational image restoration. Often they lead to linear systems of equations which a similar structure as for finite difference methods [24]. Their use is somewhat more complicated than finite differences, but they offer advantages if one is interested in using adaptive methods with grid coarsening in slowly varying regions [23].

For both types of methods it is possible to achieve significant speed-ups by means of parallelizations. For more details on parallelization strategies and a juxtaposition of finite differences and finite elements we refer to [28].

5 Extensions

PDE-based methods can be designed in such a way that they are optimized for a specific application. For instance, using somewhat more sophisticated tools for image structure analysis than a Gaussian-smoothed gradient, it is possible to design anisotropic diffusion processes that diffuse along parallel lines and flowlike structures [25]. Such a a coherence-enhancing diffusion process is used in Figure 5. This figure depicts a CT scan of a human bone. Its internal structure consists of tiny elongated bony structural elements, the *trabeculae*. Their density and orientation is an important clinical parameter in orthopedics: for instance, the trabecular structures allow to judge the recovery after surgical procedures, or to quantify he rate of progression of rheumatism and osteoporosis. From Figure 5(b) we observe that the anisotropic diffusion filter is capable of enhancing the trabecular structures in order to ease their subsequent orientation analysis.

Another application area of PDE-based methods are active contour models. They are very popular in medical imaging since they allow interactive segmentation [13]. Recent results have shown that it is possible to design specific active contour models (geodesic active contours [4, 12]) that resemble nonlinear diffusion filters. Here the evolving contour is extracted as a level line of a diffusion-like image evolution. Methods of this type offer the advantage that they can handle automatically topological changes such as splitting and merging of contours. Figure 6 shows an example.

Last but not least it should be mentioned that there exist further medical areas

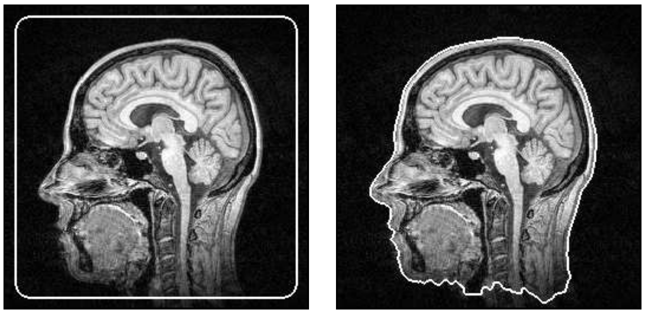

Figure 6: (a) Left: MR image from Figure 1(a) with user-specified initial contour. (b) Right: A geodesic active contour model has moved the initial curve to the object.

where PDE-based methods are relevant: related variational approaches can be used for example for the calculation of displacement fields between subsequent frames in image sequences or for medical image registration.

6 Conclusions

Recent years have witnessed a fruitful interplay between novel PDE-based image restoration techniques and medical image processing as their main application field. In this paper we have sketched the basic ideas and interrelations between two important classes of PDE-based methods (diffusion filtering and variational restoration methods) and demonstrated their use as preprocessing tools for segmentation methods. On one hand medical imaging problems have given rise to develop better image processing methods, while on the other hand progress in PDE-based denoising methods has a direct impact on medical imaging techniques, e.g. by enabling a reduction of the X-ray dose in CT image acquisition. We are confident that this fruitful relation between modern imaging techniques and medical applications is only in its first stage and that much progress is still possible in the near future.

References

[1] L. Alvarez, F. Guichard, P.-L. Lions, J.-M. Morel: Axioms and fundamental equations in image processing, Arch. Rational Mech. Anal. 123, 199–257, 1993.

[2] S.R. Arridge, A. Simmons: Multi-spectral probabilistic diffusion using Bayesian classification, In B. ter Haar Romeny, L. Florack, J. Koenderink, M. Viergever (Eds.), Scale-Space Theory in Computer Vision, Lecture Notes in Computer Science, Vol. 1252, Springer, Berlin, 224–235, 1997.

[3] I. Bajla, I. Holländer: Nonlinear filtering of magnetic resonance tomograms by geometry-driven diffusion, Machine Vision and Applications 10, 243–255, 1998.

[4] V. Caselles, R. Kimmel, G. Sapiro: Geodesic active contours, Int. J. Comput. Vision 22, 61–79, 1997.

[5] V. Caselles, J.M. Morel, G. Sapiro, A. Tannenbaum (Eds.): Special issue on partial differential equations and geometry-driven diffusion in image processing and analysis, IEEE Trans. Image Proc. 7(3), March 1998.

[6] F. Catté, P.-L. Lions, J.-M. Morel, T. Coll: Image selective smoothing and edge detection by nonlinear diffusion, SIAM J. Numer. Anal. 29, 182–193, 1992.

[7] P. Charbonnier, L. Blanc–Féraud, G. Aubert, M. Barlaud: Two deterministic half-quadratic regularization algorithms for computed imaging, In Proc. IEEE Int. Conf. Image Processing (ICIP–94, Austin, Nov. 13–16, 1994), Vol. 2, IEEE Computer Society Press, Los Alamitos, 168–172, 1994.

[8] G. Gerig, O. Kübler, R. Kikinis, F.A. Jolesz: Nonlinear anisotropic filtering of MRI data, IEEE Trans. Medical Imaging 11(2), 221–232, 1992.

[9] B. ter Haar Romeny, L. Florack, J. Koenderink, M. Viergever (Eds.): Scale-Space Theory in Computer Vision, Lecture Notes in Computer Science, Vol. 1252, Springer, Berlin, 1997.

[10] D.-S. Luo, M.A. King, S. Glick: Local geometry variable conductance diffusion for post-reconstruction filtering, IEEE Trans. Nuclear Sci. 41, 2800–2806, 1994.

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

[12] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, A. Yezzi: Conformal curvature flows: from phase transitions to active vision, Arch. Rat. Mech. Anal. 134, 275–301, 1996.

[13] T. McInerney, D. Terzopoulos: Deformable models in medical image analysis: A survey, Medical Image Analysis 1, 91–108, 1996.

[14] M. Nielsen, P. Johansen, O.F. Olsen, J. Weickert (Eds.), Scale-Space Theories in Computer Vision, Lecture Notes in Computer Science, Vol. 1682, Springer, Berlin, 1999.

[15] W.J. Niessen, K.L. Vincken, J. Weickert, B.M. ter Haar Romeny, M.A. Viergever: Multiscale segmentation of three-dimensional MR brain images, Int. J. Comput. Vision 31, 185–202, 1999.

[16] N. Nordström: Biased anisotropic diffusion – a unified regularization and diffusion approach to edge detection, Image and Vision Computing 8(4), 318–327, 1990.

[17] P.J. Olver, G. Sapiro, A. Tannenbaum: Classification and uniqueness of invariant geometric flows, C. R. Acad. Sci. Paris 319, Série I, 339–344, 1994.

[18] P. Perona, J. Malik: Scale-space and edge-detection, IEEE Trans. Pattern Anal. Mach. Intell. 12(7), 629–639, 1990.

[19] L.I. Rudin, S. Osher, E. Fatemi: Nonlinear total variation based noise removal algorithms, Physica D 60, 259–268, 1992

[20] A. Sarti, K. Mikula, F. Sgallari: Nonlinear multiscale analysis of 3D echocardiographic sequences, IEEE Trans. Medical Imaging 18, 453–466, 1999.

[21] O. Scherzer, J. Weickert: Relations between regularization and diffusion filtering, J. Math. Imag. Vision 12, 43–63, 2000.

[22] C. Schnörr: Unique reconstruction of piecewise smooth images by minimizing strictly convex non-quadratic functionals, J. Math. Imag. Vision 4, 189–198, 1994.

[23] C. Schnörr: A study of a convex variational diffusion approach for image segmentation and feature extraction, J. Math. Imag. Vision 8(3), 271–292, 1998.

[24] C. Schnörr: Variational methods for adaptive image smoothing and segmentation, In B. Jähne, H. Haußecker, P. Geißler (Eds.): Handbook on Computer Vision and Applications, Vol. 2: Signal Processing and Pattern Recognition, Academic Press, San Diego, 451–484, 1999.

[25] J. Weickert: Anisotropic Diffusion in Image Processing, Teubner, Stuttgart, 1998.

[26] J. Weickert: Fast segmentation methods based on partial differential equations and the watershed transformation, In P. Levi, R.-J. Ahlers, F. May, M. Schanz (Eds.): Mustererkennung 1998, Springer, Berlin, 93–100, 1998.

[27] J. Weickert: Nonlinear diffusion filtering, In B. Jähne, H. Haußecker, P. Geißler (Eds.), Handbook on Computer Vision and Applications, Vol. 2: Signal Processing and Pattern Recognition, Academic Press, San Diego, 423–450, 1999.

[28] J. Weickert, J. Heers, C. Schnörr, K.J. Zuiderveld, O. Scherzer, H.S. Stiehl: Fast parallel algorithms for a broad class of nonlinear variational diffusion approaches, Real-Time Imaging, in press.

[29] R.T. Whitaker: Characterizing first and second-order patches using geometry-limited diffusion, In H.H. Barrett, A.F. Gmitro (Eds.), Information Processing in Medical Imaging, Lecture Notes in Computer Science, Vol. 687, Springer, Berlin, 149–167, 1993.