Nonlinear Anisotropic Filtering of MRI Data

Guido Gerig, Olaf Kübler, Ron Kikinis, and Ferenc A. Jolesz

Abstract—Despite significant improvements in image quality over the past several years, the full exploitation of magnetic resonance image (MRI) data is often limited by low signal-to-noise ratio (SNR) or contrast-to-noise ratio (CNR). In implementing new MR techniques, the criteria of acquisition speed and image quality are usually paramount. To decrease noise during the acquisition either time averaging over repeated measurements or enlarging voxel volume may be employed. However, these methods either substantially increase the overall acquisition time or scan a spatial volume in only coarse intervals. In contrast to acquisition-based noise reduction methods we propose a post-process based on anisotropic diffusion. Extensions of this new technique support 3-D and multi-echo MRI, incorporating higher spatial and spectral dimensions. The procedure overcomes the major drawbacks of conventional filter methods, namely the blurring of object boundaries and the suppression of fine structural details. The simplicity of the filter algorithm enables an efficient implementation even on small workstations.

We demonstrate the efficient noise reduction and sharpening of object boundaries by applying this image processing technique to 2-D and 3-D spin echo and gradient echo MR data. The potential advantages for MRI, diagnosis and computerized analysis are discussed in detail.

I. INTRODUCTION

IN medical imaging we often face a relatively low SNR with good contrast, or a low contrast with good SNR. Fortunately the human visual system is highly effective in recognizing structures even in the presence of a considerable amount of noise. But if the SNR is too small or the contrast too low it becomes very difficult to detect anatomical structures because tissue characterization fails. A definition of overall image quality includes physical and perceptual criteria. Furthermore, it largely depends on specific diagnostic tasks. In some cases a high spatial resolution and a high contrast are required, whereas in other cases more perceptual criteria may be favored. For a visual analysis of medical images, the clarity of details and the object visibility are important, whereas for image processing a high SNR is required because most of the image segmentation algorithms are very sensitive to noise.

There are several ways to improve SNR, and they can be divided into subcategories according to time requirements, resolution criteria, hardware- versus software-based techniques.

- Methods affecting acquisition time or pixel (voxel) dimensions.

- Time domain averaging (averaging repeated acquisitions). Problem: inefficiency.
- Scanning with larger voxels. Problem: loss of resolution, mostly in the out-of-plane direction.
- Methods without time and/or resolution penalty.
  - Signal processing during acquisition (i.e., variable bandwidth: narrower bandwidth for the second echo of double echo spin echo acquisitions).
  - Improvement of acquisition hardware (increase signal or reduce source of noise).
  - Postprocessing of raw data or image data (filter techniques). Problem: blurring, loss of resolution, generation of artifacts.

Although the acquisition parameters can be optimized regarding SNR and contrast, methods to reduce noise (e.g., increasing the number of excitations) usually result in a significant increase in the overall acquisition time. While providing access to important new anatomical and functional information through high-speed acquisition, or high spatial resolution, advanced imaging techniques are often penalized by a decrease in image SNR.

Filtering techniques have the advantage of not affecting the acquisition process. In linear spatial filtering, the content of a pixel is given the value of the average brightness of its immediate neighbors. Simple spatial averaging, often called "low-pass filtering," does reduce the amplitude of noise fluctuations, but also degrades sharp details such as lines or edges. The filtering does not respect region boundaries or small structures, and the resulting images appear blurry and diffused. This undesirable effect can be reduced or avoided by the design of nonlinear filters, the most common technique being median filtering. Edges are retained, but the filtering results in a loss of resolution by suppressing fine details. Another approach is adaptive filtering (see [1] for a detailed survey), which entails a tradeoff between smoothing efficiency, preservation of discontinuities, and the generation of artifacts. When developing a filtering method for medical image data, image degradation by blurring or by artifacts resulting from a filtering scheme is not acceptable. The following requirements should ideally be fulfilled:

a) minimize information loss by preserving object boundaries and detailed structures,
b) efficiently remove noise in regions of homogeneous physical properties, and
c) enhance morphological definition by sharpening discontinuities.

Recent developments based on anisotropic diffusion filtering overcome the major drawbacks of conventional spatial filtering [3], and significantly improve image quality while satisfying

the main criteria stated above. Special extensions for filtering multichannel and 3-D data make the method especially appropriate to enhance various types of magnetic resonance (MR) image data.

In the present paper, the term *volume data* will be used independently of the acquisition type. With *volume data* we mean a volume-covering acquisition with isotropic or nearly isotropic voxel dimensions. Such data can be measured as a 3-D FT acquisition or as a 2-D multislice acquisition with thin slices and no gap between the slices.

II. METHOD: ITERATED SPATIAL ANISOTROPIC SMOOTHING

Perona and Malik [2], [3] developed a multiscale smoothing and edge detection scheme which is a powerful new concept for image processing. Their anisotropic diffusion filtering method is mathematically formulated as a diffusion process, and encourages intraregion smoothing in preference to smoothing across the boundaries. In their filtering method the estimation about local image structure is guided by knowledge about the statistics of the noise degradation and the edge strengths.

The following description is based on the anisotropic diffusion process as proposed by Perona and Malik.

A. Mathematical Principles of the Nonlinear Anisotropic Filtering Method

Smoothing is formulated as a diffusive process, which is suppressed or stopped at boundaries by selecting locally adaptive diffusion strengths. In any dimension this process can be formulated mathematically as follows, assuming no sinks or sources ([4, pp. 97]):

$$\frac{\partial}{\partial t} u(\overline{x}, t) = \mathrm{div}(c(\overline{x}, t)\nabla u(\overline{x}, t)). \tag{1}$$

The diffusion strength is controlled by $c(\overline{x}, t)$. In our data filtering scheme the vector $\overline{x}$ represents the spatial coordinate(s). The variable t is the process ordering parameter; in the discrete implementation it is used to enumerate iteration steps. The function $u(\overline{x}, t)$ is taken as image intensity $I(\overline{x}, t)$. The diffusion function $c(\overline{x}, t)$ depends on the magnitude of the gradient of the image intensity. It is a monotonically decreasing function $c(\overline{x}, t) = f(|\nabla I(\overline{x}, t)|)$, which mainly diffuses within regions and does not affect region boundaries at locations of high gradients.

Two different diffusion functions have been proposed:

$$c_1(\overline{x}, t) = \exp\left(-\left(\frac{|\nabla I(\overline{x}, t)|}{\mathcal{K}}\right)^2\right) \tag{2}$$

$$c_2(\overline{x}, t) = \frac{1}{1 + \left(\frac{|\nabla I(\overline{x}, t)|}{\mathcal{K}}\right)^{1+\alpha}} \ \big| \ \alpha > 0. \tag{3}$$

Fig. 1(a) shows the monotonical decrease of the diffusion coefficient with increasing gradient. If the gradient is large, a discontinuity is assumed and the diffusion is halted. The parameter $\mathcal{K}$ is chosen according to the noise level and the edge strength. To understand the relation between the parameter $\mathcal{K}$ and the discontinuity value ∇I, it is proposed
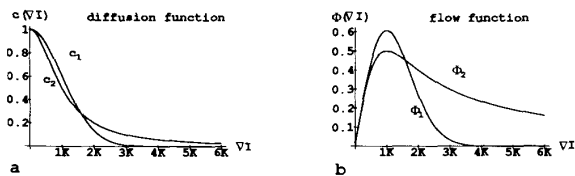

Fig. 1.   (a) Diffusion functions (diffusion strength c versus gradient ∇I). (b) Flow functions (flow $\Phi(\nabla I)$ versus gradient ∇I). ∇I is given in scales of the parameter $\mathcal{K}$.

to define $\Phi(\nabla I)$ as the product $c * \nabla I$, called flow. The flow functions related to c_1 and c_2 are plotted in Fig. 1(b). The maximum flow is generated at locations with gradient ∇I equal to $\mathcal{K}$. When decreasing below $\mathcal{K}$, the flow reduces to zero because in homogeneous regions only minimal or no flow takes place. Above $\mathcal{K}$ the flow function again decreases to zero, halting diffusion at locations of high gradients. A proper choice of the diffusion function not only preserves, but also enhances edges while being numerically stable [3], [5]. To simplify the mathematical treatment of edge sharpening, the amplification of the edge strength is usually considered at the point of inflection. It is supposed that within the near neighborhood of the point of inflection the slope of the edge will increase with time (number of iterations).

Assuming the edge is oriented along the y-axis, the variation of the edge slope $\frac{\partial I}{\partial x}$ over time becomes

$$\frac{\partial}{\partial t}\frac{\partial I}{\partial x} = \frac{\partial^2 c}{\partial x^2}\frac{\partial I}{\partial x} + 2\frac{\partial c}{\partial x}\frac{\partial^2 I}{\partial x^2} + c\frac{\partial^3 I}{\partial x^3}. \tag{4}$$

Fig. 2 illustrates the variation of the edge slope along the edge profile. By selecting a diffusion function $c(x, t)$, and utilizing the error function as an ideal model of a blurred step edge [Fig. 2(a)], the slope change was calculated along the edge profile. Simulations were done with varying parameters $\mathcal{K}$ and the two diffusion functions c_1 and c_2. The edge gradient (first derivative) is represented in Fig. 2(b). Fig. 2(c) shows the slope change along the edge profile for c_1 and for the value of $\mathcal{K}$ giving the maximum sharpening effect ($\mathcal{K} = 0.23$). The positive central portion of Fig. 2(c) represents the region of increasing slope, the negative sidelobes the region of decreasing slope. Additional positive sidelobes indicate that there is another increase of slope in the foot and the shoulder regions of the edge, which creates an initial "round off" of the edge function.

The anisotropic diffusion has the property of blurring small discontinuities and sharpening edges. Fig. 3 illustrates the filtering of a 1-D, blurred step edge (3a) and a blurred and noisy step edge (3b) as a function of the number of iterations.

B. Assumption About Signal Structure: Image Model

The anisotropic diffusion as described performs a piecewise smoothing of the original signal. The propagation of information between discontinuities results in regions of constant intensity or linear variations of low frequency. The assumption of piecewise constant or slowly varying intensities is a good
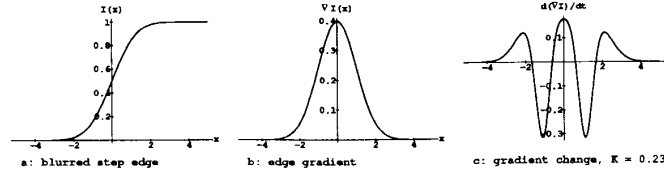
Fig. 2. Variation of edge slope (gradient) with time. (a) Blurred step edge. (b) Edge gradient. (c) Initial rate of change of gradient.
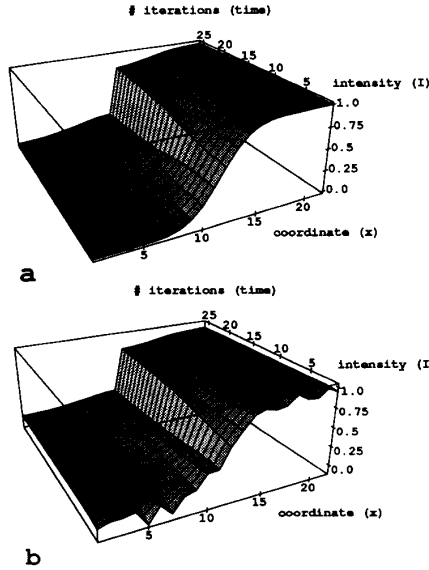

Fig. 3. Iterative edge sharpening and noise suppression. (a) Blurred step edge. (b) Blurred and noisy step edge.
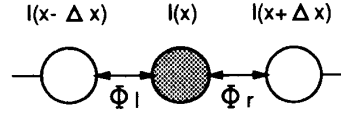

Fig. 4. One-dimensional network structure. Circles represent pixel nodes $I(x)$ linked together by arcs. Φ_l and Φ_r are the flow contributions.

C. Discrete Formulation of Anisotropic Smoothing

The filtering of discrete signals requires a reformulation of the methods defined for the continuous case: Instead of differentiation, estimates about local gradients are calculated as differences between neighboring data elements ([4, pp. 104]).

D. 1-D Filtering

In order to provide an appreciation of anisotropic nonlinear diffusive smoothing, we first present the 1-D case. The intensity change in one iteration step is defined as the sum of the flow contributions between neighboring pixel intensities. The structure is simulated as a network, wherein the center points of pixels represent nodes and are linked together by arcs, whose flow characteristics are determined by the conductivity function (Fig. 4).

The 1-D discrete implementation derives from the continuous diffusive process as follows:

$$
\begin{aligned}
\frac{\partial}{\partial t} I(x,t) &= \mathrm{div}[c(x,t) * \mathrm{grad}\ I(x,t)] \\
&= \nabla^T[c(x,t) * \nabla I(x,t)] \\
&= \frac{\partial}{\partial x}\left[c(x,t) * \frac{\partial}{\partial x} I(x,t)\right] \\
&\approx \frac{\partial}{\partial x}\left[c(x,t) * \frac{1}{\Delta x}\left(I\left(x + \frac{\Delta x}{2}, t\right)\right.\right. \\
&\qquad\left.\left. - I\left(x - \frac{\Delta x}{2}, t\right)\right)\right] \\
&\approx \frac{1}{\Delta x^2}\left[c\left(x + \frac{\Delta x}{2}, t\right)\right. \\
&\qquad * (I(x + \Delta x, t) - I(x,t)) \\
&\qquad - c\left(x - \frac{\Delta x}{2}, t\right) \\
&\qquad\left. * (I(x,t) - I(x - \Delta x, t))\right] \\
&= \Phi_{\mathrm{right}} - \Phi_{\mathrm{left}}\ \mid\ \Delta x = 1 \qquad (5)
\end{aligned}
$$

estimate to model MR image data, which comprise smooth regions separated by discontinuities, representing various tissue categories distinguished by different proton densities and relaxation properties. A tissue category is characterized by a specific intensity level which can be assumed to be approximately constant over the image plane if the MR scanner is optimally tuned and if nonlinearity artifacts are precorrected. From the anatomical point of view this piecewise constancy can only be partially true.

If the original signal were not constant or not slowly varying, a filtering method such as ours would fail. Noisy intensity surfaces of constant slope would be broken into several homogeneous patches separated by newly generated discontinuities. Saint Marc and Medioni [6] presented a solution for the filtering of range data where depth is directly encoded as gray-level intensity. They applied the adaptive filtering not to the original signal but to its first derivative, generating patches of constant slope.

The power of the anisotropic smoothing scheme proposed herein lies in the fact that it deals with local estimates of the underlying image structures, which are highly flexible. Discontinuities are preserved and their position is not affected. Intensity fields of a weak slope remain nearly unchanged if the slope falls within the monotonically increasing part of the flow function (gradient values below $\mathcal{K}$).
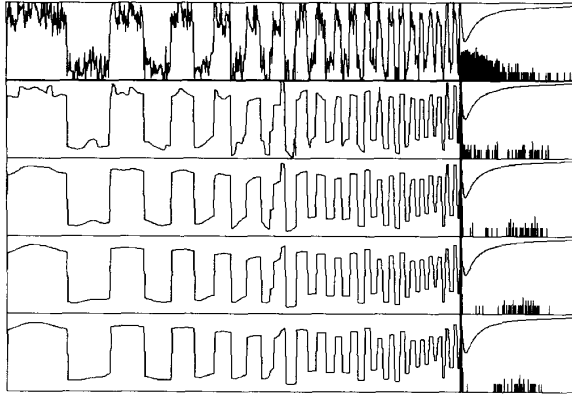
Fig. 5. One-dimensional iterated anisotropic smoothing. Top to bottom: original function and different stages of smoothing. Right-hand column: histogram of gradients and flow function (vertically flipped).

$$I(t + \Delta t) \approx I(t) + \Delta t * \frac{\partial}{\partial t} I(t)$$
$$= I(t) + \Delta t * (\Phi_r - \Phi_l).$$

$$(6)$$

Stability of the iterated processing scheme is obtained by a proper setting of the integration constant Δt (see Appendix A).

Performing the anisotropic diffusion on 1-D test functions allows systematic study of the signal change over time. Note the important capability of smoothing noise while enhancing discontinuities. Fig. 5 shows the iterated smoothing of an original test function which is generated as a rectangular function corrupted by additive white noise. In addition to the 1-D functions the histogram of gradient magnitudes and the flow function (vertically flipped) were plotted (right-hand column). The depletion of gradient values generating large flow contributions (values around the peak of the flow function) is significant, such gradients are either smoothed or sharpened (shifted to higher values).

E. 2-D Filtering

The 2-D procedure is a simple extension of the 1-D discrete implementation:

$$\frac{\partial}{\partial t} I(\overline{x}, t) = \mathrm{div}[c(\overline{x}, t) * \mathrm{grad}\ I(\overline{x}, t)]$$
$$= \nabla^T [c(\overline{x}, t) * \nabla I(\overline{x}, t)]$$
$$= \frac{\partial}{\partial x}\left[c(\overline{x}, t) * \frac{\partial}{\partial x} I(\overline{x}, t) \right]$$
$$+ \frac{\partial}{\partial y}\left[c(\overline{x}, t) * \frac{\partial}{\partial y} I(\overline{x}, t) \right]$$
$$= \frac{1}{\Delta x^2}\left[c\left(x + \frac{\Delta x}{2}, y, t \right) \right.$$
$$* (I(x + \Delta x, y, t) - I(x, y, t))$$
$$- c\left(x - \frac{\Delta x}{2}, y, t \right)$$
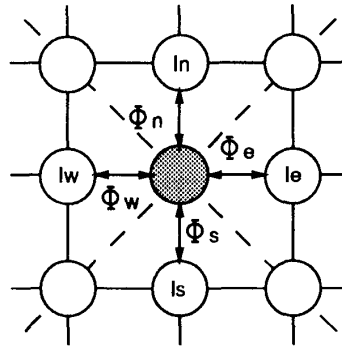$$\left. * (I(x, y, t) - I(x - \Delta, y, t)) \right]$$

Fig. 6. Two-dimensional network structure.

$$+ \frac{1}{\Delta y^2}\left[c\left(x, y + \frac{\Delta y}{2}, t \right) \right.$$
$$* (I(x, y + \Delta y, t) - I(x, y, t))$$
$$- c\left(x, y - \frac{\Delta y}{2}, t \right)$$
$$* (I(x, y, t)$$
$$\left. - I(x, y - \Delta y, t)) \right]$$
$$= \Phi_{\mathrm{east}} - \Phi_{\mathrm{west}} + \Phi_{\mathrm{north}} - \Phi_{\mathrm{south}}.$$

$$(7)$$

The 2-D discrete implementation results in simple, local operations replicated over the digital image. In the 2-D network (Fig. 6) in a first step the signal flow is calculated between neighboring nodes. In a second step the node intensities are updated by the local sum of the flow contributions:

$$I(t + \Delta t) \approx I(t) + \Delta t * \frac{\partial}{\partial t} I(t)$$
$$= I(t) + \Delta t * (\Phi_e - \Phi_w + \Phi_n - \Phi_s). \quad (8)$$

To obtain better isotropy, flow is also calculated between diagonally neighboring pixels (dashed connections Fig. 6), resulting in an eight-way-connected network. The longer distance between diagonal neighbors is taken into account by setting Δd to $\sqrt{2}$. Our experience has shown that inclusion of these grid points is sufficient and more elaborate integration schemes are not necessary. The integration constant Δt must be adjusted to the different neighborhood structures (Appendix A).

F. Convergence of Iterated Smoothing

The diffusion process as proposed by Perona and Malik [3] did not incorporate a convergence criterion. Nordström [5] argued, that at the limit of infinite time, however, the image would converge to a constant, and that remarkable impressive edges could be obtained at some stage of the iterated filtering. In the discrete application of the diffusion process, it can be stopped after a few iterations. For large gradient values, the flow function decreases to zero (limits in hardware precision)

and completely arrests any diffusion. We prespecified the maximum number of iterations to halt diffusion.

Nordström [5] unified the concepts of anisotropic diffusion and variational regularization. He claimed that his algorithm, called "biased anisotropic diffusion," combined the better properties of both concepts. The algorithm converges to a steady state solution and only requires the solution of a single boundary value problem over the entire image domain. Biased anisotropic diffusion was shown to be a method for solving a well-defined mathematical problem. Nordström's method is closely related to the anisotropic diffusion described by Perona and Malik, only differing in an additional term expressing the deviation between the original $I(t_0)$ and the filtered image functions. Equation (9) describes the 1-D discrete implementation of the biased anisotropic diffusion:

$$I(t + \Delta t) \approx I(t) + \Delta t * [(\Phi_r - \Phi_l) + (I(t_0) - I(t))]. \tag{9}$$

The bias term $(I(t_0) - I(t))$ is responsible for the mathematically sound formalism and the satisfying convergence property, but it does not influence the local decisions of enhancement versus blurring. The images are repeatedly filtered until a steady state is reached, no number of iteratons has to be prespecified.

For a small number of iterations, the results of the unbiased- and biased anisotropic filtering are supposed to be the same. In the limit, however, a nice convergence becomes very important, as it only guarantees image structures to be preserved.

G. Extensions for Optimal Filtering of 3-D and Multi-echo MR Data

The formulation of the anisotropic smoothing as a diffusion process on a regular lattice of pixel intensities easily allows a special adaptation to MR image data. The 2-D filtering scheme described above is appropriate for the enhancement of single slice data. If volume-covering data are acquired with isotropic or nearly isotropic voxels the filtering should make use of its three-dimensional nature, because the significantly enlarged neighborhood results in a much better noise reduction and in an enhancement of edges in 3-D. Another case is the acquisition of multi-echo MR image data, resulting in multiple measurements at one spatial location. Noise can be assumed to be independent between the multiple channels, while the correlation of the spatial location of discontinuities on two or more echoes helps to further increase the sensitivity of the local filtering to significant image structures. An augmented filtering scheme optimized to process the two special data types is described in the following sections.

1) Filtering of 3-D Data: The formulation of a 3-D diffusion process follows directly from the original anisotropic diffusion equation where $\overline{x}$ corresponds to (x, y, z).

The total of the flow contributions at each node is now taken from a 3-D neighborhood of volume elements (voxels) including the 6 immediate neighbors or the 26 voxels within a $3 \times 3 \times 3$ voxel window. This increased number of samples results in a much better noise reduction and in an enhancement

of 3-D discontinuities, which allows a more accurate preservation of the continuity of structures in 3-D space. When dealing with noncubic voxels (one voxel dimension significantly larger than the other two), the larger distance in one dimension is taken into account by setting the correct absolute distance into the calculation of the gradients and the diffusion coefficients.

2) Filtering of Multichannel Data: To obtain maximum information from a single clinical MR exam, usually two echoes are measured. The different measurements at one voxel location represent vector-value information, the components of which describe different physical properties. They can lead to a better discrimination of tissue characteristics if analyzed together, providing the anatomical structure can be distinguished on both channels. This permits the development of an enhancement filter working simultaneously on both channels, assuming perfect spatial coincidence. In this case, two network structures represent the two image channels of a measured original slice. The discontinuities between neighboring nodes in both nets determine the amount of diffusion. A coupling between the two networks can be achieved by combining the corresponding diffusion coefficients between equivalent nodes. We used the Euclidean Norm to combine the information of the multiple channels, which requires a little more computational effort than taking the average of the gradient magnitudes or choosing the maximum of the two gradient magnitudes, but has advantages because the former too strongly supports discontinuities observed in both channels whereas the latter does not amplify correlated structural phenomena observed in both channels.

$$\frac{\partial}{\partial t} \begin{pmatrix} I_1(\overline{x}, t) \\ I_2(\overline{x}, t) \end{pmatrix} = \begin{pmatrix} \text{div}[c_1(\overline{x}, t) * \text{grad } I_1(\overline{x}, t)] \\ \text{div}[c_2(\overline{x}, t) * \text{grad } I_2(\overline{x}, t)] \end{pmatrix}$$
$$= \begin{pmatrix} \text{div}[c_c(\overline{x}, t) * \text{grad } I_1(\overline{x}, t)] \\ \text{div}[c_c(\overline{x}, t) * \text{grad } I_2(\overline{x}, t)] \end{pmatrix} \tag{10}$$

$$c_c(\overline{x}, t) = f(|\nabla I_1(\overline{x}, t)|, |\nabla I_2(\overline{x}, t)|)$$
$$= f\left(\sqrt{\nabla I_1(\overline{x}, t)^2 + \nabla I_2(\overline{x}, t)^2} \right). \tag{11}$$

The combined diffusion coefficients $c_c(\overline{x}, t)$ are calculated by replacing the gradient magnitude $|\nabla I(\overline{x}, t)|$ in (2) and (3) by $\sqrt{\nabla I_1(\overline{x}, t)^2 + \nabla I_2(\overline{x}, t)^2}$. The coefficients c_c are used in both networks to define the amount of diffusion in each. Fig. 7 illustrates the coupling of the two diffusion arcs between corresponding nodes. The behavior of the coupled networks can be characterized as follows. If discontinuities are detected in both nets, the combined diffusion coefficient is larger than any single component, and the significance of local estimations is increased. If, on the other hand, a discontinuity is detected only in one channel, the combined coefficient responds to the discontinuity and halts the diffusion.

A nice property of this multivariate formulation of the anisotropic smoothing by coupling the diffusion coefficients is the inherent preservation of correlating and contrasting effects of image intensities among multiple channels, because the diffusion is proportional to the absolute gradient. After calculating c_c, the amount of flow and the flow direction is
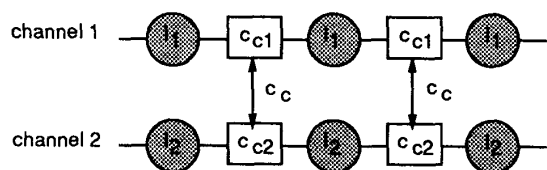
Fig. 7. One-dimensional network structure of two-channel anisotropic filtering.

calculated in each network individually. A simple averaging of different echoes to reduce noise would not have this property, because intensity steps with opposite sign could counteract each other.

III. RESULTS

The filtering scheme has been applied to 2-D and 3-D image data with one or two channels. Tests with various numbers of iterations and different parameters were made. In addition to subjective visual evaluation, it is desirable to present quantitative results of the noise reduction. A procedure that measures noise in images has been developed and is described below.

A. Measuring Noise in MR Images

The methods of measuring SNR for comparing the performance of imaging techniques is described in detail [7]. To express the noise-removing efficiency of the filtering quantitatively a measurement technique is needed.

Noise resulting from different sources is defined simply as the variation representing a deviation from the true value. It is assumed that homogeneous, non-textured image areas are representative of tissue categories. A "true" value is estimated by the mean value of a set of pixels (or voxels) expected to belong to the same tissue category, whereas noise is expressed by the standard deviation (SD). Finding uniform regions over which to measure noise is difficult. A region must be large enough to yield a significant result for the SD. On the other hand the larger the region the higher the likelihood that the statistics will be affected by small nonuniformities such as intensity gradients or structural variations. Noise can be measured in areas where there is no signal, that is through air, but is underestimated ($0.655 * SD$) due to the rectification performed when obtaining the MR magnitude images [7].

Canny [8] assumes that the amplitude distribution of the gradients tends to be different for edges and noise. In his global histogram estimate he computed a cumulative histogram of the absolute gradients and supposed the lower 80 percent to represent mainly the noise energy. Perona and Malik [3] used this "noise estimator" and set the parameter $\mathcal{K}$ equal to the 90% value of the cumulative histogram of the gradients. In the medical image data represented in our current report the gradient histogram did not show a significant separation between a compact distribution due to noise at low gradients and infrequently occurring larger values due to edge responses. The right-hand column of Fig. 5 represents typical examples of gradient histograms.

We developed a procedure to find most homogeneous regions in images automatically. After the selection of window size and grid spacing, the mean and standard deviation (SD) are calculated in windows centered at each grid point. The image intensity range is quantized into a set of intervals. The window with the minimum SD of all the regions having a mean value falling within each intensity subrange is chosen to be a representative region. It is usually not possible to find a homogeneous region for each intensity subrange. The SD estimation is biased towards lower values by selecting the window with the smallest SD. Experiments show, that the procedure allows homogeneous areas to be found efficiently, the dependence of noise upon the signal level to be calculated, and automatic estimates of noise in background and tissue regions to be made. A region size of $8 * 8$ pixels and a subdivision of the intensity range into 25 intervals allows relatively large regions which are only minimally affected by nonuniformities to be identified. From the resulting list of SD's, two values corresponding to noise in background and noise in tissue are selected.

B. Test Data of a Formalin Fixed Brain Specimen

We tested the noise reduction efficiency and the ability to preserve image structures on a series of images of a formalin fixed human brain with different numbers of excitations (NEX). Acquisitions (TR 2000 ms, TE 20/80 ms) were made with NEX of 4, 2, 1 and 0.5. The 0.5 NEX image is acquired using the half-Fourier technique. Increasing the number of NEX from 0.5 to 4 theoretically results in a decrease of noise by the factor $\sqrt{8}$. A visual analysis of the image series in selected regions of interest (ROI's) clearly shows the different noise levels (Fig. 8), and exhibits the delineation of finely detailed structures in the 4 NEX image which cannot be detected in the 0.5 NEX image, although the human visual system is extremely efficient in detecting structures even in the presence of a considerable amount of noise.

We applied the filtering scheme to the 0.5 and 1 NEX images and compared the results to the unfiltered 4 NEX acquisition. Table I displays the results of measuring noise in the most homogeneous background and tissue regions. Noise in the filtered half-Fourier image is lower than in the 4 NEX image, demonstrating the efficient noise removing capacity of the filtering. Three iterations of the 2-D filtering were carried out using the diffusion function c_1 with parameter $\mathcal{K} = 2.0$.

Fig. 8 depicts enlarged subregions of the first echo image in order to compare the detail preserving and edge-enhancement ability of the filtering, demonstrating that edges are significantly enhanced while fine details are kept.

C. 2-D Filtering of Dual-Echo SE Images

Dual echo spin echo brain images were obtained as part of routine clinical patient studies. Imaging was performed at 1.5 Tesla using a General Electric SIGNA MRI system (Milwaukee, WI) with a double echo sequence (TR 3000 ms, TE 30/80 ms) with 1 mm * 1 mm * 3 mm voxel dimensions. Contiguous slices (no gap) were obtained by combining two
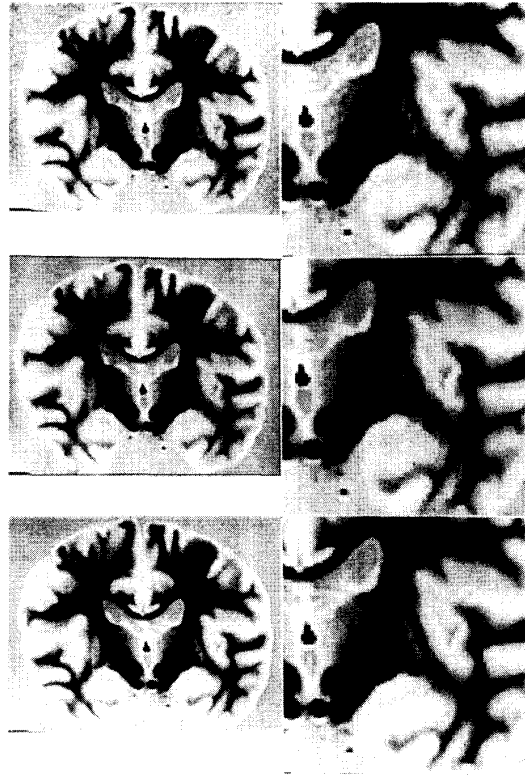
Fig. 8.   Images of a formalin fixed human brain specimen (TR 2000 ms, TE 20 ms, FOV 24.0 cm, slice thickness 5 mm). Comparison of 0.5 NEX, 0.5 NEX filtered and 4 NEX (top to bottom), original size (left) and zoomed ($2.r$) (right).

TABLE I

MRI OF FORMALIN FIXED HUMAN BRAIN: AUTOMATIC NOISE ESTIMATION ON A SERIES WITH DIFFERENT NUMBERS OF EXCITATIONS: UNFILTERED AND FILTERED

Acquisition	theoretical relative noise	Background		Tissue	
		1st echo	2nd echo	1st echo	2nd echo
0.5 NEX	1.41	0.66	0.59	1.23	1.35
1 NEX	1.00	0.50	0.51	1.03	0.98
2 NEX	0.71	0.38	0.38	0.83	0.77
4 NEX	0.50	0.28	0.29	0.56	0.59
0.5 NEX filtered		0.11	0.09	0.36	0.41
1 NEX filtered		0.10	0.09	0.28	0.29

interleaved sequences. Forty double echo slices were used to cover the brain volume. Imaging time was reduced by using half-Fourier sampling (0.5 NEX), in which the SNR is reduced by a factor of $\sqrt{2}$ compared to one NEX. The imaging time was under 8 min.

If the voxel geometry is highly anisotropic such as in this study, a 3-D instead of a 2-D filtering does not substantially improve the result because of the relatively low noise smoothing contribution of voxels on neighboring slices separated by a large distance. More helpful is the availability of multiple echoes, which are suitable for the application of the new multichannel filtering process, allowing significant noise reduction provided by the correlation of the information about image structures in several channels. The unfiltered and the filtered

images are shown in Fig. 9. The white boxes mark most homogeneous subregions of background and tissue, giving noise estimations for the first and second echo of $\sigma_n = 4.69$ and $\sigma_n = 3.21$, respectively. Filtering was performed with three iterations of the 2-D two-channel filtering with diffusion function c_1 and $\mathcal{K} = 6.0$. The intensity profile along the white line clearly demonstrates the smoothing and edge enhancing effect. The difference images between the original and the filtered image data illustrate the noise structures removed. Fig. 10 depicts enlarged parts of the original and filtered slice pair.

D. Application to 3-D MR-Angiograms ("Time-of-Flight" MRI)

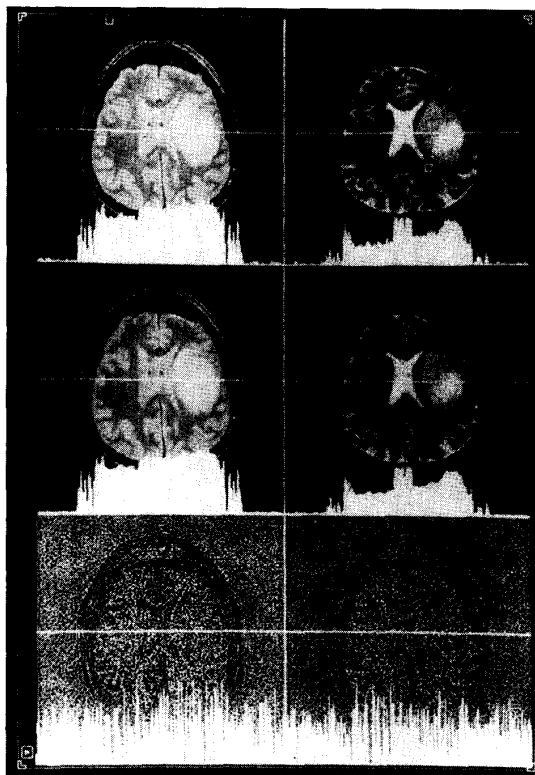Fully or nearly fully isotropic volume data may be obtained

Fig. 9. Dual echo SE (TR 3000 ms, TE 30/80 ms, FOV 24.0 cm, slice thickness 3 mm, no gap): original slice (top), filtered slice (2-D two-channel) (middle), and difference image (bottom). Left column: 1st echo; Right column: 2nd echo. The intensity profile along the white line is overlaid. White boxes indicate regions where noise was estimated.

Fig. 10. Dual echo SE: Zoomed parts ($4x$) of Fig. 9, original (top row) and filtered (bottom row) slices.

by applying special gradient echo sequences. Here, the new processing extended to 3-D exhibits its full advantage, because local estimation and smoothing is performed over a spatially much larger population of measurements. The 3-D filtering will become important for imaging techniques in which the criterion of acquisition speed takes precedence over image quality (e.g., to avoid motion artifacts), or in which a certain amount of noise degradation is linked to the physical measurement and cannot be reduced.

One example of this situation is MR-angiography techniques (MRA), which enhance the contrast of fast flowing substance and suppress the intensity of static tissues. The method used to obtain the MRA was a multiple thin-slab 3-D gradient echo technique [9]. Seventy-two gradient echo slices (TR = 50 ms, TE = 5 ms, flip angle = 80°) were acquired within 15 min. The 15 mm thick slabs were divided with 16 phase encoding steps into approximately 1 mm thick slices. The resulting image data have 3-D isotropic resolution to yield complete information about finely detailed vascular structures. Especially important in such a case is the discontinuity preserving and enhancing property of a filtering process, because the recognition of thin tubular structures is the main goal (Fig. 11). Noise was measured within the $8 * 8$ windows marked by the white boxes, yielding values for background $\sigma_n = 0.65$

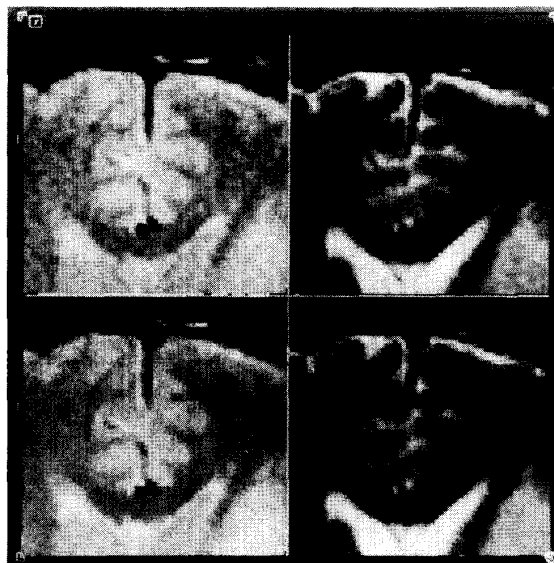and tissue region $\sigma_n = 1.18$. Three iterations of the 3-D filtering with $\mathcal{K} = 2.4$ were applied. After smoothing, noise in the tissue region was reduced to $\sigma_n = 0.42$.

E. Application to 3-D Gradient Echo Data

The 3-D filtering was applied to a spoiled grass (SPGR, gradient echo) volume acquisition consisting of 124 contiguous coronal slices with 1.5 mm $* 1$ mm $*$ mm voxels (Fig. 12). Noise in background and tissue was calculated to be $\sigma_n = 1.17$ and $\sigma_n = 2.78$ (see white boxes). After three iterations with the diffusion function c_1 and $\mathcal{K} = 5.0$ noise was significantly suppressed ($\sigma_n = 0.63$ in tissue), while the low-contrast edges between gray- and white matter were enhanced. Fig. 12 (right) demonstrates the smoothing and edge enhancement capacity of the filtering as a preprocessing step for structural image processing. A simple local differentiation was applied to the original and the filtered images. This edge detection operation serves as a basic segmentation step for a contour-based image analysis.

F. 3-D Two-Channel Filtering of Dual-Echo SE Data

Spin-echo image data with thin slices, acquired with the half-Fourier technique are especially appropriate for the application of the 3-D 2-channel filtering, because both the structural correlation in 3-D space and the correlation among the two echoes support an efficient smoothing and enhancement (Fig. 13). To obtain optimal information about 3-D anatomical structures, the filtering was applied to 120 double echo contiguous slices of a spin echo acquisition (TR = 4000 ms, TE = 30/80 ms) with nearly isotropic voxels of dimension 1.8 mm $* 1$ mm $* 1$ mm (interleaved acquisition, no gap). To speed up the measurement, a half-Fourier sampling was used. The decrease in the overall acquisition time by a

Fig. 11. MR angiogram (TR 50 ms, TE 5 ms, FOV 24.0 cm, slice thickness 1 mm, flip angle 80°): original slice (left), slice after 3-D filtering (right), zoomed parts ($2x$) of original and filtered slices (bottom left and right).

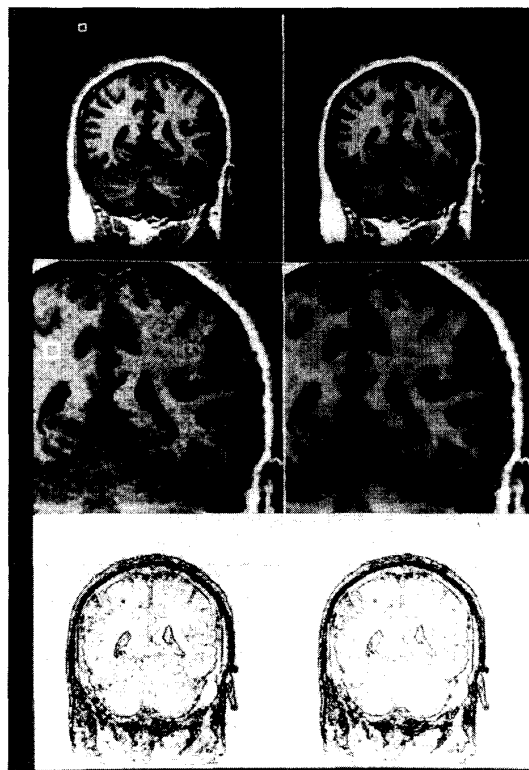

Fig. 12. Three-dimensional SPGR (TR 35 ms, TE 5 ms, FOV 24.0 cm, slice thickness 2 mm, no gap): left: original slice (top), zoomed region ($2x$) (middle), edge detection (bottom); right: 3-D filtering (top), zoomed region (middle), edge detection (bottom).

factor of two is counterbalanced by a decrease of signal to noise by $\sqrt{2}$.

Noise in the most homogeneous tissue region was reduced from $\sigma_n = 3.67$ to $\sigma_n = 0.99$ in the first and from $\sigma_n = 3.63$ to $\sigma_n = 1.05$ in the second echo, using three filter iterations with c_1 and $\mathcal{K} = 7.3$. The noise elimination is

illustrated by generating 2-D histograms (scatterplot) of dual echo pairs. Fig. 13 bottom left and bottom right demonstrate the significant noise reduction indicated by sharpened clusters. This result is important for a later segmentation of dual-echo images by statistical classification methods. Sharp clusters with only minimal overlap allow a clear discrimination of
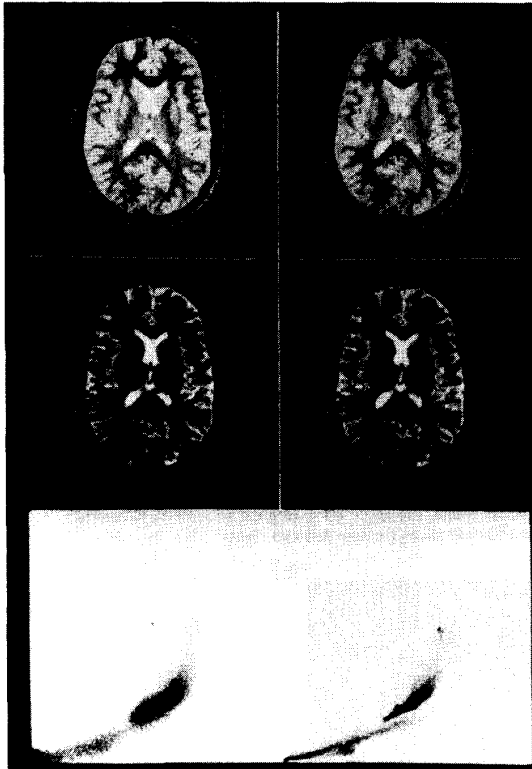
Fig. 13. Three-dimensional filtering of dual-echo SE image data (TR 4000 ms, TE 30/80 ms, FOV 24.0 cm, slice thickness 1.8 mm, nogap): left: original first echo (top), second echo (middle), and scatterplot (bottom); right: filtered first echo (top), filtered second echo (middle) and scatterplot of filtered data (bottom).

tissue categories based on multivariate pixel properties [10], [12].

IV. DISCUSSION

In 1981 Wang and Vagnucci [13] recognized the need for locally adaptive nonlinear image data filtering and proposed an iterative scheme in which the weighting coefficients of a spatial averaging mask were the normalized gradient inverse between the center pixel and its neighbors, without inferring any *a priori* knowledge about image and noise statistics. Based on the observation that intensity differences (gradients) between neighboring pixel values across boundaries are much larger than those within regions, they proposed that the contribution of a neighboring pixel should be weighted inversely proportional to its intensity difference. Their choice of normalization, however, weakens the adaptive smoothing capacity significantly. The weight of the central pixel and therefore the sum of the neighboring weight coefficients, is fixed at 0.5. The inverse linear coupling of filter mask weights with the gradient magnitude does not prevent blurring. After performing a quantitative evaluation of their smoothing scheme, Wang and Vagnucci suggested smoothing an image to no more than five iterations to obtain an acceptable compromise between smoothing and blurring.

Our filtering method reduces noise significantly while preserving important image structures. When it was applied to images with varying MR parameters, analysis clearly showed that the filtering of 0.5 NEX images achieves a noise level which is even lower than that of 4 NEX images, requiring an eight times longer acquisition. Filtered images of smaller NEX are very similar in appearing to unfiltered images with large NEX, especially when comparing region homogeneity and edge sharpness. Differences are mainly due to a lesser detail enhancing capacity of faster acquisitions and only secondarily to a loss by filtering. The sharp and clear images look somewhat unusual, but the similarities of the filtered images to 4NEX acquisitions imply that noise does not necessarily have to be an unavoidable image component in MRI.

The optimal filter parameters defining the diffusion function and the number of iterations were selected by visually comparing different results. The evaluation of the results is further supported by the generation of difference images, line profiles, edge images and scatterplots. The main parameter $\mathcal{K}$ depends on the noise and the strength of step edges to be kept. Nordström [5] calls $\mathcal{K}$ an edge enhancement threshold because $\mathcal{K}$ is directly related to the gradient magnitude to be kept or smoothed. The mathematical treatment of the edge enhancement has shown that points with gradient magnitudes close to $\mathcal{K}$ are either blurred or enhanced. In medical imaging, especially in MRI, low contrast edges are often significant. The selection of the parameter $\mathcal{K}$ is mainly based on the noise level and only secondly on the edge strength. We prefer to keep $\mathcal{K}$ as small as possible while sufficiently suppressing noise. If a reliable estimate of the noise can be made, e.g., by applying the estimation technique proposed in this paper, the parameter $\mathcal{K}$ can be selected directly. A good choice was $1.5 * \sigma_{\text{nose}} < \mathcal{K} < 2.0 * \sigma_{\text{noise}}$. If noise level is not known, a test series with different parameters must be generated and compared. Once selected, the parameter can be kept fixed for acquisitions taken under similar conditions. Several tests with diffusion functions c_1 and c_2 indicate that c_1 possesses a much stronger edge enhancing capacity, while noise smoothing is comparable.

A second parameter that must be controlled is the number of iterations. The iterated filtering process results in a sequence of diffused image functions, from which an optimal has to be selected. We prespecified the number of filter operations, as impressive improvements are already obtained after three to five iterations. The "biased anisotropic filtering" proposed by Nordström converges to a steady-state solution, requiring no stopping criteria. Our experiments have shown that with that method, smoothing efficiency is comparable, while complete convergence is obtained after approximately 50 iterations (integer calculations).

The 2-D version of the filter is appropriate for processing single slices or series of slices with anisotropic voxel dimensions. The influence of neighboring nodes falls off with $1/\Delta x^2$ where Δx is the distance between voxel centers. When processing series of slices, the decision whether to use 2-D or 3-D filtering is based on the voxel dimensions, the desired accuracy and computation time. 2-D filtering with its 3*3 mask is faster than 3-D filtering with $3*3*3$ neighbors. In practical

applications we found it reasonable to apply 2-D filtering to voxels with a noncubical edge ratio of 1 : 1 : 3 and more.

Three-dimensional filtering is useful with data sets of cubical or nearly cubical voxel dimensions. The larger neighborhood results in a better noise reduction and in a maintenance of 3-D edge and line structures. Whenever several channels are acquired a multichannel version of the filter should be chosen. Assuming independent noise among the channels the choice between noise and image structures becomes more robust.

The interactive selection of the parameter $\mathcal{K}$ is still a disadvantage of the procedure. Future improvement should include an automated selection of the optimal edge enhancement parameter. We plan to integrate the automated noise-estimation into the diffusion procedure.

The analogy of anisotropic diffusion to iterated processing on a network structure implies a simple and efficient implementation on many types of computers. A 2-D iteration on 16 bit 256 by 256 slices requires 6.2 s, whereas for 3-D iteration 11.9 s/slice are needed (SUN SPARCstation IPC, 24 mbytes memory). This performance is achieved by using a look-up-table technique which precomputes corresponding gradient magnitude—flow values. The algorithm reduces to the calculation of differences of the neighboring node intensities, to a table look up, and to a summation of the associated flow contributions.

V. CONCLUSION

Recent improvements in MRI techniques make it accessible for many new applications, but these advances do not always result in images of higher quality. High spatial resolution and high speed acquisitions can lead to image data of relatively poor quality. Methods usually applied to enhance noisy MR images clearly show that better image quality is often accompanied by a decrease in spatial resolution or by a considerably longer acquisition time.

The present paper indicates that the signal to noise ratio must not necessarily be the delimiting factor in applying special parameters or special pulse sequences in MRI because noisy images can be restored efficiently in a postprocessing stage. The spatial filtering process presented here is based on anisotropic diffusion. New extensions adapt it to the processing of multiple echoe and 3-D data, and allow the optimal filtering of arbitrary types of MR image data. The results clearly illustrate the efficient noise reduction in homogeneous regions. Object contours, boundaries between different tissues and small structures such as vessels are not only preserved, but even enhanced. The filtered images appear clearer and boundaries are much better defined, leading to an improved differentiation of adjacent regions of similar intensity characteristics.

The filtering method described herein fulfills most of the criteria defined for improvement in SNR. Besides better smoothing and enhanced visualization characteristics, the potential advantages of the filtering in clinical routine applications include the following.

- We demonstrate with the enhancement of filtered half-Fourier images that the method could support an acquisition technique which might not otherwise be the first choice because of inadequate image quality. The reduced image quality of half-Fourier acquisition in comparison to 1 NEX acquisition is compensated for by the filtering, thus supporting the usage of that time-saving acquisition technique in combination with noise-filtering in future medical routine applications. The relevance of faster acquisitions in clinical routine is significant, because it reduces artefacts due to patient movement, allows an improved MR instrument throughput, and provides benefits to patients in terms of comfort.
- Filtering applied as a postprocessing procedure does not affect the acquisition process. The processing can be carried out on existing data for retrospective image enhancement, or image data can be processed at the time of the study.
- Image processing methods specifically adapted to the multidimensional and multispectral MR data can significantly improve the analysis by extracting, analyzing, and visualizing the structural and functional properties of tissues. The anisotropic diffusion process proposed here is an important prerequisite of a segmentation process [10] to extract brain (gray and white matter), and the CSF spaces from MR data with only minimal user interaction. This segmentation procedure was applied to image data depicting various diseases of the brain (Alzheimer's disease, brain atrophy, normal pressure hydrocephalus, multiple sclerosis) in which quantitative morphometric and volumetric information, as well as spatial configuration or distribution of normal structures and/or lesions are essential [11]. The experience with filtering as a preprocessing step allowed to improve the spatial resolution (i.e., reduce slice thickness to 3 mm) while maintaining sufficient signal to noise for visual evaluation as well as segmentation.

We conclude that the filtering proposed here combines a highly efficient noise reduction and the ability to preserve and even enhance important image structures. The medical images discussed here fit the image model of piecewise slowly varying structures, supporting the assumptions made for applying the anisotropic diffusion method.

Further studies will have to be undertaken to find the lower limits of input image quality to given satisfactory filtered results. Such an evaluation will be interesting in low- and mid-field MR imaging, where the problem of sufficient signal to noise is much more pronounced.

The filtering of MR data as proposed here is part of the analysis protocol in a large NIH-study (NIH-NINDS-90-03) for investigation of MRI in multiple sclerosis (MS). The study includes over 1200 MR examinations acquired at the Brigham and Women's hospital in Boston and comprises a careful validation and reliability analysis. After the two year study period, the large data-base will allow to conclude about the clinical utility and sufficiency of the proposed filtering technique, both for visual reporting and for computerized quantitative analysis.

APPENDIX
CALCULATION OF INTEGRATION CONSTANT Δt

The integration of partial differential equations is a problem of numerical mathematics that has been dealt with extensively (e.g., [4]). A number of methods of different degrees of complexity are available. We present here a simple derivation to facilitate intuitive appreciations.

The integration constant Δt determines the iterative approximation of stability. There is no limitation for the lower bound of Δt. A small value results in a good approximation of the continuous case, but requires many iteration steps. The upper bound can be calculated using the discrete formulation of the diffusion process.

$$I(t + \Delta t) \approx I(t) + \Delta t * I_t$$
$$= I_0 + \Delta t * \sum_{i=1}^{n} c_i (I_i - I_0)$$
$$= I_0 \left(1 - \Delta t \sum_{i=1}^{n} c_i \right) + \Delta t \sum_{i=1}^{n} c_i I_i. \quad (A1)$$

To achieve a monotonic variation of the node intensities the central weight must be larger than or equal to the weights of the neighboring elements. The case of maximum blurring is considered. Setting $\mathcal{K}$ to infinity results in diffusion coefficients equal to one, independent of the intensity differences.

$$\left(1 - \Delta t \sum c_i \right) \geq \Delta t c_k | k \in \{1, \cdots, n\}$$
$$\Delta t \left(c_k + \sum c_i \right) \leq 1$$
$$\Delta t \leq \frac{1}{1 + n} | c_i = 1, i \in \{1, \cdots, n\} \quad (A2)$$

The parameter n equals the number of neighboring nodes between which a flow takes place. For diagonal elements the larger distance must be taken into account (factor $1/\Delta x^2$). Contributions from diagonal elements within planes reduce to 0.5, whereas corners of cubical neighborhoods amount to 0.33. Table II lists the integration constants for different kinds of neighborhoods. Stability analysis according to [4] yields virtually identical maximum time steps.

ACKNOWLEDGMENT

TABLE II
INTEGRATION CONSTANTS Δt FOR DIFFERENT NEIGHBORHOOD STRUCTURES

Dimension	Connectedness	$1 + n$	maximum Δt
1-D	2	3	1/3
2-D	4	5	1/5
	8	7	1/7
3-D	6	7	1/7
	26	14 2/3	3/44

REFERENCES

[1] R. T. Chin and C. L. Yeh, "Quantitative evaluation of some edge-preserving noise-smoothing techniques," in *Comput. Graph. Image Processing CGIP*, vol. 23, 1983, pp. 67–91.

[2] P. Perona and J. Malik, "Scale space and edge detection using anisotropic diffusion," in *Proc. IEEE Workshop Comput. Vision*, Miami, FL, Nov. 1987, pp. 16–22.

[3] ——, "Scale-Space and Edge Detection using Anisotropic Diffusion," *IEEE Pattern Anal. Machine Intell.*, vol. 12, pp. 629–639, July 1990.

[4] G. Sewell, *The Numerical Solution of Ordinary and Partial Differential Equations*. New York: Academic, 1988, pp. 67.

[5] N. Nordström, "Biased anisotropic diffusion—A unified regularization and diffusion appraoch to edge detection," *Image Vision Comput.*, vol. 8, no. 4, pp. 318–327, 1990.

[6] Ph. Saint-Marc, J.-S. Chen, and G. Medioni, "Adaptive smoothing: A general tool for early vision," *IEEE Pattern Anal. Machine Intell.*, vol. 13, pp. 514–525, June 1991.

[7] L. Kaufman, D. M. Kramer, L. E. Crooks, and D. A. Ortendahl, "Measuring signal-to-noise ratios in MR-imaging," *Radiology*, vol. 173, pp. 265–267, 1989.

[8] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, pp. 679–698, 1986.

[9] D. L. Parker, Ch. Yuan, and D. D. Blatter, "MR angiography by multiple thin slab 3D acquisition," *Magnet. Reson. Med.*, vol. 17, pp. 434–451, 1991.

[10] G. Gerig and R. Kikinis, "Segmentation of 3D magnetic resonance data," in *Prog. Image Anal. Processing*, Cantoni, Cordella, Levialdi, Sanniti di Baja, Eds., World Scientific, Singapore, *Proc. 5th Int. Conf. Image Anal. Processing*, 1990, pp. 602–609.

[11] R. Kikinis, F. Jolesz, G. Gerig, T. Sandor, H. Cline, W. Lorensen, M. Halle, and S. Benton, "3D morphometric and morphologic information derived from clinical brain MR images," in *3D Imaging in Medicine*, K. H. Höhne, H. Fuchs, and S. M. Pizer, Eds., NATO ASI Series, Serie F: Computer and Systems Science, vol. 60, Springer-Verlag, 1990, pp. 441–454.

[12] O. Kuebler and G. Gerig, "Segmentation and analysis of multidimensional data-sets in medicine," in *3-D Imaging in Medicine*, K. H. Höhne, H. Fuchs, and S. M. Pizer, Eds., NATO ASI Series, Serie F: Computer and Systems Science, vol. 60, Springer-Verlag, Travemünde, FRG, June 1990, pp. 63–81.

[13] D. C. C. Wang, A. H. Vagnucci, and C. C. Li, "Gradient inverse weighted smoothing scheme and the evaluation of its performance," *Comput. Graph. Image Processing CGIP*, vol. 15, 1981, pp. 167–181.