Multigrid Methods for Anisotropic Diffusion

Simon Long

Intelligent, Real-time Imaging and Sensing Group
School of Information Technology and Electrical Engineering
The University of Queensland
E-mail: simonl@itee.uq.edu.au

Abstract

Multigrid methods provide a means with which to accelerate the solution of many problems derived from multi-dimensional linear and nonlinear Partial Differential Equations. A multigrid approach is applied to anisotropic diffusion, a process that is useful for image smoothing and edge strengthening. It is demonstrated to improve the response of the diffusion process by smoothing stubborn low-frequency artefacts. Where traditional relaxation approaches are used to solve large systems of equations on high-resolution images, multigrid methods sustain superior rates of convergence to arbitrary precision and provide a computational complexity that is linear in the number of pixels of the image.

1. Introduction

Especially in the context of medical imaging, data is recorded from increasingly high-resolution sources in multiple dimensions. This expansion poses several problems for existing image processing techniques, relating to the scalability of the algorithms designed to process this data.

Anisotropic diffusion techniques were originally used for the generation of scale spaces by Perona and Malik [6] and were quickly characterised by their edge strengthening and image simplification properties [4]. These characteristics make them useful in preprocessing stages for many medical segmentation and edge detection problems. In general, anisotropic diffusion filters been shown to estimate a piecewise smooth image from a noisy one [2].

Diffusion in image processing acts much like the physical process of diffusion, causing dispersion of intensity at each point while conserving the average grey level of the image. The process acts iteratively in an explicit discretisation of the continuous Partial Differential Equation (PDE), relaxing on the estimate at each step to generate a successively smoother and simpler image. The number of iterations to be performed may form a parameter of the system, or the PDEs may contain a reaction term [8] to prevent a trivial solution. The latter diffusion-reaction allows a more finely tunable process with an analytical solution for a given image independant of the initial estimate. Like most pure relaxation methods, it is slow to resolve low-frequency artefacts and the rate of convergence decreases sharply with image size.

Different discretisations of the same equations have yielded implicit schemes that are significantly faster than the original explicit scheme. The Additive Operator Splitting (AOS) scheme is an order of magnitude faster than the explicit formulation [7] but also suffers with image size.

Further adaptation of AOS has embedded the process in a pyramid framework [8]. This acts as a simple multi-resolution approach to mitigate low-frequency artefacts and tends to increase greatly the speed of AOS. However, amending the AOS scheme in this manner is only weakly justified and a more stringent theory is desirable.

Multi-resolution schemes in general use the efficiency of an algorithm acting on a small image, by exploiting the similarities between the solutions of the process on a fine grid and a coarse grid. Multigrid approaches fall into this category, and can solve a relaxation process on a linear system of equations in optimal time complexity [3]. That is, to reach a solution of desired precision, the cost of a multigrid approach is linear in the number of pixels in the image.

The basic operation of a multigrid scheme involves the transfer of images between fine grids containing many pixels and coarse grids with fewer pixels. On coarser grids the solution error is improved at lower frequencies, while on the finer grids the solution error is improved at the higher frequencies. When applied to anisotropic diffusion [1] multigrid allows effective reduction of low-frequency artefacts at a similar rate to high-frequency artefacts, without losing the properties of edge strengthening and region smoothing.

Multigrid methods are suited to improving iterative processes on multi-dimensional data, especially where the solution may be arbitrarily precise. Since their initial development for solving naturally occurring PDEs, multigrid has seen extensions to incorporate nonlinear problems and the algebraic abstraction to problems on irregularly shaped networks, demonstrating the versatility of the methods to solving many varied forms of problems.

2. Anisotropic Diffusion

As introduced by Perona and Malik [6], anisotropic diffusion in image processing is a discretisation of the family of continuous partial differential equations that include both the physical processes of diffusion and the Laplacian.

$$\frac{\partial u}{\partial t} = \nabla \cdot (\boldsymbol{c} \nabla u) \tag{1}$$

The continuous equation in (1) describes diffusion in general on a continuous image u, where the precise nature of $\boldsymbol{c}$ determines which of the distinct kinds is to occur. Anisotropic diffusion is denoted by a tensor-valued $\boldsymbol{c}$ that prevents flow across areas of high discontinuity, restricting diffusion from smoothing across discernible object boundaries. In the explicit discretisation employed by Weickert [7], the effect of c operating on u can be expressed in the following form

$$u_i^{k+1} = u_i^k + \tau \sum_{l=1}^{m} \sum_{j \in \mathcal{N}_l(i)} \frac{g_j^k + g_i^k}{2h_l^2} (u_j^k - u_i^k) \tag{2}$$

The system in (2) represents a network in which the ith pixel intensity of the kth iteration u_k will flow towards a neighbouring pixel of lower intensity, at a rate weighted by the average of the two corresponding diffusivity coefficients g_i, g_j. Here $\mathcal{N}_l(i)$ denotes the two neighbours of pixel i along axis l. Essentially, this operation is relaxation performed on u on a grid of step h_l along the l axis. The coefficients of g will take values between 0 and 1, where a zero value denotes the presence of an edge in the image. Weickert's formulation of g is based somewhat on that of Catté [4].

$$u^{k+1} = \left(I + \tau \sum_{l=1}^{m} A_l \right) u^k = (I + \tau A)\, u^k \tag{3}$$

If the right hand side of (2) is expressed in matrix form as (3) then each element along the diagonal of A will be negative, and will equal the sum of the remaining (positive) elements in the row. This indicates that the least eigenvalue of A will be of zero value, and under the assumption that the process converges, the greatest eigenvalue of $(I - \tau A)$ will have a value of 1, characteristic of such PDEs. After many iterations, the second eigenvalue will determine the rate of convergence; its value will tend to increase nonlinearly towards one as the size of the image increases. This would cause, for instance, more than twice the computation for an image with twice as many pixels.

When the stopping time is to be only several iterations it is clear that certain components of the error being corrected by the diffusion process will respond much more quickly than others, leading to the presence of larger, spurious artefacts within the image. Adding a backwards reaction to (3) can mitigate this problem by providing a non-trivial solution to the system of equations that can be solved entirely.

$$u^{k+1} = (I + \tau A)\, u^k + \beta\big(w - u^k\big) \tag{4}$$

The final term of (4) ensures that the diffusion process does not drift too far from the original image, w.

The AOS method introduced by Weickert uses a different discretisation of (1), wherein the matrix representation of the relaxation process is given by the implicit formulation

$$u^{k+1} = \frac{1}{m} \sum_{l=1}^{m} (I - m\tau A_l)^{-1}\, u^k \tag{5}$$

This yields stability in convergence for all positive time-step τ, while the explicit method (3) is restricted in τ. With increased values of τ, Weickert [7] demonstrated a ten-fold speed gain when compared to the explicit formulation. However the cost of every improved bit of precision will still decrease dramatically as image size increases, making it unsuitable for applications of increased precision.

3. Applying Multigrid Methods

Although the original multigrid method was first applied to solve problems involving linear operators in naturally occurring systems of PDEs, it has since been developed to handle nonlinear systems of equations, such as the class described above for anisotropic diffusion. Several methods exist to apply multigrid approaches to nonlinear systems, such as the Full Approximation Storage method [5]. Other approaches assume linearity over small time-steps.

When performing anisotropic diffusion, let v be the solution to the system of diffusion equations

$$A \cdot v = f \tag{6}$$

In the case of (3), f is zero, while A is an operator containing the diffusivity coefficients generated for v. It is convenient to describe an estimate u in terms of the solution v less an error, $v - e$. Relaxation upon the estimate reduces this error until the stable solution is reached.

Denoted by Ω_h is the m-dimensional grid of step size h on which this image is sampled. A coarser grid Ω_{2h} can be defined by doubling the sampling period along each dimension. Multigrid also names the *restriction operator* $(\cdot)_\downarrow$ to transfer an image from Ω_h to Ω_{2h} and the *prolongation operator* $(\cdot)_\uparrow$to transfer an image from Ω_{2h} to Ω_h. The Galerkin condition specifies that these two (linear) inter-grid transfer operators should be transposes of each other by a factor of 2^m.

The operation of A on Ω_{2h} is in fact a reformulation of the original PDEs on the coarser grid. Equation (7) illustrates how A operates on a coarser grid.

$$A \cdot u_{2h} = (A \cdot (u_{2h})_\uparrow)_\downarrow \tag{7}$$

The residual r_h of a solution estimate u_h on a grid Ω_h is defined as

$$r_h = f - A \cdot u_h \tag{8}$$

For a relaxation scheme on a grid Ω_h, multigrid proposes a similar relaxation scheme for a system of equations (9) on a coarser grid Ω_{2h} and equates the residuals of the two (11). Most importantly, the solution to the fine grid problem is exactly a solution to the coarse grid problem – once reached in the fine grid, further relaxation in the coarse grid will cause no change. The relaxation on the coarser grid effectively solves a portion of the error in the fine grid problem.

$$A \cdot u_{2h} = f_{2h} \tag{9}$$
$$r_{2h} = f_{2h} - A \cdot u_{2h} \tag{10}$$
$$f_{2h} - A \cdot (u_h)_\downarrow = (f_h - A \cdot u_h)_\downarrow \tag{11}$$
$$\therefore f_{2h} = (r_h)_\downarrow + A \cdot (u_h)_\downarrow \tag{12}$$

Equations (9)-(12) yield the terms of (7) necessary to solve as completely as possible for e_h in Ω_{2h}, for some estimate u_h of the solution in the fine scale. Once some satisfactory value of v_{2h} has been obtained, the coarse grid error e_{2h} for u_h is

$$e_{2h} = (v_h)_\downarrow - (u_h)_\downarrow = v_{2h} - (u_h)_\downarrow \tag{13}$$
$$u_h \leftarrow u_h + (e_{2h})_\uparrow \tag{14}$$

The coarse grid error is then used as in (14) to correct the fine grid estimate.

Solving for u_{2h} could be done by performing relaxation on (9). Note however that (9) is a set of equations essentially the same as (6) – the true elegance of multigrid is that coarse grid correction can be performed hierarchically, minimising the total relaxation performed on still coarser grids. The *multigrid v-cycle* is one such recursive structure, its coarse grid correction consisting of brief relaxation before and after a still coarser grid correction.

4. Results

In Section 3, the multigrid framework was presented with the aim of applying it to anisotropic diffusion. The nonlinear approach selected assumes linearity during individual iterations. The inter-grid transfer operators used are the traditional upsample and nearest-neighbour interpolation operator and its transpose, the nearest-neighbour blur and downsample operator [3].

Given the recursive method with which multigrid augments relaxation at many grid resolutions, each full multigrid iteration will naturally have a computational cost higher than a single iteration of relaxation on the original image. If a single relaxation operation on the finest grid of m dimensions is taken as a unit cost C_{relax} of computation and the cost of performing a single relaxation iteration scales linearly with the number of pixels in the image, then the relative cost $C_{v-cycle}$ of a multigrid v-cycle iteration is approximately

$$C_{\text{v-cycle}} = 2C_{\text{relax}} \cdot \left(1 + 2^{-m} + 2^{-2m} + \dots\right) \tag{15}$$
$$C_{\text{v-cycle}} = \frac{2C_{\text{relax}}}{1 - 2^{-m}} \tag{16}$$

Compared to the relaxation iteration, each multigrid iteration is significantly more efficient. Figure 1 compares the two when operating to solve a linearised diffusion-reaction problem to convergence as in (4). The progress of each is measured by comparing the norm of the error $\|e\|_2$ in the solution estimate to the equivalent cost in relaxation operations performed. The size of the logarithm of this norm yields the number of digits to which the error is minimised.

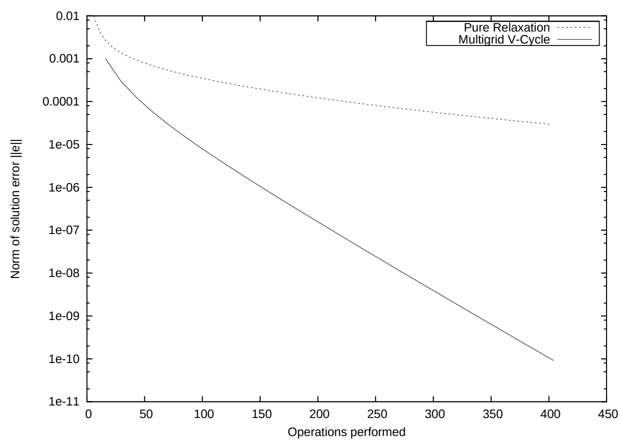


Figure 1. Comparison of cost of convergence of multigrid and pure relaxation anisotropic diffusion processes.

Well and truly before the relaxation operation reaches its range of linear convergence, it is clear that the multigrid method has reached a constant number of converged digits per iteration, and that the rate of convergence for the multigrid approach is much faster.

Figure 2 indicates the change in the cost of computation caused by image size. The major advantage of multigrid methods is that the cost of convergence is linearly related to the number of pixels in the image. In this figure is plotted the progress of multigrid anisotropic diffusion acting on an image and on a half-size representation of the same image (one quarter of the total number of pixels). The cost of converging further digits for the small image is one quarter that for the large image.

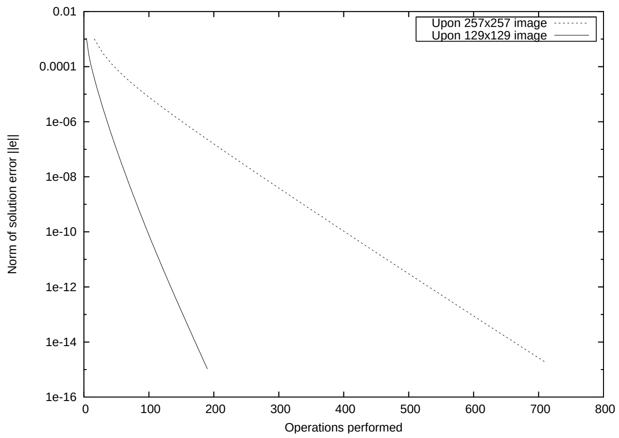
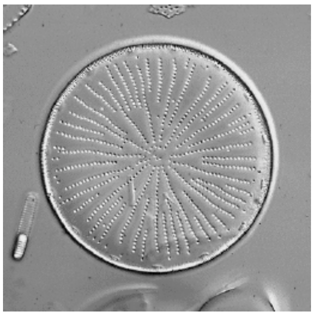


Figure 2. Comparison of cost of convergence of multigrid anisotropic diffusion on an image of two different sizes.

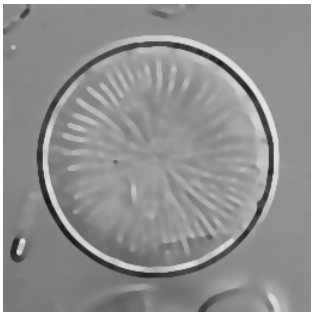
Figure 3 illustrates the efficiency of multigrid at removing many levels of noise from a solution. The image of the diatom was processed with anisotropic diffusion, both by performing three iterations of multigrid, and by using the eight iterations purely of relaxation that comprise the same computational cost by (16). The difference in remaining detail between Figures 3(b) and 3(c) is mostly comprised of larger patches of discolouration that pure relaxation failed to diffuse. Simplified in this manner, Figure 3(b) could be easily processed by a simple segmentation algorithm for deriving the boundary of the diatom.

As a denoising tool, anisotropic diffusion is generally considered quite effective. Figure 4(a) presents a clean image of a lung, corrupted in Figure 4(b) by independent and identically distributed additive Gaussian noise of standard deviation $\sigma = 0.01$. The noisy image was anisotropically diffused using relaxation to give Figure 4(c). As the noise was diffused, it formed irregularities in the image that the



(a) Original image

(b) Diffused with multigrid



(c) Diffused with relaxation

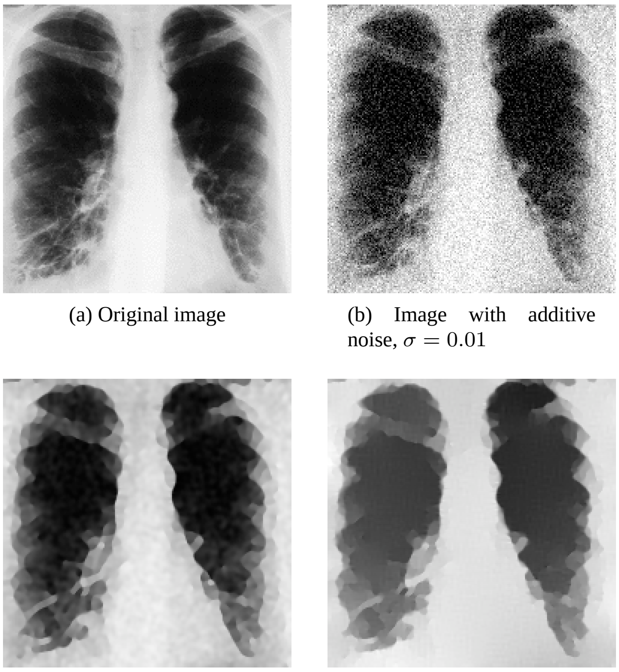
Figure 3. A comparison of anisotropic diffusion techniques acting on (a) a microscope image of Actinocyclus Actinochilus, a diatom using (b) multigrid, and (c) pure relaxation.

relaxation method was less effective at smoothing. Figure 4(d) shows that the multigrid method produced fewer spurious edges and blocks.

5. Conclusion

Multigrid methods are a means to accelerate linear and nonlinear relaxation problems derived from PDEs. They provide convergence to within a given precision that is linear in the number of pixels in an image, and can be applied to systems of equations of any number of dimensions. The hierarchical operation on an estimated solution allows for the correction of many scales of error at once, where traditional relaxation methods would perform poorly.

Anisotropic diffusion is useful as a preprocessing stage to higher levels of image processing. It smooths image interiors to accentuate boundaries for segmentation; it removes spurious detail to improve the response of edge detection algorithms; it also proves effective at removing noise from images. However, relaxation processes that implement anisotropic diffusion tend towards leaving low-frequency artefacts that are difficult to dissipate without over-processing the image.



(a) Original image

(b) Image with additive noise, $\sigma = 0.01$

(c) Denoised with relaxation

(d) Denoised with multigrid

Figure 4. Comparing diffusion methods for denoising; (a) the unaltered image; (b) with additive Gaussian noise; (c) denoising using relaxation; (d) denoising using multigrid.

Combining anisotropic diffusion with multigrid methods greatly diminishes the artefacts introduced, improving the response of the processing while reducing the computational cost. Multigrid methods can be broadly applied to many other PDEs for similarly excellent improvements in computational efficiency.

Acknowledgements

The image of the diatom in Figure 3(a) is from the ADIAC public data web page:
`http://www.ualg.pt/adiac/pubdat/pubdat.html`

References

[1] S. T. Acton. Multigrid anisotropic diffusion. In *IEEE Transactions of Image Processing.*, volume 7, pages 280–291, 1998.

[2] M. J. Black, G. Sapiro, D. H. Marimont, and D. Heeger. Robust anisotropic diffusion: Connections between robust statistics, line processing, and anisotropic diffusion. In *Proceedings of the First International Conference on Scale-Space Theory in Computer Vision*, pages 323–326. Springer-Verlag, 1997.

[3] W. L. Briggs. *A Multigrid Tutorial*. Society for Industrial and Applied Mathematics, Philadelphia, 1987.

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

[5] W. Hackbush and U. Trottenberg, editors. *Multgrid Methods*. Springer-Verlag, New York, 1982.

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

[7] J. Weickert. Recursive separable schemes for nonlinear diffusion filters. In *SCALE-SPACE '97: Proceedings of the First International Conference on Scale-Space Theory in Computer Vision*, pages 260–271. Springer-Verlag, 1997.

[8] J. Weickert. Efficient image segmentation using partial differential equations and morphology. Technical Report DIKI-TR-98/10, University of Copenhagen, Copenhagen, May 1998.