

# On Implicit Image Derivatives and Their Applications

Alexander Belyaev  
<http://www.hw.ac.uk/~belyaev>

School of Engineering & Physical Sciences  
 Heriot-Watt University, Edinburgh, UK

## Abstract

The paper demonstrates advantages of using implicit finite differences for fast, accurate, and reliable differentiating and filtering of multidimensional signals defined on regular grids. In particular, applications to image enhancement and edge detection problems are considered. The theoretical contribution of the paper is threefold. The first adapts the Fourier-Padé-Galerkin approximations approach for constructing compact implicit finite difference schemes with desirable spectral resolution properties. The second establishes a link between implicit and explicit finite differences used for gradient direction estimation. Finally, the third one consists of introducing new implicit finite difference schemes with good spectral resolution properties.

## 1 Introduction

The main goal of this paper is demonstrating advantages of using compact implicit finite difference schemes for basic image processing applications.

**Previous work on image derivative estimation.** Fast and reliable estimation of image derivatives is among the most fundamental tasks of low level image processing. Unfortunately the image processing literature often provides the reader with a set of recipes given without appropriate mathematical analysis. For example, for image gradient estimation, the image processing textbooks usually refer to the Prewitt and Sobel operators while mentioning that these operators are not so good in preserving the rotation equivariance property of the gradient [1], [5], Chapter 15].

Typically, for a 2D image defined on a regular grid with spacing  $h$ , image processing textbooks recommend to use a  $3 \times 3$  kernel

$$D_x = \frac{1}{2h(w+2)} \begin{bmatrix} -1 & 0 & 1 \\ -w & 0 & w \\ -1 & 0 & 1 \end{bmatrix} \equiv \frac{1}{2h} [-1 \ 0 \ 1] \cdot \frac{1}{w+2} \begin{bmatrix} 1 \\ w \\ 1 \end{bmatrix} \quad (1)$$

and its  $\pi/2$ -rotated counterpart  $D_y$  for estimating the  $x$ -derivative and  $y$ -derivative, respectively. Here  $w$  is a parameter: setting  $w = 1$  in (1) yields the Prewitt mask and  $w = 2$  corresponds to the Sobel mask. Searching for an optimal value of  $w$  in (1) remains to be an active research area [2, 3, 4, 13, 16, 20, 33] (see also references therein).

More than sixty years ago W. G. Bickley, a British applied mathematician, noted that

$$D_x \Big|_{w=4} = \left( 1 + \frac{h^2}{12} \Delta \right) \frac{\partial}{\partial x} + O(h^4), \quad \text{as } h \rightarrow 0, \quad (2)$$

where  $\Delta$  is the Laplacian, and, therefore, (1) with  $w = 4$  has optimal rotation-invariant properties for small grid spacing  $h$ . After its introduction in [6], the Bickley kernel was rediscovered at least twice [23, 29], was even patented [36] but nevertheless remains little known to the graphics and imaging communities.

Recently H. Scharr and co-authors [20, 33] suggested to use (1) with  $w = 10/3$ . Their analysis is based on a frequency-based optimization of rotation invariance properties of (1). The Scharr mask works very well in practice [57] and quickly gains popularity among computer vision researchers and practitioners [8, Chapter 6].

Bigger stencils for estimating image derivatives were also widely considered. In particular, E. P. Simoncelli and H. Farid proposed a general frequency-based framework for the design of discrete multidimensional differentiators [14, 32]. Sophisticated gradient estimation filters were studied in [10, 9, 19] in connection with high-quality volume visualization isosurface rendering problems.

**Previous work on compact implicit finite differences.** Although implicit finite differences had become known to a general audience of numerical mathematicians and computational physicists after Collatz' book [9], their heyday began after Lele's seminal paper [24] where a remarkable performance of implicit finite differences for computational problems with a range of spatial scales was analyzed and demonstrated. At present compact implicit finite difference schemes constitute advanced but standard tools for accurate numerical simulations of physical problems involving linear and non-linear wave propagation phenomena [11], [30, Section 5.8], [25, Chapter 5].

In the signal processing language, compact implicit finite differences and more general implicit filtering schemes are described by infinite impulse response (IIR) filters. In image processing applications, finite impulse response (FIR) filters continue to overwhelmingly dominate over IIR image filtering schemes [26, 27] in spite of existence of highly computationally efficient implementations for the latter [12]. A notable exception is given by very recent studies on accurate visualization of volumetric data [1, 19] where sophisticated implicit gradient estimation schemes were developed. The schemes combine FIR derivative filters with generalized interpolation of Blu, Thévenaz and Unser [4, 35] and similar spline-based interpolation methods. It is also worth mentioning works of W. H. Raymond with coworkers [37], in which low- and high-pass implicit filters were considered.

**Paper contribution.** We believe that the potential of implicit finite difference schemes for image processing applications is largely underestimated and consider this work as an attempt to demonstrate usefulness of implicit finite differences for basic image processing tasks. Other contributions of the paper include

- Adapting Fourier-Padé-Galerkin approximations for designing high-quality implicit image differentiation and filtering schemes (the second half of Section 4);
- Establishing a link between implicit and explicit finite differences used for gradient estimation (Section 3);
- Introducing new implicit differencing schemes and evaluating their properties (implicit Scharr and Fourier-Padé-Galerkin schemes).

## 2 Estimating the derivative for univariate signals

Consider a uniformly sampled signal  $f(x)$ . Let us recall that the Nyquist, or folding, frequency is the highest frequency that can be represented in the signal. It equals one-half of the sampling rate.

Consider the simplest central difference operator

$$f'(x) \approx \frac{1}{2h} [f(x+h) - f(x-h)] \quad (3)$$

and its corresponding mask

$$\frac{1}{2h} [-1, 0, 1] \quad (4)$$

defined on a grid with spacing  $h$ . For the sake of simplicity we assume that  $h = 1$ , which is equivalent to rescaling the  $x$ -coordinate:  $x \rightarrow x/h$ . The eigenvalues of the linear operators corresponding to the left and right hand sides of (3) and subject to periodic boundary conditions are found by setting  $f(x) = \exp\{j\omega x\}$ ,  $j = \sqrt{-1}$ ,  $-\pi < \omega < \pi$ , in (3):

$$j\omega e^{j\omega x} \approx \frac{1}{2} \left[ e^{j\omega(x+1)} - e^{j\omega(x-1)} \right] \equiv j \sin \omega e^{j\omega x}, \quad j\omega \approx j \sin \omega.$$

The frequency response function (the eigenvalue for eigenfunction  $\exp\{j\omega x\}$ )  $j \sin \omega$  corresponding to the central difference operator in (3) delivers a satisfactory approximation of the frequency response function  $j\omega$  of the ideal derivative only for sufficiently small frequencies (wavenumbers)  $\omega$  (see, for example, [13, Section 6.4]). One way to improve (3) consists in using implicit finite differences.

**Implicit finite differences for univariate signals.** One can say that (4) adds to the true derivative a certain amount of smoothing applied to non-zero wavenumbers. A natural way to compensate for smoothing introduced by (4) consists of adding approximately the same amount of smoothing to the derivative. This simple idea leads us immediately to the concept of implicit finite differences.

For example, the two simplest implicit schemes for approximating the first-order derivative of a function  $f(x)$  are based on the following relations [9, p. 538]

$$f'(x-h) + 4f'(x) + f'(x+h) = \frac{3}{h} [f(x+h) - f(x-h)] + O(h^4), \quad (5)$$

$$f'(x-h) + 3f'(x) + f'(x+h) = \frac{1}{12h} [f(x+2h) + 28f(x+h) - 28f(x-h) - f(x-2h)] + O(h^6) \quad (6)$$

The corresponding implicit finite differences

$$f'_{i-1} + 4f'_i + f'_{i+1} = 3(f_{i+1} - f_{i-1}), \quad (7)$$

$$f'_{i-1} + 3f'_i + f'_{i+1} = \frac{1}{12} (f_{i+2} + 28f_{i+1} - 28f_{i-1} - f_{i+2}) \quad (8)$$

are often called the 4th- and 6th-order tridiagonal Padé schemes, respectively, since (5) and (6) are derived using classical Padé rational approximations and since (7) and (8) lead to tridiagonal systems of linear equations. It is interesting that (7) can be obtained if the grid data  $\{f_i\}$  is first B-spline interpolated and then processed by the central difference filter (4).

Taking into account a similarity between the Bickley mask (1) with  $w = 4$  and (7) let us call the latter the implicit Bickley scheme. Note that (2) can be rewritten as

$$\frac{\partial}{\partial x} = \left( \delta + \frac{h^2}{12} \Delta \right)^{-1} D_x \Big|_{w=4} + O(h^4)$$

which immediately leads to the implicit Bickley scheme (7).

One can observe that the frequency response function

$$S_w(\omega) = \frac{w + 2 \cos \omega}{w + 2} \quad (9)$$

corresponds to the smoothing kernel

$$\frac{1}{w + 2} [1, w, 1]^T. \quad (10)$$

The frequency response functions corresponding to (7) and (8) are given by

$$j \sin \omega \cdot \frac{1}{S_4(\omega)} \quad \text{and} \quad j \sin \omega \cdot \frac{S_{28}(\omega)}{S_3(\omega)} \quad (11)$$

respectively. It is easy to verify that these two frequency response functions deliver, respectively, fourth- and sixth-order approximations of  $j\omega$ , the frequency response functions function corresponding to the ideal derivative, at  $\omega = 0$ .

Thus the amount of smoothing introduced (4) is compensated in (11) by applying (10) with  $w = 4$  to the derivative. Finite difference scheme (8) delivers a more accurate approximation of the true derivative: smoothing by (10) with  $w = 3$  applied to the derivative is balanced by the amount of smoothing introduced (4) plus smoothing by (10) with  $w = 28$  applied to the sampled function itself.

It is natural to evaluate the quality of a finite difference approximation by its *resolving efficiency*, the range of frequencies (wavenumbers)  $\omega$  over which a satisfactory approximation of the exact differentiation is achieved. Since it is not possible to get a reasonably good approximation when  $\omega$  is close to  $\pi$ , the frequency range for the optimization is often specified by  $0 \leq \omega \leq r\pi$  with some  $0 < r \leq 1$ .

Although a quantitative study of the resolving efficiency of finite difference scheme is a straightforward task [24], everywhere below we make our judgements based on a visual analysis of graphs of the corresponding frequency response functions.

In the left image of Fig. 1, we provide the reader with a visual comparison of the resolving efficiencies of  $j \sin \omega$  corresponding to central difference (3) and frequency response functions (11). In addition, we plot the graph of

$$j \sin \omega \cdot \frac{1}{S_{10/3}(\omega)} \equiv j \sin \omega \cdot \frac{w + 2}{w + 2 \cos \omega} \Big|_{w=10/3} \quad (12)$$

which we call the *implicit Scharr scheme*, since it can be considered as a counterpart of the original Scharr kernel (1) with  $w = 10/3$  introduced in [20, 53]. One can observe that the resolving efficiency of the implicit Scharr scheme is comparable with that of (8) in spite of the fact that the latter has a wider stencil.

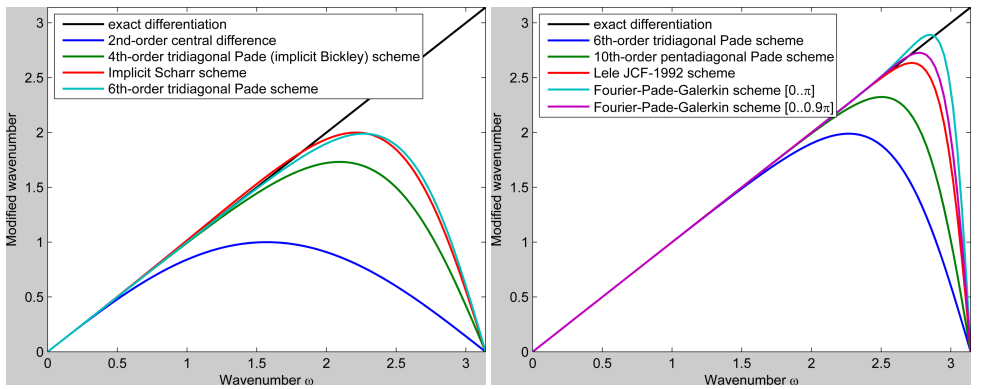


Figure 1: *Left*: Resolving efficiency of various finite difference schemes is demonstrated by plotting the graphs of their corresponding frequency response functions. Note good resolving efficiency of the implicit Scharf scheme. *Right*: A graphical comparison of the resolving efficiency of various frequency response functions corresponding to implicit finite difference schemes. A high resolving efficiency of both the Fourier-Padé-Galerkin schemes is clearly demonstrated.

### 3 Estimating image gradient

Now it becomes clear in which way (1) improves the standard central difference (4): smoothing due to the use of the central difference operator instead of the true  $x$ -derivative is compensated by adding a certain amount of smoothing in the  $y$ -direction. Thus (1) and its  $y$ -direction counterpart do a better job in estimating the gradient direction than in estimating the gradient magnitude.

If the goal is to achieve an accurate estimation of both the gradient direction and magnitude, we can combine (1) and the corresponding  $3 \times 3$  discrete Laplacian

$$L_w = \frac{1}{h^2(w+2)} \begin{bmatrix} 1 & w & 1 \\ w & -4(w+1) & w \\ 1 & w & 1 \end{bmatrix} \quad (13)$$

as follows. Let  $\delta = [0 \ 0 \ 0; 0 \ 1 \ 0; 0 \ 0 \ 0]$  be the  $3 \times 3$  identity kernel. Note that

$$\delta + \frac{h^2}{w+2} L_w \equiv \frac{1}{(w+2)^2} \begin{bmatrix} 1 & w & 1 \\ w & w^2 & w \\ 1 & w & 1 \end{bmatrix} \equiv \frac{1}{w+2} [1 \ w \ 1] \cdot \frac{1}{w+2} \begin{bmatrix} 1 \\ w \\ 1 \end{bmatrix}, \quad (14)$$

which can be considered as simultaneous smoothing (averaging) with respect to both the coordinate directions. Thus, in order to remove smoothing introduced by (1), it is natural to use

$$\left( \delta + \frac{h^2}{w+2} L_w \right)^{-1} D_x \quad (15)$$

which combines (1) with an implicit Laplacian-based sharpening. The frequency response function corresponding to (15) applied to the eigenfunction  $\exp(j(\omega_1 x + \omega_2 y))$  is given by

$$H(\omega_1, \omega_2) = j \sin \omega_1 \cdot \frac{w+2}{w+2 \cos \omega_1} \quad (16)$$

which, in its turn, corresponds to the following implicit finite difference scheme

$$\frac{1}{w+2} (f'_{i-1,j} + w f'_{i,j} + f'_{i+1,j}) = \frac{1}{2h} (f_{i+1,j} - f_{i-1,j}). \quad (17)$$

## 4 High-resolution schemes

Both (7) and (8) are particular cases of a general seven-point stencil introduced in [24]

$$\beta f'_{i-2} + \alpha f'_{i-1} + f'_i + \alpha f'_{i+1} + \beta f'_{i+2} = c \frac{f_{i+3} - f_{i-3}}{6} + b \frac{f_{i+2} - f_{i-2}}{4} + a \frac{f_{i+1} - f_{i-1}}{2}, \quad (18)$$

where  $\{f_i\}$  and  $\{f'_i\}$  are the values of a given function  $f(x)$  and its derivative at the grid points, respectively. Here the coefficients  $\alpha$ ,  $\beta$  and  $a$ ,  $b$ ,  $c$  are determined such that the amount of smoothing introduced by the right-hand side of (18) is compensated by averaging the derivatives in the left-hand side of (18).

The frequency response function corresponding to (18) is given by

$$H(\omega) = \frac{a \sin \omega + (b/2) \sin 2\omega + (c/3) \sin 3\omega}{1 + 2\alpha \cos \omega + 2\beta \cos 2\omega}. \quad (19)$$

The problem of determining the coefficients  $\alpha$ ,  $\beta$  and  $a$ ,  $b$ ,  $c$  in (19) can now be formulated in a more mathematical way: they are chosen such that  $H(\omega)$  delivers a good approximation of  $\omega$ , the frequency response function of the true derivative.

In his seminal paper [24], Lele used empirical considerations to derive the set of coefficients

$$\left. \begin{aligned} \alpha &= 0.5771439, & \beta &= 0.0896406, \\ a &= 1.302566, & b &= 0.99355, & c &= 0.03750245 \end{aligned} \right\} \quad (20)$$

which deliver an exceptional resolving efficiency to compare with many other approximations used currently in computational aeroacoustics [10, Section 4.1.1] (see also references therein). Below we use Fourier-Padé-Galerkin approximations for improving this Lele's result and establishing a general approach to frequency response function design.

**Fourier-Padé-Galerkin approximation for derivative.** Following [28] let us consider the linear space of trigonometric polynomials of degree  $N$

$$\mathcal{F}_N = \text{span}\{e^{jn\omega} \mid -N \leq n \leq N\}$$

and define a rational Fourier series by

$$R_{kl}(\omega) = P_k(\omega) / Q_l(\omega), \quad P_k \in \mathcal{F}_k, \quad Q_l \in \mathcal{F}_l.$$

Given a  $2\pi$ -periodic function  $f(\omega)$ ,  $\omega \in [-\pi, \pi]$ , the Fourier-Padé-Galerkin approximation  $f(\omega) \approx R_{kl}(\omega)$  determines the unknown coefficients in  $R_{kl}$  from orthogonality conditions

$$\int_{-\pi}^{\pi} [P_k(\omega) f(\omega) - Q_l(\omega)] \overline{g(\omega)} W(\omega) d\omega = 0 \quad \forall g(\omega) \in \mathcal{F}_{k+l}, \quad (21)$$

where  $W(\omega)$  is a properly chosen weighting function. Now (21) yields a system of  $k+l$  linear equations with  $k+l$  unknowns.

As a simple application of the above Fourier-Padé-Galerkin approach, let us use (21) with  $W(\omega) \equiv 1$  to determine the coefficients in (18). Since we approximate  $f(\omega) = \omega$  which is odd on  $[-\pi, \pi]$ , we set  $\mathcal{T}_5 = \text{span}\{\sin n\omega \mid 1 \leq n \leq 5\}$ , use

$$P_3(\omega) = a \sin \omega + (b/2) \sin 2\omega + (c/3) \sin 3\omega \quad \text{and} \quad Q_2(\omega) = 1 + 2\alpha \cos \omega + 2\beta \cos 2\omega,$$

and immediately arrive at

$$\alpha = \frac{3}{5}, \quad \beta = \frac{21}{200}, \quad a = \frac{63}{50}, \quad b = \frac{219}{200}, \quad c = \frac{7}{125}. \quad (22)$$

We call (19) with (22) a *Fourier-Padé-Galerkin scheme*. The right image of Fig. 1 demonstrates advantages of the scheme. The scheme has a good resolving efficiency but suffers from a slight Gibbs-type phenomenon. A simple yet efficient way to reduce this Gibbs-type artifact consists of choosing an appropriate weight function  $W(\omega)$  in (21). In particular, setting  $W(\omega) = 1$  for  $0 \leq \omega \leq 0.9\pi$  and  $W(\omega) = 0$  otherwise gives a quite satisfactory result, as shown in the right image of Fig. 1.

Note all the calculations involved are very simple. For this particular choice of the weighting function, the integrals can be evaluated analytically and solving a system of five linear equations with five unknowns is a simple linear algebra exercise (we use Maple for both of these tasks).

The right image of Fig. 1 provides the reader with a visual comparison of the resolving efficiencies of (20) and both the Fourier-Padé-Galerkin schemes. In addition, we also consider the 6th-order tridiagonal Padé scheme whose frequency response function is also shown in Fig. 1 and the tenth-order pentadiagonal Padé scheme

$$\alpha = \frac{1}{2}, \quad \beta = \frac{1}{20}, \quad a = \frac{17}{12}, \quad b = \frac{101}{150}, \quad c = \frac{1}{100} \quad (23)$$

whose frequency response function (19), (23), delivers the maximal approximation order of  $f(\omega) \equiv \omega$  at  $\omega = 0$  among the family of five-diagonal schemes (18).

## 5 Applications

Below we demonstrate advantages of implicit image derivatives and filters for two basic image processing tasks: feature detection and image enhancement.

**Feature detection** The classical Canny edge detection scheme is used as our first test for implicit image derivatives. In Fig. 2 we compare implicit Scharf scheme (12) and Fourier-Padé-Galerkin scheme (19), (22) with the Sobel mask, Scharf kernel, and Farid-Simoncelli 5-tap filter [4]. We use MATLAB functions from [2] with the same parameter settings for all the schemes tested. As expected, the explicit schemes add unnecessary blur and, therefore, are less sensitive to fine image details to compare with the implicit schemes.

Similar results are obtained for the Harris corner detector.

**Inverse diffusion** In our next test, we deal with image deblurring and sharpening. We take an image and add a Gaussian blur to it. Then we start a simple deblurring/sharpening process by the inverse diffusion equation

$$\partial I / \partial t = -\Delta I, \quad I(x, y, t)|_{t=0} = I_0(x, y) \quad (24)$$

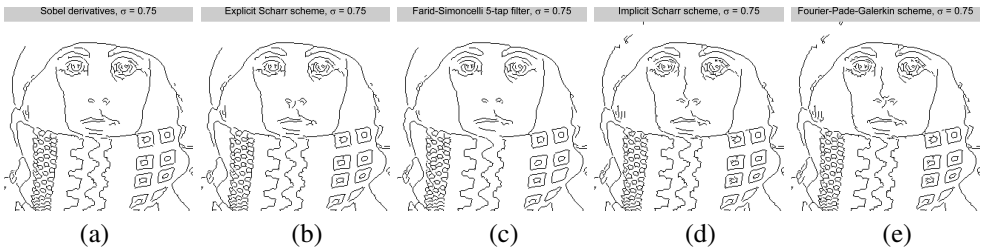


Figure 2: Canny edge detection with explicit and implicit finite differences,  $\sigma = 0.75$ . (a) Sobel mask; (b) explicit Scharr mask; (c) Farid-Simoncelli 5-tap filter; (d) implicit Scharr scheme; (e) Fourier-Padé-Galerkin scheme. Explicit schemes (a), (b), and (c) add unnecessary blur and, therefore, are less sensitive to fine image edges. For example, the edge of the nose is not sufficiently sharp and only appears with implicit schemes (d) and (e).

which we solve numerically by the forward Euler method

$$I(x, y, t + dt) = I(x, y, t) - dt \Delta_h I(x, y, t), \quad (25)$$

where  $\Delta_h$  is a discrete approximation of the Laplacian. Of course, (24) and its discrete counterpart (25) are highly unstable because they lead to an exponential growth of the non-zero image frequencies. Two discrete Laplacians are tested: the standard five-point Laplacian and the Laplacian whose each second-order derivative is estimated via the implicit Bickley schemes applied twice. The left image of Fig. 3 compares the frequency response functions for the standard second-order central difference for estimating the second-order derivative (it is used in the five-point Laplacian) with the twice-applied implicit Bickley scheme. One can see that the latter combines a very accurate approximation for the low frequencies with high quality low-pass filtering. So it is natural to expect that (25) with the second-order derivatives estimated using the twice-applied implicit Bickley scheme demonstrates better deblurring and sharpening properties. Indeed, Fig. 4 demonstrate advantages of using the twice-applied implicit Bickley scheme.

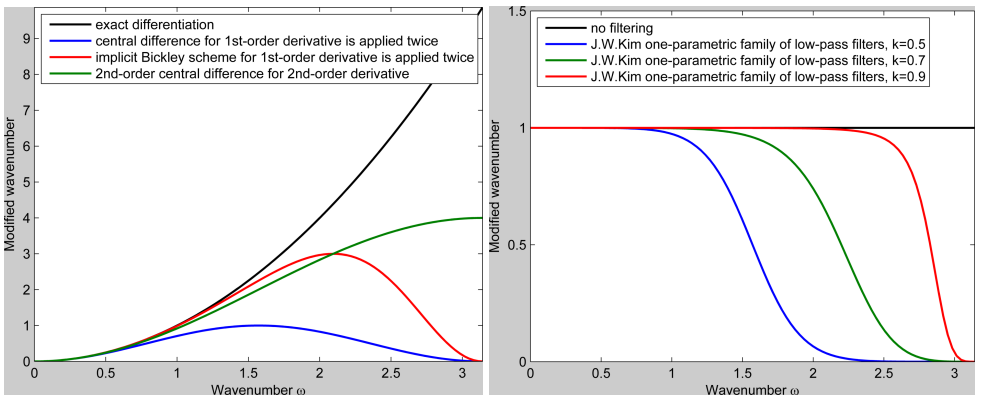


Figure 3: *Left*: Frequency response functions for the exact 2nd-order derivative and its finite-difference approximations. *Right*: The frequency response functions of three filters from the family of low-pass filters proposed in [20]. Each frequency response function from the family has a nice monotonically decreasing profiles with an inflection point at  $\omega = k\pi$  where  $k$  parameterizes the family.

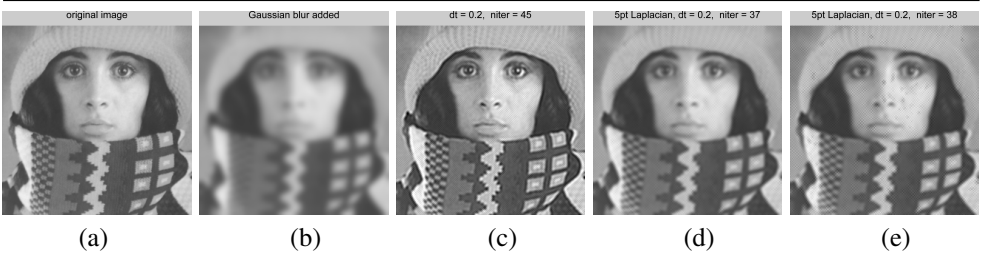


Figure 4: (a) The original Trui image. (b) The image is severely Gaussian blurred. (c) The blurred image is restored and even sharpened by 45 iterations of (25) with the Laplacian approximated via the twice-applied implicit Bickley schemes applied to each dimension (use zoom to see how well small-scale image details are restored). (d) An attempt to restore the blurred image by (25) with the standard five-point discrete Laplacian; small image defects are already observed after 37 iterations of (25) while the blur is not eliminated. (e) After one more iteration of (25): the blur is not reduced and the defects become much more apparent.

**Unsharp masking** Our last example is also devoted to image sharpening. This time we use the Laplacian subtraction filter [5] which consists of one iteration (25) and can be considered as a variant of unsharp masking. As we know, the Laplacian subtraction filter amplifies high frequencies and, therefore, may oversharpen small-scale image texture. In order to reduce such oversharpening one can combine the Laplacian subtraction filter with a low-pass filter. We already did it for deblurring when a special discrete Laplacian was employed. This time, however, stronger low-pass filtering is required. We could design a good low-pass filter using the Fourier-Padé-Galerkin approach with  $f(\omega) \equiv 1$  and a properly chosen weighting function  $W(\omega)$  in (21). Instead we have decided to use a one parametric family of low-pass filters proposed very recently by J. W. Kim [24]. The right image of Fig. 3 displays the frequency response functions of three filters from the family. Each frequency response function from the family has a nice monotonically decreasing profiles with an inflection point at  $\omega = k\pi$  where  $k$  parameterizes the family. In Fig. 5, we demonstrate advantages of our strategy to combine the Laplacian subtraction filter with low-pass filtering.

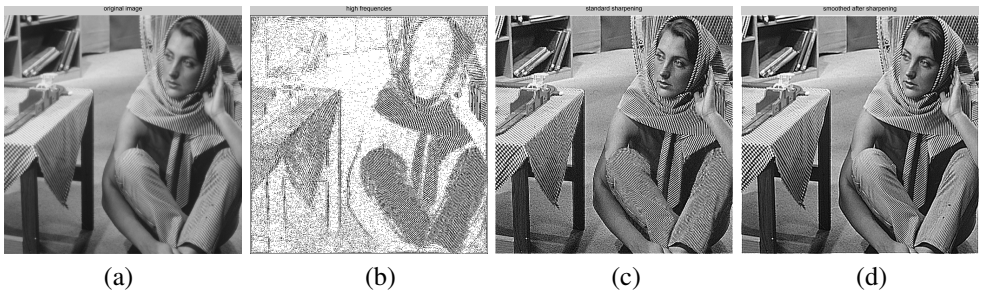


Figure 5: (a) Original Barbara image. (b) High frequency image details extracted using the  $k = 0.5$  (half the Nyquist frequency) low-pass filter from the right image of Fig. 3. (c) The original image is enhanced by the Laplacian subtraction filter (unsharp masking), high-frequency texture details are severely oversharpened. (d) The sharpened image is smoothed by the  $k = 0.5$  low-pass filter from [24], oversharpening is suppressed.

## 6 Discussion and conclusion

Implicit finite differences offer much higher spectral resolving efficiency compared to explicit finite differences while the computational effort increases only slightly. The resulting narrow-banded systems of linear equations can be efficiently solved by direct solvers (see, for example, [15]).

In this paper, we focused on demonstrating advantages of implicit schemes for basic picture processing tasks and adapted Fourier-Padé-Galerkin approximations for designing implicit image differentiation and filtering schemes with good spectral resolution properties.

Another theoretical contribution of the paper consists of establishing a simple link between implicit and explicit finite differences used for gradient direction estimation.

A potential application of our approach consists of using implicit finite differences within the highly popular HOG approach [16] where it was reported that smoothing introduced by common gradient estimation schemes decreases performance of the HOG descriptor.

We are sure that implicit image differentiation and filtering schemes deserve to be among the standard computational tools used by computer graphics and image processing researchers and practitioners.

**Acknowledgements.** I would like to thank the anonymous reviewers for thoroughly reading the paper and providing valuable and constructive comments. I am grateful to J. W. Kim for an enlightening e-mail exchange on implicit filtering schemes.

## References

- [1] U.R. Alim, T. Möller, and L. Condat. Gradient estimation revitalized. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1494–1503, 2010.
- [2] L. Alvarez. Images and PDE's. In *ICAOS '96, Lecture Notes in Control and Information Sciences, Vol. 219/1996*, pages 3–14. Springer, 1996.
- [3] S. Ando. Consistent gradient operators. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(3):252–265, 2000.
- [4] M. J. Bentum, T. Malzbender, and B. B. Lichtenbelt. Frequency analysis of gradient estimators in volume rendering. *IEEE Transactions on Visualization and Computer Graphics*, 2(3):242–254, 1996.
- [5] P. Bhat, B. Curless, M. Cohen, and C. L. Zitnick. Fourier analysis of the 2D screened Poisson equation for gradient domain problems. In *Proceedings of the 10th European Conference on Computer Vision (ECCV-08)*, pages II: 114–128, 2008.
- [6] W. G. Bickley. Finite difference formulae for the square lattice. *Quart. J. Mech. Appl. Math.*, 1:35–42, 1948.
- [7] T. Blu, P. Thévenaz, and M. Unser. Generalized interpolation: Higher quality at no additional cost. In *Proc. IEEE Int. Conf. Image Processing (ICIP 99)*, pages 667–671, Kobe, Japan, 1999.
- [8] G. Bradski and A. Kaehler. *Learning OpenCV: Computer Vision with the OpenCV Library*. O'Reilly, 2008.

- [9] L. Collatz. *The Numerical Treatment of Differential Equations*. Springer, 3rd edition, 1960.
- [10] T. Colonius and S. K. Lele. Computational aeroacoustics: progress on nonlinear problems of sound generation. *Progress in Aerospace Sciences*, 40(6):345–416, August 2004.
- [11] N. Dalai and B. Triggs. Histograms of oriented gradients for human detection. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2005)*, pages 886–893, San Diego, CA, USA, 2005.
- [12] M. Daniel and A.S. Willsky. Efficient implementations of 2-D noncausal IIR filters. *IEEE Trans. on Circuits and Systems II*, 44(7):549–563, July 1997.
- [13] E. R. Davies. Circularity - a new principle underlying the design of accurate edge orientation operators. *Image and Vision Computing*, 2(3):134–142, 1984.
- [14] H. Farid and E. P. Simoncelli. Differentiation of multi-dimensional signals. *IEEE Transactions on Image Processing*, 13(4):496–508, 2004.
- [15] G. E. Forsythe and C. B. Moler. *Computer solution of linear algebraic systems*. Prentice Hall, 1967.
- [16] W. Frei and C. C. Chen. Fast boundary detection: A generalisation and a new algorithm. *IEEE Trans. Comput.*, C-26:988–998, 1997.
- [17] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Pearson Prentice Hall, 3rd edition, 2008.
- [18] R. W. Hamming. *Digital Filters*. Dover, 3rd edition, 1998.
- [19] Z. Hossain, U. R. Alim, and T. Möller. Toward high-quality gradient estimation on regular lattices. *IEEE Transactions on Visualization and Computer Graphics*, 17(4): 426–439, 2011.
- [20] B. Jähne, H. Scharr, and S. Körkel. Principles of filter design. In *Handbook of Computer Vision and Applications, Vol. 2, Signal Processing and Applications*, pages 125–151. Academic Press, 1999.
- [21] J. W. Kim. High-order compact filters with variable cut-off wavenumber and stable boundary treatment. *Computers & Fluids*, 39:1168–1182, 2010.
- [22] P. D. Kovesi. MATLAB and Octave functions for computer vision and image processing. School of Computer Science & Software Engineering, The University of Western Australia. Available from: <<http://www.csse.uwa.edu.au/~pk/research/matlabfns/>>.
- [23] A. Kumar. Isotropic finite-differences. *Journal of Computational Physics*, 201(1): 109–118, 2004.
- [24] S. K. Lele. Compact finite difference schemes with spectral-like resolution. *Journal of Computational Physics*, 103:16–42, 1992.
- [25] J. Li and Y.-T. Chen. *Computational Partial Differential Equations Using MATLAB*. CRC Press, 2009.

- [26] J. S. Lim. *Two-dimensional signal and image processing*. Prentice-Hall, 1990.
- [27] W.-S. Lu and A. Antoniou. *Two-dimensional Digital Filters*. CRC Press, 1992.
- [28] M. S. Min, S. M. Kaber, and W. S. Don. Fourier-Padé approximations and filtering for spectral simulations of an incompressible Boussinesq convection problem. *Mathematics of Computation*, 76(259):1275–1290, July 2007.
- [29] C. Mulat, M. Donias, P. Baylou, G. Vignoles, and C. Germain. Optimal orientation estimators for detection cylindrical objects. *Signal, Image and Video Processing*, 2(1): 51–58, 2008.
- [30] T. Petrila and D. Trif. *Basics of Fluid Mechanics and Introduction to Computational Fluid Dynamics*. Springer, 2005.
- [31] W. K. Pratt. *Digital Image Processing*. John Wiley & Sons, 3rd edition, 2001.
- [32] W. H. Raymond and A. Garder. A review of recursive and implicit filters. *Monthly Weather Review*, 119:477–495, 1991.
- [33] H. Scharr, S. Körkel, and B. Jähne. Numerische isotropieoptimierung von fir-filtern mittels querglättung. In *Proc. DAGM'97*, pages 367–374, 1997.
- [34] E. P. Simoncelli. Design of multi-dimensional derivative filters. In *Proc 1st IEEE Int'l Conf. on Image Proc., Vol. I*, pages 790–794, 1994.
- [35] P. Thévenaz, T. Blu, and M. Unser. Image interpolation and resampling. In *Handbook of Medical Imaging, Processing and Analysis*, pages 393–420. Academic Press, 2000.
- [36] Y. E. Wang, Y. Luo, and M. N. Alfaraj. Enhanced isotropic 2D and 3D gradient method. U.S. Patent 7,653,258, January 2010.
- [37] J. Weickert and H. Scharr. A scheme for coherence-enhancing diffusion filtering with optimized rotation invariance. *Journal of Visual Communication and Image Representation*, 13:103–18, 2002.