DIRECTIONAL VARYING SCALE APPROXIMATIONS FOR ANISOTROPIC SIGNAL PROCESSING

Vladimir Katkovnik, Alessandro Foi, Karen Egiazarian and Jaakko Astola

Signal Processing Laboratory, Tampere University of Technology, P.O. Box 553, 33101 Tampere, Finland
e-mail: katkov, foi, karen, jta@cs.tut.fi

ABSTRACT

A spatially adaptive restoration of a multivariable anisotropic function given by uniformly sampled noisy data is considered. The presentation is given in terms of image processing as it allows a convenient and transparent motivation of basic ideas as well as a good illustration of results. To deal with the anisotropy discrete directional kernel estimates equipped with varying scale parameters are exploited. The local polynomial approximation (LPA) technique is modified for a design of these kernels with a desirable polynomial smoothness. The nonlinearity of the method is incorporated by an intersection of confidence intervals (ICI) rule exploited in order to obtain adaptive varying scales of the kernel estimates for each direction. In this way we obtain the pointwise varying scale algorithm which is spatially adaptive to unknown smoothness and anisotropy of the function in question. Simulation experiments confirm the advanced performance of the new algorithms.

1. INTRODUCTION

Points, lines, edges, textures are present in all images. They are locally defined by position, orientation and scale. Often being of small size these specific features encode a great proportion of information contained in images. To deal with these features oriented/directional filters are used in many vision and image processing tasks, such as edge detection, texture and motion analysis, etc.

The key question is, how to design a kernel for a specified direction. A good initial idea arises from a definition of the *right-hand directional derivative* for the direction defined by the angle θ, $\partial_{+\theta}y(x) = \lim_{\rho\to 0^+}(y(x_1+\rho\cos\theta,\ x_2+\rho\sin\theta)-y(x_1,x_2))/\rho$. Whenever y is a differentiable function, elementary calculations give the well known result

$$\partial_{+\theta}y(x) = \partial_\theta y(x) = \cos\theta\cdot\partial_{x_1}y(x)+\sin\theta\cdot\partial_{x_2}y(x). \quad (1)$$

Thus, in order to find the derivative for any direction θ it suffices to estimate the two derivatives on x_1 and x_2 only. This concept has been exploited and generalized by the so-called steerable filters [4].

Although continuous models of the discrete image intensity are widely used in image processing, estimates such as (1) are too rough in order to be useful for those applications where the sharpness and details are of first priority. For discrete images lacking global differentiability or continuity the only reliable way to obtain an accurate directional anisotropic information is to calculate variations of y in the desired direction θ and, say, to estimate the directional derivative by the finite difference counterpart of $\partial_{+\theta}y(x)$. In more general terms this means that the estimation or image analysis should be based on directional kernels, templates or atoms which are quite narrow and concentrated in desirable directions. Since points, lines, edges and textures can exist at all possible positions, orientations and scales one would like to use families of filters that can be tuned to all orientations, scales and positions.

Recent development shows an impressive success of methods for this sort of directional image/multivariable signal processing. In particular, narrow multidirectional items are the building blocks of the new ridgelet and curvelet transforms [14].

The nonparametric regression originated in mathematical statistics offers an original approach to signal processing problems (e.g. [5], [2], [6]). It basically results in kernel filtering with the kernels designed using some moving window local approximations. Adaptive versions of these algorithms are able to produce efficient filtering with the varying window size (scale, bandwidth) which is pointwise adaptive (see [11], [13] and references therein). This pointwise adaptive scale selection is based on the following idea known as Lepski's approach. The algorithm searches for a largest local vicinity of the point of estimation where the estimate fits well to the data. The estimates $\hat{y}_h(x)$ are calculated for a set of window sizes (scales) $h \in H$ and compared. The adaptive scale is defined as the largest of those windows in the grid which estimate does not differ significantly from the estimators corresponding to the smaller window sizes. The intersection of confidence intervals (ICI) rule [7] being one of the versions of this approach has appeared to be quite efficient for the adaptive scale image restoration [8], [9].

Cited above papers on the adaptive scale kernel estimation concern a scalar scale parameter and assume that the estimators can be ordered by their variances. Vector scale parameter kernels in d-dimensional space, $x, h \in R^d$, are of special interest for anisotropic function with highly varying properties at different directions. Imaging is one of the typical examples of such problems.

A direct generalization of the Lepski's approach to adaptive smoothing with a vector scale parameter $h \in R^d$ faces a principal obstacle as the variance of the estimates calculated for different $h \in R^d$ cannot be ordered and can be only semi-ordered as there could be many estimators with the same or similar variance [13].

The first algorithm and analysis results concerning the multivariable scale adaptive kernel algorithms have been reported in [10]. This work is concentrated on theoretical aspects of the problem and is formulated in continuous variables. It is shown, in particular, that a new proposed adaptive algorithm is able to exactly attain the minimax rate for a large class of anisotropic Besov spaces. The multidimensional kernel $g_h(x)$ used in [10] is defined as a product of the corresponding univariate ones, $\Pi_{i=1}^{d} g_i(x_i/h_i)$. This poses a basic limitation of the approach as the neighborhood used for estimation can only be scaled along the components of x.

Similar to [10] the main intention of the new approach introduced in the present paper is to obtain in a data-driven way a largest local vicinity of the estimation point in which the underlying model fit the data. We assume that this vicinity is a star-shaped set, which can be approximated by some sectorial segmentation with say K non-overlapping sectors. We use special directional kernels with supports embedded in these sectors. The kernels are equipped with univariate scale parameters defining the size of the supports in the sector. The ICI rule is exploited K times, once for each sector, in order to find the optimal pointwise adaptive scales for each sector's estimates which are then combined into the final one. In this way, we reduce the d-dimensional scale selection problem to a multiple univariate one.

Contribution of this paper can be summarized as follows. A new approach to multidimensional scale selection is proposed for a wide class of kernel estimators. The local polynomial approximation technique is modified for a design of discrete directional kernels of desirable polynomial smoothness. Fast adaptive scale selection algorithms are developed for image processing including the following problems: denoising, deblurring, edge detection. Simulation experiments confirm advanced performance of the new algorithms. Overall these results can be considered as a further development of the algorithms studied in [8], [9].

2. MOTIVATION AND IDEA

Introduce a covering of the unit sphere $\partial B^d = \{x \in R^d : \|x\| = 1\}$ with a finite family $\{D_{\theta_i}\}_{i=1,...,K}$ of non-overlapping contractible bodies (in the sphere topology) $D_{\theta_i} \subset \partial B^d$ whose baricenters have spherical angular components θ_i. For any given $h \in R^+$, $S^h_{\theta_i} = \bigcup_{0 \leq \alpha \leq h} \alpha D_{\theta_i}$ are then the corresponding positive cones constituting an alike covering of the ball $hB^d = \{x \in R^d : \|x\| \leq h\}$ with angular sectors having their vertex in the origin and oriented as θ_i. Let g_{h,θ_i} be a compactly supported kernel such that $\mathrm{supp}\, g_{h,\theta_i} = S^h_{\theta_i}$ for all values of the scalar scale parameter h. Then, the introduced directional estimator has the following generic form

$$\hat{y}(x) = \sum\nolimits_i \lambda(\theta_i)\hat{y}_{h,\theta_i}(x), \quad \hat{y}_{h,\theta_i}(x) = (g_{h,\theta_i} \circledast z)(x), \qquad (2)$$

where $\lambda(\theta_i) \geq 0$, $\sum_i \lambda(\theta_i) = 1$, and the directional kernel $g_{h,\theta_i}(x)$ satisfies vanishing moment conditions

$$(g_{h,\theta_i} \circledast 1)(0) = 1, \quad (g_{h,\theta_i} \circledast x^t)(0) = 0, \quad 0 \leq t \leq m, \ |t| \neq 0.$$

Here and in what follows a compact multi-index notation is used. A multi-index t is a d-tuple t of nonnegative integers t_j, $j = 1, ..., d$, $t = (t_1, ..., t_d)$, where $t_j \geq 0$ and $|t|$ is used to denote the length $\sum_{j=1}^d t_j$. Then $x^t = x_1^{t_1} \cdot \ldots \cdot x_d^{t_d}$ for $x \in R^d$, and $0 \leq t \leq m$ means $0 \leq t_j \leq m_j$, $j = 1, ..., d$. Although in applications and illustrations we discuss imaging and assume $d = 2$, the kernel design procedure considered later is quite general and will be given in a form applicable for d-dimensional signals.

The $\hat{y}_{h,\theta_i}(x)$ in (2) is the estimate of $y(x)$ using the observations from the sector $S^h_{\theta_i}$. Optimization of h for each of the sector estimates gives the adaptive scales $h^*(\theta_i)$ depending on θ_i. The union of the supports of $g_{h^*(\theta_i),\theta_i}$, $\bigcup_i \mathrm{supp}\, g_{h^*(\theta_i),\theta_i}$, can be therefore considered as an approximation of the best local vicinity of x in which the estimation model fit the data.

Figure 1 illustrates this concept and shows sequentially: a local best estimation neighborhood U^*, a sectorial segmentation of the unit ball, and the sectorial approximation of U^* using the adaptive scales $h^*(\theta_i)$ defining the length of the corresponding sectors. Varying size sectors enable one to get a good approximation of any neighborhood of the estimation point x provided that it is star-shaped body.

Formula (2) makes clear our basic intentions. We introduce the directional estimates $\hat{y}_{h,\theta_i}(x)$, optimize the scalar scale parameter for each of the directions (sectors) and fuse these directional estimates in the final one $\hat{y}(x)$ using the weights $\lambda(\theta_i)$. Two points are of the importance here. First, we are able to find good approximations of estimation supports which can be of a complex form. Second, this approximation is composed from the univariate scale optimizations on h, thus the complexity is proportional to the number of sectors. What follows mainly concerns applied aspects of the approach and includes:

- Design of the discrete directional kernels $g_{h,\theta}$;
- Application of the ICI rule for the adaptive varying scale selection for each direction;
- Fusing of the directional estimates into the final one using the data-driven weights $\lambda(\theta_i)$;
- Application examples proving a good performance of the presented technique.

3. DIRECTIONAL *LPA* KERNEL DESIGN

Let us start from the LPA technique. Introduce linearly independent d-dimensional polynomials $x^k/k! = x_1^{k_1}/k_1! \cdot \ldots \cdot x_d^{k_d}/k_d!$, $k_1 = 0, \ldots, m_1, \ldots, k_d = 0, ..., m_d$, $\sum_i k_i \leq \max_i(m_i)$. The vector $\phi(x)$ is composed from these polynomials starting from the zero order term $x^0/0! = 1$. The observations z are given on the

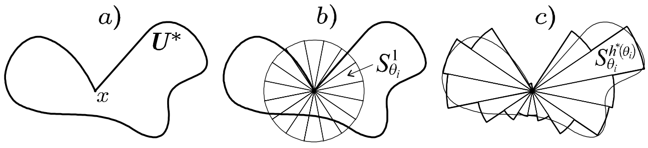

Figure 1: A neighborhood of the estimation point x: a) the best estimation set U^*, b) the unit ball segmentation, c) sectorial approximation of U^*.

d-dimensional grid $\{\tilde{x}\}$ and the estimates are needed for a desired x. The weighted least square criterion $J_h(C) = \sum_{\tilde{x}} w_h(x - \tilde{x})(z(\tilde{x}) - \bar{y}_h(x - \tilde{x}))^2$ is commonly used for design of the nonparametric regression estimates [5], [6], [2]. Here $\bar{y}_h(x) = C^T\phi_h(x)$, $\phi_h(x) = \phi(x/h)$, $h = (h_1, \ldots, h_d) \in R^d$ is a vector scale parameter, w is a window function used for localization of the estimates and $w_h(x) = w(x/h)$, where $x/h = x_1/h_1 \cdot \ldots \cdot x_d/h_d$. Thus, we produce a fit of the observations z by the model $C^T\phi_h(x)$ with unknown C.

According to the idea of the LPA the minimizing $J_h(C)$ on C, gives $\hat{C}$. Then, the estimates $\hat{y}_h^{(r)}(x)$ of the function y and its derivatives $\partial^{(r)}y(x)$, $r = (r_1, ..., r_d)$, are in the form

$$\hat{y}_h^{(r)}(x) = \hat{C}^T\phi^{(r)}(0)(-1)^{|r|}/h^r, \ \phi^{(r)} = \partial^{(r)}\phi,$$

where the estimate of the function corresponds to $r = 0$. Assuming that the grids $\{\tilde{x}\}$ and $\{x\}$ are regular, identical and unrestricted, these estimates can be given in the convolution form (e.g. [8], [9])

$$\hat{y}_h^{(r)}(x) = (z \circledast g_h^{(r)})(x), \qquad (3)$$

$$g_h^{(r)}(x) = (-1)^{|r|}h^{-r}w_h(x)\phi_h^T(x)\Phi_h^{-1}\phi^{(r)}(0),$$

$$\Phi_h = \sum\nolimits_x w_h(x)\phi_h(x)\phi_h^T(x).$$

Conventionally the estimation kernels $g_h^{(r)}$ have simple form supports (square, discs, etc.), symmetric with respect to the origin and/or the coordinate axes.

The directional version of the LPA method comprises three independent steps. First, the design of the basic window w_h oriented in the basic direction $\theta_0 = 0$. The support of the window is finite, non-symmetric, elongated and well oriented in the basic direction θ_0. Second, the rotation of the window w_h to the direction θ. Let $U(\theta)$ be a matrix of the rotation operator and $u = U(\theta)x$ be new rotated variables. Then the rotated window defined on the grid on the original variable x is given as $w_h(U(\theta)x)$. Third, the standard LPA procedure is applied using the polynomials in the rotated variables u with the weights $w_h(u)$. Finally, the LPA kernel directed to θ has the form

$$g_{h,\theta}^{(r)}(u) = (-1)^{|r|}h^{-r}w_h(u)\phi_h^T(u)x)\Phi_h^{-1}\phi^{(r)}(0), \qquad (4)$$

$$\Phi_h = \sum\nolimits_u w_h(u)\phi_h(u)\phi_h^T(u), \ u = U(\theta)x.$$

What differs this procedure from any attempt to interpolate the kernels (3) to the desirable directions is that the directional LPA (4) preserves the normalization and the polynomial smoothness of the kernels (vanishing moment conditions) as well as the directionality of the kernel support.

Let $G_{h,0}^{(r)}$ be the frequency characteristic of the basic kernel $g_{h,0}^{(r)}$, then it can be shown that the frequency characteristic $G_{h,\theta}^{(r)}(\omega)$ of the directed LPA kernel is such that $G_{h,\theta}^{(r)}(\omega) \simeq G_{h,0}^{(r)}(U(\theta)\omega)$. It is an approximate equality as the rotation of w_h makes the grid of the rotated kernel irregular in the new coordinate system.

This technique allows to design the estimates for smoothing and differentiation which are important on their own and can be used in many applications. They have a number of valuable benefits:

- Unlike many other transforms which start from the continuous domain and then discretized, this technique works directly in

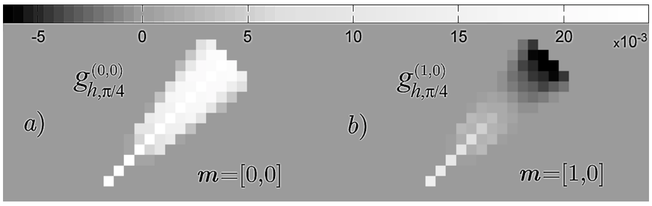

Figure 2: Directional smoothing (function estimation) kernel (a) and differentiating kernel (b) obtained by the directional LPA design with $m = [0,0]$ and $m = [1,0]$ respectively.

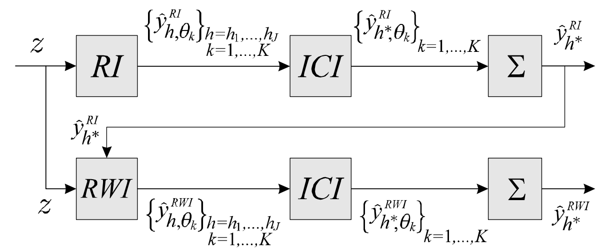

Figure 3: Directional LPA-ICI regularized Wiener inverse algorithm. In the first line of the flowchart the RI estimates are calculated for a set of scales and directions, the ICI is used to obtain the pointwise optimal scale directional estimates that are then fused into the $\hat{y}_{h^*}^{RI}$ estimate. In the second line the RWI estimates are calculated using $\hat{y}_{h^*}^{RI}$ as a reference signal in Wiener filtering, again ICI and fusing are performed to obtain the final $\hat{y}_{h^*}^{RWI}$ estimate.

the multidimensional discrete domain;

- The designed kernel are truly multivariable, non-separable and anisotropic with arbitrary orientation, width and length;
- The desirable smoothness of the kernel along and across the main direction is enabled by the corresponding vanishing moment conditions;
- The kernel support can be flexibly shaped to any desirable geometry in order to capture geometrical structural and pictorial information. In this way a special design can be done for complex form objects and specific applications;
- The smoothing and corresponding differentiating directional kernels can be designed.

4. ADAPTIVE ALGORITHM

In general, the scaling vector h controls the size as well as the shape of the kernel in (4). Consider $d = 2$ and let u_1 be the axis directed along the radius of partition sectors in Figure 1. Then, h_1 and h_2 (the last is a scale along the axis u_2 perpendicular to u_1) define the length and the width of the kernel. It has been proved by minimizing the pointwise mean squared error on h_1 and h_2 that the following anisotropic scaling law holds between the optimal length and width,

$$h_1 \propto h_2^{(m_2+1)/(m_1+1)}, \qquad (5)$$

where m_1 and m_2 are the orders of the polynomials in the LPA on u_1 and u_2, respectively. Then, assuming that $h_2(h_1)$ we can treat h as a vector function $h\,(h_1)$ of the unique scale parameter h_1 and implement the algorithm using a univariate scale parameter for each sector as in (2). Results similar to (5) can be obtained also for $d > 2$. Then, we can assume $h_j = h_j\,(h_1)$ for $j \geq 2$ and apply univariate scale estimates for a general multivariable case.

Note that provided natural assumptions that $m_2 = 0$ and $m_1 = 1$ the formula (5) gives the scaling law *width*$\propto$*length*^{2} exploited in the curvelet transform [14].

The univariate scale makes possible to apply the ICI rule [7], [8], [9] for pointwise data driven selection of its values for each of the estimates $\hat{y}_{h,\theta_i}(x)$ in (2), i.e. for each direction θ_i and for each x. It gives the adaptive scales $h^*(x,\theta_i)$ which shapes the adaptive estimation neighborhood $\bigcup_i \operatorname{supp} g_{h^*(x,\theta_i),\theta_i}$. In this way we arrive to the spatially adaptive varying scale estimation.

Let $\hat{y}_{h^*(x,\theta_i),\theta_i}(x)$ be the adaptive estimate and $\sigma_i^2(x)$ be the variance of this estimate, then all these directional estimates can be fused according to (2) in the final one as follows [8], [9]

$$\hat{y}(x) = \sum\nolimits_i \lambda(\theta_i)\hat{y}_{h^*(x,\theta_i),\theta_i}(x),\ \lambda(\theta_i) = \sigma_i^{-2}(x)/\sum\nolimits_j \sigma_j^{-2}(x). \quad (6)$$

We use a linear fusing of the estimates with the inverse variances of the estimates as the weights. Note that the weights $\lambda(\theta_i)$ in (6) are data-driven adaptive as $\sigma_i^{-2}(x)$ depend on the adaptive $h^*(x,\theta_i)$.

Concerning the algorithm complexity we note that the algorithm is fast as based on the fast convolution operations. The calculation of the estimate $\hat{y}_{h_j,\theta_i}$ for a given scale h_j is a linear convolution requiring $N_{conv} \sim n \log n$ where n is the size of the signal. This procedure is repeated $J \cdot K$ times, where K is a number of the sectors in the estimator and J is the number of the used scales h_j.

5. APPLICATIONS

To illustrate the improved performance arising from the proposed directional kernels and adaptive varying scale selection we present here some results for deblurring and edge detection.

5.1 Adaptive deblurring algorithm

We wish to recover an image y from noisy observations $z = (\upsilon \circledast y) + \sigma\eta$, where υ is the point spread function (PSF) of the blurring system. It is assumed that the PSF is known and that the noise η is standard gaussian. In the frequency domain the observation equation has the form $Z(f) = Y(f)V(f) + \sigma\eta(f)$, where f is the frequency and capital letters are used for the discrete Fourier transform of the corresponding variables. The considered technique is based on the following regularized inversion (RI) and regularized Wiener inversion (RWI) algorithms, using the directional LPA kernels $g_{h,\theta}$ [9]:

$$\hat{Y}_{h,\theta}^{RI}(f) = \frac{V(-f)G_{h,\theta}(f)}{|V(f)|^2 + \varepsilon_1^2} Z(f), \qquad (RI), \qquad (7)$$

$$\hat{Y}_{h,\theta}^{RWI}(f) = \frac{V(-f)|Y(f)|^2 G_{h,\theta}(f)}{|V(f)Y(f)|^2 + \varepsilon_2^2\sigma^2} Z(f), \qquad (RWI). \quad (8)$$

The estimate of y is given by the RWI deconvolution scheme (8) that uses the ICI based RI estimate as a reference signal Y. Thus, we arrive to two steps procedure (see Figure 3). The adaptive procedure assumes that the estimates $\{\hat{y}_{h,\theta_k}^{RI}\}_{h \in H}$ are calculated according to (7) for a set of scales H and the ICI rule selects the best scales for each direction and for each pixel. In this way we obtain the directional varying scale adaptive estimates $\hat{y}_{h^*(x,\theta_k),\theta_k}^{RI}$, $k = 1,\ldots,K$, which are fused in the final one $\hat{y}_{h^*}^{RI}$ according to (6). This $\hat{y}_{h^*}^{RI}$ serves as the reference signal in the RWI procedure (see Figure 3). The adaptive RWI algorithm is similar and gives the ICI adaptive varying scales estimates $\hat{y}_{h^*(x,\theta_k),\theta_k}^{RWI}$ for each direction and x. Then, the final estimate $\hat{y}_{h^*}^{RWI}$ is obtained by fusing these directional ones again according to (6).

The ICI adaptive scales $h^*(\cdot,\theta_k)$ represent the distribution of image features across the direction θ_k, as shown in Figure 5 (right). Exploiting the directional nature of the kernel supports, we improve the adaptive scale selection by embedding directionally-weighted order-statistics filters within the ICI algorithm. These specially designed filters effectively remove the possible outliers in $h^*(\cdot,\theta_k)$ and yet preserve accurate edge adaptation.

Table 1 presents results for four different experiments: Cameraman image, 9×9 boxcar υ, *BSNR*=40dB (Experiment 1, see Figure 4); $\upsilon(x_1,x_2) = (1 + x_1^2 + x_2^2)^{-1}$, $x_1, x_2 = -7,\ldots,7$, $\sigma^2 = 2$ (Exp.2) or $\sigma^2 = 8$ (Exp.3), and Lena image, υ is a 5×5 separable filter with the weights $[1,4,6,4,1]/16$ in horizontal and vertical

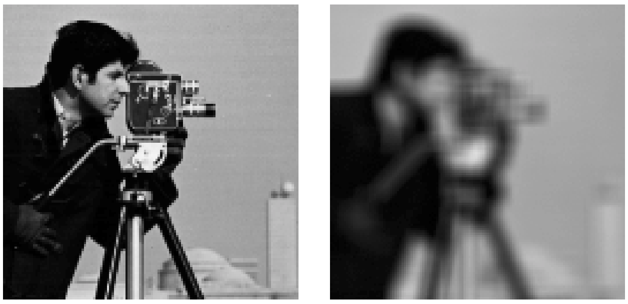

Figure 4: Original Cameraman image (left) and noisy blurred observation (Experiment 1) (right)

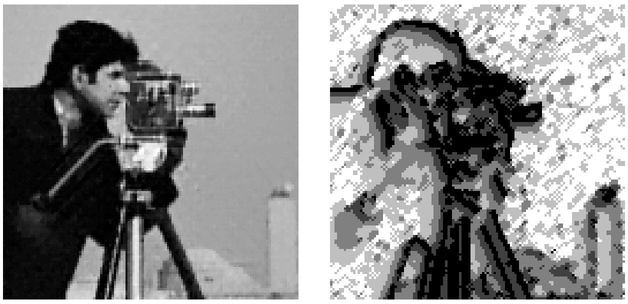

Figure 5: LPA-ICI algorithm performance: restored image, *ISNR*=8.23dB (left) and adaptive scales scales $h^*(\cdot,\pi/4)$ (right)

directions, *BSNR*=15.93dB (Exp.4). For these experiments a set of eight directions, $\{\theta_k\}_{k=1}^{8}=\{0,\pi/4,\pi/2,\ldots,7/4\pi\}$ and five scales, $\#H=5$, are used. Function estimation kernels were designed on conically-supported windows choosing the LPA orders $m=[1,0]$ and $m=[0,0]$ for the RI and RWI stages, respectively. These kernels are shown in Figure 2. For smaller scales in H the supp w_h is a 1-pixel-width line.

Overall, the *SNR* improvement (*ISNR*) in Table 1 shows that the new developed RWI algorithm demonstrates a good performance and outperforms some state-of-the-art techniques. Visual inspection is also in favor of the new algorithm. Figure 5 (left) shows a fragment of the restored Cameraman image.

The directionality of the kernels is an important element of this good performance. For example, in the same algorithm nondirectional quadrant kernels give *ISNR*=7.52dB for Exp.1 (see [9]) versus *ISNR*=8.23dB in Table 1.

5.2 Derivative estimation and edge detection

As a further illustration of the flexibility of our approach we present two examples of differentiation of y using the noisy blurred observations. Let us replace in the RWI stage of the algorithm (8) the smoothing kernels $g_{h,\theta_k}^{(0,0)}$ by the discrete differentiation kernels $g_{h,\theta_k}^{(1,0)}$ (4). Then the output $\hat{y}_{h^*,\theta_k}^{RWI}$ of the two stage algorithm gives the estimate of the directional right-hand derivative $\partial_{+\theta_k}y$. Figure 6 (left) shows the diagonal derivative estimate $\hat{\partial}_{\theta_2}$ calculated for $\theta=\pi/4$ as the mean of the two one-sided directional derivatives with $\theta_2=\pi/4$ and $\theta_5=\theta_2+\pi=5\pi/4$, $\hat{\partial}_{\theta_2}=(\hat{y}_{h^*,\theta_2}^{RWI}-\hat{y}_{h^*,\theta_5}^{RWI})/2$.

Further, for the edge detection we calculate the sum of the absolute values of these derivatives $\sum_{k=1}^{4}|\hat{\partial}_{\theta_k}|$. The image of this sum is shown in Figure 6 (right). It demonstrates a very accurate recovery of the image edges from the blurred noisy image data.

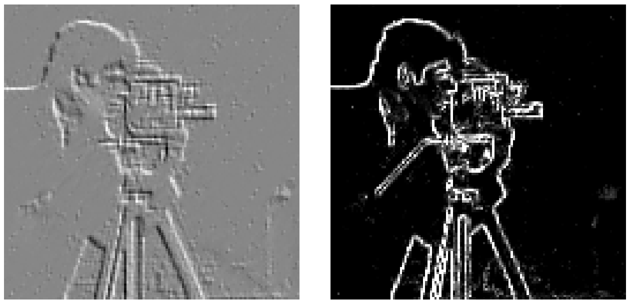

Figure 6: Directional derivative (left) and edge detection (right)

Method Experiment	1	2	3	4
LPA-ICI directional	8.23	7.78	6.04	3.76
GEM (Dias) [1]	8.10	7.47	5.17	–
EM (Figueiredo and Nowak) [3]	7.59	6.93	4.88	2.94
ForWaRD (Neelamani et al.) [12]	7.30	6.75	5.07	2.98

Table 1: *ISNR* (dB) of the proposed algorithm and of methods [1], [3] and [12] for the four experiments.

REFERENCES

[1] Dias, J.M.B., ”Fast GEM wavelet-based image deconvolution algorithm”, *Proc. Int. Conf. on Image Proc. 2003*, vol. 2.

[2] Fan J. and I. Gijbels, *Local polynomial modelling and its application*. London: Chapman and Hall, 1996.

[3] Figueiredo M. A.T. and R. D. Nowak,” An EM algorithm for wavelet-based image restoration”. *IEEE Trans Image Proc.*, vol. 12, N 8, pp. 906-916, 2003.

[4] Freeman W. T. and E. H. Adelson, ”The design and use of steerable filters”. *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 13, no. 9, pp. 891-906, 1991.

[5] Katkovnik V., “Problem of approximating functions of many variables”. *Autom. Remote Control*, vol. 32, no. 2, Part 2, pp. 336-341, 1971.

[6] Katkovnik V., *Nonparametric identification and smoothing of data (Local approximation methods)*. Nauka, Moscow, 1985 (in Russian).

[7] Katkovnik V., “A new method for varying adaptive bandwidth selection”, *IEEE Trans. on Signal Proc.,* vol. 47, no. 9, pp. 2567-2571, 1999.

[8] Katkovnik V., K. Egiazarian, and J. Astola, ”Adaptive window size image de-noising based on intersection of confidence intervals (ICI) rule”. *J. of Math. Imaging and Vision*, vol. 16, N. 3, pp. 223-235, 2002.

[9] Katkovnik V., K. Egiazarian, and J. Astola. *Adaptive varying scale methods in image processing*. Tampere International Center for Signal Processing, TICSP Series, N.19, Tampere, TTY, Monistamo, 2003.

[10] Kerkyacharian G., O. Lepski and D. Picard, " Nonlinear estimation in anisotropic multi-index denoising", *Prob. Theory and Related Fields*, vol. 121, no. 2, 137-170, 2001.

[11] Lepski, O., E. Mammen and V. Spokoiny, ”Ideal spatial adaptation to inhomogeneous smoothness: an approach based on kernel estimates with variable bandwidth selection”. *Annals of Statistics*, vol. 25, no. 3, 929–947, 1997.

[12] Neelamani R., H. Choi, and R. G. Baraniuk, ”Forward: Fourier-wavelet regularized deconvolution for ill-conditioned systems”. *IEEE Trans. on Image Proc., 2003.*

[13] Polzehl, J., Spokoiny, V., ”Image denoising: pointwise adaptive approach,” *Annals of Statistics*, vol. 31, no. 1, 2003.

[14] Starck J., E.J. Candes, and D.L. Donoho, ”The curvelet transform for image denoising”. *IEEE Trans. on Image Proc.*, vol. 11, N. 6, pp. 670-684, 2002.