ADAPTIVE-SIZE BLOCK TRANSFORMS FOR POISSONIAN IMAGE DEBLURRING

Alessandro Foi[a], Sakari Alenius[b], Mejdi Trimeche[b], and Vladimir Katkovnik[a]

[a] Institute of Signal Processing, Tampere University of Technology,
P.O. Box 553, Tampere 33101, FINLAND - firstname.lastname@tut.fi
[b] Multimedia Technologies Laboratory, Nokia Research Center, Tampere, FINLAND - firstname.lastname@nokia.com

ABSTRACT

A novel deconvolution technique for blurred observations corrupted by signal-dependent noise is presented. Deblurring is performed with a transform-domain inverse-filtering applied locally, on a sliding block of adaptively-selected pointwise varying size.

Simulation results demonstrate a good quality of the proposed method, which is versatile and which can be easily combined with other transform-domain processing.

1. INTRODUCTION AND MOTIVATION

In this paper we propose a new transform-based deblurring technique for a broad class of signal-dependent noise observations. It is based on regularized inverse filtering in local block-transform domain, where the size of the block is adaptively selected. The LPA-ICI algorithm [11, 15] is exploited in order to select the size in pointwise-adaptive manner [8, 16].

The use of an adaptive-size block is of crucial importance. Firstly, it improves the sparsity of the signal representation in the transform domain, significantly increasing the effectiveness of coefficient shrinkage (e.g. hard-thresholding, Wiener filtering). At the same time, it also enables a simpler and more direct noise modelling. In particular, special approximate calculations are developed in order to deal efficiently with the signal-dependent noise in the transform domain. The resulting image deblurring algorithm is simple yet very effective.

The presented method can be considered as a generalization of the denoising technique [8] to the following model for blurred and noisy observations.

1.1 Signal-dependent noise model

Given an original image $y : X \to \mathbb{R}$, we consider its noisy blurred observations $z(x)$, $x \in X \subset \mathbb{Z}^2$, where the expectations $E\{z(x)\} = (y \circledast v)(x) \geq 0$ are given as the convolution of y against a blurring kernel v, known as the point-spread function (PSF). The degradation in z has a stochastic and a deterministic component. The stochastic errors (noise) $\eta(x) = z(x) - (y \circledast v)(x)$ are assumed as independent and the variance of the observations $z(x)$ is modeled as

$$\sigma_z^2(x) = \mathrm{var}\{z(x)\} = \mathrm{var}\{\eta(x)\} = \rho(E\{z(x)\}), \quad (1)$$

ρ being a given positive function called the *variance function*.

This work was supported by the Finnish Funding Agency for Technology and Innovation (Tekes), AVIPA project, and by the Academy of Finland, project No. 213462 (Finnish Centre of Excellence program 2006 - 2011).

It is practical to express this observation model in the additive form

$$z = y \circledast v + \eta, \quad \eta = n\sqrt{\rho(y \circledast v)}, \quad (2)$$

where η and n are two noise processes which have zero mean and are, respectively, heteroskedastic and homoskedastic with unit variance. The actual distribution of the noise might be unknown. Typically, η and n are non-Gaussian (i.e. $n(\cdot) \not\sim \mathcal{N}(0,1)$) and, although homoskedastic, n is spatially variant (i.e. $n(x_1) \not\sim n(x_2)$ for some x_1, x_2).

The most important example of the above model is given by the Poissonian blurred observations which describe the acquisition of photons coming through a non-focused lens system:

$$\chi z \sim \mathcal{P}(\chi y \circledast v). \quad (3)$$

The parameter $\chi \in \mathbb{R}^+$ can be used to model the so-called quantum efficiency of the photonic sensor. For this particular case, we have $\sigma_z^2(x) = \mathrm{var}\{\eta(x)\} = \chi^{-1}(y \circledast v)(x)$, $\rho(\cdot) = \left(\frac{\cdot}{\chi}\right)$, and $\eta = n\sqrt{\chi^{-1} y \circledast v}$.

The observation model (2) can be represented in frequency domain as

$$Z = YV + \mathcal{F}(\eta),$$

where $\mathcal{F}$ stands for the Fourier transform operator and capital letters are used to indicate the Fourier transform of the corresponding small-case functions. Throughout the paper we will assume an orthonormal Fourier transform. For the discrete case, convolutions and translations will be implicitly assumed to be circular. Let us observe that, because of the independence of η, the noise addend $\mathcal{F}(\eta)$ has constant variance equal to $\|\rho(y \circledast v)\|_1$, but in general it is not independent (otherwise η would be also constant).

The problem is to reconstruct the true image y from the noisy blurred observations z. In this paper we restrict ourself to the non-blind case, i.e. we assume that both the PSF v and the variance function ρ are known. In the case of Poissonian observations, the problem is known as Poissonian deblurring.

1.2 Deblurring in transform domain

Given a frame $\{\psi^{(i)}\}_i$ and its dual $\{\breve{\psi}^{(i)}\}_i$, we can represent the original signal using the usual analysis-synthesis form

$$y = \sum_i \langle y, \breve{\psi}^{(i)} \rangle \psi^{(i)}. \quad (4)$$

Let us now be given not y but its blurred (noise-free) observation $y \circledast v$.

There are essentially two strategies which can be followed starting from Equation (4).

1.2.1 Synthesis with $\{\psi^{(i)}\}_i$

We would like to reconstruct y using the frame elements $\psi^{(i)}$ in the form

$$y = \sum_i c_i \psi^{(i)}, \tag{5}$$

where the coefficients c_i are computed from $y \circledast v$. Obviously, $c_i = \langle y, \breve{\psi}^{(i)} \rangle$ is a solution of the equation (and is the unique solution if the frame is a basis). For a generic y, the problem can be formally solved under the often unrealistic hyphotesis that the blur operator is invertible. In this case we have that

$$\begin{aligned} \langle y, \breve{\psi}^{(i)} \rangle &= \left\langle \mathcal{F}(y), \mathcal{F}(\breve{\psi}^{(i)}) \right\rangle = \left\langle YV, \tfrac{\mathcal{F}(\breve{\psi}^{(i)})}{\overline{V}} \right\rangle = && (6) \\ &= \left\langle y \circledast v, \mathcal{F}^{-1}\left(\tfrac{\mathcal{F}(\breve{\psi}^{(i)})}{\overline{V}}\right) \right\rangle = && (7) \\ &= \left\langle y \circledast v, \breve{\psi}^{(i)} \circledast \mathcal{F}^{-1}\left(\tfrac{1}{\overline{V}}\right) \right\rangle, && (8) \end{aligned}$$

where $\overline{V}$ is the complex conjugate of V. So the coefficients c_i can be calculated as the inner product between the blurred observations and $\breve{\psi}^{(i)} \circledast \mathcal{F}^{-1}\left(\frac{1}{\overline{V}}\right)$.

1.2.2 Analysis with $\{\breve{\psi}^{(i)}\}_i$

An alternative approach to (5) is to seek a solution of the form

$$y = \sum_i \langle y \circledast v, \breve{\psi}^{(i)} \rangle \xi^{(i)}, \tag{9}$$

where the true signal is analyzed with respect to the dual frame $\{\breve{\psi}^{(i)}\}_i$ and synthetized with an appropriate set of reconstructing functions $\{\xi^{(i)}\}_i$. In order to get an explicit form for $\{\xi^{(i)}\}_i$ we convolve the left and right-hand side of the above equation against v, obtaining

$$y \circledast v = \sum_i \langle y \circledast v, \breve{\psi}^{(i)} \rangle (\xi^{(i)} \circledast v). \tag{10}$$

This perfect reconstruction formula implies that $\{\xi^{(i)} \circledast v\}_i$ is a dual frame of $\{\breve{\psi}^{(i)}\}_i$. For example, $\xi^{(i)} \circledast v = \psi^{(i)}$, which gives (assuming invertibility of V)

$$\xi^{(i)} = \mathcal{F}^{-1}\left(\frac{\mathcal{F}(\psi^{(i)})}{V}\right). \tag{11}$$

1.2.3 Ill-posedness and regularization

Usually, blur operators are not invertible. In the case V has zeros, the simplest ("naïve") approach is to use a generalized inverse of V, i.e.

$$V^{-1}(\cdot) = \begin{cases} 0 & \text{if } V(\cdot) = 0 \\ 1/V(\cdot) & \text{if } V(\cdot) \neq 0 \end{cases}.$$

However – even when V does not have zeros – the problem can still remain ill-posed, with the inverse operator V^{-1} being unbounded. Consequently, when approaching noisy observations of the form (2), the typical strategy is to employ regularization at some stage of the inversion procedure.

The simplest form of regularization is the following (Tichonov) regularized inverse

$$T^{RI} = \frac{\overline{V}}{|V|^2 + \varepsilon^2}, \tag{12}$$

where $\varepsilon > 0$ is the regularization parameter. A larger regularization parameter corresponds to a more stable but also more biased inverse estimate. The regularization (12) has been used extensively for deblurring, in the spatial domain (e.g. [3],[18],[17],[6]), as well as in the transform domain (e.g. [21], [9]). In all these approaches, the combination of regularization with other filtering techniques (e.g adaptive smoothing, shrinkage, etc.) is shown to lead to an improved inverse estimate.

1.2.4 Intermediate cases; comments; vaguelettes

In addition to the two strategies from Sections 1.2.1 and 1.2.2, there are also infinitely many intermediate cases. Exploiting in (4) the fact that formally $\langle y, \breve{\psi}^{(i)} \rangle = \left\langle YV^{\alpha}, \mathcal{F}(\breve{\psi}^{(i)})\overline{V}^{-\alpha} \right\rangle$, it is easy to show that $\{\psi^{(i)} \circledast \mathcal{F}^{-1}(V^{\alpha})\}_i$ and $\left\{\breve{\psi}^{(i)} \circledast \mathcal{F}^{-1}\left(\overline{V}^{-\alpha}\right)\right\}_i$ are also a pair of frames in duality, for any $\alpha \in \mathbb{R}$. By using this new pair of frames instead of $\{\psi^{(i)}\}_i$ and $\{\breve{\psi}^{(i)}\}_i$ in (5) and by varying α, one can obtain the first ($\alpha = 0$) and the second strategy ($\alpha = 1$), all intermediate cases ($0 < \alpha < 1$), as well as other decompositions ($\alpha < 0$, $\alpha > 1$).

Although the two approaches might seem formally equivalent, they become in practice very different as soon as they are considered with respect to a predefined "meaningful" frame.

Equation (5) implicitly assumes that $\{\psi^{(i)}\}_i$ is a suitable frame to enable a good approximation of y. Since the inverse operator is embedded in the analysis frame, the approach works according to the paradigm "first invert the blur, and then approximate this inverse".

On the contrary, Equation (9) assumes that $\{\psi^{(i)}\}_i$ is suitable to approximate $y \circledast v$, with the inverse operator embedded in the reconstruction frame $\{\xi^{(i)}\}_i$. Thus, it essentially obeys to the paradigm "first approximate the blurred signal, and then invert this approximation".

If not all coefficients are used for the reconstruction, the two approches lead to significantly different approximations of y.

In our work we consider frames corresponding to sliding block-transforms and especially to the sliding block-DCT transform. Earlier, these sort of approaches have been studied with $\{\psi^{(i)}\}_i$ and $\{\breve{\psi}^{(i)}\}_i$ being biorthogonal wavelet bases: the representations (5) and (9) are respectively known within the wavelet community as wavelet-vaguelette [4, 14] and vaguelette-wavelet [1] decompositions. "Vaguelette" stands for the combination of the inverse of the blur with the corresponding dual wavelet, such as $\xi^{(i)}$ from (11). We refer the reader to [1] for a theoretically-oriented discussion and comparison of the asymptotical risk of estimators based on the two approaches.

Let us note that the frame $\{\psi^{(i)}\}_i$ and its dual $\{\breve{\psi}^{(i)}\}_i$ can be obviously interchanged in all the above equations.

1.2.5 Practical aspects

Pragmatically, if to be used in estimation from noisy observations, each one the two approaches has its own advantages and disadvantages. In particular, assuming a properly normalized frame and independent noise with constant variance, shrinkage of coefficients which are calculated directly from the blurred observations as in (10) is simplified. Note that shrinkage requires the knowledge of the standard-deviation of each coefficient, which can be calculated according to the simple formula

$$\mathrm{std}\{\langle f, g \rangle\} = \sqrt{\langle \mathrm{var}\{f\}, |g|^2 \rangle},$$

provided the independency of the noise in f. If $\mathrm{var}\{f\}$ is also constant, say, $\mathrm{var}\{f\} = \sigma^2$, then

$$\mathrm{std}\{\langle f, g\rangle\} = \sigma \|g\|_2 . \quad (13)$$

Thus, as far as coefficient shrinkage is concerned, the approach is very similar to transform-based denoising from standard Gaussian noise. On the other hand, the convolution against the (regularized) inverse of V can cause that $\{\xi^{(i)}\}_i$ does not enjoy any of the good decorrelation properties of the original frame $\{\psi^{(i)}\}_i$. Thus the approximation ability of the approach (9) can be seriously impaired.

In (5) the original frame is used for reconstruction, however, even in the simplest case of noise with constant variance, the shrinkage of coefficients obtained by (8) is rather involved because the combination of dual frame with the inverse operator.

Nevertheless in practice, since the main motivation to use a transform-based method is its ability of to represent the signal to be recovered with good approximation by using only few transform coefficients (i.e. sparsity), the approach from Section 1.2.1 is more appropriate, despite the calculation of the standard-deviation (required for the shrinkage) is usually more involved.

This is especially valid for the case of signal-dependant noise as in the observation (2), where calculation of the standard-deviation cannot be done using Equation (13).

1.2.6 Decoupling of the transform and the inversion

As mentioned in Section 1.2.4, Equation (5) works according to the paradigm "first invert the blur, and then approximate this inverse".

If in the term $\left\langle YV, \frac{\mathcal{F}(\breve{\psi}^{(i)})}{\overline{V}} \right\rangle$ from (6) we replace the noise-free data YV with Z and the naive inverse $\frac{1}{V}$ with a regularized inverse (12), we obtain that the coefficients c_i are calculated as

$$\begin{aligned} \left\langle Z, \mathcal{F}(\breve{\psi}^{(i)}) \tfrac{V}{|V|^2+\varepsilon^2} \right\rangle &= \left\langle Z \tfrac{\overline{V}}{|V|^2+\varepsilon^2}, \mathcal{F}(\breve{\psi}^{(i)}) \right\rangle = \\ &= \langle z^{RI}, \breve{\psi}^{(i)} \rangle, \end{aligned}$$

where $z^{RI} = \mathcal{F}^{-1}\left(\frac{Z\overline{V}}{|V|^2+\varepsilon^2}\right)$ is the regularized-inverse estimate of y. Although, in principle, different forms of regularization (e.g. different regularization parameters) could be used for different coefficients, for computational reasons it is preferable to use a unique regularized inverse z^{RI} for all i. In this way the calculation of the inner products is done exactly as for the analysis of z^{RI} with the frame $\{\breve{\psi}^{(i)}\}_i$. Thus, we come to the following approximation of y in terms of the frame $\{\psi^{(i)}\}_i$

$$y \simeq \sum\nolimits_i \langle z^{RI}, \breve{\psi}^{(i)} \rangle \psi^{(i)}.$$

Note that the coefficients $\langle z^{RI}, \breve{\psi}^{(i)} \rangle$ are noisy and some shrinkage should be performed in order to obtain a good estimate. The fact that the noise in z^{RI} is not white, but coloured through inverse filtering with T^{RI} (12), makes the accurate shrinkage a very delicate and demanding issue in the whole deblurring procedure, which will be subject of the coming sections. Additionally, in order for this shrinkage to be effective it is important that the frame $\{\psi^{(i)}\}_i$ represents well y: well-known transforms (e.g. DCT, wavelets, DFT, etc.) can be used in place of $\{\psi^{(i)}\}_i$ when natural images are considered as the original signal to be reconstructed.

Finally, the estimate has the generic form

$$\sum\nolimits_i \hat{\varphi}(i)\, \psi^{(i)}, \quad (14)$$

where $\hat{\varphi}$ are the filtered coefficients obtained by shrinkage (e.g. hard-thresholding) of $\langle z^{RI}, \breve{\psi}^{(i)} \rangle$.

1.3 The proposed approach

Our approach is based on the estimate (14) obtained not by a unique global transform, but rather by a family of sliding block transforms. It means that for each $x \in X$, we consider a localized transform whose basis elements are supported on a block located at x. Since blocks are bigger than a pixel, and since each pixel is contained (and thus processed) by different blocks, the approach is obviously overcomplete.

Similarly to our denosing method [8], we use square blocks whose size is pointwise-adaptive. This enables improved adaptivity with respect both to the non-stationarity of the noise and to the features and singularities in the image. Any orthonormal transform which provides a coefficient corresponding to the mean value of the data over its support (i.e. a DC coefficient) can be directly used in the proposed algorithm. Let us observe that the near totality (including the DCT, the DFT/FFT, and wavelets) of transforms used for image processing applications satisfy this requirement.

2. PRELIMINARIES

Before we proceed further, we present some elements and additional notation which we use throughout the remaining part of the paper: basic manipulations which depend on elementary properties of inner product and convolution; their relation with block transforms; the basics of the LPA-ICI technique for adaptive block-size selection [8, 16].

2.1 Elementary manipulations

Let ψ be a real function. We denote the "direct" and "mirror" translations of ψ as $\vec{\psi}_\tau = \psi(\cdot - \tau)$ and $\tilde{\psi}_\tau(\cdot) = \psi(\tau - \cdot)$, respectively. Trivially, $\widetilde{(\vec{\psi}_\lambda)}_\tau = \tilde{\psi}_0$. Further, given a function f we can calculate the inner product against ψ as the convolution against $\tilde{\psi}_\tau$ sampled at τ

$$\langle f, \psi \rangle = \left(f \circledast \tilde{\psi}_\tau\right)(\tau).$$

It is immediate to obtain that the inner products against all (direct) translates of ψ can be obtained with a single convolution against $\tilde{\psi}_0$: first we have $\langle f, \vec{\psi}_\lambda \rangle = \left(f \circledast \widetilde{(\vec{\psi}_\lambda)}_\tau\right)(\tau)$, for any λ and τ; then, by setting $\tau = \lambda$, we have

$$\langle f, \vec{\psi}_\lambda \rangle = \left(f \circledast \tilde{\psi}_0\right)(\lambda). \quad (15)$$

2.2 Block transforms

We denote by $B_{x,h}$ a square block of size $h \times h$ located at x. Since in practice we always deal with discrete data it is important to clarify what we mean by "located at x": if h is odd it simply means "centered at x", whereas if h is even it means that x coincides with the point at position $\left(\frac{h}{2}, \frac{h}{2}\right)$.

Given a block $B_{x,h}$ we indicate the corresponding block transform by its basis elements $\boldsymbol{\psi}_{B_{x,h}} = \left\{\psi^{(i)}_{B_{x,h}}\right\}_{i=1}^{h^2}$. These elements are supported on $B_{x,h}$, $\psi^{(i)}_{B_{x,h}}(u) = 0$ $u \notin B_{x,h}$. For simplicity, in what follows we consider exclusively real orthonormal block transforms, $\langle \psi^{(i)}_{B_{x,h}}, \psi^{(j)}_{B_{x,h}} \rangle = \delta(i-j)$, δ

being the Kronecker delta, and in particular, we use the B-DCT and the block wavelet transforms (however, essentially the same algorithm can be used with biorthogonal as well as redundant block-transforms). Therefore, given a function f, its transform coefficients $\left\{\varphi_{f,B_{x,h}}(i)\right\}_{i=1}^{h^2}$ are calculated using the inner product against the transform's basis elements

$$\varphi_{f,B_{x,h}}(i) = \left\langle f, \psi^{(i)}_{B_{x,h}} \right\rangle, \quad i = 1,\ldots,h^2.$$

Traditionally, block-transforms are used in sliding manner with a fixed size of the block: for all $x \in X$ a block-transform $\boldsymbol{\psi}_{B_{x,h}}$ is used for the reconstruction of the signal within the block $B_{x,h}$ and all block-transforms used for blocks at different locations coincide upon translation. In particular, they can all be interpreted as a translation of a unique set of elements located in the origin,

$$\psi^{(i)}_{B_{x,h}} = \overrightarrow{\left(\psi^{(i)}_{B_{0,h}}\right)}_x, \quad i = 1,\ldots,h^2, \quad \tau \in \mathbb{Z}^2. \tag{16}$$

Combining Equations (15) and (16) manifest the relation between convolutional and sliding-transform signal processing:

$$\left\langle f, \psi^{(i)}_{B_{x,h}} \right\rangle = \left\langle f, \overrightarrow{\left(\psi^{(i)}_{B_{0,h}}\right)}_x \right\rangle = \left(f \circledast \widetilde{\left(\psi^{(i)}_{B_{0,h}}\right)}_0 \right)(x). \tag{17}$$

We assume that a DC coefficient exists and that it is the first one of the (ordered) transform's coefficients, i.e. $\varphi_{f,B_x^+}(1)$ denotes the DC coefficient.

2.3 Adaptive block-size: LPA-ICI technique

In our approach we proceed in sliding manner, using localized blocks of a varying size $h(x) \times h(x)$ which is pointwise adaptive with respect to the signal. We exploit the LPA-ICI (local polynomial approximation - intersection of confidence intervals) [15, 11] approach to define the adaptive block-sizes.

For a fixed point x, the LPA-ICI is implemented in the following way [8]. Given a set of pointwise LPA estimates $\left\{\hat{y}^{\mathrm{LPA}}_{h_j}(x)\right\}_{j=1}^{J}$ calculated on a nested sequence $\left\{B_{x,h_1} \subset \cdots \subset B_{x,h_J}\right\}$ of blocks located at x, we determine a sequence of confidence intervals

$$\mathcal{D}_j = \left[\hat{y}^{\mathrm{LPA}}_{h_j}(x) - \Gamma\sigma_{\hat{y}^{\mathrm{LPA}}_{h_j}(x)}, \hat{y}^{\mathrm{LPA}}_{h_j}(x) + \Gamma\sigma_{\hat{y}^{\mathrm{LPA}}_{h_j}(x)}\right], \tag{18}$$

where $\Gamma > 0$ is a threshold parameter and $\sigma_{\hat{y}^{\mathrm{LPA}}_{h_j}(x)}$ is the standard-deviation of $\hat{y}^{\mathrm{LPA}}_{h_j}(x)$. The set $H = \{h_1 < \cdots < h_J\}$ is commonly called the *set of scales*.

The ICI rule can be stated as follows: *Consider the intersection of confidence intervals $\mathcal{I}_j = \bigcap_{i=1}^{j} \mathcal{D}_i$ and let j^+ be the largest of the indexes j for which $\mathcal{I}_j$ is non-empty, $\mathcal{I}_{j^+} \neq \varnothing$ and $\mathcal{I}_{j^+1} = \varnothing$. Then the adaptive scale $h^+(x)$ at x and the adaptive block B_x^+ located at x are defined as $h^+(x) = h_{j^+}$ and as $B_x^+ = B_{x,h^+(x)}$, respectively.*

This basic procedure is repeated for every $x \in X$, thus defining an adaptive block B_x^+ and an adaptive scale $h^+(x)$ for each location in the image.

3. ADAPTIVE-SIZE BLOCK TRANSFORM DECONVOLUTION: THE ALGORITHM

3.1 Regularized inverse

In accordance with the decoupling the regularized inverse from the transform – discussed in Section 1.2.6 – we first produce a regularized inverse (RI) estimate z^{RI} of y. It is computed in the frequency domain as

$$z^{RI} = \mathcal{F}^{-1}\left(T^{RI} Z\right), \qquad T^{RI} = \frac{\overline{V}}{|V|^2 + \varepsilon_1^2 \Phi_\eta}, \tag{19}$$

where T^{RI} is a regularized-inversion operator, $\varepsilon_1 > 0$ is a regularization parameter, and Φ_η is the power spectrum of the noise η. It is easy to show that for the observations (2) the spectrum Φ_η is constant, equal to $\mathcal{F}\left(\sigma_z^2\right)(0)$. Up to a factor, which depends on the particular normalization of Fourier transform, it coincides with $\left\|\sigma_z^2\right\|_1$.

The variance $\sigma_z^2 = \rho(y \circledast \upsilon)$, since y is unknown, is estimated here directly from the noisy observations, i.e. $\hat{\sigma}_z^2 = \rho(|z|)$. Note that for the Poissonian observations (3) the variance function is linear and thus $\rho(z) = z/\chi$ is actually the simplest possible unbiased estimate of σ_z^2, $E\{\rho(z)\} = E\{z\}/\chi = \mathrm{var}\{z\}$.

In what follows we make explicit use of the impulse response t^{RI} of the regularized-inversion operator, $t^{RI} = \mathcal{F}^{-1}(T^{RI})$. Although this inversion is naturally computed globally, there are a number of efficient "sectioning" techniques (e.g. [2], [24]) which allow to produce accurate approximations of this inverse by localized processing. Whenever t^{RI} has a fast enough decay, such sectioning techniques can be used without producing any visible distortions.

3.2 Adaptive block-size selection

As in other LPA-ICI based deconvolution algorithms (e.g. [18], [17], [6]), the adaptive scale selection is implemented directly onto the regularized inverse z^{RI}. In order for the approach to be practical, it is fundamental to be able to calculate the estimates $\left\{\hat{y}^{\mathrm{LPA}}_{h_j}(x)\right\}_{j=1}^{J}$ and their standard-deviations for all $x \in X$ efficiently. We achieve this by exploiting the zero-order local polynomial model, using uniform weighting over nested square blocks $\left\{B_{x,h_j}\right\}_{j=1}^{J}$. In practice, this means that the estimates $\hat{y}^{\mathrm{LPA}}_{h_j}$ are calculated as the convolution

$$\hat{y}^{\mathrm{LPA}}_{h_j} = z^{RI} \circledast 1_{B_{h_j}} \tag{20}$$

where $1_{B_{h_j}}$ is a uniform kernel equal to h_j^{-2} on a square support (block) of size $h_j \times h_j$. The kernel, which is separable, is assumed to be located at the origin 0. For a fixed x, Equation (20) is equivalent to

$$\hat{y}^{\mathrm{LPA}}_{h_j}(x) = \frac{1}{\left|B_{x,h_j}\right|} \sum_{\upsilon \in B_{x,h_j}} z^{RI}(\upsilon),$$

where B_{x,h_j} is a square block of size $h_j \times h_j$ located at x and $\left|B_{x,h_j}\right| = h_j^2$.

These convolutions are computed for a set $H = \{h_1 < \cdots < h_J\}$ of increasing h_j and with nested supports of the kernels $1_{B_{h_j}}$, $j = 1,\ldots,J$. Using established terminology, such h_j's are called *scales*, and H is thus the *set of scales*.

Calculation of the standard-deviations $\sigma_{\hat{y}^{\mathrm{LPA}}_{h_j}}$ is the most delicate issue here. Exploiting the independence of the noise and the associativity of the convolution, one can ob-

tain from $\hat{y}^{\mathrm{LPA}}_{h_j} = z \circledast t^{RI} \circledast 1_{B_{h_j}}$ that

$$\sigma^2_{\hat{y}^{\mathrm{LPA}}_{h_j}} = \sigma_z^2 \circledast \left(t^{RI} \circledast 1_{B_{h_j}}\right)^2 = \quad (21)$$

$$= \mathcal{F}^{-1}\left(\Sigma_z^2 \cdot \mathcal{F}\left(\left(t^{RI} \circledast 1_{B_{h_j}}\right)^2\right)\right), \quad (22)$$

where $\Sigma_z^2 = \mathcal{F}\left(\sigma_z^2\right)$ is the Fourier transform of the space-varying variance of z. Again, we estimate σ_z^2 directly from the observations z, thus Σ_z^2 in the above formula is replaced by $\hat{\Sigma}_z^2 = \mathcal{F}(\rho(|z|))$.

The expression on the right-hand side of (22) is purposely in a mixed space-frequency domain. It leads to a significant reduction of complexity using first a convolution (against $1_{B_{h_j}}$ which has a very small support), and then a multiplication in Fourier domain (t^{RI} is an IIR filter, and so is $\left(t^{RI} \circledast 1_{B_{h_j}}\right)^2$). Let us remark that (20) and (22) give the estimates $\hat{y}^{\mathrm{LPA}}_{h_j}(x)$ and their variance $\sigma^2_{\hat{y}^{\mathrm{LPA}}_{h_j}}(x)$ for all $x \in X$.

Adaptive scale (i.e. block-size) selection is achieved by the ICI rule, where the confidence intervals $\mathcal{D}_j$ are constructed as in Equation (18), with $\hat{y}^{\mathrm{LPA}}_{h_j}(x)$ and $\sigma_{\hat{y}^{\mathrm{LPA}}_{h_j}(x)}$ given by (20) and (22), respectively. This yields, for every $x \in X$, an adaptive-size block $B_x^+ = B_{x,h^+(x)}$.

3.3 Local transform-domain deblurring

It was shown in [6], that the above zero-order LPA-ICI estimate can be interpreted as nonparametric local maximum-likelihood estimate. In such analysis, the zero-order modeling was a formal necessity to ensure local stability of the noise variance and thus give way to a linear inverse of the form (19). Nevertheless, in [6] it was also shown that by mildly relaxing the zero-order modeling and in particular by exploiting first-order or mixed-order LPA estimators, one can achieve a better restoration performance.

In this work, we relax the polynomial modeling even further, and utilize a (locally) complete transform to model the signal on the adaptive-size block obtained by the zero-order LPA-ICI. Therefore, we can preserve the assumption of stationarity of the noise, and at the same time we can reconstruct image details which cannot be represented accurately by the basic local-constant modeling.

It is important to remind that the LPA estimate $\hat{y}^{\mathrm{LPA}}_{h_j}(x)$ is simply the mean of the regularized inverse z^{RI} over B_{x,h_j}, and – up to a factor $\sqrt{|B_{x,h_j}|}$ – it coincides with the corresponding DC coefficient of a DCT or DFT transform or with the coarsest approximation coefficient of a full wavelet decomposition (with the transforms supported on B_j). Provided that the DC-term of the transform is preserved, the reconstructed estimate from transform-domain processing is then equivalent to the local zero-order LPA estimate plus a number of transform subband estimates.

3.3.1 Local estimate

We construct a local estimate $\hat{y}^{\mathrm{TRA}}_{B_x^+}$ of the true (noise and blur free) image y by synthesis with the transform elements $\boldsymbol{\psi}_{B_x^+} = \left\{\psi^{(i)}_{B_x^+}\right\}_i$

$$\hat{y}^{\mathrm{TRA}}_{B_x^+} = \sum_{i=1}^{h^2} \hat{\varphi}_{y,B_x^+}(i)\, \psi^{(i)}_{B_x^+} \quad (23)$$

where $\hat{\varphi}_{y,B_{x,h}}$ are estimates of the transform coefficients $\varphi_{y,B_x^+} = \left\langle y, \psi^{(i)}_{B_x^+}\right\rangle$ of y. By *local* we mean that the estimate $\hat{y}^{\mathrm{TRA}}_{B_x^+}$ is supported on the block B_x^+.

The estimated coefficients $\hat{\varphi}_{y,B_{x,h}}$ are obtained from the noisy coefficients $\varphi_{z^{RI},B_x^+}(i) = \left\langle z^{RI}, \psi^{(i)}_{B_x^+}\right\rangle$ using hard-thresholding. It is performed in the form

$$\hat{\varphi}_{y,B_x^+}(i) = \begin{cases} 0 & \text{if } \left|\varphi_{z^{RI},B_x^+}(i)\right| < \tau\sigma_{\varphi_{z^{RI},B_x^+}(i)} \text{ and } i > 1 \\ \varphi_{z^{RI},B_x^+}(i) & \text{if } \left|\varphi_{z^{RI},B_x^+}(i)\right| \geq \tau\sigma_{\varphi_{z^{RI},B_x^+}(i)} \text{ or } i = 1 \end{cases},$$

where $\tau = \gamma\sqrt{2\ln\left|B_x^+\right| + 1}$ is a size-dependent threshold parameter which is essentially the so-called "universal" threshold. The factor γ is a fixed constant, invariant with respect to the block-size. Only AC coefficients are thresholded, hence the DC is always preserved when the transform is inverted.

3.3.2 Standard-deviation of the coefficients (accurate)

While for the calculation of the coefficient $\varphi_{z^{RI},B_x^+}(i)$ it is possible to decouple the regularized inverse from the analyis element $\psi^{(i)}_{B_x^+}$, this cannot be done for the calculation of its standard deviation $\sigma_{\varphi_{z^{RI},B_x^+}(i)}$. The inverse operator must be considered here explicitly. Following the manipulations from Sections 2.1 and 2.2, and in particular Equation (17), we have

$$\varphi_{z^{RI},B_x^+}(i) = \left\langle z^{RI}, \psi^{(i)}_{B_x^+}\right\rangle = \left\langle z^{RI}, \overrightarrow{\left(\psi^{(i)}_{B^+_{0,h^+(x)}}\right)_x}\right\rangle = \quad (24)$$

$$= \left(z^{RI} \circledast \left(\widetilde{\psi^{(i)}_{B^+_{0,h^+(x)}}}\right)_0\right)(x) = \quad (25)$$

$$= \left(z \circledast \left(t^{RI} \circledast \left(\widetilde{\psi^{(i)}_{B^+_{0,h^+(x)}}}\right)_0\right)\right)(x). \quad (26)$$

Here, $\psi^{(i)}_{B^+_{0,h^+(x)}}$ is the i-th element of the transform supported on the block $B^+_{0,h^+(x)}$ of size $h^+(x) \times h^+(x)$ and located at 0 and $\left(\widetilde{\psi^{(i)}_{B^+_{0,h^+(x)}}}\right)_0$ is its copy mirrored about 0. From (26) it is straighforward to obtain the variance of $\varphi_{z^{RI},B_x^+}(i)$ as

$$\sigma^2_{\varphi_{z^{RI},B_x^+}(i)} = \left(\sigma_z^2 \circledast \left(t^{RI} \circledast \left(\widetilde{\psi^{(i)}_{B^+_{0,h^+(x)}}}\right)_0\right)^2\right)(x). \quad (27)$$

Observe that the same function in (27) yields also the variance of the i-th transform coefficient corresponding to the same transform localized at any other position $\xi \neq x$,

$$\sigma^2_{\varphi_{z^{RI},B_{\xi,h^+(x)}}(i)} = \left(\sigma_z^2 \circledast \left(t^{RI} \circledast \left(\widetilde{\psi^{(i)}_{B^+_{0,h^+(x)}}}\right)_0\right)^2\right)(\xi).$$

Calculation of this function can be done more efficiently in Fourier domain as

$$\sigma_z^2 \circledast \left(t^{RI} \circledast \left(\widetilde{\psi^{(i)}_{B^+_{0,h^+(x)}}}\right)_0\right)^2 = \quad (28)$$

$$= \mathcal{F}^{-1}\left(\Sigma_z^2 \cdot \mathcal{F}\left(\left(t^{RI} \circledast \left(\widetilde{\psi^{(i)}_{B^+_{0,h^+(x)}}}\right)_0\right)^2\right)\right). \quad (29)$$

The convolution (28) gives simultaneously the variance of the coefficients corresponding to all possible translated copies of the same basis function $\left(\overrightarrow{\psi^{(i)}_{B^+_{0,h^+(x)}}}\right)_u$, where u is an arbitrary translation parameter. Only the size of the block $h^+(x)$ and the index i of the basis element are fixed. There are only a very limited number of different possible block-sizes, so these variances can be precalculated.

3.3.3 Standard-deviation of the coefficients (approximate)

Further, important simplification can be obtained if we assume that the variance σ_z^2 is locally constant. In particular, we can assume that the variance of the observations z used for the estimation within the adaptive block, denoted as $\sigma^2_{z(B_x^+)}$, is constant and equal to

$$\rho\left(\left|\varphi_{z,B_x^+}(1)\right|/\sqrt{\left|B_x^+\right|}\right).$$

Thus, the standard-deviation of each transform coefficient is simply

$$\sigma_{\varphi_{zRI},B_x^+(i)} = \sigma_{z(B_x^+)}\left\| t^{RI} \circledast \left(\widetilde{\psi^{(i)}_{B^+_{0,h^+(x)}}}\right)_0 \right\|_2. \qquad (30)$$

We can approximate $\sigma_{z(B_x^+)}$ as $\sigma_{\hat{y}^{\mathrm{LPA}}_{h^+(x)}} / \left\| t^{RI} \circledast 1_{B_{h_j}} \right\|_2$ (see Equation (21)) and obtain the following approximation for $\sigma_{\varphi_{zRI},B_x^+(i)}$:

$$\sigma_{\varphi_{zRI},B_x^+(i)} \approx \frac{h^+(x)\cdot\sigma_{\hat{y}^{\mathrm{LPA}}_{h^+(x)}}}{\left\| t^{RI}\circledast\left(\widetilde{\psi^{(1)}_{B^+_{0,h^+(x)}}}\right)_0\right\|_2}\left\| t^{RI}\circledast\left(\widetilde{\psi^{(i)}_{B^+_{0,h^+(x)}}}\right)_0\right\|_2. \qquad (31)$$

Finally, as it is shown in [9], the convolutions which appear in (30) and (31) can be computed very efficiently in a downsampled Fourier domain (e.g. of size 32x32), with no significant loss of accuracy. Note also that in the same convolutions the translation and mirroring of the convolution kernel do not affect to the ℓ^2-norm.

3.4 Aggregation of the local estimates

For every $x \in X$, inverse transformation of the hard-thresholded coefficients $\hat{\varphi}_{y,B_x^+}$ yields a local estimate of the signal $\hat{y}^{\mathrm{TRA}}_{B_x^+} : B_x^+ \to \mathbb{R}$ supported on the adaptive-size block B_x^+. Usually, blocks corresponding to adjacent pixels overlap, hence for a single pixel there can be multiple estimates available.

3.4.1 Averaging with adaptive weights

In order to combine all these *local* estimates into a single *global* estimate $\hat{y}^{RI}$, we do averaging using adaptive weights which depends on the local statistics of the signal. How this averaging is done can have a dramatic impact on the quality of the final estimate. As in [8], we use the following convex combination with adaptive weights w_x which depend on the variance as well as on the size of the reconstructed local estimates:

$$\hat{y}^{RI} = \frac{\sum_{x\in X} w_x \hat{y}^{\mathrm{TRA}}_{B_x^+}}{\sum_{x\in X} w_x \chi_{B_x^+}}, \quad w_x = \frac{1}{\sigma^2_{\hat{y}^{\mathrm{TRA}}_{B_x^+}}\left|B_x^+\right|}, \qquad (32)$$

where $\chi_{B_x^+}$ is the characteristic function of B_x^+. In (32) we implicitly assume that the local estimates $\hat{y}^{\mathrm{TRA}}_{B_x^+}$ are zero-padded outside their domain B_x^+.

We estimate the variance $\sigma^2_{\hat{y}^{\mathrm{TRA}}_{B_x^+}}$ as the sum of the variances of all filtered coefficients $\left\{\hat{\varphi}_{y,B_x^+}(i)\right\}_i$,

$$\sigma^2_{\hat{y}^{\mathrm{TRA}}_{B_x^+}} = \sum_i \sigma^2_{\varphi_{zRI},B_x^+(i)}.$$

The particular form of the weights w_x is justified with essentially the same arguments as for the adaptive block-size denoising algorithm from [8]. In particular, the total variance $\sigma^2_{\hat{y}^{\mathrm{TRA}}_{B_x^+}}$ can be seen as an upper bound for the pointwise residual-noise variance of the local estimate $\hat{y}^{\mathrm{TRA}}_{B_x^+}$ (such pointwise variance is not necessarily uniform over B_x^+), and the extra factor $\left|B_x^+\right|$ addresses the correlation that exists between overlapping blocks (the number of overlapping blocks is loosely proportional to their size).

Qualitatively speaking, the weights (32) favour estimates which correspond to sparser representations (little energy is preserved after shrinkage) and at the same time avoid that estimates with a small support (thus representing image details) are oversmoothed by other overlapping estimates which have a large support (which usually are strongly correlated among themselves and outnumber estimates of a smaller support).

3.5 Wiener stage

The estimate $\hat{y}^{RI} : X \to \mathbb{R}$ obtained from (32) can already be considered as a satisfactory estimate of the unknown signal y. Alternatevely, it can be used as a reference estimate for the second part of the algorithm, which is based on Wiener filtering. Although this second stage is optional, it usually provides further improvement to the restoration performance.

The reference estimate is used in several ways: first the regularized inverse (19) is replaced by a more accurate regularized-Wiener inverse, then for the local-transform domain shrinkage we use Wiener filtering instead of hard-thresholding, and finally the noise estimates can be improved by using $\hat{y}^{RI} \circledast \upsilon$ instead of z inside the variance function ρ. Let us present all these changes more in detail.

3.5.1 Regularized Wiener Inverse

The regularized Wiener inverse (RWI) z^{RWI} is computed as

$$z^{RWI} = \mathcal{F}^{-1}(T^{RWI} Z), \qquad T^{RWI} = \frac{\overline{V}|Y|^2}{|VY|^2 + \varepsilon_2^2 \Phi_\eta}, \qquad (33)$$

where $\varepsilon_2 > 0$ is a regularization parameter. Since the spectrum $|Y|^2$ of the true image is unknown, the estimate $\hat{y}^{RI}$ from the RI stage is used quite naturally as a "pilot" estimate in the Wiener filtering. It means that $|Y|^2$ in (33) is replaced by $|\hat{Y}^{RI}|^2$.

3.5.2 Adaptive block-sizes

The size of the adaptive blocks is calculatd as in Section 3.2. However, in the expressions we replace z^{RI}, T^{RI}, t^{RI} with z^{RWI}, T^{RWI}, t^{RWI}, respectively. The estimate of the variance σ_z^2, used in the formulas for the standard devations (21-22), can now be computed as $\hat{\sigma}_z^2 = \rho(\hat{y}^{RI} \circledast \upsilon)$.

3.5.3 Wiener filtering in transform domain

The local estimates $\hat{y}^{\mathrm{TRA}}_{B^+_x}$ are obtained as in (23), with the estimates of the transform coefficients $\{\hat{\varphi}_{y,B^+_x}(i)\}_i$ given by

$$\begin{aligned}\hat{\varphi}_{y,B^+_x}(i) &= \omega_x \varphi_{z^{RI},B^+_x}(i),\\ \omega^{(i)}_x &= \frac{\left|\varphi_{\hat{y}^{RI},B^+_x}(i)\right|^2}{\left|\varphi_{\hat{y}^{RI},B^+_x}(i)\right|^2+\sigma^2_{\varphi_{z^{RWI},B^+_x}(i)}},\end{aligned} \quad (34)$$

where $\{\omega^{(i)}_x\}_i$ are the attenuation coefficients of the Wiener filter.

3.5.4 Variances, adaptive weights

Calculation of the coefficient variances $\sigma^2_{\varphi_{z^{RWI},B^+_x}(i)}$ from (34) can be done with the same formulas as in Sections 3.3.2-3.3.3, provided that z^{RI} is changed to z^{RWI}.

These variances are needed also for the adaptive weights w_x used for the aggregation of the local estimates. The weights are defined analogously to (32), as the reciprocal of an estimate of the total variances of the local estimate, times the size of the adaptive block:

$$w_x = \frac{1}{\left|B^+_x\right|\sum_i\left(\omega^{(i)}_x\right)^2\sigma^2_{\varphi_{z^{RWI},B^+_x}(i)}}.$$

By aggregating all the local Wiener estimates, we obtain a global Wiener estimate, which we denote by $\hat{y}^{RWI}: X \to \mathbb{R}$.

4. EXPERIMENTAL RESULTS

The above algorithm has been tested with various choices of transforms. For the hard-thresholding (regularized inverse) stage, it appeared that the DCT and the Daubechies wavelets are quite effective. For the Wiener filtering (regularized-Wiener inverse) stage, the DCT produced the best results in terms of objective criteria (e.g. PSNR). However, in some cases it is possible to obtain visually more pleasant results if the DCT is replaced by some orthogonal or biorthogonal wavelet. Nevertheless, the differences are not significant and in practice the considered application and algorithm's implementation shall determine the most suitable choice of transforms.

For the simulation experiments presented in this paper, we selected the algorithm parameters and transforms following objective criteria rather than subjective inspection. In particular, we use Daubechies wavelets with three vanishing moments for the hard-thresholding and the DCT for the Wiener filtering. Other parameters are as follows: H={4,8,16}, $\Gamma = 1.5$, $\gamma = 1$, ε_1 =0.014 and ε_2 =0.040. Standard-deviations are calculated as in Section 3.3.3.

Firstly, we compare our newly proposed method against the *Anisotropic LPA-ICI RI-RWI* deconvolution algorithm developed for Poissonian observations [6], and consider the deblurring experiment where an image is heavily blurred by a 9×9 "boxcar" uniform PSF and its noisy Poissonian distributed observation are created using χ=17600, as in Equation (3). This corresponds to a *BSNR* (blurred-*SNR*) of 32.5dB. Table 1 presents the results of the two algorithms for a few standard test images: in terms of *ISNR* (improvement of *SNR*), the new transform-based method consistently outperforms the Poissonian Anisotropic LPA-ICI.

Figure 1 shows fragments of the *Cameraman* image from this comparison. While both methods demonstrate similar the ability in restoring the details of the image, the estimate of the new method is more clear, with fewer periodic patterns showing in the background. Image contrast is also improved.

Figure 2 further illustrates the performance of the new method. For this Figure, the observations are generated using a separable kernel v=[1,4,6,4,1]$^{\mathrm{T}}$[1,4,6,4,1]/256 and accurately simulating the noise of the raw-data from a Nokia cameraphone's CMOS sensor (dim lighting conditions, with analog gain set to 8dB). The estimates are clean, with very few artifacts and good restoration of fine details.

5. CONCLUSIONS AND PERSPECTIVE

The proposed method produces estimates which are good, in terms of objective criteria as well as of visual quality. The method is quite versatile, and can be used for a very general class of observations with signal-dependent noise. Any transform which provides a DC term can be utilized for the filtering (including fast transforms).

The transform-domain approach allows to introduce very easily other image processing operations (e.g. compression, sharpening, etc.) within a unique framework. Current work concentrates on embedding a color interpolation (demosaicking) algorithm within the presented deconvolution approach.

REFERENCES

[1] Abramovich, F., and B.W. Silverman, "Wavelet Decomposition Approaches to Statistical Inverse Problems", *Biometrika*, vol. 85, no. 1., pp. 115-129, March, 1998.

[2] Aghdasi, F., R.K. Ward, "Reduction of boundary artifacts in image restoration", *IEEE Transactions on Image Processing*, vol. 5, no. 4, pp. 611-618, April 1996.

[3] Bertero, M., and P. Boccacci, *Introduction to inverse problems in imaging*. Inst. of Physics Publishing, 1998.

[4] Donoho, D., "Nonlinear solution of linear inverse problems by wavelet–vaguelette decomposition", *Appl. Comput. Harmon. Anal.*, vol. 2, pp. 101-126, 1995.

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

[6] Foi, A., S. Alenius, M. Trimeche, V. Katkovnik, and K. Egiazarian, "A spatially adaptive Poissonian image deblurring", *Proc. IEEE 2005 Int. Conf. Image Processing, ICIP 2005,* pp. 925-928, September 2005.

[7] Foi, A., R. Bilcu, V. Katkovnik, and K. Egiazarian, "Anisotropic local approximations for pointwise adaptive signal-dependent noise removal", *Proc. of XIII European Signal Proc. Conf., EUSIPCO 2005*, Antalya, September 2005.

[8] Foi, A., R. Bilcu, V. Katkovnik, and K. Egiazarian, "Adaptive-Size Block Transforms for Signal-Dependent Noise Removal", *Proc. 7th Nordic Signal Processing Symposium, NORSIG 2006*, Reykjavik, Iceland, June 2006.

[9] Foi, A., K. Dabov, V. Katkovnik, and K. Egiazarian, "Shape-adaptive DCT for denoising and image reconstruction", *Proc. SPIE El. Imaging 2006, Image Process.: Algorithms and Systems V*, 6064A-18, January 2006.

[10] Foi, A., D. Paliy, V. Katkovnik, and K. Egiazarian, "Anisotropic nonparametric image restoration demobox" (MATLAB software), *LASIP (Local Approximations in Signal and Image Processing) Project*, `http://www.cs.tut.fi/~lasip/` , 2005.

[11] Goldenshluger, A., and A. Nemirovski, "On spatial adaptive estimation of nonparametric regression", *Math. Meth. Statistics*, vol. 6, pp. 135-170, 1997.

Algorithm	*Cameraman*	*Peppers*	*Lena*	*Boats*
Adaptive Block-Size Transform Deblurring	**6.63**	**7.88**	**5.74**	**6.20**
Poissonian Anisotropic LPA-ICI RI-RWI	6.55	7.72	5.54	5.95

Table 1: *ISNR* (dB) for *Cameraman* image for Poissonian image deblurring.

Figure 1: *Cameraman* image, from left to right: blurred noisy observation (9x9 box-car PSF, Poissonian noise χ=17600), adaptive block-sizes (dark gray h_{j^+}=4, white h_{j^+}=16), deblurred estimate $\hat{y}^{\mathrm{RWI}}$ (*ISNR* 6.63dB), and Poissonian Anisotropic LPA-ICI RW estimate (*ISNR* 6.55dB).

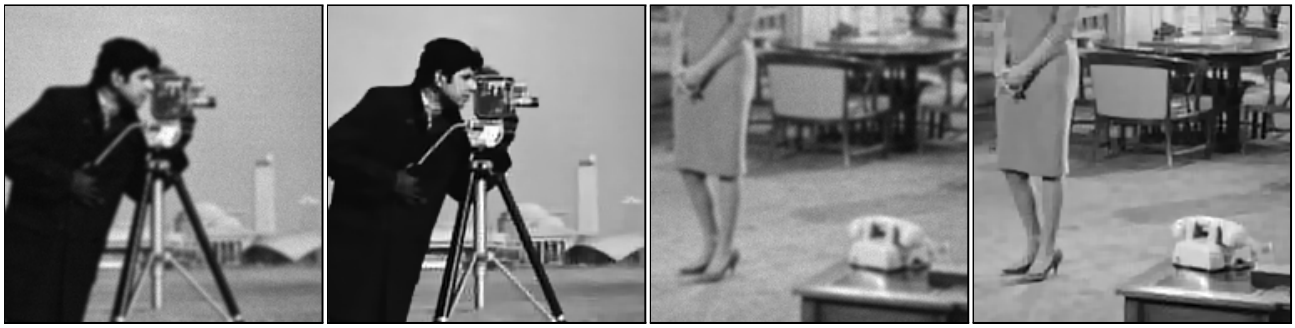

Figure 2: Fragments of *Cameraman* and *Couple* images blurred with a separable spline kernel and corrupted by simulated noise as in the raw-data from the CMOS sensor of a Nokia cameraphone and the corresponding estimates reconstructed using the proposed algorithm.

[12] Jiang, S.S., and A.A. Sawchuk, "Noise updating repeated Wiener filter and other adaptive noise smoothing filters using local image statistics", *Appl. Opt.*, vol. 25, pp. 2326-2337, 1986.

[13] Kalifa, J., and S. Mallat, "Thresholding Estimators for Linear Inverse Problems and Deconvolutions", *Annals of Statistics*, vol. 31, no. 1, pp. 58-109, 2003.

[14] Lee, N., *Wavelet-Vaguelette Decompositions and Homogeneous Equations*, Ph.D. thesis, Purdue University, December 1997.

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

[16] Katkovnik, V., K. Egiazarian, and J. Astola, "Adaptive window size image de-noising based on intersection of confidence intervals (ICI) rule", *J. Math. Imaging and Vision*, vol. 16, no. 3, pp. 223-235, 2002.

[17] Katkovnik, V., K. Egiazarian, and J. Astola, "A spatially adaptive nonparametric regression image deblurring", *IEEE Trans. on Image Processing*, vol. 14, no. 10, pp. 1469-1478, October 2005.

[18] Katkovnik, V., A. Foi, K. Egiazarian, and J. Astola, "Directional varying scale approximations for anisotropic signal processing", *Proc. XII European Signal Proc. Conf., EUSIPCO 2004,* Vienna, pp. 101-104, September 2004.

[19] Katkovnik, V., A. Foi, K. Egiazarian, and J. Astola, "Anisotropic local likelihood approximations", *Proc. of Electronic Imaging 2005*, 5672-19, 2005.

[20] Lu, H., Y. Kim, and J.M.M. Anderson, "Improved Poisson intensity estimation: denoising application using poisson data," *IEEE Trans. Image Processing,* vol. 13, no. 8, pp. 1128-1135, 2004.

[21] Neelamani, R., H. Choi, and R. Baraniuk, "Forward: Fourier-wavelet regularized deconvolution for ill-conditioned systems", *IEEE Trans. on Image Proc.,* vol. 52, no. 2, 2004.

[22] Öktem, H., V. Katkovnik, K. Egiazarian, and J. Astola, "Local adaptive transform based image de-noising with varying window size", *Proc. IEEE Int. Conference on Image Process., ICIP 2001,* Thessaloniki, Greece, pp. 273-276, October 2001.

[23] Rangarayyan, R.M., M. Ciuc, and F. Faghih, "Adaptive-neighborhood filtering of images corrupted by signal-dependent noise", *Appl. Opt.*, vol. 37, pp. 4477-4487, 1998.

[24] Reeves, S.J., "Fast image restoration without boundary artifacts", *IEEE Transactions on Image Processing*, vol. 14, no. 10, pp. 1448-1453, October 2005.

[25] Timmermann, K.E., and R. Nowak, "Multiscale modeling and estimation of Poisson processes with application to photon-limited imaging", *IEEE Trans. Inf. Theory,* vol. 45, no. 3, pp. 846-862, 1999.

[26] Willett, R.M., and R. Nowak, "Platelets: a multiscale approach for recovering edges and surfaces in photon-limited medical imaging", *IEEE Trans. Medical Imaging,* vol. 22, no. 3, pp. 332-350, 2003.