

Shape-Adaptive DCT for Denoising and Image Reconstruction

Alessandro Foi, Kostadin Dabov, Vladimir Katkovnik, Karen Egiazarian

Institute of Signal Processing, Tampere University of Technology, 33101, Tampere, Finland.

ABSTRACT

The shape-adaptive DCT (SA-DCT) can be computed on a support of arbitrary shape, but retains a computational complexity comparable to that of the usual separable block DCT. Despite the near-optimal decorrelation and energy compaction properties, application of the SA-DCT has been rather limited, targeted nearly exclusively to video compression. It has been recently proposed by the authors⁸ to employ the SA-DCT for still image denoising. We use the SA-DCT in conjunction with the directional LPA-ICI technique, which defines the shape of the transform's support in a pointwise adaptive manner. The thresholded or modified SA-DCT coefficients are used to reconstruct a local estimate of the signal within the adaptive-shape support. Since supports corresponding to different points are in general overlapping, the local estimates are averaged together using adaptive weights that depend on the region's statistics. In this paper we further develop this novel approach and extend it to more general restoration problems, with particular emphasis on image deconvolution. Simulation experiments show a state-of-the-art quality of the final estimate, both in terms of objective criteria and visual appearance. Thanks to the adaptive support, reconstructed edges are clean, and no unpleasant ringing artifacts are introduced by the fitted transform.

Keywords: denoising, deconvolution, deblurring, restoration, transforms, shape-adaptive, DCT, anisotropic.

1. INTRODUCTION

The two-dimensional separable DCT, computed on a square or rectangular support, is a well established and very efficient transform in order to achieve a sparse representation of image blocks. For natural images, its decorrelating performance is close to that of the optimum Karhunen-Loève transform. Thus, the DCT has been successfully used as the key element in many compression and denoising applications. However, in presence of singularities or edges such near-optimality fails. Because of the lack of sparsity, edges cannot be coded or restored effectively, and ringing artifacts arising from the Gibbs phenomenon become visible.

In the last decade, significant research has been made towards the development of region-oriented, or *shape-adaptive*, transforms. The main intention is to construct a system (frame, basis, etc.) that can efficiently be used for the analysis and synthesis of arbitrarily shaped image segments, where the data exhibit some uniform behavior.

Initially, Gilge^{11,12} proposed to consider the orthonormalization of a (fixed) set of generators restricted to the arbitrarily shaped region of interest. These generators could be a basis of polynomials or – for example – a 2D block DCT basis, thus yielding a “shape-adapted” DCT transform. Orthonormalization can be performed by the standard Gram-Schmidt procedure and the obtained orthonormal basis is supported on the region. Because the region-adapted basis needs to be recalculated for each differently shaped region and because the basis elements are typically non-separable, the overall method presents a rather high computational cost. While even today it is considered as one of the best solutions to the region-oriented transforms problem, Gilge's approach is clearly unsuitable for real-time applications, and faster transforms were sought.

A more computationally attractive approach, namely the *shape-adaptive DCT* (SA-DCT), has been proposed by Sikora et al.^{29,31}. It is based on a cascaded application of standard one dimensional varying-length DCT transforms first on the rows and then on the columns. Thus, the SA-DCT does not require costly matrix inversions or iterative orthogonalizations and can be interpreted as a direct generalization of the classical 2D block DCT transform. In particular, the SA-DCT and the block DCT (separable) have the same computational complexity and in the special case of a square the two transforms exactly coincide. Therefore, the SA-DCT has received considerable interest from the MPEG community, eventually becoming part of the MPEG-4 standard²³. Fast and low-power hardware implementations of the SA-DCT are currently becoming available²².

The SA-DCT has been shown^{1,19,29,30} to provide a compression efficiency comparable to those of more computationally complex transforms, such as¹². The good decorrelation and energy compaction properties on

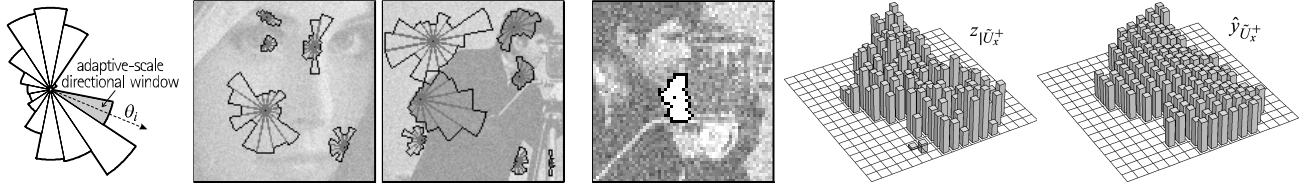


Figure 1. From left to right: sectorial structure of the anisotropic neighborhood achieved by combining a number of adaptive-scale directional windows; some of these windows selected by the *ICI* for the noisy *Lena* and *Cameraman* images; a detail of the noisy *Cameraman* showing an adaptive-shape neighborhood \tilde{U}_x^+ determined by the anisotropic *LPA-ICI* procedure, and the image intensity corresponding to this region before and after hard-thresholding in SA-DCT domain.

which this efficiency depends are also the primary characteristics sought for any transform-domain denoising algorithm. In this sense, the SA-DCT features a remarkable potential not only for video compression applications, but also for image and video denoising.

However, this potential has been apparently ignored by the image denoising and restoration community. While this indifference may seem rather surprising, there are sound reasons that can justify it.

The shape-adaptive coding of *noise-free* video objects always assumes that the boundary of these objects is known. This information can be obtained either from a-priori knowledge (e.g. motion estimation, chroma keying, layered structure), or it can be estimated from the data with one of the many automated segmentation algorithms suitable for this purpose. On the contrary, obtaining an accurate and robust segmentation of *noisy* data constitutes an extremely more complex task than the region-oriented coding itself. Unlike in MPEG-4 video coding, such a segmentation cannot be reasonably assumed to be known a-priori. This very aspect may be identified as the principal reason why the SA-DCT had not been used for the restoration of noisy images.

A breakthrough in the use of SA-DCT for image filtering was reported by the authors in⁸. An image denoising algorithm was proposed which exploits the anisotropic LPA-ICI technique^{6,16}, in order to obtain – directly from the noisy data – an appropriate shape on which to apply the shape-adaptive transform.

In this paper we further develop the approach and extend it to image restoration, proposing an image deblurring algorithm based on adaptive regularized deconvolution in overcomplete SA-DCT domain. Furthermore, the use of Gilge’s “shape-adapted” DCT is considered as an alternative to the faster Sikora’s SA-DCT. Efficient implementations based on this more complex transform are presented.

The paper is organized as follows: first, we recall the main points of the anisotropic LPA-ICI technique and present various aspects of the shape-adaptive DCT transform; the developed LPA-ICI-driven SA-DCT denoising algorithm is then introduced, followed by its extension for image deblurring; we conclude with a number of experimental results that demonstrate the advanced performance of the proposed algorithms.

2. OBSERVATION MODEL AND NOTATION

Let us introduce the considered observation model and some notation. First, we consider noisy observations z of the form

$$z(x) = y(x) + \eta(x), \quad x \in X, \quad (1)$$

where y is the original image, $\eta(x) \sim \mathcal{N}(0, \sigma^2)$ is white Gaussian noise, x is a spatial variable belonging to the image domain $X \subset \mathbb{Z}^2$. Given a function $f : X \rightarrow \mathbb{R}$, a subset $U \subset X$, and a function $g : U \rightarrow \mathbb{R}$, we denote by $f|_U : U \rightarrow \mathbb{R}$ the restriction of f on U , $f|_U(x) = f(x) \forall x \in U$, and by $g^{lX} : X \rightarrow \mathbb{R}$ the zero-extension of g to X , $(g^{lX})|_U = g$ and $g^{lX}(x) = 0 \forall x \in X \setminus U$. The characteristic (indicator) function of U is defined as $\chi_U = 1|_U^{lX}$. We denote by $|U|$ the cardinality (i.e. the number of its elements) of U .

3. ANISOTROPIC LPA-ICI

The approach is based on the Intersection of Confidence Intervals (ICI) rule, a method originally developed for pointwise adaptive estimation of 1D signals^{13,15}. The idea has been generalized for 2D image processing, where adaptive-size quadrant windows have been used¹⁷. Significant improvement of this approach has been

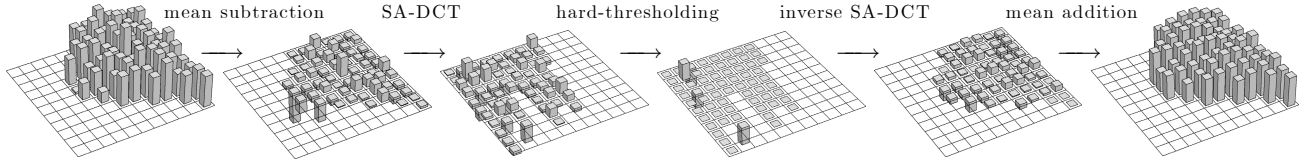


Figure 2. Hard-thresholding in SA-DCT domain. The image data on an arbitrarily shaped region is subtracted of its mean. The zero-mean data is then transformed and thresholded. After inverse transformation, the mean is added back.

achieved on the basis of anisotropic directional estimation^{6,16}. Multidirectional sectorial-neighborhood estimates are calculated for every point and the ICI rule is exploited for the adaptive selection of the size of each sector. Thus, the estimator is anisotropic and the shape of its support adapts to the structures present in the image. In Figure 1 we show some examples of these anisotropic neighborhoods for the *Lena* and *Cameraman* images. The developed anisotropic estimates are highly sensitive with respect to change-points, and allow to reveal fine elements of images from noisy observations.

Let us present the overall method in more detail. For every specified direction θ_k , $k = 1, \dots, K$, a varying-scale family of directional-LPA (local polynomial approximation) convolution kernels¹⁶ $\{g_{h,\theta_k}\}_{h \in H}$ is used to obtain a corresponding set of directional varying-scale estimates $\{\hat{y}_{h,\theta_k}\}_{h \in H}$, $\hat{y}_{h,\theta_k} = z \otimes g_{h,\theta_k}$, $h \in H$. These estimates are then compared according to the ICI rule^{13,15}, and as a result an adaptive scale $h^+(x, \theta_k) \in H$ is defined for every $x \in X$ and for every direction θ_k . The corresponding adaptive-scale estimates $\hat{y}_{h^+(x,\theta_k),\theta_k}(x)$ are then “fused” together in an adaptive convex combination^{6,16} in order to yield the final anisotropic LPA-ICI estimate.

However, in this paper we are not interested in this anisotropic estimate. Instead, we consider only the adaptive neighborhood U_x^+ , constructed as the union of the supports of the directional adaptive-scale kernels $g_{h^+(x,\theta_k),\theta_k}$, $U_x^+ = \bigcup_{k=1}^K \text{supp } g_{h^+(x,\theta_k),\theta_k}$, which we use as the support for a shape-adaptive transform. Observe that, being convolution kernels, the LPA kernels g_{h,θ_k} are always “centered” at the origin, therefore U_x^+ is a neighborhood of the origin. The actual adaptive neighborhood of x , which contains the observations that are used for estimation, is instead $\tilde{U}_x^+ = \{v \in X : (x - v) \in U_x^+\}$. The neighborhoods shown in Figure 1 are in fact examples of \tilde{U}_x^+ for a few points $x \in X$.

Let us remark that there is a substantial difference between image segmentation, in which the image is decomposed in a limited number ($\ll |X|$) of non-overlapping subsets (image segments), and the anisotropic LPA-ICI, which for every $x \in X$ provides an adaptive neighborhood \tilde{U}_x^+ of x . In particular, because of the nonparametric nature of the procedure, neighborhoods corresponding to adjacent points do usually overlap.

4. SHAPE-ADAPTIVE DCT

The SA-DCT^{29,31} is computed by cascaded application of one dimensional varying-length DCT transforms first on the rows and then on the columns that constitute the considered region. Several improvements over its original definition have been proposed. In this paper, we concentrate only on the most significant¹⁹, which concern with the normalization of the transform and with the subtraction of the mean and which have a fundamental impact on the use of the SA-DCT for image filtering.

4.1. Orthonormal Shape-Adaptive DCT

The normalization of the SA-DCT is obtained by normalization of the individual one-dimensional transforms that are used for transforming the rows and the columns. In terms of their basis elements, they are defined as:

$$\psi_{L,m}^{\text{1D-DCT}}(n) = c_m \cos\left(\frac{\pi m(2n+1)}{2L}\right), \quad m, n = 0, \dots, L-1, \quad c_0 = \sqrt{1/L}, \quad c_m = \sqrt{2/L}, \quad m > 0. \quad (2)$$

Here L stands for the length of the row or column to be transformed. The normalization in (2) is indeed the most natural choice, since in this way all the transforms used are orthonormal and the corresponding matrices belong to the orthogonal group. Therefore, the SA-DCT – which can be obtained by composing two orthogonal matrices – is itself an orthonormal transform. A different normalization of the 1D transforms would produce, on an arbitrary shape, a 2D transform that is non-orthogonal (for example as in^{29,31}, where $c_0 = \sqrt{2}/L$ and $c_m = 2/L$ for $m > 0$).

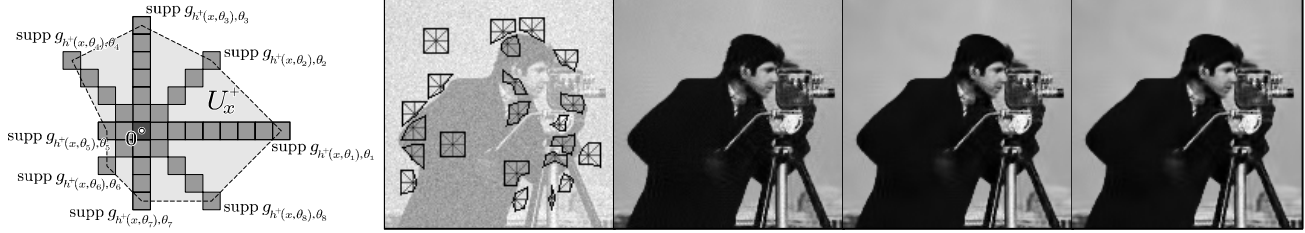


Figure 3. Fast implementation of the *LPA-ICI* anisotropic neighborhoods. “Linewise” one-dimensional directional *LPA* kernels are used for 8 directions. The anisotropic neighborhood U_x^+ is constructed as the polygonal hull of the adaptive-scale kernels’ supports (left). Some examples of the anisotropic neighborhoods \tilde{U}_x^+ used for SA-DCT filtering of the noisy *Cameraman* image (right), $\sigma=25$. Only the adaptive scales h^+ are needed to construct the neighborhoods. The corresponding Anisotropic LPA-ICI SA-DCT hard-thresholding estimate \hat{y} (7) and Wiener estimate \hat{y}^{wi} (12), and Gilge’s “shape-adapted” DCT Wiener estimate (*PSNR*=28.87, 29.10 dB, and 29.22 dB, respectively). In all the examples $h=1,2,3,5,7,9$.

Let us denote by $T_U : \mathcal{U} \rightarrow \mathcal{V}_U$ the orthonormal SA-DCT transform obtained for a region $U \subset X$, where $\mathcal{U} = \{f : U \rightarrow \mathbb{R}\}$ and $\mathcal{V}_U = \{\varphi : V_U \rightarrow \mathbb{R}\}$ are function spaces and $V_U \subset \mathbb{Z}^2$ indicates the domain of the transform coefficients. Let $T_U^{-1} : \mathcal{V}_U \rightarrow \mathcal{U}$ be the inverse transform of T_U . We indicate the thresholding (or quantization) operator as Υ . Thus, the SA-DCT-domain processing of the observations z on a region U can be written as $\hat{y}_U = T_U^{-1}(\Upsilon(T_U(z|_U)))$, $\hat{y}_U : U \rightarrow \mathbb{R}$. From the orthonormality of T and the model (1) follows that $T_U(z|_U) = T_U(y|_U) + \bar{\eta}$, where $\bar{\eta} = T_U(\eta|_U)$ is again Gaussian white noise with variance σ^2 and zero mean.

4.2. Mean subtraction

There is, however, an adverse consequence of the normalization (2). Even if the signal restricted to the shape $z|_U$ is constant, the reconstructed \hat{y}_U is usually non-constant. In¹⁹ this behavior is termed as “mean weighting defect”, and it is proposed there to attenuate its impact by applying the orthonormal SA-DCT on the zero-mean data which is obtained by subtracting from the initial data z its mean. After the inversion, the mean is added back to the reconstructed signal $\hat{y}_U : U \rightarrow \mathbb{R}$:

$$\hat{y}_U = T_U^{-1}(\Upsilon(T_U(z|_U - m_U(z)))) + m_U(z), \quad (3)$$

where $m_U(z) = \frac{1}{|U|} \sum_{x \in U} z(x)$ is the mean of z on U .

Although this operation – which is termed “DC separation” – is not fully justified from the approximation theory standpoint (because $m_U(z)$ is calculated from the noisy data, and by subtracting it the noise in the coefficients is no longer white), it produces visually superior results without affecting to the objective restoration performance.

An illustration of the SA-DCT domain hard-thresholding, performed according to (3) is given in Figure 2.

5. ANISOTROPIC *LPA-ICI*-DRIVEN SA-DCT DENOISING

We use the anisotropic adaptive neighborhoods \tilde{U}_x^+ defined by the *LPA-ICI* as supports for the SA-DCT, as shown in Figure 1. In this way, we ensure that data are represented sparsely in the transform domain, significantly improving the effectiveness of thresholding.

In practice, we do not need a variety of different shapes as broad as in the examples of Figure 1 and a much simplified neighborhood structure is used in our implementation. Narrow one-dimensional “linewise” directional *LPA* kernels $\{g_{h, \theta_k}\}_{h \in \{1,2,3,5,7,9\}}$ are used for $K = 8$ directions, and after the *ICI*-based selection of the adaptive-scales $\{h^+(x, \theta_k)\}_{k=1}^8$ the neighborhood U_x^+ is the octagon constructed as the polygonal hull of $\{\text{supp } g_{h^+(x, \theta_k), \theta_k}\}_{k=1}^8$. Such neighborhoods are shown in Figure 3. Although the supports obtained in this way have relatively simple shapes when compared to the more general examples of Figure 1, we found that this is not a significant restriction. On the contrary, a more regular boundary of the transform’s support usually improves the efficiency of the SA-DCT¹.

5.1. Local estimates

For every point $x \in X$, we construct a *local* estimate $\hat{y}_{\tilde{U}_x^+} : \tilde{U}_x^+ \rightarrow \mathbb{R}$ of the signal y by thresholding in SA-DCT domain as in (3),

$$\hat{y}_{\tilde{U}_x^+} = T_{\tilde{U}_x^+}^{-1} \left(\Upsilon_x \left(T_{\tilde{U}_x^+} \left(z|_{\tilde{U}_x^+} - m_{\tilde{U}_x^+}(z) \right) \right) \right) + m_{\tilde{U}_x^+}(z), \quad (4)$$

where Υ_x is a hard-thresholding operator based on the threshold

$$\sigma \sqrt{2 \ln |\tilde{U}_x^+| + 1}. \quad (5)$$

An estimate of the variance $\bar{\sigma}_{\hat{y}_{\tilde{U}_x^+}}^2$ of $\hat{y}_{\tilde{U}_x^+}$, calculated as mean on \tilde{U}_x^+ , is given in the form

$$\bar{\sigma}_{\hat{y}_{\tilde{U}_x^+}}^2 = \sigma^2 (1 + N_x^{\text{har}}) / |\tilde{U}_x^+|, \quad (6)$$

where N_x^{har} is the number of non-zero coefficients after thresholding (which corresponds to the number of SA-DCT harmonics used for reconstruction) and the unit addend accounts for the addition of the mean after the inversion of the transform.

Since the anisotropic neighborhoods corresponding to nearby points are usually overlapping, and since the SA-DCT is a complete basis for an individual support \tilde{U}_x^+ , the overall approach is obviously overcomplete.

5.2. Global estimate as aggregation of local estimates

In order to obtain a single *global* estimate $\hat{y} : X \rightarrow \mathbb{R}$ defined on the whole image domain, all the local estimates (4) are averaged together using adaptive weights $w_x \in \mathbb{R}$ in the following convex combination:

$$\hat{y} = \frac{\sum_{x \in X} w_x \hat{y}_{\tilde{U}_x^+} |X|}{\sum_{x \in X} w_x \chi_{\tilde{U}_x^+}}. \quad (7)$$

It is a standard approach to use weights w_x that are inversely proportional to the mean variance (6) of $\hat{y}_{\tilde{U}_x^+}$. As shown in¹⁴ for the case of sliding 8×8 block DCT denoising, such a simple weighting enables to attain the same performance achievable with much more involved models of the blocks' statistics.

However, this simple approach is inadequate when instead of fixed-size blocks one is considering adaptive regions with arbitrary shape and size. In particular, not only the size of the regions may vary, but also the number of overlapping shapes may be different for different points. If the inverse of the average variances are used as weights, it can be observed that when regions of significantly different sizes overlap (this may happen along edges or transitions), then the local estimates corresponding to larger regions will inevitably "submerge" the finer details restored by smaller regions.

We found experimentally that in such cases crucial compensation of these oversmoothing effects can be obtained by dividing the weights by the square of the size of the support, and we define w_x as

$$w_x = \frac{\bar{\sigma}_{\hat{y}_{\tilde{U}_x^+}}^{-2}}{|\tilde{U}_x^+|^2} = \frac{1}{(1 + N_x^{\text{har}}) |\tilde{U}_x^+|}. \quad (8)$$

Let us observe that in areas where the size of the adaptive neighborhood is nearly constant (e.g. within smooth parts of the image) the weights (8) are inversely proportional to the mean variances of the corresponding local estimates, $w_x \propto \bar{\sigma}_{\hat{y}_{\tilde{U}_x^+}}^{-2}$. Thus, we can use the weights (8) for such areas also.

5.3. Wiener filtering in SA-DCT domain

Using the same approach as for thresholding, we introduce an empirical Wiener filter in the SA-DCT domain. It assumes that an estimate \hat{y} of y is known (in practice, we obtain this estimate using the above thresholding technique). For every $x \in X$, let $\varphi_{\hat{y},x} : V_{\tilde{U}_x^+} \rightarrow \mathbb{R}$ and $\varphi_{z,x} : V_{\tilde{U}_x^+} \rightarrow \mathbb{R}$ be – respectively – the SA-DCT (on \tilde{U}_x^+) coefficients of \hat{y} and z , where the mean $m_{\tilde{U}_x^+}(z)$ is subtracted from the both before applying the transform:

$$\varphi_{\hat{y},x} = T_{\tilde{U}_x^+} \left(\hat{y}|_{\tilde{U}_x^+} - m_{\tilde{U}_x^+}(z) \right), \quad \varphi_{z,x} = T_{\tilde{U}_x^+} \left(z|_{\tilde{U}_x^+} - m_{\tilde{U}_x^+}(z) \right). \quad (9)$$

The *local* Wiener estimate $\hat{y}_{\tilde{U}_x^+}^{\text{wi}}$ is defined as

$$\hat{y}_{\tilde{U}_x^+}^{\text{wi}} = T_{\tilde{U}_x^+}^{-1} (\omega_x \varphi_{z,x}) + \varpi_x m_{\tilde{U}_x^+}(z), \quad (10)$$

where $\omega_x \in \mathcal{V}_U$ and $\varpi_x \in \mathbb{R}$ are, respectively, the Wiener attenuation factors for the SA-DCT coefficients and for the subtracted mean value,

$$\omega_x = \frac{\varphi_{\hat{y},x}^2}{\varphi_{\hat{y},x}^2 + \sigma^2}, \quad \varpi_x = \frac{m_{\tilde{U}_x^+}^2(\hat{y})}{m_{\tilde{U}_x^+}^2(\hat{y}) + \sigma^2}. \quad (11)$$

The *global* estimate \hat{y}^{wi} can be obtained analogously as in (7), using the convex combination with the adaptive weights w_x^{wi} :

$$\hat{y}^{\text{wi}} = \frac{\sum_{x \in X} w_x^{\text{wi}} \hat{y}_{\tilde{U}_x^+}^{\text{wi}} |X|}{\sum_{x \in X} w_x^{\text{wi}} \chi_{\tilde{U}_x^+}}, \quad w_x^{\text{wi}} = \frac{1}{(\varpi_x^2 + \sum_{V_{\tilde{U}_x^+}} \omega_x^2) |\tilde{U}_x^+|}. \quad (12)$$

Similarly to (8), the term $\varpi_x^2 + \sum_{V_{\tilde{U}_x^+}} \omega_x^2$ in the adaptive weights corresponds to an estimate of the variance of $\hat{y}_{\tilde{U}_x^+}^{\text{wi}}$.

5.4. Gilge's "shape-adapted" DCT by orthogonalization

The shape-adaptive DCT proposed by Sikora et al. can be implemented very efficiently with a negligible computational overhead with respect to the usual separable 2D DCT on a rectangular support. However, for video-coding applications Gilge's "shape-adapted" DCT^{11,12} based on orthonormalization with respect to shape is known to achieve superior results than the approach by Sikora et al. It is considered³⁰ as a kind of upper bound on the decorrelation ability of the faster method. Thus, if computational complexity is not of primary concern, Gilge's method may be utilized as an alternative method to achieve further improvement in denoising.

We remark that the implementation of such "shape-adapted" DCT is not trivial. If general enough shapes are allowed*, precalculation of all possible shape-adapted bases is not feasible because the resulting look-up-table would be huge. Hence, the orthogonalization must be performed "on-line". There are two more problems. First, when restricting a basis to a region smaller than the initial domain of its elements, linear independence is certainly lost. This makes orthonormalization ill-posed. Second, the resulting shape-adapted orthonormal basis significantly depends on the order in which the initial elements are orthonormalized. In particular, since an increasing number of orthogonal complements are successively subtracted during the orthogonalization, the similarity (and thus the common decorrelating features) between the resulting orthonormal elements and the initial ones diminishes as the procedure progresses. The combination of these two problems can have a devastating impact, because the attempt to orthonormalize any linearly-dependent element results in a systematic introduction of amplified rounding-errors in all subsequent elements, which are thus distorted. A rather profound study of these problems is given in²⁷, providing criteria for the pruning of the starting basis so to obtain a set of elements which are linearly independent on the smaller region. Unfortunately, such criteria do not fully apply to arbitrary-shaped regions and additional numerical tests for independence must be employed.

In our implementation of Gilge's "shape-adapted" DCT we rely on a *QR*-decomposition to perform the orthonormalization. In order to overcome the above difficulties, we use the criteria from²⁷ to prune the starting rectangular DCT basis, sorted according to a zig-zag scan. Additionally, to make sure that no distorted elements are eventually used for the shape-adapted basis, we retain only the first $N_x \leq |\tilde{U}_x^+|$ columns of the orthogonal *Q* matrix for which the magnitude of the corresponding diagonal entry of the upper-triangular matrix *R* exceeds a predefined threshold. If $N_x < |\tilde{U}_x^+|$, then the resulting set of generators is obviously not complete. However, from the practical point of view it is still rich-enough to represent most considered signals which – thanks to the LPA-ICI-driven adaptive support – enjoy a particularly sparse decomposition.

Let us remark that since the same shape-adapted basis is used for processing all neighborhoods which share the same shape, the above orthonormalization procedure needs to be performed only once for each unique shape. We process neighborhoods with identical shape one after another, simply by following the lexicographic order

*In our implementation – for example – despite the rather simplified neighborhood structure (based on $|H| = 6$ scales and $K = 8$ directions, as described in Figure 3), a total of $6^8 = 1679616$ different shapes are possible.

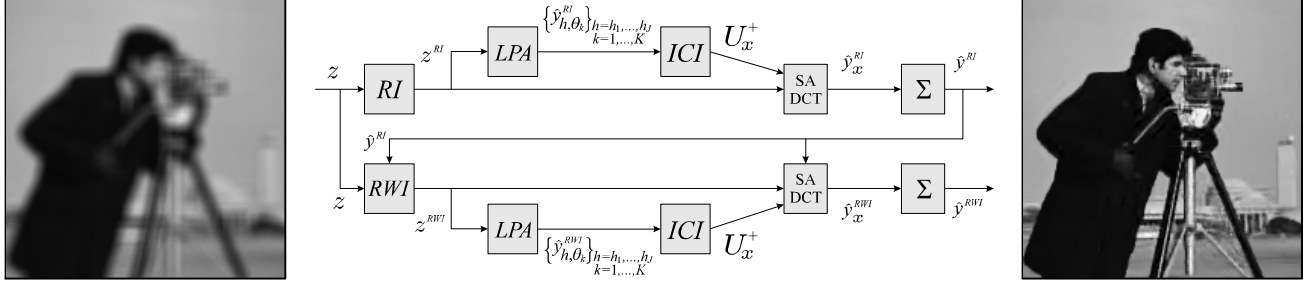


Figure 4. Fragment from the blurred and noisy observation of the *Cameraman* image (Exp. 1) (left). Flowchart of the SA-DCT regularized Wiener inverse algorithm (center). In the upper part of the flowchart a deblurred estimate \hat{y}^{RI} is obtained by combining the regularized-inversion (RI) operator with the SA-DCT-domain hard-thresholding. This estimate is used as reference estimate for the regularized-Wiener stage of the lower part of the flowchart, where a regularized Wiener inverse (RWI) operator is combined with Wiener filtering in SA-DCT domain. The intermediate \hat{y}^{RI} and the final \hat{y}^{RWI} estimates are respectively obtained by averaging with adaptive weights (represented as “ Σ ”) the local SA-DCT estimates \hat{y}_x^{RI} and \hat{y}_x^{RWI} , whose supports are defined by the anisotropic LPA-ICI technique on the regularized inverse z^{RI} and regularized Wiener inverse z^{RWI} . The deblurred RWI estimate \hat{y}^{RWI} of the *Cameraman* image using Gilge’s “shape-adapted” DCT ($ISNR=8.58\text{dB}$) (right).

induced by the directional adaptive scales $h^+(x, \theta_k)$ which define \tilde{U}_x^+ . Therefore only a single shape-adapted basis needs to be kept in memory. Since the number of distinct shapes found in a image is usually much less than the total number of pixels, the computational drawback of Gilge’s against Sikora’s method is mostly due to the fact that the former transform is performed as multiplication against a $N_x \times |\tilde{U}_x^+|$ matrix, whereas the latter uses transformation by rows and by columns, just like a separable transform.

The denoising algorithm described in Sections 5.1-5.3 is valid also if Gilge’s “shape-adapted” DCT is used instead of the SA-DCT. However, since these shape-adapted bases have indeed a DC term which truly corresponds to a constant basis element, there is no mean-weighting defect (as described in Section 4.2) and the subtraction of the mean is not necessary. It means that for Gilge’s “shape-adapted” DCT, formulas (3-12) can take a simpler form, where the mean terms $m_{\tilde{U}_x^+}(z)$ and $m_{\tilde{U}_x^+}(y)$ are replaced by zeros.

6. SA-DCT REGULARIZED DECONVOLUTION

In this section we extend the above SA-DCT denoising approach to image deblurring. First, let us generalize the observation model (1) given in Section 2, introducing an optical blurring into the observation process. This distortion is commonly modeled by the convolution $y \otimes v$ of the true image y with the point-spread function (PSF) v of the optical system. Thus, we wish to recover y from the blurred and noisy observations

$$z(x) = (v \otimes y)(x) + \eta(x), \quad x \in X. \quad (13)$$

In the frequency domain the above equation becomes $Z = YV + \tilde{\eta}$, where capital letters are used for the discrete Fourier transform of the corresponding variables and $\tilde{\eta}(\cdot) \sim \mathcal{N}(0, \sigma^2)$. It is assumed that the PSF v is known.

Note that the observation model (1) is a particular case of (13) where v is a Dirac function.

The proposed deblurring algorithm is based on the anisotropic *regularized inverse* (RI) and *regularized Wiener inverse* (RWI) LPA-ICI deconvolution scheme developed in^{16, 18}. The same deconvolution technique has also been used for inverse halftoning⁷ and Poissonian deblurring⁹.

The idea of the proposed SA-DCT domain deblurring is to combine the SA-DCT transform with the regularized inversion operators, using the adaptive supports defined by the anisotropic LPA-ICI regularized deconvolution. The flowchart of the overall deblurring algorithm is shown in Figure 4.

6.1. Regularized inverse

The regularized inverse (RI) z^{RI} is computed in the frequency domain as

$$z^{RI} = \mathcal{F}^{-1}(T^{RI}Z), \quad T^{RI} = \frac{\bar{V}}{|V|^2 + \varepsilon_1^2}, \quad (14)$$

where T^{RI} is a (Tichonov) regularized-inversion operator, $\varepsilon_1 > 0$ is a regularization parameter, and \bar{V} is the complex conjugate of V . In what follows we make explicit use of the impulse response t^{RI} of the regularized-inversion operator, $t^{RI} = \mathcal{F}^{-1}(T^{RI})$.

6.2. Adaptive anisotropic transform support

The anisotropic neighborhood U_x^+ is constructed from the adaptive scales $\{h^+(x, \theta_k)\}_{k=1}^K$ obtained from to the following procedure. The LPA is performed in the spatial domain as a convolution of the regularized inverse z^{RI} against the LPA kernel g_{h, θ_k} , $\hat{y}_{h, \theta_k}^{RI} = z^{RI} \otimes g_{h, \theta_k}$, for each specified direction θ_k , $k = 1, \dots, K$, and for every scale $h \in H$. All the varying-scale LPA estimates $\{\hat{y}_{h, \theta_k}^{RI}\}_{h \in H}$ obtained for each θ are fed (together with their standard deviations $\{\sigma_{\hat{y}_{h, \theta}^{RI}} = \sigma \|t^{RI} \otimes g_{h, \theta}\|_2\}_{h \in H}$) into the ICI algorithm, which selects the pointwise-adaptive scale $h^+(x, \theta_k)$. The adaptive scale selection is done *independently* for each direction θ_k .

6.3. SA-DCT deblurring

A non-linear adaptive deblurring operator is obtained by combining the regularized inversion (14) with the shape-adaptive DCT domain shrinkage. While formally it is enough to replace z by z^{RI} in (4), several further modifications to the basic denoising procedure of Section 5 are required in order to properly take into account the initial regularized inversion.

The principal difference arises from the fact that, because of the blur inversion (14), the noise in z^{RI} is colored. Therefore the noise variance in the transform domain is not constant and instead can vary significantly from a coefficient to the other. Consequently, the threshold (5) cannot be used in order to determine which coefficients should be used for the reconstruction of the deblurred estimate.

6.3.1. Thresholding with adaptive variance in transform domain

We consider the individual variance of the transform coefficients $\varphi_{z^{RI}, x} = T_{\tilde{U}_x^+}(z_{|\tilde{U}_x^+}^{RI} - m_{\tilde{U}_x^+}(z^{RI}))$. Let $i \in V_{\tilde{U}_x^+}$ be coordinates in SA-DCT domain. The SA-DCT basis functions $\psi_{\tilde{U}_x^+}^{(i)} : \tilde{U}_x^+ \rightarrow \mathbb{R}$, which are defined implicitly by the cascaded row-column structure of the SA-DCT, can be obtained explicitly by inverse transformation of the standard basis $\{e_i\}_{i \in V_{\tilde{U}_x^+}}$ as $\psi_{\tilde{U}_x^+}^{(i)} = T_{\tilde{U}_x^+}^{-1}(e_i)$, $i \in V_{\tilde{U}_x^+}$.

Neglecting the mean subtraction (“DC separation”), the variance $\sigma_{\varphi_{z^{RI}, x}}^2(i)$ of the i -th transform coefficient $\varphi_{z^{RI}, x}(i)$ is

$$\sigma_{\varphi_{z^{RI}, x}}^2(i) = \sigma^2 \|t^{RI} \otimes \psi_{\tilde{U}_x^+}^{(i)}\|_2^2 = \sigma^2 \|T^{RI} \mathcal{F}(\psi_{\tilde{U}_x^+}^{(i)})\|_2^2. \quad (15)$$

The appropriate threshold level for each transform coefficient is obtained by replacing the coefficient-invariant σ in the threshold (5) with the varying standard deviation $\sigma_{\varphi_{z^{RI}, x}} : V_{\tilde{U}_x^+} \rightarrow \mathbb{R}$. In this way, we realize an adaptive hard-thresholding rule for the joint deblurring and filtering in SA-DCT domain.

Analogously to (4), a local deblurred estimate $\hat{y}_{\tilde{U}_x^+}^{RI}$ is obtained by inverse transformation of the thresholded coefficients.

6.3.2. Adaptive weights and global estimate

Just as in Section 5.2, in order to obtain a global estimate $\hat{y}^{RI} : X \rightarrow \mathbb{R}$ all the local estimates $\hat{y}_{\tilde{U}_x^+}^{RI}$ are averaged together using adaptive weights $w_x^{RI} \in \mathbb{R}$ in a convex combination analogous to (7). However, because of the different variances of the coefficients used in the reconstruction, the adaptive weights are defined by a different formula. Ignoring the correlation between the noise in the transform domain, the corresponding expression of (8) for deblurring is

$$w_x^{RI} = \frac{1}{\left(\sigma_{m_{z^{RI}, x}}^2 + \sum_{j \in S} \sigma_{\varphi_{z^{RI}, x}}^2(j)\right) |\tilde{U}_x^+|}$$

where $S \subseteq V_{\tilde{U}_x^+}$ indicates the coefficients which are non-zero after thresholding and $\sigma_{m_{z^{RI}, x}}^2 = \frac{\sigma^2}{|\tilde{U}_x^+|} \|t^{RI} \otimes \chi_{\tilde{U}_x^+}\|_2^2 = \frac{\sigma^2}{|\tilde{U}_x^+|} \|T^{RI} \mathcal{F}(\chi_{\tilde{U}_x^+})\|_2^2$ is an extra term which accounts for the “DC separation”.



Figure 5. A fragment of *Lena*. From left to right: original, noisy observation ($\sigma=25$, $PSNR=20.18$ dB), BLS-GSM estimate²⁸ ($PSNR=31.69$ dB), and the proposed Anisotropic LPA-ICI-driven SA-DCT and Gilge’s “shape-adapted” DCT Wiener estimates ($PSNR=31.66$ and 31.76 dB, respectively).

6.4. Regularized Wiener inverse

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

$$z^{RWI} = \mathcal{F}^{-1}(T^{RWI} Z), \quad T^{RWI} = \frac{\overline{V}|Y|^2}{|VY|^2 + \varepsilon_2^2 \sigma^2}, \quad (16)$$

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 (16) is replaced by $|\hat{Y}^{RI}|^2$.

Analogously to Section 6.2, the adaptive neighborhoods \tilde{U}_x^+ are defined by the anisotropic LPA-ICI technique applied to z^{RWI} .

On these neighborhoods we perform empirical Wiener filtering in SA-DCT domain in a fashion similar to Section 5.3, where the attenuation coefficients for the filter (10) are

$$\omega_x = \frac{\varphi_{\hat{y}^{RI},x}^2}{\varphi_{\hat{y}^{RI},x}^2 + \sigma_{\varphi_{z^{RWI},x}}^2}, \quad \varpi_x = \frac{m_{\tilde{U}_x^+}^2(\hat{y}^{RI})}{m_{\tilde{U}_x^+}^2(\hat{y}^{RI}) + \sigma_{m_{z^{RWI},x}}^2}.$$

The variances $\sigma_{\varphi_{z^{RWI},x}}^2$ and $\sigma_{m_{z^{RWI},x}}^2$ are defined as in (15), but using the regularized Wiener inverse operator T^{RWI} or its impulse response $t^{RWI} = \mathcal{F}^{-1}(T^{RWI})$ instead of T^{RI} or t^{RI} , respectively. These variances are also used to define the adaptive weights used for the aggregation of the local estimates:

$$w_x^{RWI} = \frac{1}{\left(\varpi_x^2 \sigma_{m_{z^{RWI},x}}^2 + \sum_{V_{\tilde{U}_x^+}} \omega_x^2 \sigma_{\varphi_{z^{RWI},x}}^2 \right) |\tilde{U}_x^+|}.$$

6.5. Fast computation of coefficients’ variance

The calculation of the variance of the transform coefficients as in (15) is a rather computationally intense procedure. In order to reduce complexity to an acceptable level, it is necessary to employ some approximations.

Our implementation of (15) is based on processing in an undersampled (“coarse”) FFT domain of size 32×32 . Denoting by $\tilde{\mathcal{F}}$ the corresponding transform, by \tilde{X} a 32×32 domain, and by \mathcal{L} a decimator with prefiltering, we have the following approximation of the variance of the i -th transform coefficient $\varphi_{z^{RI},x}(i)$:

$$\sigma_{\varphi_{z^{RI},x}}^2(i) \simeq \sigma^2 \left\| \mathcal{L}\{(\mathcal{F}(t^{RI}))^2\}(\tilde{\mathcal{F}}(\psi_{\tilde{U}_x^+}^{(i)}|\tilde{X}))^2 \right\|_1. \quad (17)$$

An obvious modification to the above formula allows to compute approximations of the variances $\sigma_{\varphi_{z^{RWI},x}}^2$.

Such approximations have a significant impact on the computational performance of the algorithm. The current MATLAB implementation can process a 512×512 image in about a minute. This includes RI and RWI stages, comprehensive of LPA-ICI and transform-domain filtering. If (15) would be used instead of the approximate (17), the processing time would be of a few hours, without any noticeable improvement in the restoration quality.

Method	σ	<i>Lena</i> 512×512			<i>Boats</i> 512×512			<i>Peppers</i> 256×256			<i>House</i> 256×256		
		15	20	25	15	20	25	15	20	25	15	20	25
Anis. LPA-ICI + Gilge’s “SA”-DCT		34.00	32.75	31.76	31.85	30.54	29.52	32.62	31.19	30.07	34.26	33.01	32.00
Anis. LPA-ICI + SA-DCT		33.86	32.62	31.66	31.79	30.49	29.47	32.44	31.04	29.93	34.14	32.92	31.92
BLS-GSM (Portilla et al.) ²⁸		33.90	32.66	31.69	31.70	30.38	29.37	31.74	30.31	29.21	33.63	32.39	31.40
Exemplar-based (Kervrann and Boulanger) ²¹		33.70	32.64	31.73	31.44	30.12	29.20	32.13	30.59	29.73	34.08	32.90	32.22
MGGD (Cho and Bui) ²		33.70	32.46	31.48	31.46	30.14	29.12	–	–	–	–	–	–
Recursive Anisotropic LPA-ICI ^{6,10}		32.72	31.44	30.43	30.87	29.58	28.58	31.78	30.30	29.16	33.18	31.82	30.73

Table 1. *PSNR* (dB) comparison table for the denoising of the *Lena*, *Boats*, *Peppers*, and *House* test images with different levels of Gaussian noise.

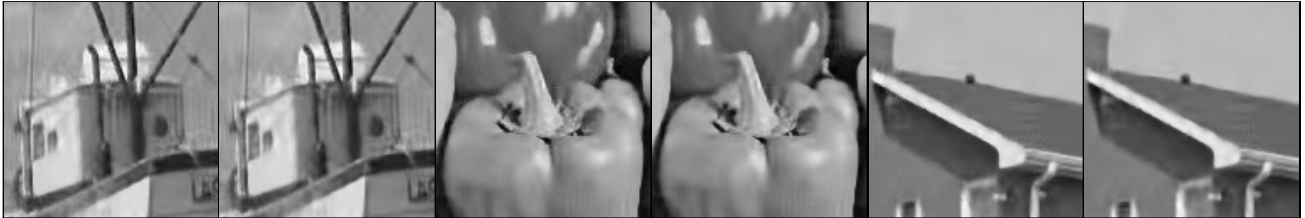


Figure 6. Fragments of the Wiener estimates of the *Boats*, *Peppers*, and *House* images, $\sigma=25$. For each test image are shown the two different final estimates, one based on the low-complexity SA-DCT (left), the other based on Gilge’s “shape-adapted” DCT. From these images it can be seen that the visual quality of the two methods is essentially the same.

6.6. “Shape-adapted” DCT deblurring

The described algorithm can be used also with Gilge’s “shape-adapted” DCT, provided that simple modifications, analogous to those described in Section 5.4, are made.

Let us remark that because of the need to go through the explicit expression of each one of the SA-DCT basis elements $\psi_{\tilde{U}_x^+}^{(i)}$, $i \in V_{\tilde{U}_x^+}$, the computational advantage of the Sikora’s SA-DCT over Gilge’s “shape-adapted” DCT is substantially reduced. In fact, our implementations of the proposed deblurring algorithm based on Gilge’s “shape-adapted” DCT and on the lower-complexity SA-DCT perform in about the same time.

7. EXPERIMENTS

We conclude the paper with a few experimental results which demonstrate state-of-the-art performance of the developed algorithms. Further simulation results and the MATLAB software which has been used to perform the presented experiments are available at <http://www.cs.tut.fi/~foi/SA-DCT/>.

The results are presented for both the SA-DCT and Gilge’s “shape-adapted” DCT, using the same transform for the thresholding and Wiener filtering stages of the denoising and deblurring algorithms.

7.1. Denoising

In Table 1 we compare our results against those reported by other authors. In terms of *PSNR*, the results of both our estimates are high, often outperforming all other methods. We emphasize the outstanding preservation of sharp details which we demonstrate in Figures 5 and 6, while almost no visible artifacts are present. Other transform-based estimates, such as those from²⁸, often display noticeable overshootings on the edges and unpleasant spurious oscillations. These artifacts, which are characteristic of all oscillatory transforms (including the SA-DCT), do not appear in our estimates thanks to the adaptive selection of the transform support.

Let us note that Gilge’s method provides only a marginal improvement over the lower-complexity SA-DCT, thus the latter is a more likely choice for practical implementations.

Method	Experiment	1	2	3	4	Exp.	Image, Point-Spread Function v , noise
Anis. LPA-ICI RI-RWI + Gilge's "SA"-DCT		8.58	8.29	6.34	4.55	1	<i>Cameraman</i> 256×256 $BSNR=40\text{dB}$ 9×9 boxcar uniform v ($\sigma^2 \simeq 0.308$)
Anis. LPA-ICI RI-RWI + SA-DCT		8.48	8.21	6.26	4.47		
Anisotropic LPA-ICI RI-RWI ^{10,16}		8.29	7.82	5.98	3.90	2	<i>Cameraman</i> 256×256 $\sigma^2=2$ $v(x_1, x_2)=1/(1+x_1^2+x_2^2)$, $x_1, x_2=-7, \dots, 7$
BOA (Figueiredo and Nowak) ⁵		8.16	7.46	5.24	2.84		
GEM (Dias) ³		8.10	7.47	5.17	2.73	3	<i>Cameraman</i> 256×256 $\sigma^2=8$ $v(x_1, x_2)=1/(1+x_1^2+x_2^2)$, $x_1, x_2=-7, \dots, 7$
Segmentation-based Regularization (Mignotte) ²⁴		8.04	7.23	–	1.34		
EM (Figueiredo and Nowak) ⁴		7.59	6.93	4.88	2.94	4	<i>Lena</i> 512×512 $\sigma^2=49$ $v=[1,4,6,4,1]^T [1,4,6,4,1]/256$
ForWaRD (Neelamani et al.) ²⁵		7.30	6.75	5.07	2.98		

Table 2. Left: *ISNR* (dB) of the proposed deblurring algorithm and of other state-of-the-art methods for four standard experiments. Right: description of the observation parameters for the four experiments.



Figure 7. Fragments of the observations z and estimates \hat{g}^{RWI} for the deblurring experiments 3 and 4 ($ISNR=6.34$ and 4.55 dB, respectively) .

7.2. Deblurring

In order to evaluate the restoration ability of the proposed deblurring algorithm, we consider the standard set of four blurred and noisy observations which has been used by many other authors. Table 1 presents the description of the observations' parameters and the improvement-in-*SNR* (*ISNR*) results for these four experiments. Figures 4 and 7 show details of the observations and the corresponding restored images. Similarly to denoising, both the objective and subjective quality of our estimates are high, outperforming the best methods known to date.

REFERENCES

1. Bi, M., S.H. Ong, and Y.H. Ang, "Comment on Shape-Adaptive DCT for generic coding of video", *IEEE Trans. Circuits and Systems for Video Technology*, vol. 6, no. 6, pp. 686-688, December 1996.
2. Cho, D., and T.D. Bui, "Multivariate statistical modeling for image denoising using wavelet transforms", *Signal Processing: Image Communications*, vol. 20, no. 1, pp. 77-89, January 2005.
3. Dias, J.M.B., "Fast GEM wavelet-based image deconvolution algorithm", *Proc. Int. Conf. Image Process., ICIP 2003*, vol. 2, 2003.
4. Figueiredo, M.A.T., and R.D. Nowak, "An EM algorithm for wavelet-based image restoration", *IEEE Trans. Image Process.*, vol. 12, no. 8, pp. 906-916, 2003.
5. Figueiredo, M.A.T., and R.D. Nowak, "A bound optimization approach to wavelet-based image deconvolution", *Proc. IEEE 2005 Int. Conf. Image Process., ICIP 2005*, vol. 2, pp. 782-785, Genova, September 2005.
6. Foi, A., V. Katkovnik, K. Egiazarian, and J. Astola, "A novel anisotropic local polynomial estimator based on directional multiscale optimizations", *Proc. 6th IMA Int. Conf. Math. Signal Process.*, pp. 79-82, Cirencester, 2004.
7. Foi, A., V. Katkovnik, K. Egiazarian, and J. Astola, "Inverse halftoning based on the anisotropic LPA-ICI deconvolution", *Proc. 2004 Int. TICSP Workshop Spectral Meth. Multirate Signal Process., SMMSP 2004*, pp. 49-56, Vienna, September 2004.
8. Foi, A., V. Katkovnik, and K. Egiazarian, "Pointwise shape-adaptive DCT as an overcomplete denoising tool", *Proc. 2005 Int. TICSP Workshop Spectral Meth. Multirate Signal Process., SMMSP 2005*, pp. 164-170, Riga, June 2005.
9. Foi, A., S. Alenius, M. Trimeche, V. Katkovnik, and K. Egiazarian, "A spatially adaptive Poissonian image deblurring", *Proc. IEEE 2005 Int. Conf. Image Process., ICIP 2005*, pp. 925-928, Genova, September 2005.

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. Gilge, M., T. Engelhardt, and R. Mehlan, "Coding of arbitrarily shaped image segments based on a generalized orthogonal transform", *Signal Processing: Image Communication*, vol. 1, no. 2, pp. 153-180, October 1989.
12. Gilge, M., "Region oriented transform coding (ROTC) of images", *Proc. of Int. Conf. on Acoustics, Speech, and Signal Processing, ICASSP-90*, Albuquerque, USA, vol. 4, pp. 2245-2248, April 1990.
13. Goldenshluger, A., and A. Nemirovski, "On spatial adaptive estimation of nonparametric regression", *Math. Meth. Statistics*, vol. 6, pp. 135-170, 1997.
14. Guleryuz, O.G., "Weighted overcomplete denoising", *Proc. Asilomar Conference on Signals and Systems*, Pacific Grove, CA, November 2003.
15. Katkovnik V., "A new method for varying adaptive bandwidth selection", *IEEE Trans. Signal Process.*, vol. 47, no. 9, pp. 2567-2571, 1999.
16. Katkovnik, V., A. Foi, K. Egiazarian, and J. Astola, "Directional varying scale approximations for anisotropic signal processing", *Proc. XII European Signal Process. Conf., EUSIPCO 2004*, pp. 101-104, Vienna, September 2004.
17. 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, no. 3, pp. 223-235, 2002.
18. Katkovnik, V., K. Egiazarian, and J. Astola, "A spatially adaptive nonparametric regression image deblurring", *IEEE Trans. on Image Process.*, vol. 14, no. 10, pp. 1469-1478, October 2005.
19. Kauff, P., and K. Schuur, "Shape-adaptive DCT with block-based DC separation and Δ DC correction", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 8, no. 3, pp. 237-242, 1998.
20. Kaup, A., and S. Panis, "On the performance of the shape adaptive DCT in object-based coding of motion compensated difference images", *Proc. of 1997 Picture Coding Symposium*, pp. 653-657, 1997.
21. Kervrann, C., and J. Boulanger, "Local adaptivity to variable smoothness for exemplar-based image denoising and representation", Research Report no. 5624, INRIA, July 2005.
22. Kinane, A., V. Muresan, and N. O'Connor, "Optimal adder-based hardware architecture for the DCT/SA-DCT", *Proc. SPIE Visual Comm. and Image Process. Conf., VCIP 2005*, Beijing, China, 12-15 July 2005.
23. Koenen, R., "Overview of the MPEG-4 Standard", ISO/IEC JTC1/SC29/WG11 Doc. N3536, July 2000.
24. Mignotte, M., "An adaptive segmentation-based regularization term for image restoration", *Proc 12th IEEE Int. Conf. Image Process., ICIP'05*, vol. 1, pp. 901-904, Italy, Genova, September 2005.
25. Neelamani, R., H. Choi, and R. Baraniuk, "Forward: Fourier-wavelet regularized deconvolution for ill-conditioned systems", *IEEE Trans. Image Process.*, vol. 52, no. 2, 2004.
26. Ostermann, J., E.S. Jang, J. Shin, and T. Chen, "Coding of arbitrarily shaped video objects in MPEG-4", *Proc. Int. Conf. Image Process., ICIP 1997*, pp. 496-499, 1997.
27. Philips, W., "Orthogonal base functions on a discrete two-dimensional region", *ELIS Technical Report DG 91-20*, Dept. of Electronics and Inform. Syst., Universiteit Gent, Belgium, November 1993.
28. Portilla, J., V. Strela, M. Wainwright, and E.P. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain", *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1338-1351, November 2003.
29. Sikora, T., "Low complexity shape-adaptive DCT for coding of arbitrarily shaped image segments", *Signal Processing: Image Communication*, vol. 7, pp. 381-395, 1995.
30. Sikora, T., S. Bauer, and B. Makai, "Efficiency of shape-adaptive 2-D transforms for coding of arbitrarily shaped image segments," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 5, no. 3, pp. 254-258, June 1995.
31. Sikora, T., and B. Makai, "Shape-adaptive DCT for generic coding of video", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 5, no. 1, pp. 59-62, 1995.