POINTWISE SHAPE-ADAPTIVE DCT AS AN OVERCOMPLETE DENOISING TOOL

Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian

Institute of Signal Processing, Tampere University of Technology, P.O. Box 553, 33101 Tampere, Finland
e-mail: firstname.lastname@tut.fi

ABSTRACT

A novel approach to image-denoising based on the shape-adaptive DCT (SA-DCT) is presented. The anisotropic *LPA-ICI* technique is used in order to define the shape of the transform's support in a pointwise adaptive manner. It means that for each point in the image an adaptive estimation neighborhood is found. For each one of these neighborhoods a SA-DCT is performed. The thresholded SA-DCT coefficients are used to reconstruct a local estimate of the signal within the adaptive-shape region. Since regions corresponding to different points are in general overlapping (and thus generate an overcomplete representation of the signal), the local estimates are averaged together using adaptive weights that depend on the region's statistics. A Wiener filtering procedure in SA-DCT domain is also proposed. Simulation experiments confirm the advanced quality of the final estimate. Not only objective criteria scores are high, but also the visual appearence of the estimate is superior: edges are clean, and no unpleasant ringing artifacts are introduced by the fitted transform.

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 the presence of edges such near-optimality fails. For this reason, other transforms with better edge adaptation capabilities (e.g. wavelets [17], curvelets [28], etc.) have been used in denoising, and post-processing (deringing, deblocking) filters are commonly used in MPEG-video decoders [23].

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 could efficiently be used for the analysis and synthesis of arbitrarily shaped image segments, where the data exhibits some uniform behaviour.

In [7] it was initially proposed to consider the orthonormalization of a (fixed) set of generators restricted to the arbitrarily shaped region of interest. These generators could be – for example – a basis of polynomials or a 2D block DCT basis. Orthormalization can be performed by the standard Gram-Schmidt procedure and the obtained orthonormal basis is supported on the region. Faster algorithms have been proposed in order to compute this orthonormalization (e.g. [21]). Nevertheless, since for each differently shaped region one has to recalculate a new region-adapted basis, and since each one of the basis element is typically non-separable, the overall method still presents an extremely high computational cost, making it unsuitable for any practical application.

A more computationally attractive approach, namely the *shape-adaptive DCT* (SA-DCT), has been proposed by Sikora et al. [25, 27] and it is based on a cascaded application of standard one dimensional n-point DCT transforms first on the rows and then on the colums. Thus, the SA-DCT doesn't require costly matrix-inversions or iterative orthogonalization and can be interpreted as a direct generalization of the classical 2D block DCT transform. In particular, the SA-DCT and the block DCT have the same computational complexity and in the special case of an 8×8 square the two transform exactly coincide. Therefore, the SA-DCT has received increasing interest from the MPEG community, eventually becoming part of the MPEG-4 standard [16]. In the context of MPEG-4, where arbitrarly shaped video-objects are introduced, the SA-DCT is employed for the coding of image segments that lie on the video-object's boundary, while the shape itself is processed separately [20]. Fast and low-power hardware implementations of the SA-DCT are becoming now available [15].

The SA-DCT has been shown [25, 26, 3, 13] to provide a compression efficiency comparable to those of more computationally complex transforms, such as [7]. The good decorrelation and energy compaction properties on 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.

It is then rather surprising that no attempt of exploiting this potential has been documented so far. We explain this apparent discrepancy as follows.

The shape-adaptive coding of video objects always assumes that the boundary of these object 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. In fact, several automated segmentation algorithms suitable for this purpose have been proposed (e.g. [18]). On the contrary, obtaining an accurate and robust segmentation of *noisy* data constitutes an extremely more complex task than the region-oriented coding itself, and – 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 shape-adaptive DCT has not been used for the restoration of noisy images.

In this paper we aim at resolving this discrepancy and we propose an image denoising algorithm which makes use of a recent nonparametric estimator, namely the anisotropic *LPA-ICI* technique [11, 6], in order to obtain – directly from the noisy data – an appropriate shape on which to apply the
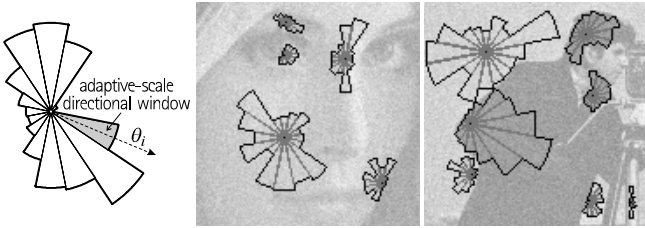
Figure 1: Anisotropic local approximations achieved by combining a number of adaptive-scale directional windows. The examples show some of these windows selected by the directional *LPA-ICI* for the noisy *Lena* and *Cameraman* images.

shape-adaptive transform. Examples of such sets are shown in Figure 1.

The proposed method can be interpreted as a generalization of the so-called sliding-window DCT denoising algorithm [29, 9] and its variants with adaptive block-size [19, 12].

The paper is organized as follows: first, we present various aspects of the shape adaptive DCT and recall the main points of the anisotropic *LPA-ICI* technique; the proposed *LPA-ICI*-driven SA-DCT denoising algorithm is then introduced, followed by a number of experimental results that validate the advanced performance of the algorithm; we conclude with a brief discussion on some of the issues that will need further research effort in order to fully exploit the potential of SA-DCT and – more generally – region-oriented transforms.

Before proceeding, let us introduce the considered observation model and some notation. Throughout the paper we consider noisy observations z of the form

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

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 $f : X \to \mathbb{R}$, a subset $U \subset X$, and a function $g : U \to \mathbb{R}$, we denote by $f_{|U} : U \to \mathbb{R}$ the restriction of f on U, $f_{|U}(x) = f(x) \; \forall x \in U$, and by $g^{|X} : X \to \mathbb{R}$ the zero-extension of g to X, $\left(g^{|X}\right)_{|U} = g$ and $g^{|X}(x) = 0 \; \forall x \in X \setminus U$.

2. SHAPE-ADAPTIVE DCT

2.1 Orthonormal Shape-Adaptive DCT

Several improvements to the original definition of the shape-adaptive DCT transform [25, 27] have been proposed. The most significant [14, 13] concerns with the normalization of the individual one-dimensional transforms that are used for transforming the rows and the columns:

$$DCT_{L,m}(n) = c_m \cos\left(\frac{\pi m (2n+1)}{2L}\right), \; m,n = 0,...,L-1, \quad (2)$$

$$\text{with } c_0 = \sqrt{1/L} \text{ and } c_m = \sqrt{2/L} \text{ for } m > 0. \qquad (3)$$

Here L stands for the length of the row or column to be transformed. The normalization (3) 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 orthonor-

mal transform. A different normalization of the 1D transforms would produce, on an arbitrary shape, a 2D transformation that is non-orthogonal (for example as in [25, 27], where $c_m = \sqrt{2}/L$ for $m = 0$ and $c_m = 2/L$ for $m > 0$).

Let us denote by $T_U : \mathcal{U} \to \mathcal{V}_U \subset \mathbb{Z}^2$ the orthonormal SA-DCT obtained for a region $U \subset X$, where $\mathcal{U} = \{f : U \to \mathbb{R}\}$ and $\mathcal{V}_U = \{f : V_U \to \mathbb{R}\}$ are function spaces and $V_U \subset \mathbb{Z}^2$ indicates the domain of the transform coefficients. Let $T_U^{-1} : \mathcal{V}_U \to \mathcal{U}$ 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 $\hat{y}_U = T_U^{-1}\left(\Upsilon\left(T_U\left(z_{|U}\right)\right)\right)$, $\hat{y}_U : U \to \mathbb{R}$. The orthonormality of T make so that, according to (1), $T_U\left(z_{|U}\right) = T_U\left(y_{|U}\right) + \bar{\eta}$, where $\bar{\eta} = T_U\left(\eta_{|U}\right)$ is again Gaussian white noise with variance σ^2 and zero mean.

2.2 Mean subtraction

There is, however, an unpleasant consequence of the normalization (3). Even for a constant $z_{|U}$, the reconstructed $\hat{y}_U$ is generally non-costant. In [13] this behaviour is termed as "mean weighting defect", and it is there proposed 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 ("DC separation"). After the inversion, the mean is added back to the reconstructed signal $\hat{y}_U : U \to \mathbb{R}$:

$$\hat{y}_U = T_U^{-1}\left(\Upsilon\left(T_U\left(z_{|U} - m_U(z)\right)\right)\right) + m_U(z), \qquad (4)$$

where $m_U(z) = \frac{1}{\#U}\sum_{x \in U} z(x)$ is the mean of z on U, $\#U$ denoting the cardinality (i.e. the number of pixels) of U.

Although this operation is not fully justified from the approximation theory standpoint ($m_U(z)$ is calculated from the noisy data, and by subtracting it the noise present on the coefficients is no longer white), it produces visually superior results without affecting to the objective restoration performance.

2.2.1 ΔDC correction

Since for the estimate $\hat{y}_U$ resulting from (4), the mean $m_U(\hat{y}_U)$ is in general very close but not exactly equal to zero, in [13] it is suggested to compensate this small difference by subtracting it from the reconstructed signal ("ΔDC correction"). Although for video-coding this special correction may be useful, we found that for denoising purposes it is in fact completely negligible. Therefore we do not implement it.

2.3 Coefficient alignment

To further improve the efficiency of the SA-DCT, it has been proposed to align the coefficients obtained after the first 1D transformation along the rows in such a way as to maximize their vertical correlation before applying the second transform along the colums. Different strategies, based on different models of the underlying signal y, have been suggested (e.g. [2], [1]). Although they can provide a significant improvement when the data agrees with the assumed signal's model, in practice when dealing with real data only marginal improvement can be achieved over the basic alignment [25, 27], where coefficients with the same index m (i.e. all DC terms, all first AC terms, etc.) are aligned in the same columns, regardless of the length L of the current row.
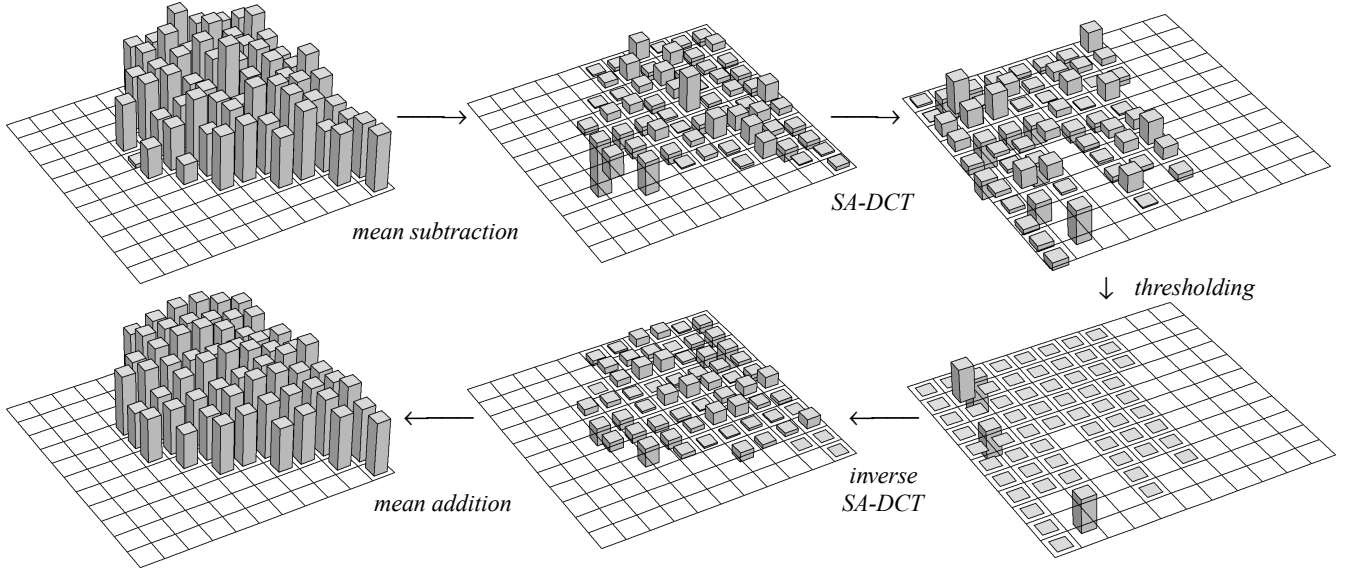
Figure 2: Thresholding in *SA-DCT* domain. The image data on the adaptive-shape region is subtracted of its mean. The zero-mean data is then transformed and thresholded. After inverse transformation, the mean is added back.

In our implementation we use the following alignment formula, denoting by m and m' the old (i.e. the one coming from (2)) and new coefficient index, respectively: $m' = \lfloor mL_{\max}/L \rfloor$, where L is the length of the current row, $m = 0, \ldots, L-1$, $L_{\max}$ is the length of the longest row in U, and the $\lfloor \cdot \rfloor$ brackets indicate the rounding to the nearest integer smaller or equal to $(\cdot)$.

An illustration of the SA-DCT domain thresholding, performed according to (4) and to the above coefficient alignment formula, is given in Figure 2.

3. ANISOTROPIC *LPA-ICI* TECHNIQUE

Our approach is based on the Intersection of Confidence Intervals (*ICI*) rule for pointwise adaptive estimation. Originally, the method has been developed for 1D signals [8, 10]. The idea was generalized for 2D image processing, where adaptive-size quadrant windows have been used [12]. Significant improvement of this approach was achieved on the base of anisotropic directional estimation [11, 6]. 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; as a result, the estimator is anisotropic and the shape of its support can adapt 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 more in detail. For every specified direction θ_k, $k = 1, \ldots, K$, a varying-scale family of directional-*LPA* (local polynomial approximation) convolution kernels [11] $\{g_{h,\theta_k}\}_{h \in H}$ is used to obtain a corresponding set of directional varying-scale estimates $\{\hat{y}_{h,\theta_k}\}_{h \in H}$, $\hat{y}_{h,\theta_k} = z \circledast g_{h,\theta_k}$, $h \in H$. These estimates are then compared according to the *ICI* rule [8, 10], and 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 esti-

mates $\hat{y}_{h^+(x,\theta_k),\theta_k}(x)$ are then "fused" together in an adaptive convex combination [11, 6] 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^+ which is used for its estimation. It is constructed as the union of the supports of the directional adaptive-scale kernels $g_{h^+(x,\theta_k),\theta_k}$:

$$U_x^+ = \bigcup_{k=1}^K \operatorname{supp} g_{h^+(x,\theta_k),\theta_k}.$$

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 $\tilde{U}_x^+ = \{x \in X : (x - v) \in U_x^+\}$. 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. *LPA-ICI*-DRIVEN SA-DCT DENOISING

We use the anisotropic adaptive neighborhoods $\tilde{U}_x^+$ defined by the *LPA-ICI* as supports for the shape-adaptive DCT, as shown in Figure 3.

4.1 Local estimates

For every point $x \in X$, we construct a *local* estimate $\hat{y}_{\tilde{U}_x^+} : \tilde{U}_x^+ \to \mathbb{R}$ of the signal y by thresholding in SA-DCT domain as in (4),

$$\hat{y}_{\tilde{U}_x^+} = T_{\tilde{U}_x^+}^{-1}\left(\Upsilon_x\left(T_{\tilde{U}_x^+}\left(z_{|\tilde{U}_x^+} - m_{\tilde{U}_x^+}(z)\right)\right)\right) + m_{\tilde{U}_x^+}(z), \quad (5)$$

where Υ_x is a hard-thresholding operator based on the universal threshold $\sigma\sqrt{2\ln(\#\tilde{U}_x^+)}$ (e.g. [17]).

An estimate of the mean variance $\bar{\sigma}^2_{\hat{y}_{\tilde{U}_x^+}}$ of $\hat{y}_{\tilde{U}_x^+}$ can be given in the form

$$\bar{\sigma}^2_{\hat{y}_{\tilde{U}_x^+}} = \sigma^2 \left(1 + N_x^{\mathrm{har}}\right)/\#\tilde{U}_x^+, \qquad (6)$$

where N_x^{har} is the number of non-thresholded coefficients (which corresponds to the number of SA-DCT harmonics that are going to be used to reconstruct the block). Formula (6) is accurate in the case when the SA-DCT basis element associated to the DC-term is constant over $\tilde{U}_x^+$ (which happens when the rows that compose $\tilde{U}_x^+$ have all the same length). Indeed, in this special case the mean is orthogonal to all SA-DCT components but to the DC-term, which is equal zero. The addition of the mean after reconstruction accounts for the unit addend in (6).

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.

4.2 Global estimate

In order to obtain a single *global* estimate $\hat{y}: X \to \mathbb{R}$ defined on the whole image domain, all the local representations are avaraged together using adaptive weights $w_x \in \mathbb{R}$ as

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

where $\chi_{\tilde{U}_x^+} = 1_{|\tilde{U}_x^+}^{|X}$ is the characteristic (indicator) function of $\tilde{U}_x^+$. 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 this would correspond to maximum-likelihood global estimate in the ideal case of independent and unbiased local estimates $\hat{y}_{\tilde{U}_x^+}$). As shown in [9] 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 seems to be inadequate when, instead of fixed-size blocks, one is considering adaptive regions with arbitrariry 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 mean variances are used as weights, it can be observed that when regions of significantly different sizes overlap (this may happen along transitions), then the local estimates corresponding to larger regions will inevitably "submerge" the finer details restored by smaller regions.

We found experimentally that a reasonable compensation can be obtained by dividing the weights by the square of the size of the block, and we define w_x as

$$w_x = \frac{1}{\left(1 + N_x^{\mathrm{har}}\right)\#\tilde{U}_x^+}. \qquad (8)$$

Let us observe that in regions where the size of the adaptive neighboorhood is nearly constant (such as in uniform areas), then the weights (8) are in fact inversely proportional to the mean variances of the corresponding blocks, $w_x \propto \bar{\sigma}^{-2}_{\hat{y}_{\tilde{U}_x^+}}$.

Figure 4: *Cameraman*: noisy observation (σ=25), and sliding 8×8 block DCT estimate (*PSNR*=28.36dB).

4.3 Empirical Wiener filtering in SA-DCT domain

Using the same approach described for thresholding, we introduce the following empirical Wiener filter in shape-adaptive DCT domain. It assumes that an estimate $\hat{y}$ of the signal y is known (this reference estimate can be obtained by the method described above or, in principle, by any other image denoising algorithm). For every $x \in X$, let $\varphi_{\hat{y},x}: V_{\tilde{U}_x^+} \to \mathbb{R}$ and $\varphi_{z,x}: V_{\tilde{U}_x^+} \to \mathbb{R}$, be – respectively – the SA-DCT (on $\tilde{U}_x^+$) coefficients of $\hat{y}$ and z, where the same mean $m_{\tilde{U}_x^+}(\hat{y})$, calculated from $\hat{y}$, is subtracted before applying the transform:

$$\begin{aligned} \varphi_{\hat{y},x} &= T_{\tilde{U}_x^+}\left(\hat{y}_{|\tilde{U}_x^+} - m_{\tilde{U}_x^+}(\hat{y})\right), \\ \varphi_{z,x} &= T_{\tilde{U}_x^+}\left(z_{|\tilde{U}_x^+} - m_{\tilde{U}_x^+}(\hat{y})\right). \end{aligned}$$

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

$$\hat{y}^{\mathrm{wi}}_{\tilde{U}_x^+} = T_{\tilde{U}_x^+}^{-1}\left(\omega\varphi_{z,x}\right) + m_{\tilde{U}_x^+}(\hat{y}), \quad \omega = \frac{\varphi^2_{\hat{y},x}}{\varphi^2_{\hat{y},x} + \sigma^2}. \qquad (9)$$

A *global* estimate $\hat{y}^{\mathrm{wi}}$ can be obtained analogously as in (7), using weights that depend on the supports' size (and thus adaptive):

$$\hat{y}^{\mathrm{wi}} = \frac{\sum_{x \in X} w_x^{\mathrm{wi}} \hat{y}^{\mathrm{wi}}_{\tilde{U}_x^+}{}^{|X}}{\sum_{x \in X} w_x^{\mathrm{wi}} \chi_{\tilde{U}_x^+}}, \quad w_x^{\mathrm{wi}} = \frac{1}{\#\tilde{U}_x^+}. \qquad (10)$$

5. EXPERIMENTS

We present a few simulation results in order to assess the restoration performance of the proposed approach. The Wiener filtering (10) is implemented using (7) as the reference estimate.

Figure 4 shows a fragment of the noisy *Cameraman* image (σ=25), and of the estimate obtained using a non-adaptive sliding DCT block with size 8×8 [29, 9]. A comparison with Figure 5, in which the estimates (7) and (10) based on the adaptive anisotropic neighborhoods $\tilde{U}_x^+$ confirms the remarkable advantange of using a transform with adaptive support.

In Table 1 we compare our results against those reported for some of the best denoising techniques appeared recently in the literature, including the Gaussian Scale Mixture model (GSM) [22], and the Multivariate Generalized Gaussian distribution (MGGD) [5]. We refer the reader to [5] for a rather complete compilation of *PSNR* results obtained by methods from other authors; we do not quote those results, which are
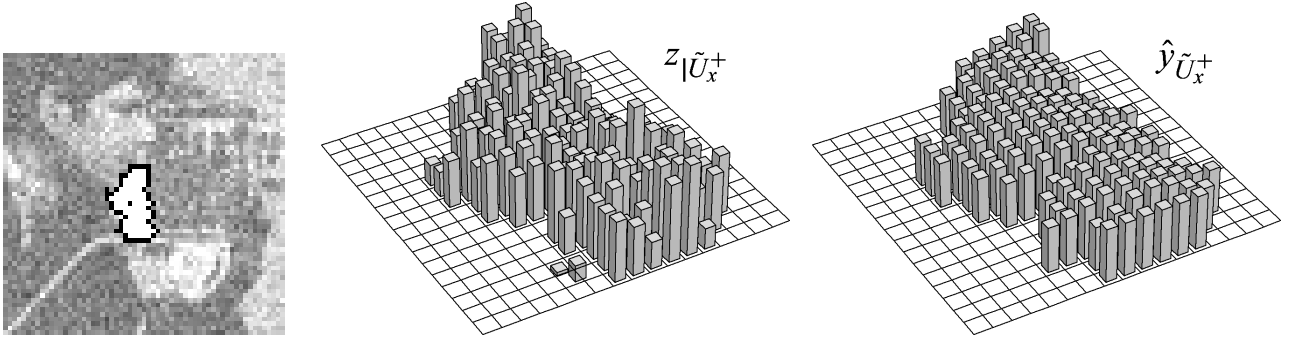
Figure 3: A detail of the noisy *Cameraman* image showing the adaptive-shape neighborhood $\tilde{U}_x^+$ determined by the Anisotropic *LPA-ICI* procedure (left), and the image intensity corresponding to this region before (center) and after (right) thresholding in the SA-DCT domain.

Figure 6: A fragment of *Lena*. From left to right: original image, noisy observation (σ=25), GSM estimate [22] (*PSNR*=31.69dB), and Anisotropic *LPA-ICI* SA-DCT estimates $\hat{y}$ (7) and $\hat{y}^{\text{wi}}$ (10) (*PSNR*=31.24 and 31.65 dB, respectively).

Figure 5: Anisotropic *LPA-ICI* SA-DCT estimate (7) (*PSNR*=28.96dB), and Wiener estimate (10) (29.09 dB). The same observation from Figure 4 is used (σ=25).

Figure 7: *House*: two estimates obtained from the same noisy observation (not shown) with σ=25: (left) the Anisotropic *LPA-ICI* SA-DCT Wiener estimate, *PSNR*=31.92dB, and (right) the GSM estimate [22], 31.40dB. The *PSNR* of the reference estimate (7) is 31.55dB.

anyway lower than those already included in Table 1. The numerical results of our Wiener algorithm are very close to the best results by [22], and outperform all other methods. The quality of the reference estimate (7) is also quite good. In fact, as long as only the visual appearence is concerned, there is no apparent advantage in the additional Wiener filtering stage. One can observe that all our estimates are very clean, with no visible artifacts introduced in the reconstruction. Despite the slightly better criteria values, the estimates obtained by [22] present visible overshootings on the edges and unpleasant spurious oscillations, as it can be seen in Figure 7 and in Figure 6.

These sorts of 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.

All the experiments presented in this paper have been

		Lena		*Boats*	
Approach σ		15	25	15	25
Anis. *LPA-ICI* + SA-DCT (7)		33.45	31.24	31.28	29.05
Anis. *LPA-ICI* + SA-DCT Wiener		33.81	31.65	31.53	29.25
GSM, Steer. Pyram. [22]		33.90	31.69	31.70	29.37
MGGD, CWT [5]		33.70	31.48	31.46	29.12
Local bi-shrink, CWT [24]		33.64	31.37	31.33	29.03
SI-adaptShr, SI Symm. [4]		33.37	31.07	—	—

Table 1: *PSNR* (dB) comparison table for the 512×512 *Lena* and *Boats* test images with different levels of Gaussian noise.

performed using a unique set of algorithm parameters and without any manual adjustment.

6. CONCLUSIONS AND FUTURE WORK

In this paper we have introduced an original image denoising strategy, based on the combined use of the SA-DCT and the anisotropic *LPA-ICI* technique. The presented experimental results not only demonstrate the state-of-the-art performance of this novel approach, but also highlight the versatility of the shape-adaptive DCT as a transform for denoising. Nevertheless, we believe that the potential of this transform is not fully exploited in our implementation. In particular, the adaptive weighting (8)-(10) is rather simple, and further improvement may be achieved by replacing the scalar weights w_x by some windowing functions, as suggested in [9]. The anisotropic *LPA-ICI* estimate of y, which we have not considered at all, could be used in order to control the adaptive averaging of the local estimates (5) and (9).

7. ACKNOWLEDGEMENTS

REFERENCES

[1] Acocella, E.C., and A. Alcaim, "Alignment by Phase of Vertical Coefficients in SA-DCT", *IEEE Signal Processing Letters*, vol. 8, no. 2, pp. 42-44, February 2001.

[2] Bi, M., S.H. Ong, Y.H. Ang, "Coefficient grouping method for shape-adaptive DCT", *Electronic Letters*, vol. 32, no. 3, pp. 201-202, February 1996.

[3] Bi, M., S.H. Ong, 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.

[4] Chang, S.G., B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression", *IEEE Trans. Image Process.*,.vol. 9, pp. 1532-1546, September 2000.

[5] 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.

[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. in Signal Processing*, Cirencester (UK), pp. 79-82, 2004.

[7] Gilge, M., T. Engelhardt, and R. Mehlan, "Coding of arbitrarily shaped image segments based on a generalized orthogonal transform," *Signal Processing: Image Communications*, vol. 1, no. 2, pp. 153-180, October 1989.

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

[9] Guleryuz, O.G., "Weighted Overcomplete Denoising", *Proc. Asilomar Conference on Signals and Systems*, Pacific Grove, CA, November 2003.

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

[11] 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,* Vienna*,* pp. 101-104, September 2004.

[12] 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.

[13] Kauff, P., and K. Schuur, "Shape-Adaptive DCT with Block-Based DC Separation and ΔDC Correction", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 8, no. 3, pp. 237-242, 1998.

[14] 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.

[15] Kinane, A., V. Muresan, and N. O'Connor, "Optimal Adder-Based Hardware Architecture for the DCT/SA-DCT", (to appear) *Proc. SPIE Visual Comm. and Image Process. Conf, VCIP 2005*, Beijing, China, 12-15 July 2005.

[16] Koenen, R., "Overview of the MPEG-4 Standard", ISO/IEC JTC1/SC29/WG11 Doc. N3536, July 2000.

[17] Mallat, S., *A wavelet tour of signal processing*, 2nd edition, Academic Press, 1998.

[18] O'Connor, N., S. Sav, T. Adamek, V. Mezaris, I. Kompatsiaris, T.Y. Lu, E. Izquierdo, C. Bennström, and J. Casas, "Region and object segmentation algorithms in the Qimera segmentation platform", *Proc. Third Int. Workshop on Content-Based Multimedia Indexing (CBMI03)*, Rennes, pp. 381-388, 2003.

[19] 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.

[20] Ostermann, J., E.S. Jang, J. Shin, and T. Chen, "Coding of arbitrarily shaped video objects in MPEG-4", *Proc. Int. Conf. on Image Process., ICIP '97*, pp. 496-499, 1997.

[21] Philips, W., "A Fast Algorithm for Orthogonalizing Polynomials on an Arbitrarily Shaped Region", *Multidimensional Systems and Signal Processing*, vol. 8, no. 4, pp. 409-421, 1997.

[22] Portilla, J., V. Strela, M. Wainwright, and E.P. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain", *IEEE Trans. on Image Process.*, vol. 12, no. 11, pp. 1338-1351, November 2003.

[23] Shen, M.-Y., and C.-C.J. Kuo, "Review of Postprocessing Techniques for Compression Artifact Removal", *Journal of Visual Communication and Image Representation*, vol. 9, no. 1, pp. 2-14, March, 1998.

[24] Şendur, L., and I.W. Selesnick, "Bivariate shrinkage with local variance estimation", *IEEE Signal Processing Letters*, vol. 9, no. 12, pp. 438-441, 2002.

[25] Sikora, T., "Low complexity shape-adaptive DCT for coding of arbitrarily shaped image segments", *Signal Processing: Image Communication*, vol. 7, pp. 381-395, 1995.

[26] 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.

[27] 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.

[28] Starck, J.L., E.J. Candès, and D.L. Donoho "The curvelet transform for image denoising", *IEEE Transactions on Image Processing*, vol. 11, pp. 670-684, 2000.

[29] Yaroslavsky, L., K. Egiazarian, and J. Astola, *Transform Domain Image Restoration Methods, Review, Comparison and Interpretation*, TICSP Series no. 9, Tampere, 2000.