Image Restoration Using Space-Variant Gaussian Scale Mixtures in Overcomplete Oriented Pyramids

Jose A. Guerrero-Colón, Luis Mancera, and Javier Portilla

Abstract—In recent years, Bayes least squares-Gaussian scale mixtures (BLS-GSM) has emerged as one of the most powerful methods for image restoration. Its strength relies on providing a simple and, yet, very effective local statistical description of oriented pyramid coefficient neighborhoods via a GSM vector. This can be viewed as a fine adaptation of the model to the signal variance at each scale, orientation, and spatial location. Here, we present an enhancement of the model by introducing a coarser adaptation level, where a larger neighborhood is used to estimate the local signal covariance within every subband. We formulate our model as a BLS estimator using space-variant GSM. The model can be also applied to image deconvolution, by first performing a global blur compensation, and then doing local adaptive denoising. We demonstrate through simulations that the proposed method, besides being model-based and noniterative, it is also robust and efficient. Its performance, measured visually and in L2-norm terms, is significantly higher than the original BLS-GSM method, both for denoising and deconvolution.

Index Terms—Bayesian estimation, Gaussian scale mixtures (GSM), image denoising, image restoration, overcomplete oriented pyramids.

I. INTRODUCTION

Image restoration is typically formulated as the estimation of an image given a linearly filtered version of the original corrupted by additive noise. This is a difficult, ill-posed, inverse problem, even if, as assumed in this work, the degradation parameters are known. A more tractable problem arises when we only consider additive noise in the degradation model (denoising problem). During the last few years, the state of the art in image denoising has experienced a formidable advance. The possibility of using image denoising techniques within more generic image restoration frames is, thus, becoming more and more attractive. In this paper, we present a new denoising-based approximation to deal with both degradation problems (just noise, and blur plus noise). For previous instantiations of this work, see [1] and [2].

Wiener filtering is one of the simplest model-based restoration methods, and yet it provides optimal results when dealing with Gaussian signals and noise. Modeling noise as Gaussian, as we do in this work, is reasonable for many image processing situations. However, it is well known that image statistics are far from Gaussian [3]. One can, instead, assume that images are locally Gaussian. Spatially adaptive models have widely used this assumption for image denoising. In 1980, Lee [4] proposed an adaptive scalar Wiener method in the pixel domain, based on estimating the local variance at every spatial location of the image. With the introduction of wavelets (in the broad sense of the word, i.e., multiscale, multiorientation linear pyramidal representations), local adaptation in the new domain became much more powerful. Whereas many authors explored other estimation approaches in the new domain, which resulted in different types of shrinkage functions to be applied to the wavelet coefficients (e.g., [5]–[8]), others continued developing the local Wiener idea. Mihçak et al. [9] included a prior on the local variance, proposing a maximum a posteriori (MAP) estimation scheme. A similar Bayesian approach was followed in [10]. Portilla et al. [11] refined the Bayesian model by including the full signal and noise covariance in the model, and simplified the formulation of the problem by using a Gaussian scale mixture (GSM) [12], [13] frame [14]. These Wiener adaptive methods followed an empirical Bayes strategy: first, estimate the local signal variance, and then, apply a Wiener filter using that estimation. In a further work, Portilla et al. [15], instead, introduced a single-step Bayes least squares (BLS) estimator, which resulted in a significant performance increase. Up to now, all Wiener-adaptive methods either consider just the local variance, disregarding the covariance structure for each subband [9], [10] or consider the covariance fixed, up to a scale factor [11], [15]. In both cases, they perform a fine adaptation, in the sense that they estimate the local variance using small neighborhoods. On the other hand, there is broad literature on local covariance estimation and its applications (e.g., [16]). Covariance estimation allows for a selective characterization of salient local spectral features. However, as it requires a much larger window than variance estimation, it loses the ability to represent rapid variance spatial fluctuations in small areas. As a result, for typical images, the extra complexity of a covariance-adaptive model, by itself, does not pay off in performance terms with respect to simpler variance-adaptive models. Here, we describe a new approximation that provides both kinds of adaptation in the same model [1]. Hammond and Simoncelli [17] have also proposed an interesting extra adaptation step for the BLS-GSM model, in their case constrained to the local dominant orientation.
BLS-GSM [15] has provided the state-of-the-art in image Gaussian denoising for the last few years. During the last year, however, some new methods have achieved impressive results, both visually and in L2 norm terms, some of them clearly surpassing the BLS-GSM performance. Many exploit an important characteristic of typical image: they present a high spatial redundancy in the sense that a local feature is likely to be found in several (possibly disconnected) spatial locations within an image (e.g., a corner of a window, in a building façade, a straw bar, or, more simply, part of an edge along the edge line). An important part of the differences between these repeated local features may be caused by noise, so this fact can be exploited for noise removal. The usual approach is to consider that every neighborhood of pixels can be written as a linear combination of other neighborhoods. Buades et al. [18] do this matching for every pixel location, using the contextual similarities among neighborhoods around the pixels. Kervrann and Boulanger [19] improve this method by introducing an adaptive neighborhood as well as a statistical estimation of the model parameters. Hiraoka and Parks [20] use a total least squares formulation for square neighborhoods of different sizes. Dabov et al. [21] use block-matching and filter the matched blocks in 3-D. Once these blocks are 3-D filtered, they are denoised by a combination of hard-thresholding and local Wiener filtering. To the best of our knowledge these authors achieve the current state of the art in performance terms. However, their method is not based on a single model for the estimation, that is, it applies re-estimation. We refer here to re-estimation when the result of an estimation method is used as the input data for a different method with the same estimation purpose. Though possibly effective and convenient in some practical situations, re-estimation is not conceptually clean practice, as it implies the overlapping of two mutually incoherent models. Another successful path to image denoising is to adapt to the geometry of the image. Foi et al. [22] have presented a method based on a shape-adaptive discrete cosine transform (SA-DCT), which adapts the local support for the data being processed to the image contents. They use the same re-estimation method as in [21]. Elad et al. [23] achieve impressive results as well by creating an image-adaptive sparse representation. They used the K-SVD algorithm to obtain a dictionary that describes the image content very effectively. A big part of the high performance of this latter method is due to using a nonlinearly adaptive image representation, which requires a heavy computational load.

In contrast to the methods referred above, the one we propose here can use any linear multiresolution representation (though best results are obtained with multioriented translation invariant pyramids) and it does not make use of re-estimation. The underlying model is very simple, and the resulting method is noniterative, efficient and robust, inheriting and extending on the good properties of the original BLS-GSM, such as its ability to deal with any spectral noise density.

When approaching the deconvolution problem, we can see that most methods are iterative and global, e.g., [24]–[27]. Among the global methods, some authors (e.g., [28]–[33]) have approached the deblurring problem by decomposing it in two steps (see Fig. 1), a) a linear global compensation of the blur in the frequency domain and b) a nonlinear local denoising method. We emphasize that this procedure is not a form of re-estimation, because the problem addressed in the first step is different from the one addressed in the second step. The advantage of this decomposition is that it converts the deconvolution problem into an easier denoising problem, opening the door to recent powerful noniterative methods whenever they are able to represent and remove noise with arbitrary spectral density (PSD). Note that, by default, classical denoising assumes white noise. In contrast, here, even if the original noise is white, it gets colored by the blur compensation filtering in the first step.

Few authors instead, e.g., [34] and [35], have approached the deconvolution problem using local models, both for the image and for the degradation, with the advantage of simplifying the formulation of the problem. Portilla and Simoncelli [35] used a local GSM model within an efficient noniterative method which provided good performance in most cases. However, the local blur compensation used in this method has some intrinsic limitations (see Section III-B). Also, its performance depends critically on an accurate estimation of the power spectrum of the original image.

In this paper, we propose a noniterative two-step deblurring method based on the previously described spatially adaptive de-noising. It provides a suboptimal, but very robust and efficient, choice for the prefiltering kernel. This approach preserves the advantages of using a local model for image denoising (simple and noniterative) whereas it overcomes the problems associated to locally compensate for the effects of convolution. It also increases both the performance and the robustness of its predecessors BLS-GSM restoration methods.

The outline of this paper is as follows. Section II describes our image statistical modeling, including a description of the image representations used. Section III explains how to use our model to do both denoising and deblurring. Implementation details are reported in Section IV, and Section V presents and discusses a broad set of simulations and their corresponding results. Section VI concludes the paper.

II. IMAGE MODELING

A. Image Representation

Image representation is a key issue for image processing, and its choice depends on the particular application. For image restoration, it has been widely demonstrated (see, e.g., [15], [36], and [37]) that translation invariant, overcomplete representations provide far better results than nonredundant representations, such as critically sampled wavelets. Among them, those based on a multiscale analysis of the image (usually in a

![Fig. 1. Two-step scheme.](image-url)
pyramid format), with orientation selectivity, typically provide the best results. Examples are the steerable pyramid [38], the curvelets [39], or the complex wavelet [40], among others. However, it has also been observed that each representation scheme fits best for processing certain image features, but none seems superior in absolute terms (see, e.g., [41]). Some authors have proposed to combine them by averaging their results [42], [1] although more powerful techniques are based on a nonlinear basis selection using a sparseness criterion [41]. A different approach is to use image content adaptive basis functions such as wedgelets [43], bandlets [44] or K-SVD overcomplete dictionaries [23], among others. In this work, we have used two representations exhibiting complementary features: the full steerable pyramid (FSP) [15] and a translation invariant version of the Haar wavelet [1] that we have termed translation invariant haar pyramid (TIHP).

1) Full Steerable Pyramid: The FSP [15] is an extension of Simoncelli and Freeman’s original steerable pyramid [38] that splits the original high-pass residual into the same number of orientations as the rest of the pyramid levels. Some good properties of the FSP are that a) their filters can be made highly selective in frequency and b) it is flexible in its number of orientations. Its drawbacks are coupled with these advantages: a) due to their high spectral selectivity, its filters have a large spatial support and b) due to the possibly very high number of subbands, it has a relatively high computational cost (best denoising results with BLS-GSM are reported for eight orientations [15], which results in a redundancy factor close to 19). According to these features, FSP is more suitable for processing texture-rich images.

2) Translation Invariant Haar Pyramid: The TIHP is an overcomplete version of the Haar wavelet [45]. It is translation invariant and it provides almost perfect reconstruction. For a detailed description, see the Appendix. In contrast to FSP, TIHP’s kernels are very small, providing excellent spatial localization. Its moderate redundancy factor (7) and reduced computational cost, make it fast and easy to operate with. TIHP’s main weaknesses are its poor oblique orientations selectivity and the low spectral localization of its kernels. These features made this representation suitable for images with little texture, most of the energy being concentrated along edges and lines (especially if oriented horizontally or vertically).

B. Image Statistics and Gaussian Scale Mixtures

Image statistics were first described in terms of its spectral properties. Empirical studies showed that typical behavior followed a $f^{-2}$ law, being $f$ a small correction factor [46], [47]. Thus, usual approximations, 50 years ago, were to consider images as samples from highly correlated Gaussian random fields. However, in 1989, Mallat [3] showed with empirical data that bandpass responses to typical images are far from Gaussian: its marginal probability density functions have a pronounced peak at zero, and heavy tails. This caused the sparse behavior of the wavelet coefficients: only a small proportion of them concentrates a high proportion of the total energy, and this occurs not only for the whole representation, but also for every subband. The histograms reflect high amplitude responses caused by most relevant localized image features (such as edges, lines, corners, etc.), whereas spatially dominant responses are typically small in amplitude, mainly corresponding to low-contrast texture within homogeneous patches in the scene. Therefore, the dominant source for non-Gaussianity of image statistics is the inhomogeneity of the image contents, which causes abrupt changes and localized high amplitude responses at the objects’ boundaries. In this way, it was definitely assessed the inadequacy of using global Gaussian distributions to model typical images. In 1997, Simoncelli [48] made another critical observation about image statistics: there is a strong coupling among neighbor wavelet coefficients in amplitude. This revealed explicitly a fundamental non-Gaussian joint statistical behavior that was coherent with the marginal densities studied so far. This feature has been exploited for image compression [49], [50] texture modeling and synthesis [51], and image restoration (all the predecessors of this work), among other applications. Two years later, Wainwright and Simoncelli found that Gaussian scale mixtures, GSM [12], was the proper mathematical tool to capture both the sparseness and the amplitude coupling of wavelet coefficients [14]. GSM’s have typically been used as a local description for neighbor coefficients in a pyramid (e.g., [15]), but they can also be used in a more powerful (but also more complex) global statistical description of the image in the transformed domain [52].

The GSM model in the wavelet domain [14] is able to represent the observed strong local variance fluctuations in neighborhoods of coefficients at the different spatial locations, scales and orientations. This spatial inhomogeneity of the local variance explains, on its own, the two basic nonlinear statistical patterns mentioned above, namely, sparseness of the responses and amplitude coupling among neighbors. The original GSM model assumed a fixed normalized covariance for all the neighborhoods at a given subband. However, it is clear that typical images have different spatial correlation patterns at different locations in a subband, even if the filtering associated with each subband greatly reduces the spatial variability of the local covariance in the image domain. We illustrate this in Fig. 2, where, using a TIHP vertical subband of Barbara, we have compared the local auto-covariance samples estimated for the whole subband (assuming a $3 \times 3$ neighborhood) with that computed from two different regions, both having high-contrast features. In this

Fig. 2. Autocovariance variability within Barbara TIHP vertical band. We assume a $3 \times 3$ neighborhood and compute each autocovariance patch using samples from the associated bounded region.
case, the corresponding auto-covariance samples reveal significant differences in local orientation. Note that the auto-covariance of the whole subband reflects the average behavior, reducing the potential of a more spatially selective signal/noise discrimination. Thus, to further increase the realism and performance of our model, it is justified to use a Gaussian scale mixture with a spatially adaptive covariance for the local signal, that is, a space-variant GSM.

Let us consider an image decomposed into oriented subbands at multiple scales. We denote as $x_j(n, m)$ the coefficient corresponding to the subband $j$, centered at spatial location $(n, m)$. We denote as $x_j(n, m)$ a neighborhood of coefficients clustered around this reference coefficient (see neighborhood structure in Section IV). We model these neighborhoods as Gaussian scale mixtures (GSMs). A random vector $\mathbf{x}$ is a GSM [12] if it can be expressed as the product of two independent random variables

$$
\mathbf{x} = \sqrt{N} \mathbf{u}
$$

where $z$ is a positive scalar and $\mathbf{u}$ is a zero-mean Gaussian vector. We model the desired space variant behavior by defining $C_{\mathbf{u}}^{(n,m)}$ as the (space variant) covariance of $\mathbf{u}^{(n,m)}$. The density of $\mathbf{x}$ is determined uniquely by $p_z(z)$ and $C_{\mathbf{u}}^{(n,m)}$ as follows:

$$
p_{\mathbf{x}}^{(n,m)}(\mathbf{x}) = \int p_{\mathbf{z}}^{(n,m)}(\mathbf{z}|\mathbf{x})p_{\mathbf{z}}(\mathbf{z})d\mathbf{z} = \int \frac{\exp\left(-\frac{1}{2} \mathbf{x}^T C_{\mathbf{u}}^{(n,m)}^{-1} \mathbf{x}/2 \right)}{(2\pi)^{N/2}|C_{\mathbf{u}}^{(n,m)}|^{1/2}} p_{\mathbf{z}}(\mathbf{z})d\mathbf{z}
$$

where $N$ is the dimensionality of $\mathbf{x}$ and $\mathbf{u}$ (the size of the neighborhood). Note that $x_j(n, m)$ is conditionally Gaussian for a given $z$. There are different possibilities for the mixing density $p_z(z)$, which are described in [15]. As in the latter work, we have assumed it constant in log $z$ for all the subbands in the working range, which corresponds to a noninformative Jeffrey’s prior.

III. IMAGE RESTORATION

A. Denoising Using Space-Variant GSMs

We aim to solve the classical denoising problem, where an image is corrupted by additive zero-mean independent Gaussian noise of known (but arbitrary) spectral density. As many other methods, we perform the denoising in the pyramid domain and obtain the image estimation by reconstructing the image from the estimated coefficients. The observation model is, in this case

$$
\mathbf{y} = \mathbf{x} + \mathbf{w}
$$

where $\mathbf{y}$, $\mathbf{x}$, and $\mathbf{w}$ are vectors (in a lexicographic order) of observed, original, and noise coefficients, respectively.

1) Two-Level Adaptive Denoising: Note that the model in [15] used a fixed signal covariance at each subband, providing only (fine) adaptation to the local variance (through the GSM hidden multiplier). Here we include another level of (coarse) adaptation by also letting the signal covariance to be a function of the spatial position at each subband. Following the GSM model presented in Section II-B, we can express a neighborhood of $N$ noisy coefficients as

$$
\mathbf{y} = \sqrt{N} \mathbf{u} + \mathbf{w}
$$

where $\mathbf{w}$ is the zero mean Gaussian noise vector with covariance matrix $C_{\mathbf{w}}$. Note that $C_{\mathbf{w}}$ (unlike the signal covariance $C_{\mathbf{u}}$) is assumed constant for all neighborhoods within a subband (but different for each subband). By GSM definition $z$ and $\mathbf{u}$ are independent, and by assumption $\mathbf{w}$ is independent, too. Hence, the density of the observed neighborhood vector conditioned on $z$ is Gaussian (zero-mean), with covariance $C_{\mathbf{y}z} = zC_{\mathbf{u}} + C_{\mathbf{w}}$

$$
p(z|\mathbf{y}) = \frac{\exp\left(-\frac{1}{2} \mathbf{y}^T (zC_{\mathbf{u}} + C_{\mathbf{w}})^{-1} \mathbf{y}/2 \right)}{(2\pi)^{N/2}|zC_{\mathbf{u}} + C_{\mathbf{w}}|^{1/2}}.
$$

2) Parameter Estimation: Given (1), signal and noise covariances estimations are needed. The neighborhood noise covariance, $C_{\mathbf{w}}$, is computed at each subband by applying the pyramidal representation to a deterministic function with the same sample autocorrelation as the noise (for instance, a scaled delta function for white noise). The elements of $C_{\mathbf{w}}$ are computed by averaging the products of pairs of coefficients over all the neighborhoods of the subband. Note that, unlike noise, this deterministic function yields a result that is free from random fluctuations.

The neighborhood observation covariance, $C_{\mathbf{y}z}$, can be computed from $C_{\mathbf{y}z}$ by taking the expectation over $z$

$$
C_{\mathbf{y}} = E\{C_{\mathbf{y}z}\} = E\{z\}C_{\mathbf{u}} + C_{\mathbf{w}}.
$$

We can set, without loss of generality, $E\{z\} = 1$ obtaining $C_{\mathbf{y}} = C_{\mathbf{u}} + C_{\mathbf{w}}$. To estimate the spatially variant signal covariance $C_{\mathbf{u}}$ we assume that it changes smoothly over the image, and, thus, that it may be sampled by estimating it into a set of nonoverlapping regions. We divide each subband into $B \times B$ blocks, which are assumed uniform in covariance. For each of these blocks we estimate $C_{\mathbf{u}}$ by subtracting $C_{\mathbf{w}}$ from the sample covariance of its coefficients, $C_{\mathbf{y}z}$.

3) Coefficient Estimation: For each neighborhood $\mathbf{y}$ of observed coefficients, we estimate $x_c$, the reference coefficient, at the center of the neighborhood. The Bayes least squares (BLS) estimate can be expressed in this case [15] as

$$
E\{x_c|\mathbf{y}\} = \int_{0}^{\infty} p(z|\mathbf{y})E\{x|\mathbf{y}, z\}dz.
$$

Thus, the solution is written as the average of the Bayes least squares estimate of $x_c$ when conditioned on $z$, weighted by the posterior density of $z$, $p(z|\mathbf{y})$. The advantage of the GSM model is that the coefficient neighborhood vector $\mathbf{x}$ is Gaussian when conditioned on $z$. This fact, coupled with the assumption of ad-

---

2For notational simplicity from here on we drop the superscript $j$ and the indices $(n, m)$. Note that the following equations are the same as in [15], but now $C_{\mathbf{w}}$ represents a local covariance, as it depends on the spatial position.

3We have empirically tested that the negligible performance improvement achieved by overlapping the blocks does not pay off the large increase of computation.
Additive Gaussian noise means that the expected value inside the integral of (2) is simply a Wiener estimate

$$E[x|y,z] = zC_u (z C_u + C_w)^{-1} y.$$ 

We can simplify the dependence of this expression on $z$ by diagonalizing the matrix $z C_u + C_w$ (see [15]). Note that, unlike in [15], now this diagonalization depends on the spatial location (as $C_u$ is a function of $(r_s, m)$), and, thus, it needs to be computed once per each considered spatial location of each subband.

### B. Deblurring-by-Denoising Using Space-Variant GSMs

We face the problem of image restoration, where prior to noise addition, the signal has been degraded by (typically) a low-pass filter. The observation is modeled as follows:

$$y(n, m) = h(n, m) * x(n, m) + w_0(n, m).$$  \hspace{1cm} (2)

where $h$ is a convolution kernel (assumed known) applied to the original image $x$, usually called Point Spread Function, PSF. $w_0$ is Gaussian noise of known (but arbitrary) power spectral density. The same BLS-GSM local approach used in [15] for image denoising was used in [35] for image restoration. However, in this case, the fact of having a local description of image and degradation caused problems when trying to compensate for certain convolution kernels. Let us analyze this problem. Given a small neighborhood for the local model, we can only aspire to compensate for smoothly varying filters in the frequency domain having no zeros (so their regularized inverse filters will also be smooth in frequency, and, as a consequence, small in the image domain). A fixed absolute neighborhood size, thus, would very severely limit the deconvolution power of the method. A better option is to use a fixed neighborhood size for all the scales in a multiresolution representation. In such a way, we are able to represent larger effective neighborhoods as we go up in the pyramid. We explain this with a 1-D example in Fig. 3.

We show the Fourier representation, in magnitude, of a spatially limited signal, which corresponds to the same neighborhood size, measured in number of coefficients, at different scales. As we go up in the scale, the frequency support becomes smaller, and, thus, more rapidly varying responses are obtained. So, thanks to multiresolution, we can use (and compensate for, whenever there are no zeros) smooth filters at high frequencies, less smooth at medium frequencies, and rapidly varying at low frequencies. Such constraint is compatible with many low-pass kernels (e.g., Laplacian, Gaussian, etc., see bottom left of Fig. 3). However, there are some practically important kernels, such as those integrating uniformly in their spatial supports, whose behavior in frequency is not flat at all at high frequencies, but present strong oscillations around zero. Trying to compensate for those zeros requires abrupt responses for the regularized inverse filtering, which implies necessarily a large spatial support (see bottom right of Fig. 3). Therefore, uniform blurring cannot be properly treated by using this kind of local multiresolution schemes. Another drawback of the referred method [35] is that, to work properly, it requires a reasonably accurate estimation of the cross covariance among the original and the filtered image. This implies, in turn, the need of estimating the power spectrum of the original image ($P_x$) from the degraded observation, which is a difficult problem. Using a general model for $P_x$ translates in this case into a significant performance fall.

1) Deblurring-by-Denoising: In order to keep the advantages of using a local model, while overcoming its drawbacks, we use here the two-step scheme depicted in Fig. 1. This deconvolution strategy was proposed more than ten years ago [28], and it has been followed by many authors [29]-[33]. The two steps are: a) blur compensation through global linear filtering, and b) local denoising of the blur-compensated observation.

We can express the observation model of (2) in the frequency domain as

$$Y_0(u, v) = H(u, v) X(u, v) + W_0(u, v).$$

where $H$ is the Fourier transform of the blur kernel $h$. Let $G$ be the filter applied in step 1 (here in after, prefilter). In order to compute the optimal prefilter in MSE sense, $G_{opt}$, we must couple both steps

$$G_{opt} = \arg \min_G \left\{ E \left\{ \| X - f(G Y_0) \| ^2 \right\} \right\} \hspace{1cm} (3)$$

4For notational simplicity from here on we drop the frequency indices $(u,v)$.
where $f$ represents here the denoising operator (step 2), so $f(GY_0)$ represents the output of the whole process. The expectation inside (3) can be expressed as

$$E\left\{\|X - f(GY_0)\|^2\right\} = E\left\{\|(X - GHX) + (GHX - f(GY_0))\|^2\right\} \approx E\left\{||(X - GHX)||^2\right\} + E\left\{||(GHX - f(GY_0))||^2\right\}$$

where the r.h.s. of the latter approximation represents the sum of the quadratic errors coming from the deblurring and the denoising steps. This assumes that both errors are approximately decorrelated, which is true, at least under usual conditions.\textsuperscript{5} Thus, to compute the optimal prefilter $G$ we need an explicit estimation error (image and pixels), the higher is the overall linear distortion, but is strictly positive on the GSM.

We represent absolute frequency providing the highest value inside (3) can be expressed as

$$\frac{H^*}{|H|^2 + \alpha \cdot \left(\frac{P_{\eta 0}}{P_x}\right)}$$

where $P_{\eta 0}$ and $P_x$ are the power spectral densities of noise and signal respectively. This regularized inversion compensates for the filter phase response. The balance between noise and blur suppression is controlled by the regularization parameter $\alpha$. Low values yield in noisy estimations ($\alpha \approx 0$ is just an inverse filter)\textsuperscript{6} whereas high values tend to suppress too much signal. The higher is $\alpha$, the higher is the overall linear distortion, but the lower it is, the noisier is the estimation. $\alpha = 1$ yields the Wiener filter, which is a far too high value when using typical images and nonlinear denoising methods in the second step, as we show in Section III-B2.

Step 2: Local Denoising: After step-1, the observation yields

$$Y_1 = GY_0 = GHX + GW_0 = H_x X + W_1.$$ 

Because $G$ has been optimized in the first step, no further attempt to correct the remaining linear distortion $H_x = GH$ is made, and what is left is to denoise the prefiltered image.

2) Prefiltering Parameter Estimation: Based on the typical images power spectrum [46], [47], we have used a simple model for $P_x$, namely $P_x = k/f^2$, where $f$ represents absolute frequency and $k$ an arbitrary constant. In order to force the model to be consistent with the estimated original variance ($\delta_x^2 = \sigma_{90}^2 - \sigma_{la0}^2$), the following expression must hold:

$$\delta_x^2 = \int_{f>0} |H(f)|^2 \frac{P_x f df}{k df}$$

Hence, $k$ is set to

$$k = \delta_x^2 / \int_{f>0} |H(f)|^2 / f df.$$ 

Despite its simplicity, this model is effective and, in contrast to other methods (such as [35]), the associated $P_x$ estimation error does not seem to have a critical impact on the overall performance.

Unlike other authors [28], [30] that used a model to choose the regularization parameter $\alpha$, we have computed it empirically, training with a set of standard test images (House, Cameraman, Barbara, Boat, and Lena), using Gaussian convolution kernels and white Gaussian noise of 0.4 to 3.2 and 1 to 16 standard deviations, respectively (in a geometric scale). For each combination of degradation parameters, we have computed the optimal $\alpha$, through golden search. We also computed a quality decrease tolerance interval around the optimal value, according to a performance decrease of 0.05 dB with respect to the maximum. Finally, we chose the common $\alpha$ value providing the highest proportion of experiments within the quality tolerance intervals. Fig. 4 shows the optimality ratio for each alpha value, showing a peak for $\alpha = 0.3$, for which value around 84% of the experiments provide quasi-optimal results. We have also tested the method with very different PSFs [2] (such as a uniform kernel), and the maximum of the histogram was still very close to 0.3, although the percentage dropped slightly (73% for the uniform kernel). We have also tested that the results are fairly independent on the particular images chosen for the training, by comparing these results with those obtained by excluding each time the image to be processed in the training set.

IV. IMPLEMENTATION

A. Boundary and Block Handling

As demonstrated in [15], careful boundary handling is crucial for achieving a high denoising performance. In this case, it is even more critical than for [15], because of the image partition into blocks. To illustrate this, let us consider a $L \times L$ GSM neighborhood in a $N_s$-scale pyramid. In the upper part of Fig. 5, we can see the corner of a subband at the the highest scale in the pyramid ($L = 3$ and $N_s = 4$ in this example). All GSM neighborhoods with reference coefficients at the boundaries of the subband are incomplete, lacking those coefficients marked by the dashed lines. To solve this, the observation is extended, by mirror reflection with no repetition. $2^{(N_s-1)}(L-1)/2$ pixels at each side, which yields the required extra coefficients for the subbands at the top of the pyramid (see dashed square in Fig. 5). When dealing with block-based methods, the boundary artifacts are potentially stronger, because of the larger proportion of boundary coefficients compared to the whole image. The
same as before, each block must be extended \((L - 1)/2\) coefficients at each side. The inner blocks are enlarged with true signal, whereas boundary blocks use the mirror extension made for the whole subband. This produces a small overlapping between extended blocks. We have tested that increasing this overlapping does not translate into any significant improvement.

Another practical issue is the block arrangement within subbands. In order to have complete blocks in all subbands, we should ensure that all the subbands have dimensions that are integer multiples of the block size, i.e., \( \mathbf{I}_{xy} /(2^s B) \) is an integer vector for \( s = 0 \ldots N_x - 1 \), being \( \mathbf{I}_{xy} \) the image dimension vector. As original image dimensions will not comply, in general, with this requirement (see dark shaded area in Fig. 5), they will generally require to be extended (by mirror reflection) to also fit this constraint (see light shaded areas in Fig. 5). In practice, we apply a relaxed version of this criterion, processing the whole subbands when they are smaller than the chosen block size, so effectively reducing the range of \( s \) for imposing this constraint.

**B. Denoising Experiments**

We have used the TIHP and the FSP as image representations. We chose the number of scales to 4 and 5 for 256 × 256 and 512 × 512 images, respectively. Note that the number of orientations is fixed for the TIHP (3), whereas we used eight orientations for the FSP (as in [15]). We used a 3 × 3 GSM neighborhood around each coefficient (same as in [15]). In the case of FSP, we also include a coefficient at the same location and orientation at the next coarser scale (a parent). Different block sizes \( B \) were tested for estimating the signal covariance \( (B = 16, 32, 64, 128, F, \) where \( F \) means the whole image). Both the GSM neighborhood and the block size are the same in all pyramid levels, except for the subbands at the top of the pyramid, which have no parent, and those subbands with a size smaller than the block.

**C. Deblurring Experiments**

We have implemented the global preprocessing (step 1) in the Fourier domain, using \( \alpha = 0.3 \). For the denoising (step 2), we have selected manually TIHP or FSP based on the relative amount of texture in the image: FSP was chosen for texture-rich and TIHP for poorly textured images. The number of orientations were kept the same as for denoising whereas the chosen number of scales was 4 for all images sizes. The GSM neighborhood was also 3 × 3, but with no parent now in either representations. The block size was fixed to 32 × 32.

**D. Computational Cost**

As a guide, Table I shows the computation time (in seconds) of our unoptimized denoising and deblurring Matlab 7.0 code, on Windows XP over a 3.0-GHz Intel Pentium IV CPU with 1-Gb RAM, for different representations, blocks, and images sizes. For comparison, we also include the computation time of the original BLS-GSM at the bottom part of the table. We extract three basic patterns: a) the cost scales approximately inversely with \( B \), roughly a factor of 3 when we halve the block; b) the cost of the first step in the deblurring algorithm is negligible when compared to the cost of the second step; c) the computation time is significantly higher, but still comparable to that of original BLS-GSM. The main difference (in cost terms) between the original BLS-GSM and the proposed method is the increase of calls to functions. The proposed method usually makes one call per block, whereas the original BLS-GSM makes only one per subband. Note also that the number of calls are multiplied by a factor of 4 when we halve the block size. This has a big impact due to using unoptimized interpret code in our current implementation. Finally, the cost scales roughly with a factor of 8 when we double the image size.

**V. RESULTS AND DISCUSSION**

We have used five standard test images to evaluate our algorithms: Barbara, Lena, House, Cameraman, and Peppers. We
have chosen Barbara and House for all visual comparisons, as two examples with very different amount of texture.

A. Denoising

We selected four standard images: Barbara, Lena (both 512 × 512), House and Peppers (both 256 × 256). We have degraded them by adding additive white Gaussian noise of different standard deviations (σ = 5, 10, 15, 20, 25, 50).

1) Model Features: In Table II, we present the results of our proposed algorithm, in terms of peak signal-to-noise ratio (PSNR), 10 \log_{10}(255^2/\sigma^2) in decibels, where \sigma^2 is the mean square error. We include results for the representations used, TIHP and FSP, and compare our method (column labeled two-level Adaptive, 2LA) with the standard version of BLS-GSM [15] (labeled one-level Adaptive, 1LA). We also include the block size, B, for which the method provided the best results. In Fig. 6, we show some visual results (cropped and zoomed, for artifacts visibility) on Barbara and House images. We include results for TIHP and FSP (second and third rows) and a comparison with standard BLS-GSM (by columns). We now present a discussion along the followings directions:

Comparison to Standard BLS-GSM: The enhancement of the model proposed in Section II-B provides us a general improvement of the results over standard BLS-GSM method, both visually and in MSE sense. Comparing the columns 1LA-2LA of Table II, we can see a significant performance increase. The new results are better (with no exception) for both representations TIHP and FSP. As we can see, the improvement is more important for images having local spectral peaks (like Barbara), whereas for other images (e.g., Lena), the improvement is more modest. This behavior is related to the amount of covariance variability among different regions in the image (see Fig. 2 for details). We can compare visually the differences between our method and the standard BLS-GSM, for Barbara and House images, in the columns of Fig. 6. Note how artifacts on the roof line of House disappear with the new method (second column). For Barbara we can see a very strong improvement, the new results having much less artifacts, and much more texture is recovered.

Representation comparison: FSP Versus TIHP: As we can see in Table II, the results confirm what we pointed out in Section II-A: Different representations fit better different images, depending on their contents. House and Peppers, two poorly textured images, are best denoised using the TIHP, whereas for Lena and Barbara, the best results are achieved with FSP. We can also see the significant difference in visual terms, by inspecting House and Barbara images in the second (FSP) and third (TIHP) rows of Fig. 6. In the case of House, the visual difference is significant (less artifacts in general), but not dramatic. However, in the case of Barbara the difference is striking: FSP provides a much better result than the one obtained with TIHP. If one of them should be chosen basing only on results (ignoring computational cost), under a minimax risk criterion, FSP would be the choice.

Another issue is the different relative gains between the new method when using TIHP versus FSP, with respect to the standard BLS-GSM. By comparing in Table II the performance difference between our method and the original BLS-GSM for each image and noise level (by rows), we see that the improvement is bigger when we use the TIHP. The main reason is that whereas FSP kernels are highly selective in frequency, TIHP kernels are not, and, hence, there is more spatial variability of the covariance inside each subband in the latter representation. This increases the difference between local covariances and the global one (see details in Section II-B). We have also tested the benefit of averaging FSP and TIHP results (in the same way as

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>FSP (8 Orientations)</th>
<th>TIHP (3 Orientations)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5/14.15</td>
<td>38.49</td>
<td>38.55</td>
</tr>
<tr>
<td>10/28.13</td>
<td>35.60</td>
<td>35.66</td>
</tr>
<tr>
<td>15/24.61</td>
<td>33.90</td>
<td>33.96</td>
</tr>
<tr>
<td>20/22.11</td>
<td>32.67</td>
<td>32.71</td>
</tr>
<tr>
<td>25/20.17</td>
<td>31.69</td>
<td>31.72</td>
</tr>
<tr>
<td>50/14.15</td>
<td>28.61</td>
<td>28.61</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>FSP (8 Orientations)</th>
<th>TIHP (3 Orientations)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5/14.15</td>
<td>38.65</td>
<td>38.88</td>
</tr>
<tr>
<td>10/28.13</td>
<td>35.29</td>
<td>35.44</td>
</tr>
<tr>
<td>15/24.61</td>
<td>33.54</td>
<td>33.64</td>
</tr>
<tr>
<td>20/22.11</td>
<td>32.30</td>
<td>32.34</td>
</tr>
<tr>
<td>25/20.17</td>
<td>31.32</td>
<td>31.30</td>
</tr>
<tr>
<td>50/14.15</td>
<td>28.22</td>
<td>28.18</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>FSP (8 Orientations)</th>
<th>TIHP (3 Orientations)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5/14.15</td>
<td>37.30</td>
<td>37.43</td>
</tr>
<tr>
<td>10/28.13</td>
<td>33.74</td>
<td>33.81</td>
</tr>
<tr>
<td>15/24.61</td>
<td>31.7</td>
<td>31.72</td>
</tr>
<tr>
<td>20/22.11</td>
<td>30.27</td>
<td>30.24</td>
</tr>
<tr>
<td>25/20.17</td>
<td>29.17</td>
<td>29.10</td>
</tr>
<tr>
<td>50/14.15</td>
<td>25.94</td>
<td>25.94</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>FSP (8 Orientations)</th>
<th>TIHP (3 Orientations)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5/14.15</td>
<td>37.30</td>
<td>37.43</td>
</tr>
<tr>
<td>10/28.13</td>
<td>33.74</td>
<td>33.81</td>
</tr>
<tr>
<td>15/24.61</td>
<td>31.7</td>
<td>31.72</td>
</tr>
<tr>
<td>20/22.11</td>
<td>30.27</td>
<td>30.24</td>
</tr>
<tr>
<td>25/20.17</td>
<td>29.17</td>
<td>29.10</td>
</tr>
<tr>
<td>50/14.15</td>
<td>25.94</td>
<td>25.94</td>
</tr>
</tbody>
</table>
other authors [42]), obtaining an average increase of 0.15 dB over the best one.

Block Size Considerations: The optimal block size depends on the image contents in the first place, but also on the noise level. The lower the SNR, the larger the optimal size. This makes sense, as local signal covariance estimation stops being reliable for small block sizes in the presence of high noise, so covariance spatial adaptation no longer pays off. In Fig. 7, we show, for Barbara and House, a plot of PSNR (with range adjusted between 0 and 1) when using $B = 16, 32,$ and 128 versus the noise standard deviation. Here one corresponds to the best result obtained with the set of block sizes tested, and zero to the worst result. We can state, as expected, that the best result for low noise levels are achieved with the smallest size block ($B = 16$), whereas $B = 128$ provides better performance for high noise variance. Overall, it seems that a finer adaptation ($B = 16$) provides better results than coarser covariance adaptations. However, differences are, in general, small, and, from a computational cost view, the smaller is the block the higher becomes CPU time. As we can see in Fig. 7, the curve associated to a block size of 32 is a good tradeoff in practice, providing best results for moderate noise levels and not dramatically decreasing when in the extremes.

2) Comparison to State-of-the-Art Methods: In Fig. 8, we show a visual comparison of our method and, what to the best of our knowledge, are the three ones with highest PSNR performance available in the literature [21]–[23]. For House, our method behaves the worst in PSNR terms, mainly because the other methods recover the wall texture and ours not. Its visual appearance is, nevertheless, convincing. For instance, the roof line is sharper and less contaminated by artifacts in our result than in [22]. For Barbara, our method behaves similarly to the other three, both visually and in MSE terms. Fig. 9 shows, also for Barbara and House, a graphical comparison, in terms of input—output PSNR, among our method and the top three reporting the best results [21]–[23]. Our method is represented by a solid line, whereas the standard BLS-GSM is included as a reference using a dash line. As we can see, our method behaves competitively to current state-of-the-art [21], especially considering its moderate computational cost. We also refer the reader to [53] for a recent perceptual comparison of state-of-the-art denoising methods.
In this paper, we have presented two methods, for image denoising and deblurring. The denoising algorithm, based on space variant Gaussian scale mixtures, is applicable to any overcomplete wavelet representation. Although the proposed model is just a conceptually simple extension of the one presented in [15], it provides a significant improvement in performance terms, still with a moderate computational cost. We have shown that the reason of the improvement is that, by modeling the local covariance spatial variability, it provides a more realistic model and a higher noise filtering selectivity. The improvement depends on both the image content and the representation used.

We have tested two complementary representations concluding that one of them is more suitable for highly textured images (FSP), whereas the other (TIHP) is better for processing images for which texture is not so relevant.

Our deblurring method is based on a two-step decoupling scheme: linear global blur compensation and nonlinear local nonwhite noise removal. This strategy provides flexibility to use any denoising method able to deal with nonwhite Gaussian noise, as the one presented here. The resulting algorithm overcomes the severe limitations imposed by local approaches to image deconvolution while keeping their advantages. In addition, and in contrast to some state-of-the-art methods, it provides completely fails (due to reasons explained in Section III-B). Compared to [33], note that, in general, our method achieves better results with Barbara, and worse with House and Cameraman. Being [33] a two-step method with a similar first global deconvolution step, the different behavior mainly comes from the denosing step [22] (based on the SA-DCT transform). This method is specially fitted to preserve sharp edges and contours (present in House and Cameraman), whereas it is not as powerful as ours to characterize local spectra. This translates into a poorer performance when dealing with highly texturized images (as Barbara).

In Fig. 10, we show a visual comparison among our method, the state-of-the-art methods [24], [33] and BLS-GSM standard restoration [35]. For Barbara, the degradation used is PSF1 with \( \sigma = \sqrt{2} \). In this case our method provides the best result, both visually and in PSNR terms. Particularly, the diagonal lines of the table cloth are better recovered by our method than by our competitors. For the House image, we have used PSF2 (uniform kernel) with \( \sigma = \sqrt{0.308} \). Here, we can see how the previous BLS-GSM restoration method [35] completely fails, whereas the new version provides results competitive to the state-of-the-art [24], [33].

Finally, we also tried our deblurring algorithm with real degradations. We restored a micro-display image captured under controlled laboratory conditions. We used a tilted lens to produce severe blur and the noise was reduced before processing by averaging 30 captures. In Fig. 11, we show the degraded image and the result of our method using an estimation of the degradation parameters. As we can see, the method behaves very well, substantially improving the visible detail in the image.

VI. CONCLUSIONS AND FUTURE WORK

In Table III, we show averaged Increment in signal-to-noise ratio (ISNR) results of applying our deblurring method to the training data set (see Section III-B2 for details). As we can see, the results exhibit a significant improvement for the whole range of tested blur and noise degradations. In Table IV, we show the results for six deconvolutions experiments with House, Cameraman (using THP) and Barbara (using FSP) comparing our method with the previous BLS-GSM restoration [35], and with those which, to the best of our knowledge, represent the current state-of-the-art [24], [33]. In the first two columns, we have replicated the experiments from [32]. The kernel used (PSF1) is \( h_{i,j} = (1 + i^2 + j^2)^{-1} \), for \( i, j = -7 \ldots 7 \). The next column reproduces the experiment in [28], [30], which uses a \( 9 \times 9 \) uniform kernel (PSF2). The fourth column uses a \( 5 \times 5 \) separable kernel (PSF3) with coefficients \([1,4,6,4,1]/16\) as in [54]. The last two columns use Gaussian kernels with 1.6 and 0.4 standard deviation (PSF4 and PSF5, respectively). The noise standard deviations are shown in the second row of the table. The results show that [33] provides the best results in a 44.4% of the experiments, [24] + [25] in a 16.7% and the proposed method in a 38.9% of the experiments. Comparing with the BLS-GSM restoration original method [35], the new method provides a very significant improvement for all experiments, especially for the uniform degradation (PSF2), where the local BLS-GSM approach...
Fig. 8. Visual comparison of denoising results on House and Barbara images cropped to 80 x 80 pixels (starting in pixel [1,78] and [1,221], respectively, [column, row]). Initial PSNR 28.13 dB in both cases $\sigma_\omega = 10$. From left to right and top to bottom (House/Barbara PSNR): results for [22] (35.98/33.48); results for [21] (36.71/34.98); results for [23] (35.94/34.42); our results using a 16 x 16 block with TIHP/FSP representation (35.56/34.60).

Fig. 9. PSNR output versus PSNR input, in decibels. We use TIHP for House and FSP for Barbara. Our results, for $\sigma_\omega = [5, 10, 15, 20, 25, 50]$, are in Table II. We compare to three state-of-the-art methods: diamonds for [21], crosses for [22], circles for [23], down triangles for [15], and squares for our new method. The latter two results are represented with dashed and solid lines, respectively.

TABLE III
AVERAGED RESULTS AS ISNR, IN DECIBELS, OF THE TRAINING DATA SET (SEE SECTION III-B2 FOR DETAILS)

<table>
<thead>
<tr>
<th>NOISE</th>
<th>$\sigma$</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>0.4</td>
<td>2.86</td>
<td>1.86</td>
<td>2.81</td>
<td>5.03</td>
<td>7.63</td>
</tr>
<tr>
<td>L</td>
<td>0.8</td>
<td>6.65</td>
<td>4.84</td>
<td>3.73</td>
<td>3.92</td>
<td>5.85</td>
</tr>
<tr>
<td>U</td>
<td>1.6</td>
<td>3.87</td>
<td>3.18</td>
<td>2.72</td>
<td>3.01</td>
<td>4.73</td>
</tr>
<tr>
<td>R</td>
<td>3.2</td>
<td>3.21</td>
<td>2.85</td>
<td>2.56</td>
<td>2.66</td>
<td>3.87</td>
</tr>
</tbody>
</table>

TABLE IV
PERFORMANCE COMPARISON WITH OTHER STATE OF THE ART METHODS IN TERMS OF ISNR, IN DECIBELS. FIRST ROW SHOWS THE BLURRING KERNEL USED (SEE TEXT FOR DETAILS), AND THE SECOND DENOTES NOISE STANDARD DEVIATION. BEST RESULTS ARE HIGHLIGHTED

<table>
<thead>
<tr>
<th>Blur</th>
<th>PSF 1</th>
<th>PSF 2</th>
<th>PSF 3</th>
<th>PSF 4</th>
<th>PSF 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_\omega$</td>
<td>$\sqrt{2}$</td>
<td>$\sqrt{8}$</td>
<td>$\sqrt{0.308}$</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>HOUSE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSNR</td>
<td>25.62</td>
<td>25.47</td>
<td>24.11</td>
<td>28.10</td>
<td>27.83</td>
</tr>
<tr>
<td>[22]</td>
<td>8.47</td>
<td>6.63</td>
<td>10.71</td>
<td>4.22</td>
<td><strong>4.49</strong></td>
</tr>
<tr>
<td>[21]</td>
<td>8.46</td>
<td>6.93</td>
<td>-0.44</td>
<td>4.37</td>
<td>4.34</td>
</tr>
<tr>
<td>[33]</td>
<td><strong>9.05</strong></td>
<td>7.64</td>
<td>10.71</td>
<td>5.10</td>
<td>4.03</td>
</tr>
<tr>
<td>Ours</td>
<td>8.64</td>
<td>7.03</td>
<td>9.04</td>
<td>4.30</td>
<td>4.11</td>
</tr>
<tr>
<td>CAMERAMAN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSNR</td>
<td>22.23</td>
<td>22.16</td>
<td>20.77</td>
<td>24.03</td>
<td>23.36</td>
</tr>
<tr>
<td>[22]</td>
<td>7.46</td>
<td>5.24</td>
<td>8.16</td>
<td>2.84</td>
<td>3.18</td>
</tr>
<tr>
<td>[26]</td>
<td>7.40</td>
<td>5.15</td>
<td>8.10</td>
<td>2.85</td>
<td>-/-</td>
</tr>
<tr>
<td>[35]</td>
<td>6.84</td>
<td>5.29</td>
<td>-1.61</td>
<td>2.56</td>
<td>2.83</td>
</tr>
<tr>
<td>[33]</td>
<td><strong>8.25</strong></td>
<td><strong>6.34</strong></td>
<td><strong>8.57</strong></td>
<td>2.56</td>
<td>3.05</td>
</tr>
<tr>
<td>Ours</td>
<td>7.45</td>
<td>5.55</td>
<td>7.33</td>
<td>2.73</td>
<td>3.25</td>
</tr>
<tr>
<td>BARBARA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSNR</td>
<td>23.33</td>
<td>23.26</td>
<td>22.49</td>
<td>24.21</td>
<td>23.77</td>
</tr>
<tr>
<td>[24]</td>
<td>3.76</td>
<td>1.99</td>
<td>3.98</td>
<td>0.9</td>
<td>0.92</td>
</tr>
<tr>
<td>[35]</td>
<td>5.70</td>
<td>3.28</td>
<td>-0.27</td>
<td>1.44</td>
<td>0.95</td>
</tr>
<tr>
<td>[33]</td>
<td>5.73</td>
<td>3.01</td>
<td>4.88</td>
<td>1.58</td>
<td>0.91</td>
</tr>
<tr>
<td>Ours</td>
<td><strong>6.85</strong></td>
<td><strong>3.80</strong></td>
<td><strong>5.07</strong></td>
<td>1.94</td>
<td>1.36</td>
</tr>
</tbody>
</table>

As shown, the proposed empirical link between the two steps of the method (parameter setup) is suboptimal. It would have been much more elegant to obtain the absolute optimal prefilter by using an error model of the denoising method.

Authorized licensed use limited to: IEEE Xplore. Downloaded on January 27, 2009 at 04:19 from IEEE Xplore. Restrictions apply.
Fig. 10. Visual comparison of deblurring results on House and Barbara images cropped to $80 \times 80$ pixels (starting in pixel $[1,78]$ and $[1,221]$, respectively, column, row). We include PSNR values, in decibels (in agreement with ISNR values of Table IV). From left to right and top to bottom (House/Barbara): original; degraded image (24.11/23.33) using PSF2 and $\sigma_w = \sqrt{0.308}$ for House and PSF1 and $\sigma_w = \sqrt{2}$ for Barbara; results for [33] (34.76/29.05); results for [24] (34.76/27.09); results of standard BLS-GSM [35] with FSP (23.68/29.02); our new method (House with TIHP, Barbara with FSP) (33.11/30.17).

Unfortunately, that is a difficult goal, given the nonlinear character of high performance denoising methods. Nevertheless, the proposed empirical approach for choosing the prefiltering parameters is simple and robust: we have tested that the same prefiltering parameters can be used for different degradations without provoking a serious performance fall. This strategy is also very flexible, because the parameters can be easily tuned to any kind of degradation by training with kernels of the desired shape. All these features, coupled with the moderated (and image-content independent) computational cost, make this method suitable to restore images in real situations [55].

Summarizing, the solutions presented here are model based, with few ad-hoc components that do not compromise a general conceptual coherence. Despite their simplicity, they provide state-of-the-art performance, both in denoising (slightly worse results than the best published ones) and deblurring (similar to the best ones on average).

A conceptually simple extension of the application of the methods presented here is to associate each spatial location with a different blurring kernel and/or with a different noise power spectral density. That would allow us to restore images with spatially variant degradations. This is interesting when dealing with low-quality optics, which cannot be accurately modeled as convolutions, because they are not translation invariant. Although this implies some technical challenges (mainly related to block’s boundary handling), we believe that it is a feasible goal, with a potential impact on the imaging industry.

We believe that BLS-GSM-based models still have a strong potential for image processing. This work can be seen as a first
Table V:
Top: 1-D Building Functions, in spatial and frequency domain. The underlined coefficient represents the one at the origin. Middle: Description of the 1-D representation responses. Bottom: TIHP Analysis filters in terms of external products of the 1-D building functions. The synthesis filters are their complex conjugates.

<table>
<thead>
<tr>
<th>Spatial Domain</th>
<th>Fourier Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi_1 = 1,1/2$</td>
<td>$\Phi_1(\omega) = (1 + e^{-j2\pi\omega})/2$</td>
</tr>
<tr>
<td>$\psi_1 = 1, -1/2$</td>
<td>$\Psi_1(\omega) = (1 - e^{-j2\pi\omega})/2$</td>
</tr>
<tr>
<td>$\phi_2 = 1,0,1/2$</td>
<td>$\Phi_3(\omega) = \cos(2\pi\omega)$</td>
</tr>
<tr>
<td>$\psi_2 = 1,0,-1/2$</td>
<td>$\Psi_3(\omega) = \sin(2\pi\omega)$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Spatial Domain</th>
<th>Fourier Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_1(\omega) = \psi_1(\omega)$</td>
<td>$H_2(\omega) = \Phi_1(\omega)\psi_1(\omega)$</td>
</tr>
<tr>
<td>$L_1(\omega) = \Phi_1(\omega)$</td>
<td>$L_2(\omega) = \Phi_1(\omega)\Phi_2(\omega)$</td>
</tr>
</tbody>
</table>

| $H_n(\omega) = L_{n-1}(\omega)\sin(2^{n-1}\pi\omega)$, $(\omega \leq 2^{n-1})$ | $L_n(\omega) = L_{n-1}(\omega)\sin(2^{n-1}\pi\omega)$, $(\omega \leq 2^{n-1})$ |

APPENDIX
Translation Invariant Haar Pyramid

We present an overcomplete version of the Haar orthogonal wavelet [45], which is translation invariant, due to the lack of aliasing, and which provides almost perfect reconstruction. Basically it is obtained by decimating its à trous [56] subbands from the third scale on, counting upwards. For simplicity, we start with the 1-D case, and after that, the 2-D extension is described.

At the top of Table V, the 1-D representation building functions are described (in spatial and frequency domains). Note how the samples at the origin are underlined. The middle part of Table V shows the filter responses at the first and second scales in terms of the building functions. Fig. 12 shows the magnitude responses for four scales. The first two scales (two upper panels) coincide with those of the à trous representation (full resolution, no decimation). Further scales are decimated each time by a factor of 2, after applying an ideal low-pass filter to avoid intra-subband aliasing, as shown in the bottom half of Fig. 12. Dashed curves represent the corresponding à trous response. The vertical dashed lines depict the cutoff frequency at each scale. See middle part of Table V for filters descriptions.

The extension to 2-D is done following the system diagram depicted in Fig. 13. The filters ($A(t_x,t_y)$ and $B(t_x,t_y)$) boxes are built by properly doing external products of the 1-D filters, as shown at the bottom of Table V. The pyramid is implemented by recursively splitting the image into a set of oriented subbands (vertical, horizontal, and double diagonal) and a low pass band, which is subsampled by a factor of two along both axes. The reconstruction is achieved hierarchically by convolving each
subband with its complex-conjugated filter, and adding the outputs.

**ACKNOWLEDGMENT**

The authors would like to thank Prof. M. Elad, Dr. A. Foi, Dr. K. Dabov, and Prof. M. Figueiredo for providing them with their results and software for testing. They would also like to thank C. Dorronsoro and CIDA (Spain) for facilitating the capturing of real degraded images in their laboratories.

**REFERENCES**


**Jose A. Guerrero-Colón** was born in Málaga, Spain. He received the M.S. degree in computer science from the Escuela Técnica Superior de Ingeniería Informática, Universidad de Málaga, in 2003. He is currently pursuing the Ph.D. degree and working on several projects related to his main topic: new visual-statistical models applied to the restoration of photographic images corrupted by blur and noise. He is a grant holder of an FPU contract, funded by the Ministerio de Educación y Ciencia, from the Spanish Government, at the Department of Computer Science and Artificial Intelligence, Universidad de Granada, Spain.

**Luis Maneuera** received the M.S. degree in computer science from the Escuela Técnica Superior de Ingeniería Informática, Universidad de Granada, Spain, in 2001. He is currently pursuing the Ph.D. degree in nonlinear sparse image representation methods and their applications.

For two years, he was with the European Centre for Nuclear Research (CERN), Geneva, Switzerland. He is a grant holder of an FPI contract at the Department of Computer Science and Artificial Intelligence, Universidad de Granada, Spain, funded by Ministerio de Educación y Ciencia, from the Spanish Government.

**Javier Portilla** received the M.S. and Ph.D. degrees from the E.T.S.I. Telecommunicación, Universidad Politécnica de Madrid, Spain, in 1994 and 1999, respectively.

After a postdoctoral stay at New York University, he was an Assistant Professor at the Computer Science and Artificial Intelligence Department, Universidad de Granada, Spain. He obtained a “Ramón y Cajal” research position, which he still enjoys, now in the Instituto de Óptica, Consejo Superior de Investigaciones Científicas, Madrid.

He is interested in image statistical modeling and its degradation sources for designing Bayesian methods, as well as new, general-purpose nonlinear image representation schemes, inspired by biological vision, and their application to image processing.