# Regularization Operators for Natural Images Based on Nonlinear Perception Models

Juan Gutiérrez, Francesc J. Ferri, and Jesús Malo

***Abstract*—Image restoration requires some *a priori* knowledge of the solution. Some of the conventional regularization techniques are based on the estimation of the power spectrum density. Simple statistical models for spectral estimation just take into account second-order relations between the pixels of the image. However, natural images exhibit additional features, such as particular relationships between local Fourier or wavelet transform coefficients. Biological visual systems have evolved to capture these relations.We propose the use of this biological behavior to build regularization operators as an alternative to simple statistical models. The results suggest that if the penalty operator takes these additional features in natural images into account, it will be more robust and the choice of the regularization parameter is less critical.**

***Index Terms*—Early vision models, image restoration, natural image statistics, regularization.**

## I. Introduction

IMAGES are produced to record or display useful information regarding some phenomenon of interest. When the process of image formation and recording is not perfect, only a degraded version $i_d(x)$ of the original image $i_o(x)$ is available. The whole distortion process is usually modeled by means of a linear operator acting upon the original image, plus additive noise

$$i_d(x) = h(x, x') \otimes i_o(x') + n(x). \tag{1}$$

The purpose of image restoration is to estimate the original image from the distorted and noisy one. This is a nontrivial problem because, even if the operator is known, a naive inversion will amplify the noise [1].

Multiple techniques of different nature have been proposed to solve this problem. Most of them require (and critically depend on) some image model or some *a priori* assumptions about the image or the noise [1]–[3]. The most exploited feature of natural images for building an image model is the specific low-pass behavior of the power spectrum [4], [5]. This feature can be easily described in the spatial domain by an autocorrelation function or the coefficients of an autoregressive (AR) model [1], [6]. In more recent approaches, the image model includes additional features. In particular, certain relationships between neighboring coefficients in a transformed domain are considered [7], [8].

In the context of regularization techniques, the restored image is a tradeoff between closeness to the observed data and the absence of some *unwanted features* in the solution [9]. The problem is again to use an appropriate image model to design the regularization operator so as to preserve the *suitable features* in the solution.

In this paper, we put forward and analyze a class of penalty operators [10] for image regularization which are not based on a model of the image or the noise, but on the current model of low-level human visual perception [11], [12]. In this way, there is no need to make any *a priori* statistical assumptions.

The use of perception models to design regularization operators relies on the Barlow hypothesis [13], in which he argued that biological vision systems have evolved for an optimal processing of natural images. A number of results support this hypothesis. First, it has been shown that the early processing mechanisms in the visual cortex perform a linear wavelet-like transform using a set of filters similar to those obtained by applying *independent component analysis* (ICA) to a set of natural images [14]. Second, biological vision systems exhibit nonlinear interactions between the responses of the linear wavelet-like stage. In these nonlinearities, each coefficient is normalized by a combination of neighboring coefficients [11], [12], [15]. It has been shown that this behavior gives rise to signals with increased independence (see [15], [16], or [17]). By maximizing independence, biological visual systems capture the most salient features of natural images and, in a way, discriminate between useful (*suitable*) and negligible (*unwanted*) information.

From this ecological point of view, current models of early vision implicitly include the generic statistical properties of natural images (as for instance their rough spectral properties and the relationships between the coefficients in a local frequency domain [15], [17]). Therefore, it makes sense to use the perceptual response in the regularization context for restoration purposes, because it can tell us what features should be present in the solution and what features should be removed from it.

The use of simple human vision properties in regularization is not new [1], [2]. However, these approaches used simpler (linear) models and, of greater importance, the rationale was

Manuscript received May 5, 2004; revised November 4, 2004. This work was supported in part by the CICYT under under Grants TIC2003-01504 and TIC2003-08496-C04-01, in part by the Universitat de València under Grant UV-AE-03-29, in part by the Generalitat Valenciana under Grant Grupos04-08, and in part by the program DN8644. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Vicent Caselles.

J. Gutiérrez is with the Department d'Informàtica and the VISTA Laboratory, Universitat de València, 50. 46100 Burjassot,València, Spain (e-mail: juan.gutierrez@uv.es).

F. J. Ferri is with the Department d'Informàtica, Universitat de València, 50. 46100 Burjassot,València, Spain (e-mail: francesc.ferri@uv.es).

J. Malo is with the Department d'Òptica and the the VISTA Laboratory, Universitat de València, 50. 46100 Burjassot, Valencia, Spain (e-mail:jesus.malo@uv.es).

Digital Object Identifier 10.1109/TIP.2005.860345

very application oriented (similar to the one which is popular in coding literature [18], [19]): As the images are addressed to human viewers, features which are not subjectively relevant should be penalized.

Apart from using a more advanced model, the main idea of this paper is different: We want to stress the fact that the information about natural images implicitly included in the human visual response can be used as a successful alternative to explicit image models in the penalty operator design. Moreover, if the penalty operator takes into account these additional features in natural images, it may be more robust and the choice of the regularization parameter is less critical. It is important to stress that the statistical benefits of using this advanced perceptual model are not limited to its application in restoration: see [17] for an application in image coding.

This paper is organized as follows. In Section II, we review the regularization framework. In Section III, we stress that the $1/f$ spectrum shape is not the only common feature shared by natural images. In particular, we show that the coefficients of local Fourier representations are related. These relationships are equivalent to those recently reported in other local frequency domains [7], [8], [15], [17], [20]–[22]. In Section IV, we propose the use of the current nonlinear early vision response to define regularization operators. We show that this response implicitly takes into account both band limitation and the additional features illustrated in Section III. Our experimental results appear in Section V, followed by our concluding remarks.

## II. Regularization and Image Features

In the Tikhonov regularization framework [9], the image estimate $i'$ is the one that simultaneously minimizes both the departure from the observed data $f_1(h(x,x') \otimes i(x'), i_d(x))$, as well as a penalty functional $f_2(p(x,x') \otimes i(x'))$, which measures the presence of some *unwanted feature* in the solution

$$i'(x) = \arg\min_{i(x)} f_1\left(h(x,x') \otimes i(x'), i_d(x)\right) + \lambda^2 f_2\left(p(x,x') \otimes i(x')\right) \quad (2)$$

where $p(x,x')$ is a linear operator that extracts the features that should be minimized in the solution, and $\lambda$, referred to as the regularization parameter, establishes the tradeoff between both terms. If Euclidean distance is used to measure these departures $f_1$ and $f_2$, we have [9]

$$i'(x) = \arg\min_{i(x)} \left( \left| h(x,x') \otimes i(x') - i_d(x) \right|^2 + \lambda^2 \left| p(x,x') \otimes i(x') \right|^2 \right). \quad (3)$$

Many different choices exist for $f_1$, $f_2$ and the penalty operator $p(x,x')$. Common choices for the measures are $l_p$ norms [23], usually $l_1$ and $l_2$ norms. In particular, the effect of a small exponent $(p < 2)$ in $f_2$ relatively decreases the penalization of large variations in luminance, thus allowing the presence of sharp discontinuities in the solution [24]. On the other hand, the penalty operator is usually taken as a fixed high-pass filter (e.g., first or second derivative) or can be adaptively computed [1]. In this paper, we propose an adaptive operator and analyze its behavior in the particular case of (3).

Assuming a local approach in (3), $h$ and $p$ may be considered shift invariant, and, hence, $\otimes$ is a convolution; therefore, the solution in a region $x_0$ of a block-Fourier domain is [2]

$$I'(x_0,f) = \frac{H^*(x_0,f)}{\left|H(x_0,f)\right|^2 + \frac{\lambda(x_0)^2}{\left|P(x_0,f)^{-1}\right|^2}} I_d(x_0,f). \quad (4)$$

The meaning of the regularization operator in the Fourier domain $P(x,f)$ in terms of a simple image model can be inferred by analogy from the result using the optimal Wiener filter [2]

$$I'(x_0,f) = \frac{H^*(x_0,f)}{\left|H(x_0,f)\right|^2 + \frac{\left|N(x_0,f)\right|^2}{\left|I_o(x_0,f)\right|^2}} I_d(x_0,f). \quad (5)$$

In the Wiener solution, the inverse operator $H^*/|H|^2$ is modified using information about the power spectrum of the original signal and the noise (which are not available in general). Equation (5) suggests that the penalty operator $P$ and the parameter $\lambda$ should be related to the power spectrum density of the signal and the energy of the noise, respectively. Therefore, the conventional operator design is usually based on a model of the autocorrelation of the image [1], [6]. The smoothness assumption (limited band signal) gives rise to high-pass penalty operators; so, in this case the *unwanted feature* is the energy in the high-frequency region.

A critical issue in the regularization framework is the choice of the regularization parameter $\lambda$. This problem gets worse if the operator is not specifically matched to the signal. For instance, when a generic high-pass operator is used, the appropriate restoration of high-frequency components in the signal becomes very *sensitive* to $\lambda$. However, the importance of the regularization parameter may be reduced if the operator not only includes second-order relations, but also some other features which are common in natural images.

## III. Additional Features in Natural Images Beyond $1/f$ Spectrum

Natural images present a strong correlation between the luminance of pixels in the spatial domain. This is the reason for assuming Gaussian models for natural images [20], [25], [26]. Under the Gaussian assumption, a simple second-order approach fully describes the signal. This simple characterization has been successfully exploited in applications where removing the statistical dependence of the samples is required [27], [28]. Principal component analysis (PCA) on a set of natural images gives rise to Fourier-like eigenvectors, i.e., oscillating functions extending all over the spatial domain. Moreover, the energy (the square of the eigenvalues) is concentrated in the low-frequency PCA components [20], [25], [26]. Specifically, natural images exhibit a $1/f^{\gamma}$ power spectrum [4]. This justifies the use of high-pass regularization operators (such as first and second derivatives of the signal [1]) in image restoration.

However, natural images are not that simple. In order to consider higher order interactions, ICA techniques [29], [30] have been developed as an alternative to PCA. When applied to natural images, wavelet-like representations emerge, i.e., spatially

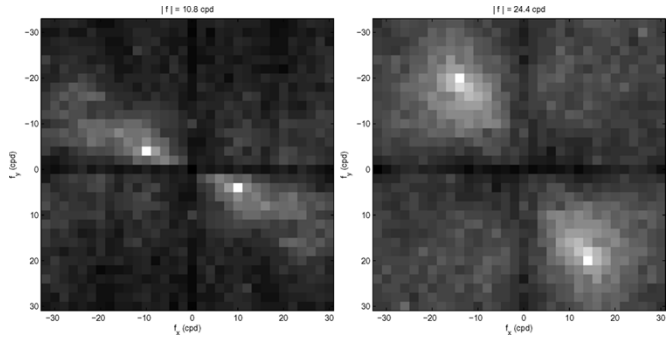


Fig. 1. Empirical interaction of two particular coefficients of the local Fourier transform with their neighbors. The absolute value of the frequency of these two coefficients is $|f| = 10.8$ and $|f| = 24.4$ cycles/degree (cpd). A sampling frequency of 64 cpd is assumed.

localized oscillating functions. Nevertheless, when representing images in these local frequency domains, not all higher order interactions are taken into account [7], [15], [17], [20]–[22]. Specifically, intriguing relationships among the coefficients of these linear frequency transforms still remain. Linear transforms cannot remove all the statistical dependence because natural images do not come from a linear combination of signals from independent sources (e.g., due to occlusion). Modeling these statistical relations led to improvements in image coding applications [21], [31].

Instead of modeling these relations explicitly, in this paper, we are going to take them into account by using a nonlinear perceptual model [11], [12]. The use of these perception models rather than the explicit statistical modeling of the relations in a local Fourier domain has been successfully applied in image coding applications (see [32], [33], or [17]).

In this section, we describe a way to show these relations in the particular local frequency domain which will be used to define the regularization operators. This is intended to clarify how the perceptual response reviewed in next section takes these relationships into account.

We have computed the cross-correlation between the absolute value of the $32 \times 32$ block-Fourier coefficients of the widely used natural images database [34]. As the local Fourier spectrum of natural images is not stationary, a direct comparison between coefficients at very different frequencies is biased. As stated above, natural images exhibit a $1/f^{\gamma}$ power spectrum, meaning that the comparison of a high-frequency coefficient with a low frequency coefficient is biased by the high energy of the latter. Therefore, instead of a direct comparison, we first divide each coefficient by $1/f$ (assuming that $\gamma \approx 2$ [5]). In that way, a unit mean process is obtained and a fair computation of the cross-correlation can be done.

Fig. 1 shows an estimate of the cross-correlation for two particular coefficients. Obviously, for any given coefficient, the cross-correlation is maximum for itself and its symmetrical. When compared to other coefficients, the cross-correlation decreases as the distance in frequency increases. The relationships illustrated in this example are consistent with those reported in other local frequency transform domains (see [7], [8], [15], [20]–[22], or [17]).

In summary, natural images have a low-pass spectrum (as assumed by conventional statistical models used for regularization), but they also exhibit the particular behavior illustrated by these relationships in a local frequency domain. This additional feature of natural images should not be neglected.

## IV. Perceptual Regularization Operators

The standard model of human low-level image analysis [11], [12], [15] has two basic stages in which the input image $i(x)$ is first transformed into a vector $I(x, f)$ in a local frequency domain (using a linear filter bank $T$) and then it is nonlinearly transformed into a response vector $r(x, f)$

$$i \xrightarrow{T} I \xrightarrow{R} r.$$

A local (block) Fourier transform will be used here as a model of the perceptual linear transform $T$. The response to each coefficient of the transform $R(I(x_0, f))$ is given by a nonlinearity in which the energy of each coefficient is normalized by a linear combination of the neighbors (*masking* behavior). This nonlinear response is commonly referred to as *divisive normalization* [11], [12]. For a particular region $x_0$ of a local Fourier transform, the response for each frequency component is

$$r(x_0, f) = R\left(I(x_0, f)\right) = \frac{\alpha(f)\,|I(x_0, f)|^2}{\beta(f) + \left(k(f, f') \otimes |I(x_0, f')|^2\right)} \tag{6}$$

where the parameters $\alpha(f)$, $\beta(f)$, and $k(f, f')$, are chosen to fit the psychophysical response of human viewers to periodic functions [16]. The vector $r$ is constituted by the responses of the set of perceptual mechanisms that analyze the image. Equation (6) describes the fact that the response of a particular mechanism tuned to $(x_0, f)$, not only depends on the energy of the linear output, $I(x_0, f)$, but is also influenced by its neighbors, $I(x_0, f')$, with $f' \neq f$ [11], [12], [15]. In particular, the sensitivity of the mechanism tuned to $(x_0, f)$ is decreased when the input image excites high responses in the neighboring coefficients of the linear stage. The kernel $k(f, f')$ determines the extent of this neighborhood in frequency and orientation. It is usually assumed to be Gaussian so that the influence of closer neighbors is bigger [11], [12], [15]. Fig. 2 shows the frequency dependence of the parameters in (6) and some examples of the response for three particular mechanisms. The effect of the denominator in (6) is to saturate the response for high-energy inputs in the local frequency region determined by $k$: Such input stimuli *mask* the response of the considered mechanism. Note that the overall behavior of the response is mediated by the band-pass filter, $\alpha(f)$, which is similar to the classical *contrast sensitivity function* (CSF) [35].

As stated in the introduction, the responses in (6) capture the most salient independent features of natural images and discriminate between useful and negligible information from the perceptual point of view. Therefore, the perceptual responses for a particular region of the observed data may tell us what features should be present in the solution and what features can be removed from it. This can be seen as a perceptual spectral estimation adapted to the most common features present in natural images. According to this perception-based heuristic, we may use

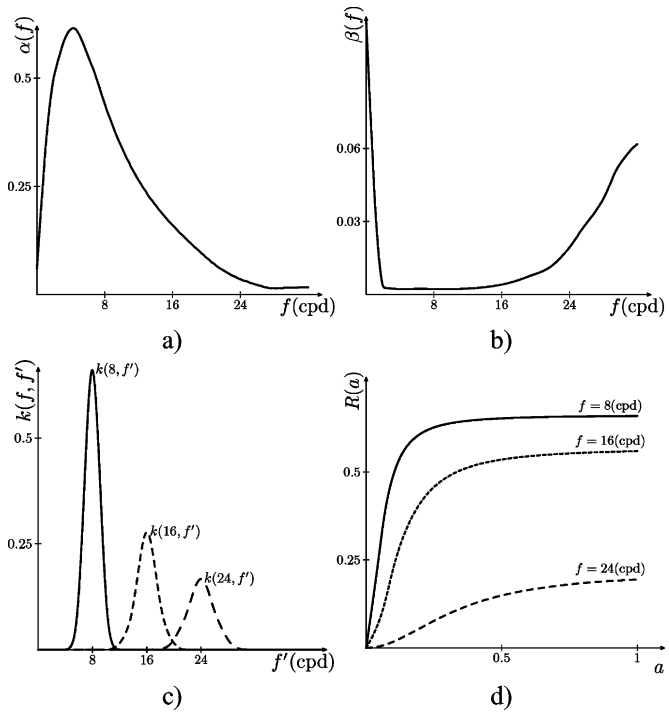


Fig. 2. Parameters of the vision model and nonlinear response functions. The values in d) assume that the Fourier coefficients (denoted by $a$) are expressed in contrast (amplitude over mean luminance). The responses of d) show the basic sigmoid behavior of (6), but they are just particular examples for isolated sinusoids. In general, these response curves will depend on the neighbor coefficients. Note that the parameters are slightly different from those reported in [17] because we are using here a local Fourier transform instead of a local DCT and a slightly different model. However, the final qualitative behavior in d) is the same.

the response (6) to estimate the (perceptually relevant) spectrum in (5). We, therefore, propose to define the regularization operator $P$ of (4) as inversely proportional to the response. Thus, each restored region in the local frequency domain $I'(x_0, f)$ will be obtained as

$$I'(x_0, f) = \frac{H^*(x_0, f)}{|H(x_0, f)|^2 + \frac{\lambda^2}{|R(I_d(x_0, f))|^2}} I_d(x_0, f). \quad (7)$$

Previous simpler perceptual approaches based on the linear CSF model [2] can be seen as a particular case of the more general class of operators presented here.

The perception model just described includes the behavior exhibited by natural images, as illustrated in Section III. Note that the Gaussian kernels in Fig. 2(c) follow the trend shown in Fig. 1: The interaction neighborhood increases with frequency. Therefore, the response implicitly includes two facts: rough low-pass behavior through the function $\alpha(f)$ and the particular relationships between the energy of neighboring coefficients through the kernels $k(f, f')$.

From the point of view of the Barlow hypothesis, the goal of the divisive normalization is to make those parts of the signal that cannot be predicted from the neighborhood more explicit. In addition, it has been reported that this energy normalization acts as a sort of divisive DPCM [16]: The energy in each transform coefficient $I(f)$ is predicted from the energy of the coefficients in the Gaussian neighborhood $k(f, f')$. This is how the saliency of the relevant coefficients of the linear transform is increased in the response [15], [16].

This predictive behavior is what makes normalization interesting for regularization operator design. Using the analogy between (5) and (7), we see that the role of the response model is the extraction of the relevant features of the image to estimate the spectrum of the original signal from its degraded version. Relevant features are those that cannot be predicted from the neighborhood. Wide-band noise introduces two nonnatural features: a noticeable change in the shape of the spectrum and increased predictability in the frequency domain. This is true if the bandwidth of the noise is bigger than the kernels in (6). In this case, the energy of the noise can be predicted from the energy of the neighbors. This represents a potential limitation of the proposed operators because they would consider a high-energy noise of too narrow a bandwidth as a relevant feature.

As in this framework, an adaptive penalty operator has to be obtained from degraded images, the operator should be robust to noise. *Robustness* in this paper is understood as the ability to preserve the same components of the signal that would be preserved with an ideal operator obtained from the original signal. The predictive behavior of the divisive normalization may be useful to achieve robustness. If an image is degraded by adding wide-band noise, the contribution of the noise to each transform coefficient can be predicted from the neighboring coefficients. In this way, the proposed operator will detect the relevant components despite the noise.

A real example of the ability of the response $r$ to capture relevant features in the presence of noise is shown in Fig. 3. Fig. 3(a) shows a block taken from the image Barbara with a texture (a salient feature at a particular frequency). The same block blurred and corrupted with white noise is shown in Fig. 3(b). Fig. 3(c) and (d) shows the corresponding regularization operators using the proposed method together with the standard second-derivative operator. As discussed above, the relevant feature at frequency $f \approx 18$ cpd is preserved despite the noise, while a second-derivative operator penalizes this feature.

## V. Results and Discussion

In this section, the properties of the proposed operator will be compared to other penalty operators taking three aspects into account: a) the sensitivity of the solution to $\lambda$, b) the robustness of the operators, and c) their performance in the restoration task using a variety of synthetic degradations. Finally, we also analyze the performance of the proposed method with naturally degraded images.

Five standard images have been used for the experiments: Barbara, Lena, Einstein, boats, and peppers. We assume that these $256 \times 256$ images have a physical extent of $4^\circ \times 4^\circ$, i.e., the sampling frequency is 64 cpd. The proposed regularization operator has been tested in all images with different degradations. The images were blurred using known linear shift invariant low-pass filters with normalized cutoff frequencies of 0.1, 0.5, and 0.85 (3.2, 16, and 27.2 cpd) and were corrupted with additive white Gaussian noise of different variances (from $\sigma_n^2 = 5$ to $\sigma_n^2 = 400$). These restoration problems range from

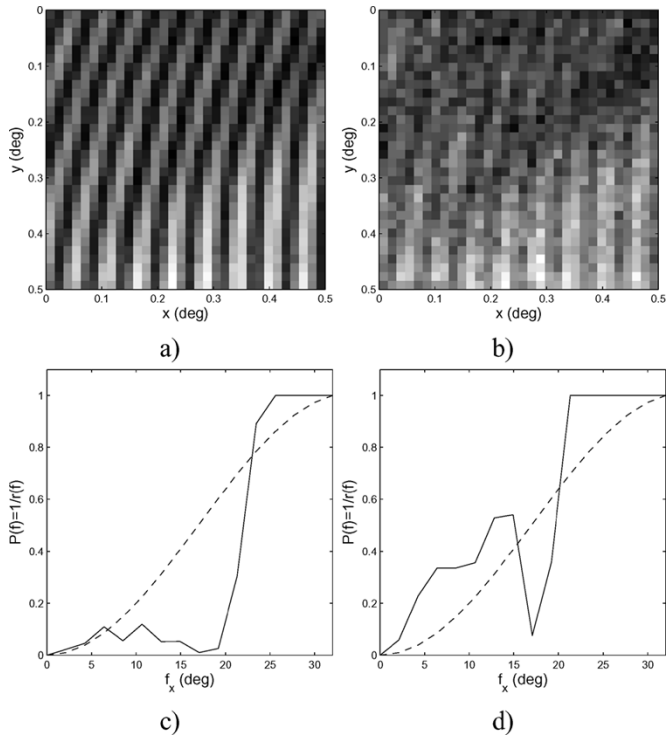


Fig. 3. Perceptual penalty operators (solid curves) in c) and d) obtained from a particular block a) without noise and b) with noise, compared with the operator based on (dashed curves) the second derivative. Notice how, with the proposed approach, the relevant feature (texture of 18 cpd) is not penalized.

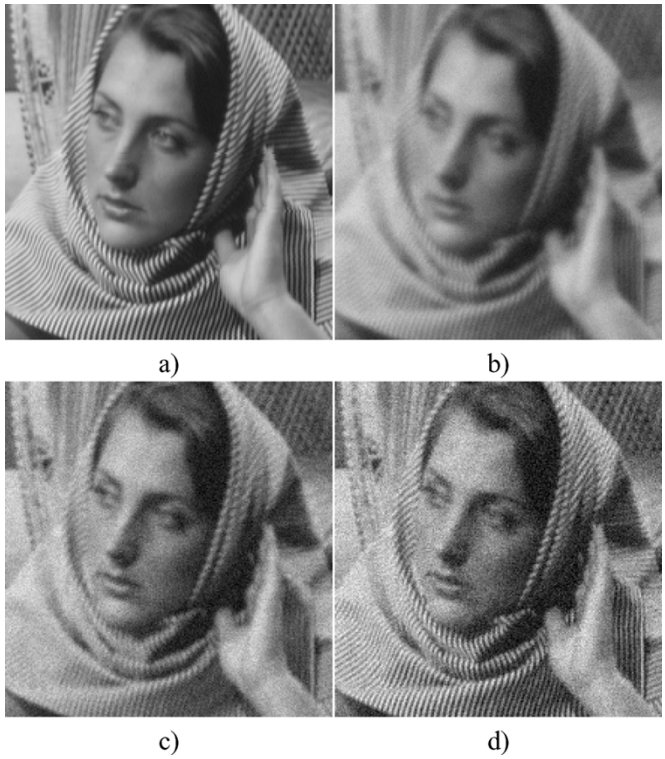


Fig. 4. Zoom of original and representative degraded images. a) Original Barbara image. b) Distorted image (almost deblurring problem): $f_c = 3.2$ cpd, $\sigma_n^2 = 15$. c) Distorted image: $f_c = 16$ cpd, $\sigma_n^2 = 100$. d) Distorted image (almost denoising problem): $f_c = 27.2$ cpd, $\sigma_n^2 = 200$.

almost pure deblurring (very low cutoff frequency and low energy noise) to almost pure denoising (high cutoff frequency and high-energy noise). Fig. 4 shows illustrative examples of these degradations using the standard image Barbara.

The different restoration methods have been applied to 50 realizations of the above mentioned corruption processes over each original image. In the experiments, the regularization parameter $\lambda$ has been obtained using the standard L-curve method [36] for each block of the local Fourier transform. In all cases, the borders of neighboring $16 \times 16$ blocks have been overlapped to reduce the blocking artifacts common to all methods.

Euclidean distortion measures may not take into account subjective aspects related to the empirical mean opinion score given by an average human observer [37]. Therefore, a perceptually meaningful metric square-root integral (SQRI) measure [38] was also used along with the standard peak signal-to-noise ratio (PSNR). However, the same trends revealed by PSNR were obtained for SQRI, as can be seen by subjective inspection of the images (which is the really important issue).

In this paper, we will compare the proposed operator with other operators based on spectral estimation, so that they can be used in (4). In particular, we considered three AR models [1], [6] of increasing neighborhoods as shown in Fig. 5. The coefficients of these AR models have been locally fitted to capture the nonstationarity of the signal. From those coefficients a local estimation of the power spectrum is obtained [39], to be used in (4).

Additionally, two classical spatially invariant operators are included for reference purposes: the second derivative and the

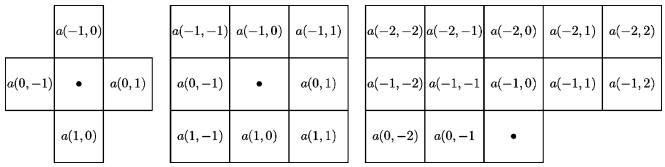


Fig. 5. Neighborhoods of the particulart two-dimensional AR models used here: (left) AR(4), (center) AR(8), and (right) AR(12). The dotted pixel is linearly predicted from the corresponding neighborhood weighted by $a(m, n)$.

CSF-based operator reported in [40]. On the one hand, the second derivative operator captures the generic $1/f$ spectral behavior of natural images. In fact, when AR models are trained on data sets which are large enough, operators which are quite similar to the second derivative are obtained. On the other hand, the CSF-based operator represents a simple example of the use of perceptual information in regularization.

### *A. Sensitivity to the Regularization Parameter*

As discussed in Section II, the sensitivity of the solution to the regularization parameter depends on the ability of the operator to capture the properties of the signal. Roughly speaking, a solution is more sensitive if small variations in the regularization parameter lead to big changes in the quality of the solution.

Given a particular operator $P$, a solution $\hat{i}_P(\lambda)$ can be computed for each value of $\lambda$. A particular $\lambda$ selection method, such as the L-curve [36], obtains a particular value $\lambda_L$, which leads to a particular solution $\hat{i}_P(\lambda_L)$. Assuming that the original image

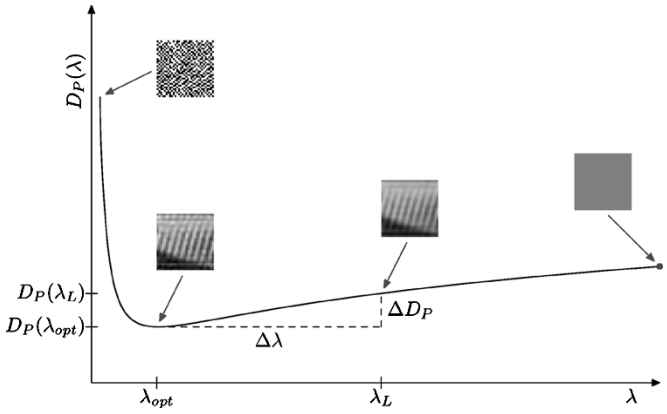


Fig. 6. Distortion of the set of solutions $\hat{i}_P(\lambda)$ for a particular operator $P$. The solution ranges from noise ($\lambda \ll$) to over-smoothed images ($\lambda \gg$). Representative points have been highlighted. Note that this sensitivity is related to the concavity of this curve which depends on $P$.

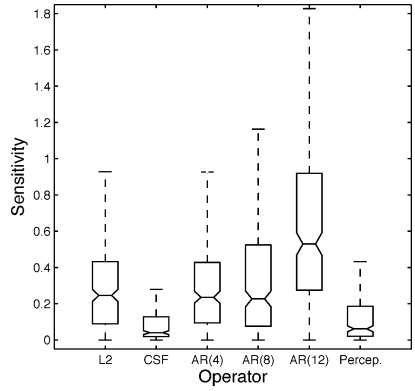


Fig. 7. Sensitivity of the different operators using image Barbara.

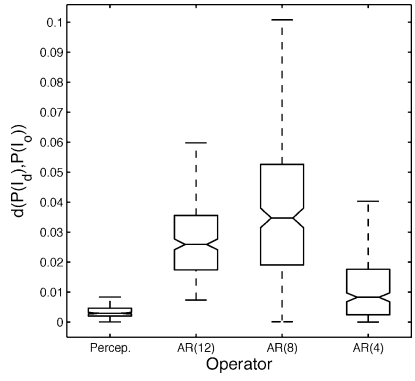


Fig. 8. Robustness of the different adaptive operators.

$i_o$ is available, and given a distortion measure $d(\cdot,\cdot)$, the optimal value of $\lambda$ could be computed as

$$\lambda_{\text{opt}} = \arg\min_{\lambda} D_P(\lambda)$$

where $D_P(\lambda) = d(i_o, \hat{i}_P(\lambda))$. The sensitivity of the solution to $\lambda$, when a particular operator $P$ is used, can be defined as the change in the image distortion with regard to the departure from the optimal $\lambda_{\text{opt}}$. That is

$$S(P) = \frac{\Delta D_P}{\Delta\lambda} = \frac{|D_P(\lambda_{\text{opt}}) - D_P(\lambda_L)|}{|\lambda_{\text{opt}} - \lambda_L|}.$$

The meaning of this sensitivity measure is illustrated in Fig. 6.

The particular sensitivity values considering different operators have been measured at each block of the Barbara image using both mean-square error (MSE) and SQRI. A standard boxplot (omitting outliers for clarity) of the corresponding results

TABLE I
PSNR FOR DIFFERENT OPERATORS AND NOISE LEVELS FOR FIVE IMAGES FOR A BLURRING WITH A CUTOFF FREQUENCY OF 3.2 cpd

| | $\sigma_n^2 = 5$ | $\sigma_n^2 = 10$ | $\sigma_n^2 = 15$ |
|---|---|---|---|
| L2 | 28.55 | 28.74 | 28.55 |
| CSF | 29.25 | 28.49 | 28.04 |
| AR(4) | 30.26 | 29.62 | 29.08 |
| AR(8) | 30.66 | 29.76 | 29.23 |
| AR(12) | 30.44 | 29.50 | 29.02 |
| **PER** | **30.71** | **29.75** | **29.20** |

TABLE II
PSNR FOR DIFFERENT OPERATORS AND NOISE LEVELS FOR FIVE IMAGES FOR A BLURRING WITH A CUTOFF FREQUENCY OF 16 cpd

| | $\sigma_n^2 = 25$ | $\sigma_n^2 = 50$ | $\sigma_n^2 = 75$ | $\sigma_n^2 = 100$ |
|---|---|---|---|---|
| L2 | 30.60 | 29.49 | 28.82 | 28.33 |
| CSF | 30.04 | 29.33 | 28.76 | 28.25 |
| AR(4) | 30.53 | 29.66 | 29.04 | 28.51 |
| AR(8) | 30.78 | 29.78 | 29.19 | 28.67 |
| AR(12) | 31.09 | 29.98 | 29.23 | 28.63 |
| **PER** | **31.15** | **30.43** | **29.87** | **29.34** |

TABLE III
PSNR FOR DIFFERENT OPERATORS AND NOISE LEVELS FOR FIVE IMAGES FOR A BLURRING WITH A CUTOFF FREQUENCY OF 27.2 cpd

| | $\sigma_n^2 = 100$ | $\sigma_n^2 = 200$ | $\sigma_n^2 = 300$ | $\sigma_n^2 = 400$ |
|---|---|---|---|---|
| L2 | 29.37 | 28.00 | 27.05 | 26.19 |
| CSF | 29.14 | 28.40 | 27.77 | 27.20 |
| AR(4) | 29.62 | 28.14 | 27.00 | 26.20 |
| AR(8) | 30.26 | 28.92 | 27.98 | 27.04 |
| AR(12) | 27.16 | 23.98 | 22.59 | 21.40 |
| **PER** | **30.67** | **29.77** | **29.23** | **28.364** |

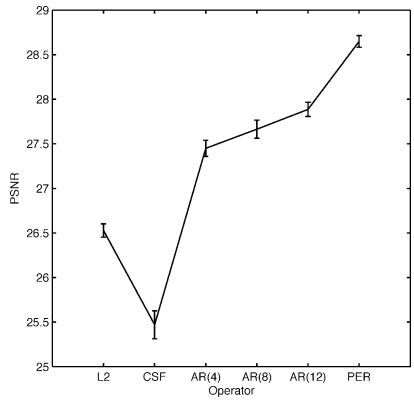


Fig. 9. Example of the results of a particular restoration experiment. Average distortions and standard deviations over 50 realizations of the degradation for a particular image (Barbara) and a particular combination of cutoff frequency (16 cpd) and variance of the noise ($\sigma_n^2 = 75$).


Fig. 10. Restoration results for the degraded image in Fig. 4(b) ($f_c = 3.2$ and $\sigma_n^2 = 15$). a) Second-derivative based result. b) CSF-based result. c) Image restored using adaptive AR(4). d) Image restored using adaptive AR(8). e) Image restored using adaptive AR(12). f) Image restored using the proposed method.

for the MSE case is shown in Fig. 7. Similar results are obtained when SQRI is used. The proposed operator and the CSF lead to smaller and less scattered sensitivity values. As a consequence, using perceptual operators makes the choice of $\lambda$ less critical.

### *B. Robustness of the Operators*

As discussed in Section IV, robustness in this paper is understood as the ability of an adaptive operator to preserve the same components of the signal that would be preserved with an ideal operator obtained from the original signal. Thus, this property only applies to adaptive operators.

If the original image $i_o$ is available, any adaptive operator $P$ estimated from the degraded image can be compared to the same operator $P_o$ obtained from the original one. The corresponding MSE between the operators can be used to assess their similarity. Consequently, we define the robustness of a particular adaptive operator as

$$R(P) = \mathrm{MSE}(P, P_o).$$

The robustness of the different adaptive operators considered for the Barbara image is shown as a boxplot in Fig. 8. As the proposed operator captures the complexity of the signal better, it is more robust to the presence of degradations.

### *C. Restoration Results*

Two different sets of restoration tasks have been considered to assess the performance of the proposed method: first, different synthetic degradations and, second, naturally degraded images.

Figs. 10–12 show representative restored images corresponding to the set of operators considered in this paper applied to the synthetic degraded images. These images correspond to the restoration of the degraded Barbara images shown in Fig. 14. Tables I–III show the distortion results for all the experiments carried out. These results are the average over the five considered images. Fig. 9 shows the details of one of the distorted images used to compute the averages in Tables I–III. As can be seen, the variances in the results due to the 50 realizations of the degradation are small. Similar plots are obtained for the remaining images and degradations.

From these results, it is obvious that the performance of spatially invariant approaches [a) and b) in Figs. 10–12] is worse than the performance of the adaptive ones [c)–d) in Figs. 10–12]. In particular, spatially invariant approaches give rise to oversmoothed areas in all cases. Similar results are obtained when

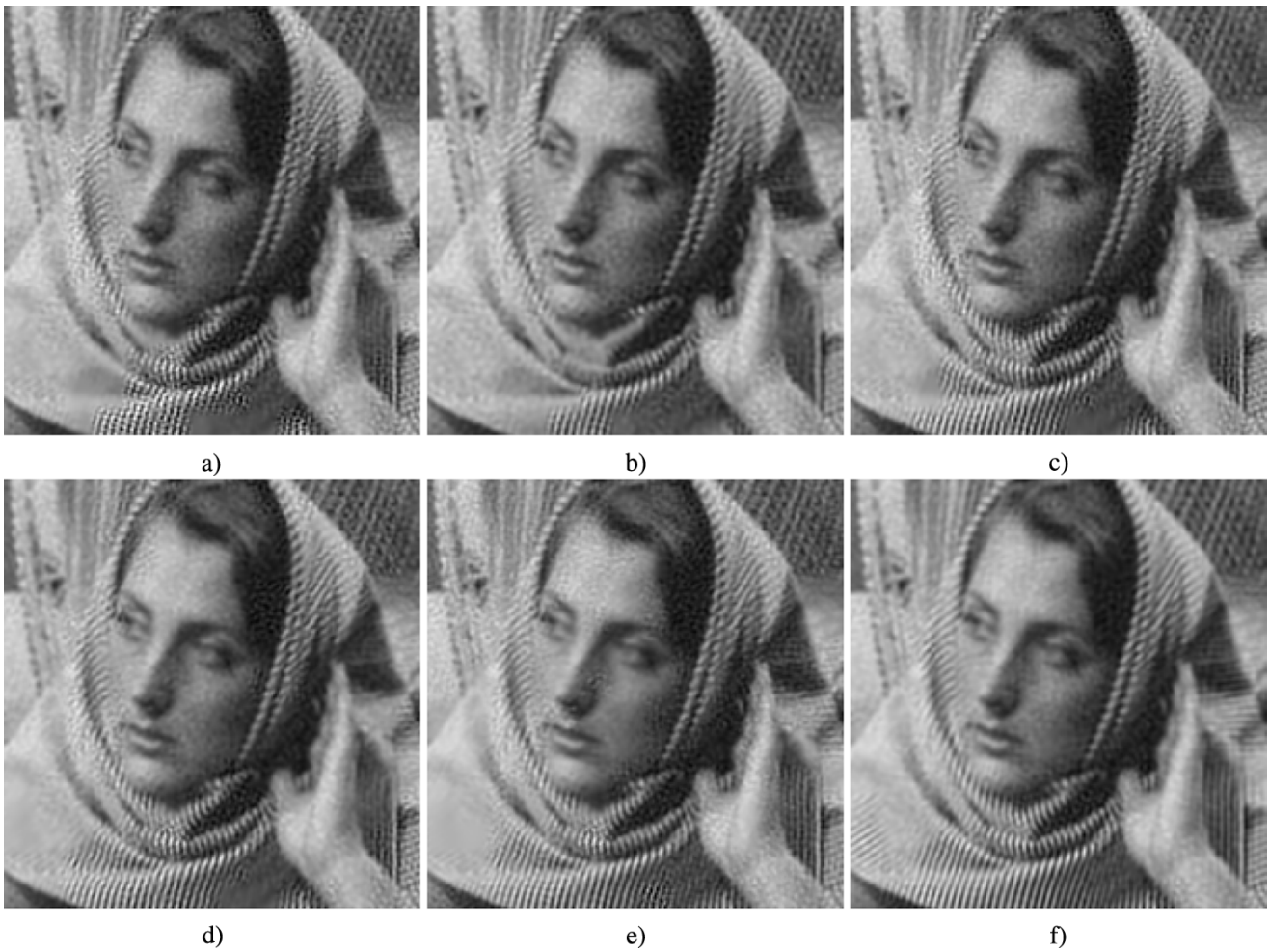


Fig. 11. Restoration results for the degraded image in Fig. 4(c) ($f_c = 16$ and $\sigma_n^2 = 100$). a) Second-derivative based result. b) CSF-based result. c) Image restored using adaptive AR(4). d) Image restored using adaptive AR(8). e) Image restored using adaptive AR(12). f) Image restored using the proposed method.

other stationary models for the autocorrelation are used (e.g., AR models trained with a wide image data set).

As expected, adaptive approaches are superior to stationary ones. If the complexity of the model is increased, AR(4) to AR(8), the power spectrum is more accurately estimated. Consequently, a better adaptation to each block is obtained and the performance improves (Tables I–III). This trend is broken if the complexity of the AR model keeps increasing, because overfitting occurs. This effect is particularly noticeable when the energy of the noise is high [Table III and Fig. 12, where the AR(12) model is learning the noise].

On the other hand, the proposed approach constitutes a tradeoff between the overall noise and the preservation of specific details, as can be observed in f) in Figs. 10–12. The artifacts obtained with this kind of penalty operators are different from those obtained using the AR methods. Our approach removes wide band noise because it is predictable (in the local Fourier domain) from the neighbors. This means that this wide band noise is not considered to be a relevant feature. However, unpredictable features in the transform domain, as, for instance, salient textures, are preserved as discussed in Section IV.

The differences in performance are less noticeable in the first experiment (almost a deblurring problem; see Table I and Fig. 10). In this case, the low energy of the noise makes the problem easier and all the adaptive methods perform similarly. However, when the energy of the noise is increased, the differences are more noticeable, as can be seen in the last experiment (almost a denoising problem; see Table III and Fig. 12). The proposed operator works better than the other operators considered in a wider range of degradations, which is consistent with the sensitivity and robustness results described above.

Finally, the proposed method is applied to the restoration of naturally degraded images. The pictures were taken using a SONY DSC P8 digital camera. In this experiment, all the images were on focus, so that the degradation comes from eventual poor illumination conditions (unknown photon noise), and from the artifacts due to the (unknown) processing in the acquisition and storage of the images (e.g., demosaicing, compression, etc.). Therefore, in this case, the proposed method is compared to a standard denoising method [41].[1]

In realistic conditions, the original image is not available in general, so the computation of quantitative distortions is not straightforward. In order to give indicative distortion measures, we took two versions of each image: The first one was taken in good illumination conditions, and the second one was taken re-

[1]We used the Matlab implementation of Lee's algorithm in the function `wiener2.m`.


Fig. 12. Restoration results for the degraded image in Fig. 4(d) ($f_c = 27.2$ and $\sigma_n^2 = 200$). a) Second-derivative based result. b) CSF-based result. c) Image restored using adaptive AR(4). d) Image restored using adaptive AR(8). e) Image restored using adaptive AR(12). f) Image restored using the proposed method.

ducing the numerical aperture or the exposure time so that the integrated irradiance at the CCD was reduced. The histogram of the second image was linearly corrected to reproduce the mean and variance (average luminance and contrast) of the first one. In this way, the second image is a naturally distorted version of the first image. We will refer to the first one as *reference image* and to the second one as *naturally degraded image*.

Figs. 13–15 show representative results of the above distortion procedure in a range of illuminations and the corresponding results obtained using Wiener denosing [41] and the proposed method. Subplots e) and f) in these figures show the difference between the restored and the degraded images. This difference represents the part of the degraded image identified as noise by each method. It is worth noting that while the standard algorithm removes features that may be relevant (letters, edges and structures of the objects), the proposed algorithm preserves these structures focusing in removing the noise.

The above qualitative results are consistent with the distortion measures given in Table IV. As can be seen, both restoration techniques improve the estimated PSNR, and the proposed method provides a better result using this measure. However, it has to be stressed that these measures are just an approximation to the true PSNR because they have been computed using a reference image which is not the (unknown) original.

## VI. Conclusions and Further Work

In this paper, the information about natural images implicitly included in the current model of early human vision (approximate $1/f$ behavior and relations between coefficients in a local frequency transform domain) has been used to define regularization operators for restoration purposes. This class of operators shows several advantages with regard to the adaptive operators considered based on AR models: They are less sensitive to the regularization parameter, more robust to the degradation, and give rise to less noticeable artifacts than the other methods. Besides, in naturally degraded images, they preserve the relevant structure of objects when compared with a standard Wiener denoising.

These results suggest that the proposed method, which does not require statistical *a priori* assumptions on the image, can be used as a successful alternative to operators based on explicit models of the spectrum of the image.

The proposed operator has been used in a $l_2$ norm penalty functional. An interesting extension of this paper could explore the behavior of the proposed operator with more general norms. For instance, using the $l_1$ norm with the proposed operator would improve edge preservation, as already reported using stationary operators such as the first derivative [24].

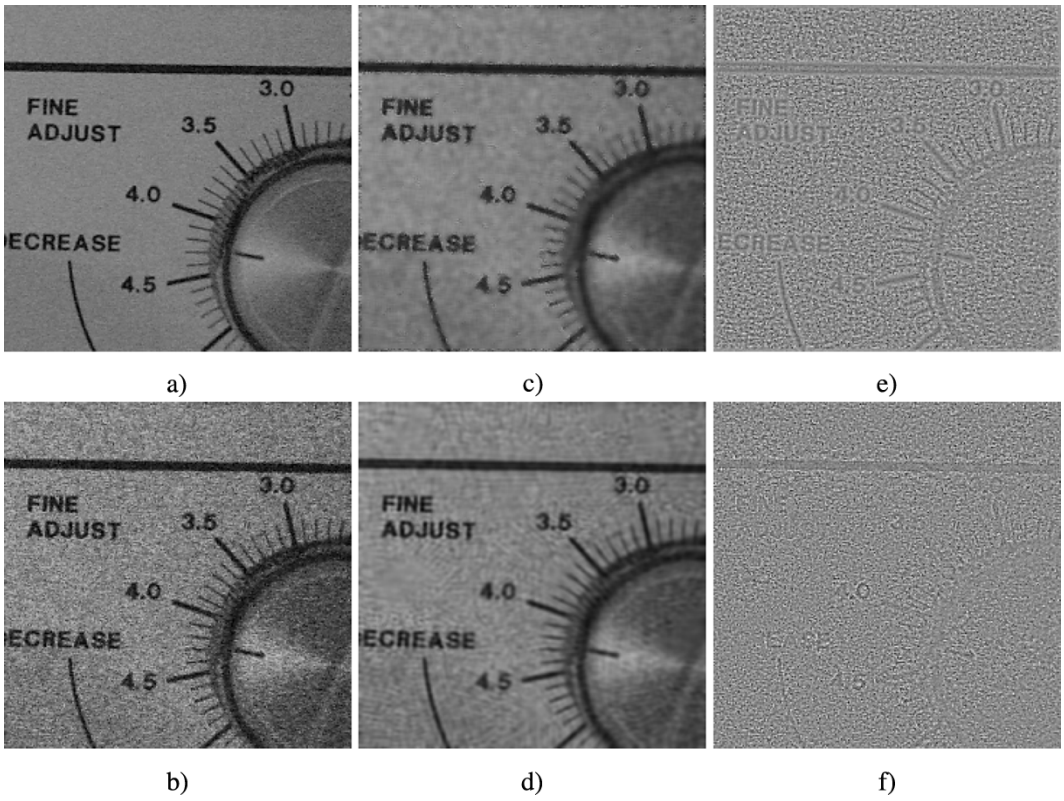


Fig. 13. Restoration results for a naturally degraded image (device). a) Reference image (exposure time 1/6 sec, aperture $\mathrm{N} = 5.2$). b) Naturally degraded image (exposure time 1/25 s, aperture $\mathrm{N} = 5.2$). c) Restored image using Wiener denoising. d) Restored image using the proposed method. e) Information removed by the Wiener method. f) Information removed by the proposed method.

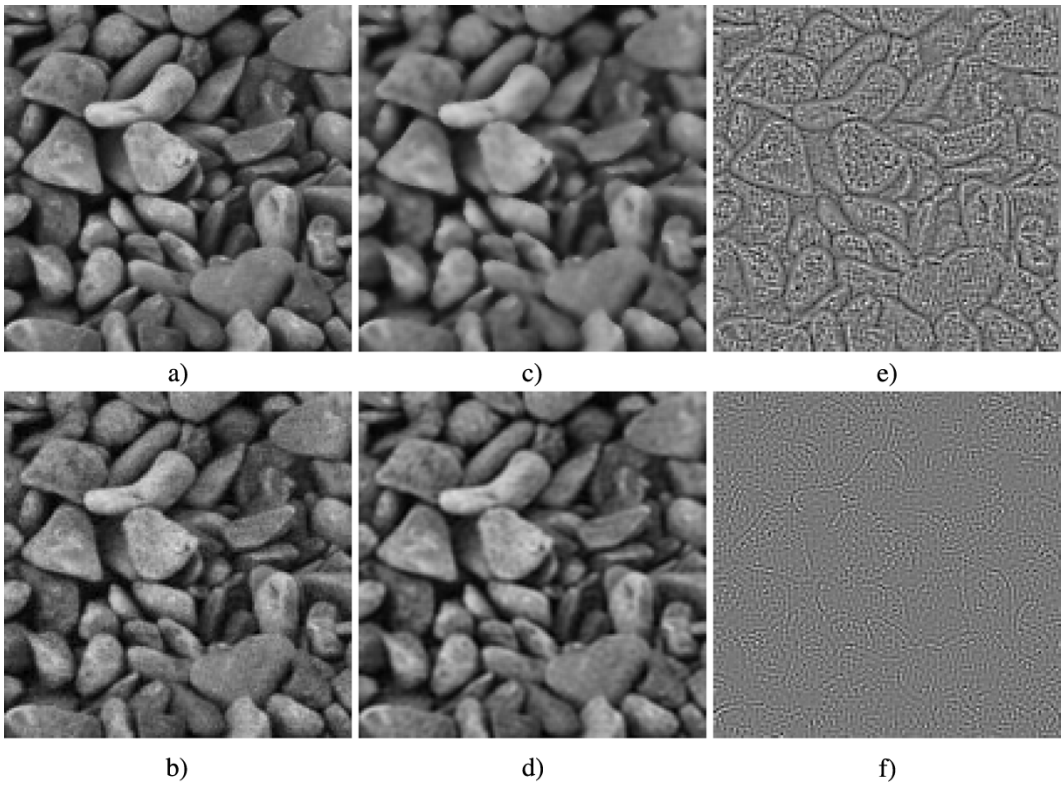


Fig. 14. Restoration results for a naturally degraded image (stones). a) Reference image (exposure time 1/200 sec, aperture $\mathrm{N} = 9$). b) Naturally degraded image (exposure time 1/800 s, aperture $\mathrm{N} = 9$). c) Restored image using Wiener denoising. d) Restored image using the proposed method. e) Information removed by the Wiener method. f) Information removed by the proposed method.

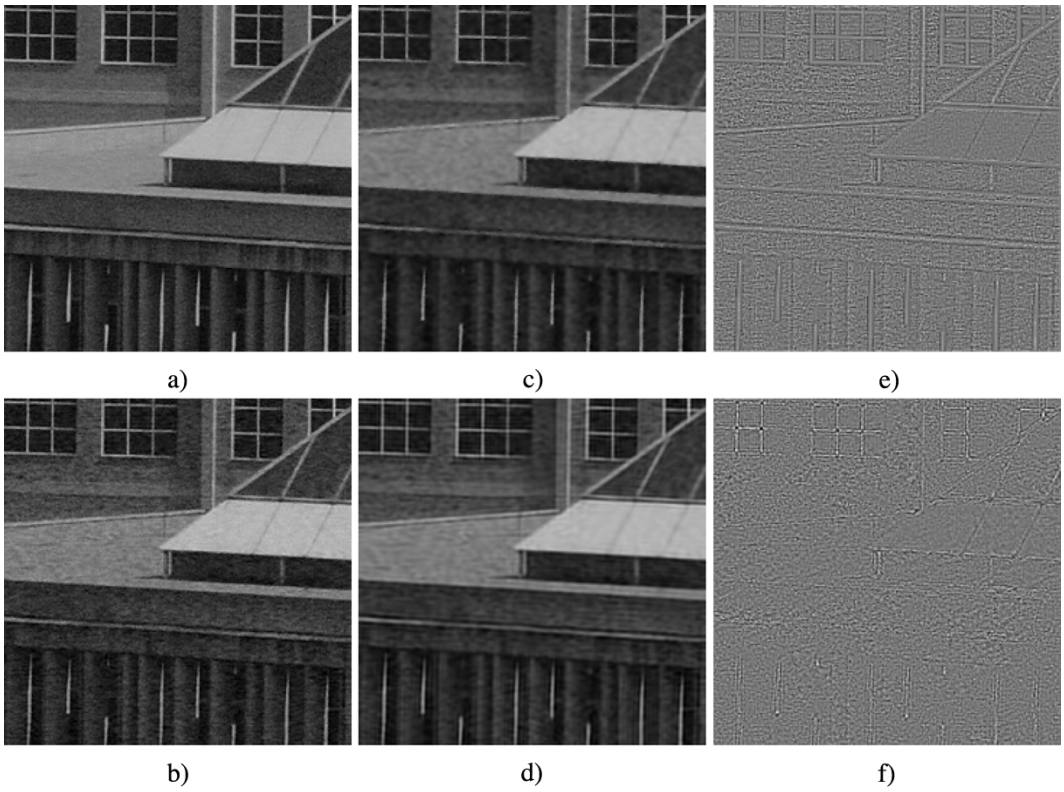


Fig. 15. Restoration results for a naturally degraded image (building). a) Reference image (exposure time 1/1250 sec, aperture $\mathrm{N} = 10$). b) Naturally degraded image (exposure time 1/2000 s, aperture $\mathrm{N} = 10$). c) Restored image using Wiener denoising. d) Restored image using the proposed method. e) Information removed by the Wiener method. f) Information removed by the proposed method.

TABLE IV
ESTIMATED PSNR FOR THE NATURALLY DEGRADED IMAGES

| | Device | Stones | Building |
|---|---|---|---|
| Degraded | 20.80 | 28.14 | 23.39 |
| Wiener | 25.33 | 28.45 | 23.96 |
| **PER** | **25.64** | **29.07** | **24.08** |

## ACKNOWLEDGMENT

The authors would like to thank Prof. P. C. Hansen for fruitful comments on the L-curve method [36], as well as Prof. E. Simoncelli for his criticism on the used transform (even though the block Fourier transform is nice enough to illustrate the reported behavior, the better way to do it is using his Steerable Pyramid [15]). They would also like to thank Dr. G. Camps for useful discussions on AR models, Dr. J. Viru for his insight on natural images features, Prof. J. M. Artigas, Dr. A. B. Watson, Dr. A. Ahumada, and Dr. A. M. Pons for their comments on the psychophysical meaning of the kernels $k(f, f')$ [12], as well as Dr. P. Mulet for his help in using his Total Variation code [42] during the reviewing process.

## REFERENCES

[1] M. R. Banham and A. K. Katsaggelos, "Digital image restoration," *IEEE Signal Process. Mag.*, vol. 14, no. 1, pp. 24–41, Jan. 1997.
[2] H. C. Andrews and B. R. Hunt, *Digital Image Restoration*. Englewood Cliffs, NJ: Prentice-Hall, 1977.
[3] A. K. Katsaggelos, "Iterative image restoration algorithms," *Opt. Eng.*, vol. 28, no. 7, pp. 735–748, 1989.
[4] D. L. Ruderman and W. Bialek, "Statistics of natural images: scaling in the woods," *Phys. Rev. Lett.*, vol. 73, no. 6, pp. 814–817, 1994.
[5] D. C. Knill, D. Field, and D. Kersten, "Human discrimination of fractal images," *J. Opt. Soc. Amer.*, vol. 7, no. 6, pp. 1113–1123, 1990.
[6] R. Molina, J. Núñez, F. J. Cortijo, and J. Mateos, "Image restoration in astronomy," *IEEE Signal Process. Mag.*, no. 3, pp. 11–29, Mar. 2001.
[7] E. P. Simoncelli, "Statistical models for images: compression, restoration and synthesis," in *Proc. 31st Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, 1997, pp. 673–678.
[8] J. Portilla, V. Strela, M. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of gaussians in the wavelet domain," *IEEE Trans. Image Process.*, vol. 12, no. 11, pp. 1338–1351, Nov. 2003.
[9] M. Bertero, T. A. Poggio, and V. Torre, "Ill-posed problems in early vision," *Proc. IEEE*, vol. 76, no. 8, pp. 869–889, Aug. 1988.
[10] J. Gutierrez, J. Malo, and F. Ferri, "Perceptual regularization functionals for natural image restoration," in *Proc. IEEE Int. Conf. Image Processing*, vol. 2, 2003, pp. 989–992.
[11] M. Carandini and D. Heeger, "Summation and division by neurons in visual cortex," *Science*, vol. 264, no. 5163, pp. 1333–1336, 1994.
[12] A. Watson and J. Solomon, "A model of visual contrast gain control and pattern masking," *J. Opt. Soc. Amer. A*, vol. 14, pp. 2379–2391, 1997.
[13] H. B. Barlow, "What is the computational goal of the neocortex?," in *Large Scale Neuronal Theories of the Brain*, C. Koch, Ed. Cambridge, MA: MIT Press, 1994, pp. 1–22.
[14] B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, vol. 381, pp. 607–609, 1996.
[15] O. Schwartz and E. P. Simoncelli, "Natural signal statistics and sensory gain control," *Nature Neurosci.*, vol. 4, no. 8, pp. 819–825, 2001.
[16] J. Malo, R. Navarro, I. Epifanio, F. Ferri, and J. Artigas, "Non-linear invertible representation for joint statistical and perceptual feature decorrelation," *Lecture Notes Comput. Sci.*, vol. 1876, pp. 658–667, 2000.
[17] J. Malo, I. Epifanio, R. Navarro, and E. Simoncelli, "Non-linear image representation for efficient perceptual coding," *IEEE Trans. Image Process.*, vol. 15, no. 1, pp. 68–80, Jan. 2006.
[18] G. Wallace, "The JPEG still picture compression standard," *Commun. ACM*, vol. 34, no. 4, pp. 31–43, 1991.

[19] J. Malo, J. Gutierrez, I. Epifanio, F. Ferri, and J. Artigas, "Perceptual feed-back in multigrid motion estimation using an improved dct quantization," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1411s–1427s, Oct. 2001.
[20] E. Simoncelli and B. Olshausen, "Natural image statistics and neural representation," *Annu. Rev. Neurosci.*, vol. 24, pp. 1193–1216, 2001.
[21] R. Buccigrossi and E. Simoncelli, "Image compression via joint statistical characterization in the wavelet domain," *IEEE Trans. Image Process.*, vol. 8, no. 12, pp. 1688–1701, Dec. 1999.
[22] J. Liu and P. Moulin, "Information-theoretic analysis of interscale and intrascale dependencies between image wavelet coefficients," *IEEE Trans. Image Process.*, vol. 10, no. 11, pp. 1647–1658, Nov. 2001.
[23] Y. Li, "A globally convergent method for $l_p$ problems," *SIAM J. Optim.*, vol. 3, no. 3, pp. 609–629, 1993.
[24] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, pp. 259–268, 1992.
[25] R. Clarke, "Relation between the Karhunen-Loeve transform and cosine transforms," *Proc. Inst. Elect. Eng., F*, vol. 128, no. 6, pp. 359–360, 1981.
[26] P. J. B. Hancock, R. J. Baddeley, and L. S. Smith, "The principal components of natural images," *Network: Comput. Neural Syst.*, vol. 3, no. 1, pp. 61–70, 1992.
[27] A. Gersho and R. Gray, *Vector Quantization and Signal Compression*. Boston, MA: Kluwer Academic, 1992.
[28] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York: Wiley, 2000.
[29] J. Cardoso, "Source separation using higher order moment," in *Proc. Int. Conf. Acoustics Speech Signal Processing*, 1989, pp. 2109–2112.
[30] P. Comon, "Independent component analysis: a new concept?," *Signal Process.*, vol. 36, no. 3, pp. 287–314, 1994.
[31] A. Said and W. A. Pearlman, "An image multiresolution representation for lossless and lossy image compression," *IEEE Trans. Image Process.*, vol. 5, no. 9, pp. 1303–1310, Sep. 1996.
[32] I. Epifanio, J. Gutierrez, and J. Malo, "Linear transform for simultaneous diagonalization of covariance and perceptual metric matrix in image coding," *Pattern Recognit.*, vol. 36, pp. 1799–1811, 2003.
[33] J. Malo, "Normalized image representation for efficient coding," in *Proc. IEEE 37th Asilomar Conf. Signals, Systems, Computers*, vol. 2, 2003, pp. 1408–1412.
[34] J. van Hateren and A. van der Schaaf, "Independent component filters of natural images compared with simple cells in primary visual cortex," *Proc. Roy. Soc. Lond. B*, vol. 265, pp. 359–366, 1998.
[35] F. Campbell and J. Robson, "Application of Fourier analysis to the visibility of gratings," *J. Physiol.*, vol. 197, pp. 551–566, 1968.
[36] P. C. Hansen and D. P. O'Leary, "The use of the L-curve in the regularization of discrete ill-posed problems," *SIAM J. Sci. Comput.*, vol. 14, pp. 1487–1503, 1993.
[37] B. Girod, "What's wrong with mean squared error?," in *Digital Images and Human Vision*, A. B. Watson, Ed. Cambridge, MA: MIT Press, 1993, pp. 207–220.
[38] P. Barten, "Evaluation of subjective image quality with the square root integral method," *J. Opt. Soc. Amer. A*, vol. 7, pp. 2024–2031, 1990.
[39] J. Lim, *Two-Dimensional Signal and Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1990.
[40] B. R. Hunt, "Digital image processing," *Proc. IEEE*, vol. 63, no. 4, pp. 693–708, Apr. 1975.
[41] J. Lee, "Digital image enhancement and noise filtering by use of local statistics," *IEEE Pattern Anal. Mach. Intell.*, vol. PAMI-2, no. 1, pp. 165–168, Jan. 1980.
[42] T. F. Chan, G. H. Golub, and P. Mulet, "A nonlinear primal-dual method for total variation-based image restoration," *SIAM J. Sci. Comput.*, vol. 20, no. 6, pp. 1964–1977, 1999.

**Juan Gutiérrez** received the Licenciado degree in physics (electricity, electronics, and computer science) from the Universitat de València, València, Spain, in 1995.

He is with the Computer Science Department, Universitat de València, where he is an Assistant Professor. Currently, he is with the Visual Statistics Group at the Universitat de València (http://www.uv.es/vista/vistavalencia). His current research interests include regularization theory, models of low-level human vision, and representation and analysis of images.

**Francesc J. Ferri** received the Licenciado degree in physics (electricity, electronics, and computer science) and the Ph.D. degree in pattern recognition from the Universitat de València, València, Spain, in 1987 and 1993, respectively.

He has been with the Computer Science and Electronics Department, Universitat de València, since 1986, first as a Research Fellow and as a Teacher of computer science and pattern recognition since 1988. He has been involved in a number of scientific and technical projects on computer vision and pattern recognition. He joined the Vision, Speech, and Signal Processing Research Group at the University of Surrey, Surrey, U.K., during a sabbatical in 1993, where he was working in feature selection and statistical pattern recognition methodology. He has authored or coauthored about 100 technical papers in international conferences and well-established journals in his fields of interest. He has also helped in refereeing for several of these journals and major conferences. His current research interests include feature selection, nonparametric classification methods, inductive learning, computational geometry, image analysis, and motion understanding.

Dr. Ferri is a member of the ACM and the IAPR (AERFAI in Spain) where he has helped several times in organizing the biennal Spanish Conference on Pattern Recognition and Image Analysis.

**Jesús Malo** received the M.Sc. and Ph.D. degrees in physics from the Universitat de València, València, Spain, in 1995 and 1999, respectively.

In 2000 and 2001, he was a Fulbright Postdoctorate at the Vision Group of the NASA Ames Research Center (A.B. Watson) and at the Laboratory of Computational Vision of the Center for Neural Science, New York University (E. P. Simoncelli). Currently, he is with the Visual Statistics Group at the Universitat de València (http://www.uv.es/vista/vistavalencia). He is interested in models of low-level human vision, their relations with information theory, and their applications to image processing and vision science experimentation.

Dr. Malo was the recipient of the Vistakon European Research Award in 1994. He is member of the Asociación de Mujeres Investigadoras y Tecnólogas (AMIT).