

PERCEPTUAL REGULARIZATION FUNCTIONALS FOR NATURAL IMAGE RESTORATION

Juan Gutiérrez, Jesús Malo[†] and Francesc Ferri

Universitat de València, Dept. d'Informàtica, [†]Dept. d'Òptica
Dr. Moliner 50. Burjassot 46100, SPAIN
juan.gutierrez@uv.es, jesus.malo@uv.es

ABSTRACT

Regularization constraints are necessary in inverse problems such as image restoration, optical flow computation or shape from shading to avoid the singularities in the solution. Conventional regularization techniques are based on some *a priori* knowledge of the solution: usually, the solution is assumed to be smooth according to simple statistical image or motion models.

Using the fact that human visual perception is adapted to the statistics of natural images and sequences, the class of regularization functionals proposed in this work are not based on an image model but on a model of the human visual system. In particular, the current non-linear model of early human visual processing is used to obtain locally adaptive regularization functionals for image restoration without any *a priori* assumption on the image or the noise. The results show that these functionals constitute a valid alternative to those based on the local autocorrelation of the image.

1. INTRODUCTION

Images are produced to record or display useful information about some phenomenon of interest. As 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 modelled by means of a linear operator acting upon the original image plus additive noise:

$$i_d(x) = h(x, x') \otimes i_o(x') + n(x) \quad (1)$$

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

Multiple techniques of different nature have been proposed to solve this problem, but most of them require (and critically depend on) some image model or some *a priori* assumptions on the image or the noise [1, 2, 3]. Usually the solution is assumed to have a particular smoothness in the spatial domain described by its autocorrelation function or

the coefficients of an autoregressive model [3, 4]. In other cases, the required features of the solution are given in a transform domain [5]. In the particular context of regularization techniques, the restored image is the one that minimizes the departure from the observed data and the presence of some *unwanted features* in the solution [6]. The problem again is to define the image model or the *suitable features* in the solution.

In this work we propose a class of penalty functionals 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 [7, 8]. In this way there is no need to make any *a priori* assumption.

The use of advanced perception models to design the regularization functionals relies on the Barlow hypothesis [9]: the biological vision systems have evolved for an optimal processing of natural images. In particular, it has been shown that the early linear processing mechanisms in the visual cortex perform a sort of linear *Independent Component Analysis* [10]. Besides, the post transform non-linearities [7, 8] increase the independization between the transform coefficients [11, 12]. This means that these mechanisms have evolved to capture the most salient features of the natural images, and, in a way, discriminate between useful and negligible information.

According to these facts, the generic statistical properties of natural images (as for instance their smoothness and the relations between the coefficients in a local Fourier domain [11, 5]) are implicitly included in the response model. Therefore, it makes sense to use the perceptual response to design the regularization 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 models in the regularization is not new [1, 3]. However, those approaches used simpler (linear) models and, what is more important, the rationale was very application-oriented (similar to the one which is popular in the coding literature): as the images are addressed to human viewers, the functional should penalize the features which are not subjectively relevant.

In addition to the use of a more advanced model, the main idea of this work is different: here we want to stress the fact that the information about the natural images implicitly included in the human visual response models can be used as a successful alternative to explicit image models in the regularization functional design.

2. REGULARIZATION AND IMAGE MODELS

In the Tikhonov regularization [6] the image estimate, i' , is the one that simultaneously minimizes the departure from the observed data, $|h(x, x') \otimes i'(x') - i_d(x)|^2$, and a penalty functional, $|p(x, x') \otimes i'(x')|^2$, that measures the presence of some *unwanted feature* in the solution:

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

where λ establishes the trade-off between both terms. Assuming that h and p are shift invariant, the solution to eq. 2 in the Fourier domain is [1]:

$$I'(f) = \frac{H^*(f)}{|H(f)|^2 + \frac{\lambda^2}{|P(f)^{-1}|^2}} I_d(f) \quad (3)$$

The meaning of the regularization functional, p , in terms of a simple image model can be inferred from the result using the optimal Wiener filter [1]:

$$I'(f) = \frac{H^*(f)}{|H(f)|^2 + \frac{|N(f)|^2}{|I_o(f)|^2}} I_d(f) \quad (4)$$

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 is not available in general). Eq. 4 suggests that the regularization penalty functional, p , and the parameter, λ , should be related to the spectrum of the signal and the energy of the noise. This is why the conventional functional design is based in a model of the autocorrelation of the image [4, 3]. The smoothness assumption (limited band signal) gives rise to high-pass regularization functionals (the *unwanted feature* is the energy in the high frequency region).

3. PERCEPTUAL REGULARIZATION FUNCTIONALS

The standard model of human low-level image analysis [8] has two basic stages, in which the input image, $i(x)$ is first transformed into a vector $I(f, x)$, in a local frequency domain (using a linear filter bank, T) and finally it is nonlinearly transformed into a response vector, $r(f, x)$

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

The perceptual transform, T , is similar to the class of transforms employed in image coding. 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 is given by a linear term which dominates when the amplitude is small (near-threshold behavior) and a non-linear term in which the energy of each coefficient is normalized by a linear combination of the neighbors (supra-threshold or masking behavior) [8, 7, 12]. For a particular region of a local Fourier transform, the response for each frequency component is:

$$r(I(f)) = \frac{\alpha(f)|I(f)|^2}{\beta(f) + (k(f, f') \otimes |I(f')|^2)} \quad (5)$$

Figure 1 shows the frequency dependence of the parameters in eq. 5. The unit-volume Gaussian kernels, k , relate each coefficient with its neighbors in frequency and orientation [8]. The parameters α and β are chosen to fit the psychophysical response of the human viewer to periodic functions of frequency, f [12]. Note that the overall behavior is mediated by the band-pass filter $\alpha(f)$, which is similar to the classical *Contrast Sensitivity Function* (CSF).

It has been reported that the energy normalization in the non-linear term acts as a sort of divisive DPCM (the energy in each transform coefficient is predicted from the energy of the neighbors). In this way, the independence between the coefficients of the response is increased [12]. This means that the basic relations between the local Fourier coefficients of natural images are described by the Gaussian kernels, $k(f, f')$. Summarizing, the implicit information about the images in this model is *smoothness* (limited band behavior of α) but also a *particular relationship between the transform coefficients* (note the frequency dependent neighborhood defined by the kernels).

Assuming that the responses in eq. 5 capture the most salient independent features of the natural images, and discriminate between useful and negligible information, the perceptual responses for a particular region of the observed data tell us what features should be present in the solution and what features should be removed from it. According to this, we define the perceptual regularization functional to be inversely proportional to the response. Therefore, each restored region will be given by,

$$I'(f) = \frac{H^*(f)}{|H(f)|^2 + \frac{\lambda^2}{|r(I_d(f))|^2}} I_d(f) \quad (6)$$

Using the analogy with eq. 4 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.

As stated in the introduction, previous perceptual ap-

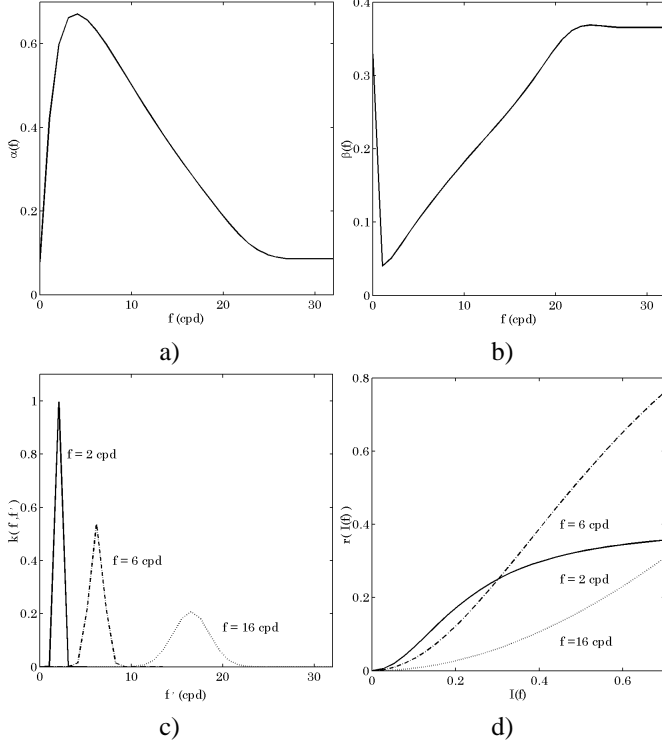


Fig. 1. Parameters of the vision model and non-linear response functions. The values in these figures assume that the Fourier coefficients are expressed in contrast (amplitude over mean luminance). The responses of figure 1.d illustrate the basic sigmoid behavior of eq.5, but they are just particular examples for isolated sinusoids. In general these response curves will depend on the neighbor coefficients.

proaches based in the linear CSF model [1]:

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

can be seen as a particular case of the more general class of functionals presented here.

4. RESULTS AND DISCUSSION

The regularization functionals proposed have been tested in several natural images with different levels of degradation. Some illustrative results are presented using the standard image Barbara (figure 2).

We assume that the 256×256 images have a physical extent of 4×4 degrees, i.e. a sampling frequency of 64 cpd. The original image has been degraded using a linear shift invariant low-pass filter with cutoff frequency of 16 cpd and Gaussian white noise with variance $\sigma_n^2 = 185$. The total distortion is $BSNR \approx 12$ dB (figure 2.b). In the experiments, the regularization parameter, λ , has been obtained using the

standard L-curve method [13] for each block of the local Fourier Transform. To objectively test the performance of the image restoration algorithms, the Improvement in Signal to Noise Ratio (ISNR) will be used.

First, two classical restoration methods are presented for reference purposes. When the second derivative is used as a model for the autocorrelation, a result like the one shown in figure 2.c is obtained. This image shows over-smoothed areas and noise in the band 10-24 cpd. Similar results are obtained when other stationary models for the autocorrelation are used. The restored image using the approach proposed by Hunt (eq.7) is shown in figure 2.d. The use of this very simple perceptual model (with a fixed regularization functional) does not substantially changes the results based on a stationary model of the autocorrelation.

The procedures with a spatially-invariant regularization functional can be improved by using adaptive approaches: in the procedures based on a model of the image it is possible a local estimation of the power spectrum to update the regularization functional. In the proposed approach, considering an input-dependent perceptual model (eq. 6) naturally implies a change in the functional in each region.

It is possible to compare the image-based and the perceptual-based approaches in the best (idealized) way provided the original image is available to design the functionals. On one hand, eq.4 can be exactly computed for each block. And, on the other hand, the response in eq.6 can be computed using I_o instead of I_d . These results (figs. 2.e and 2.f) constitute an upper bound of what can be obtained in each case. Both locally adaptive approaches certainly improve the fixed functional results: e.g., specific details are preserved. However, even if the best possible image model (the *true* autocorrelation) and the *true* energy noise are used in eq. 4, the corresponding result displays annoying noise in the smooth regions which is not present in the perceptual case.

In practice, a model needs to be locally fitted using the distorted image in order to properly apply the first approach, and results are very dependent on the accuracy and robustness of the model. On the other hand, the proposed approach directly uses the degraded image and there is no need of any model. In this example, the result obtained is the one shown in figure 2.g. The restored image constitutes a trade-off between the overall noise and the preservation of specific details. The loss in figure 2.f with regard to the upper bound result in figure 2.g is small given the big differences between i_o and i_d .

5. CONCLUDING REMARKS

In this paper, the information about natural images implicitly included in the current model of early human vision (smoothness and relations between coefficients in a trans-

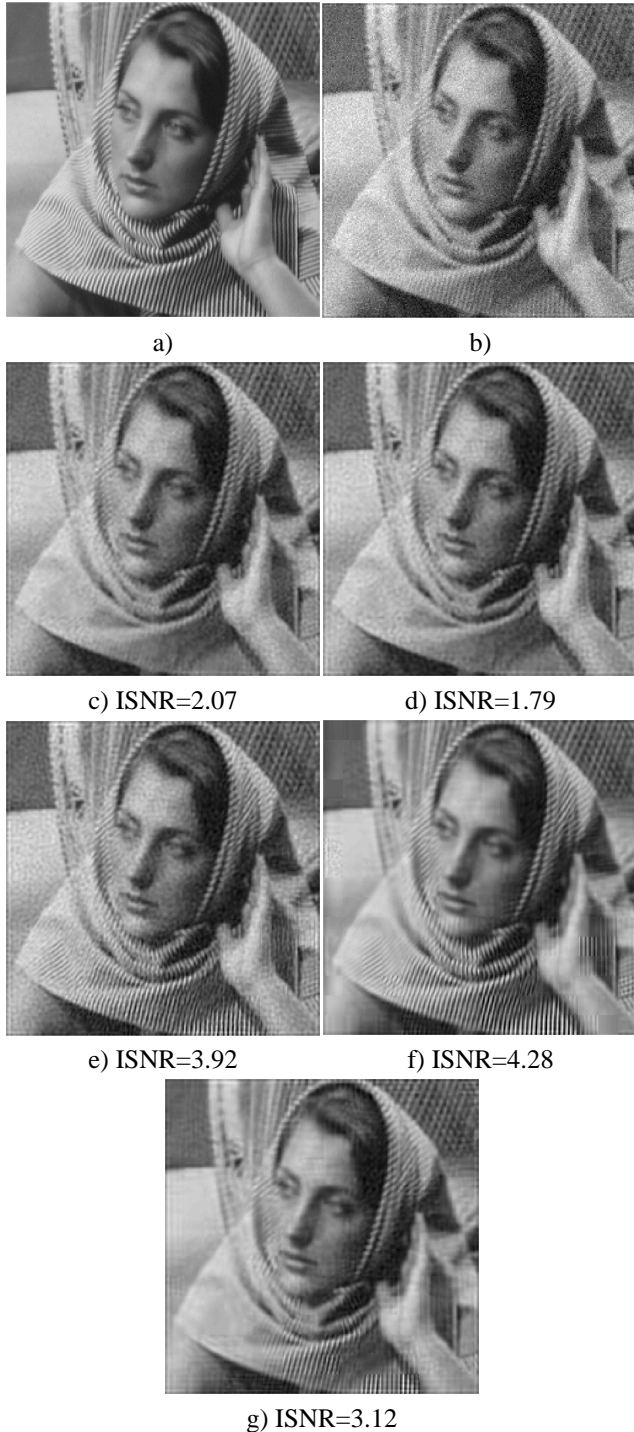


Fig. 2. a) Original image, b) Distorted image, c) Second-derivative based result, d) CSF-based result e) Idealized local Wiener result, f) Idealized local non-linear perceptual result, g) Proposed method.

form domain) has been used to define regularization functionals for restoration purposes. Results suggest that the

proposed method, which does not require *a priori* assumptions on the image, can be used as a successful alternative to explicit image models.

6. REFERENCES

- [1] H. C. Andrews and B. R. Hunt, *Digital Image Restoration*, Prentice Hall, NJ, 1977.
- [2] A. K. Katsaggelos, "Iterative image restoration algorithms," *Opt. Eng.*, vol. 28, no. 7, pp. 735–748, 1989.
- [3] M. R. Banham and A. K. Katsaggelos, "Digital image restoration," *IEEE Sig. Proc. Mag.*, vol. 14, pp. 24–41, 1997.
- [4] R. Molina, J. Núñez, F. J. Cortijo, and J. Mateos, "Image restoration in astronomy," *IEEE Sig. Proc. Mag.*, pp. 11–29, March 2001.
- [5] J. Portilla, V. Strela, M. Wainwright, and E. P. Simoncelli, "Adaptive wiener denoising using a gaussian scale mixture model in the wavelet domain.," in *8th IEEE ICIP*, 2001, pp. 37–40.
- [6] M. Bertero, T. A. Poggio, and V. Torre, "Ill-posed problems in early vision," *Proc. IEEE*, vol. 76, no. 8, pp. 869–889, 1988.
- [7] M. Carandini and D. Heeger, "Summation and division by neurons in visual cortex," *Science*, vol. 264, no. 5163, pp. 1333–6, 1994.
- [8] A. B. Watson and J. A. Solomon, "A model of visual contrast gain control and pattern masking," *JOSA A*, vol. 14, pp. 2379–2391, 1997.
- [9] H. B. Barlow, "What is the computational goal of the neocortex?," in *Large scale neuronal theories of the brain*, C. Koch, Ed., pp. 1–22. MIT Press, 1994.
- [10] 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.
- [11] O. Schwartz and E. P. Simoncelli, "Natural signal statistics and sensory gain control," *Nature Neuroscience*, vol. 4, no. 8, pp. 819–825, 2001.
- [12] J. Malo, R. Navarro, I. Epifanio, F. Ferri, and J.M. Artigas, "Non-linear invertible representation for joint statistical and perceptual feature decorrelation," *Lect. Not. Comp. Sci.*, vol. 1876, pp. 658–667, 2000.
- [13] P. C. Hansen and D. P. O'Leary, "The use of the L-curve in the regularization of discrete ill-posed problems," *SIAM Journal of Sci. Comp.*, vol. 14, pp. 1487–1503, 1993.