BIO-INSPIRED FRAMEWORK FOR AUTOMATIC IMAGE QUALITY ENHANCEMENT

Andrea Ceresi, Francesca Gasparini, Fabrizio Marini and Raimondo Schettini

Department of Informatics, Systems and Communication, University of Milano-Bicocca, Viale Sarca 336, 20126 Milano, Italy

ABSTRACT

We propose a bio-inspired framework for automatic image quality enhancement. Restoration algorithms usually have fixed parameters whose values are not easily settable and depend on image content. In this study, we show that it is possible to correlate no-reference visual quality values to specific parameter settings such that the quality of an image could be effectively enhanced through the restoration algorithm. In this paper, we chose JPEG blockiness distortion as a case study. As for the restoration algorithm, we used either a bilateral filter, or a total variation denoising detexturer. The experimental results on the LIVE database will be reported. These results will demonstrate that a better visual quality is achieved through the optimized parameters over the entire range of compression, with respect to the algorithm default parameters.

Keywords: image quality assessment, no reference methods, image quality enhancement

1. INTRODUCTION

Restoration algorithms usually have fixed parameters whose values are not easily settable and depend on image content. Generally, these parameters are chosen empirically. When reference images are available, full-reference quality metrics (such as MSE or SSIM[1]) could be used to optimize these parameters. Unfortunately, in most practical applications, reference images are not available. Some restoration algorithms exploit analytical methods to set the parameter values.[2–7] These methods not only suffer a high computational complexity, but also are based on measures that do not represent subjective visual quality. Zhu and Milanfar,[8] proposed a measure (not properly a quality metrics) that can be used to automatically set the parameters of image denoising algorithms. These parameters are estimated for each processed image with an iterative procedure.

In this paper we propose a bio-inspired framework for automatic image quality enhancement. We show that it is possible to correlate no-reference visual quality values to specific parameter settings of a restoration algorithm such that the quality of an image could be effectively enhanced. We chose JPEG blockiness distortion as a case study. Deblocking could be seen as a twice problem, as far as one must face:

- a *localization problem*, or how to identify which part of the signal has to be considered as noise with no other hints than the corrupted signal itself;
- a *recovery problem*, or how to sieve the noise from the signal losing as few licit information (edges, textures,...) as possible.

Many methods have been proposed, ranging from spatial to frequency domain processing, and involving plenty of original techniques.[9–19] In our work we have employed the bilateral filter[20] and a detexturer based on a total variation method.[21] Both these methods are general purpose and widely used in a variety of applications. In this paper we want to prove that our framework makes it possible to achieve a better visual quality through optimized parameters, with respect to the algorithm default settings.

Send correspondence to: Francesca Gasparini E-mail: francesca.gasparini@disco.unimib.it, Telephone: 0039 0264487856

Digital Photography VIII, edited by Sebastiano Battiato, Brian G. Rodricks, Nitin Sampat, Francisco H. Imai, Feng Xiao, Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 8299, 82990G · © 2012 SPIE-IS&T · CCC code: 0277-786X/12/$18 · doi: 10.1117/12.910712

2. BIO-INSPIRED FRAMEWORK

The underline idea of the proposed bio-inspired framework develops as follows. Given a generic degraded image (whose original undistorted version is not available) a no-reference metric evaluates its visual quality. This value is correlated to a tuple $\widehat{\beta}_i$ of optimized algorithm parameters determined a priori and gathered in a table (OQP - Optimized Quality Parameters table). This tuple of parameters is passed to the restoration algorithm that processes the degraded input image and produces a visually enhanced output. In Figure 1 the flux diagram of this restoration process is reported. The association between the no-reference quality value and the optimized restoration parameters is found exploiting, on turn, a full-reference image quality metric.

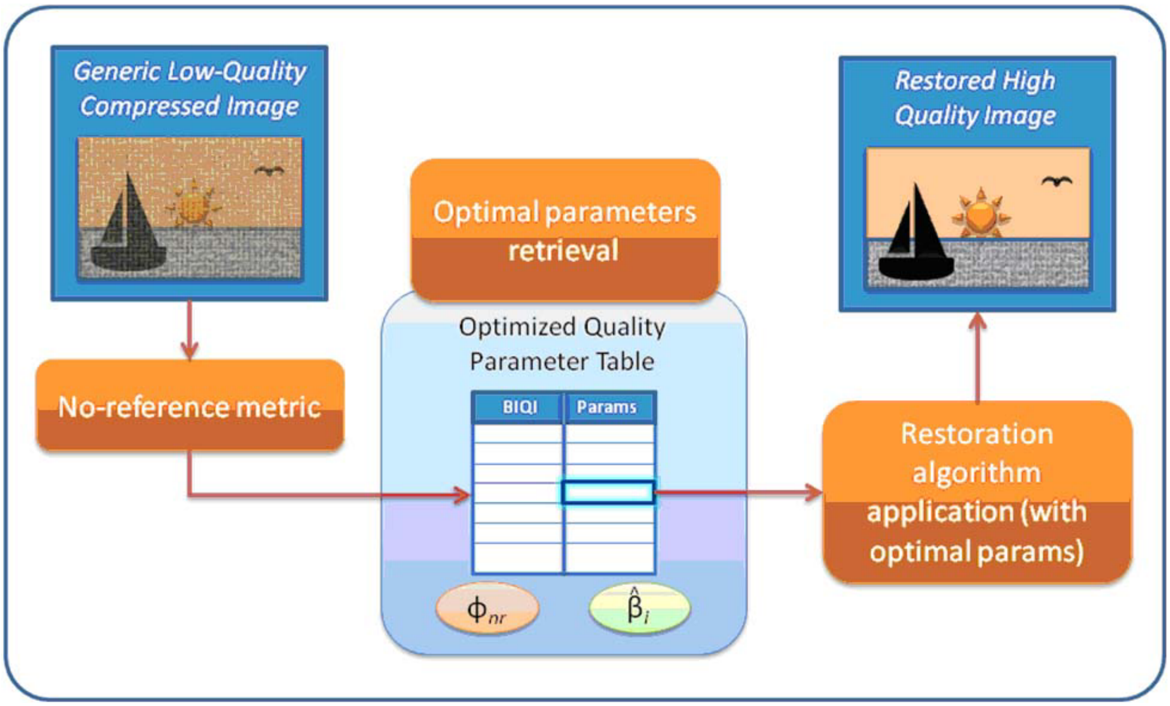


Figure 1. The Restoration Process, through the Optimizes Quality Parameter Table

The OQP table, matching visual quality values and optimized restoration parameters, is the outcome of an off-line processing shown in Figure 2. For a given artifact, this processing involves a restoring algorithm, a full-reference metric and a database of reference images that have been degraded with a wide range of distortion. Each degraded image is processed by the restoration algorithm that, by means of the given parameters β_i, produces an enhanced version of the image. This, along with its original reference, is evaluated by the full-reference quality metric, generating a score φ_{fr}. These restoration and evaluation operations are repeated within a genetic optimization algorithm. The genetic evolution is guided by the full-reference metric as the fitness function: the more the restored image score φ_{fr} is high, the more the parameters β_i that produced it are deemed to fit the visual quality criterion. When the evolution is over, the best so-found parameters $\widehat{\beta}_i$ are associated to a no-reference quality measurement φ_{nr} of the initial degraded image. These data, as previously stated, form an entry of the OQP table.

3. A CASE STUDY: JPEG

In this paper, we chose JPEG blockiness distortion as a case study. Degradation was then obtained by JPEG compression with different Q-factors. As for the restoration algorithm, we used either a fast bilateral filter[20] and the Rudin-Osher-Fatemi detexturer based on total variation denoising,[21] hereafter called TVD-ROF. The full-reference metric we adopted (that is FRI blockiness index from[22]) featured human visual system mechanisms, allowing a more precise estimation of perceived quality. Concerning the no-reference metric, we adopted the Blind Image Quality Index (BIQI) described in[23] to evaluate blockiness.

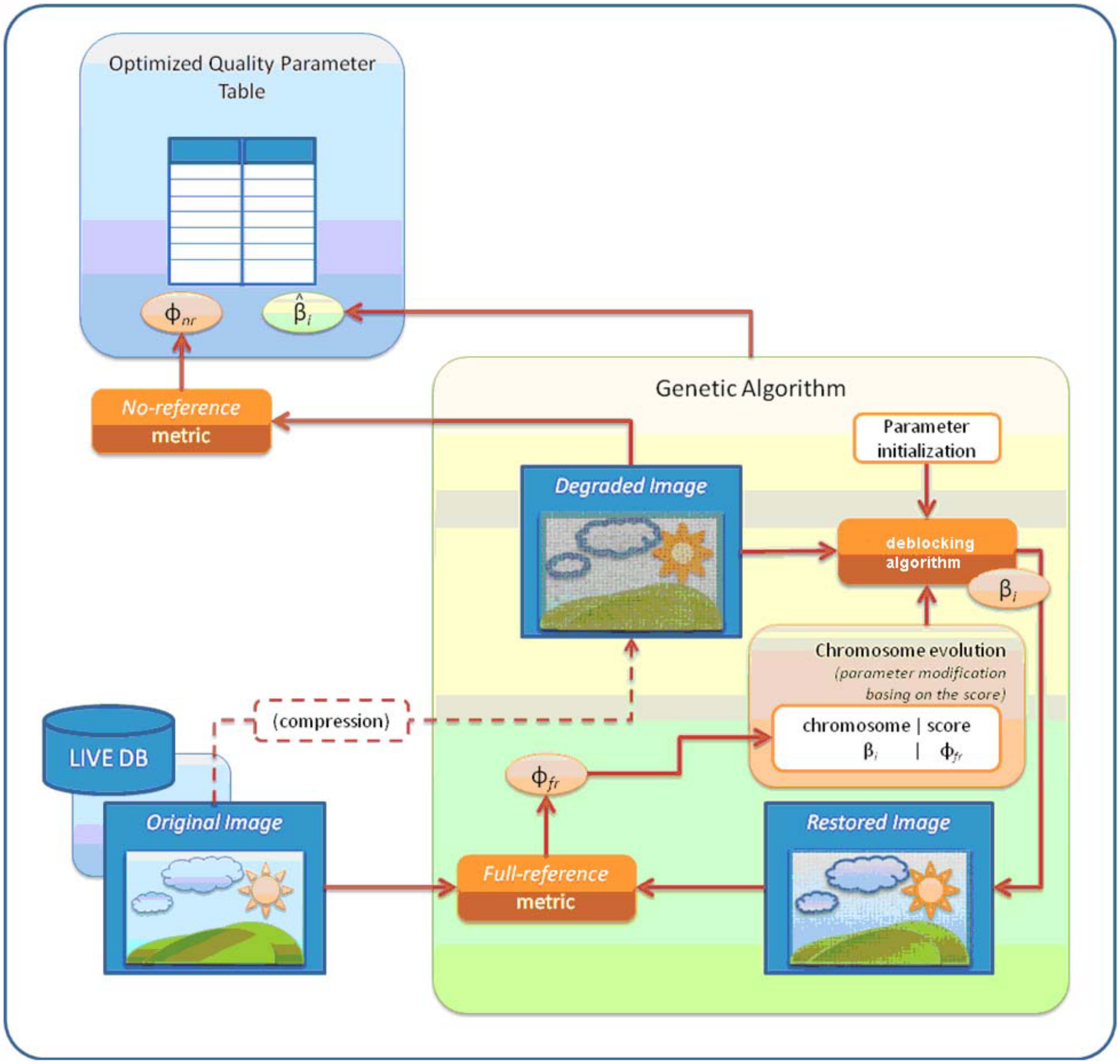


Figure 2. The psychovisually-driven genetic evolution that produces the OQP-Tables.

3.1 Restoration algorithms

3.1.1 Bilateral Filiter: BF

A bilateral filter is a Gaussian-based smoothing filter whose kernel is able to adapt itself to the edges of the scene. In other words, the processing depends on the image content: thanks to its adaptivity, the filtering action will preserve those detected as main edges, whereas other minor details will be smoothed out. Technically, when the bell-shaped function approaches the convolution, it can sense the presence of an edge and automatically adapt its own shape reducing the trans-border part to a zero plateau. The Gaussian envelope results to be split into two parts in order to longitudinally fit the contour, having its smoothing potential on just one side. This technique allows to attenuate small variations while preserving great steps: other than a Gaussian function for the weighted average of the neighborhood, it also relies on a Gaussian function to establish how similar are the neighborhood and the center samples, giving larger weight to the most similar. Formally the kernel follows:

$$BF\left(i_p\right) = \frac{1}{w_p} \sum G_{\sigma_s}\left(\|p-q\|\right) G_{\sigma_r}\left(i_p - i_q\right) i_q \qquad (1)$$

where i_p is the intensity value of the pixel at position $p = (x, y)$ of image I, or the center, S is the set called neighborhood of all the pixels q surrounding the center p, G_σ are the 2-D Gaussian functions and $\frac{1}{W_p}$ is a normalization factor to preserve the local mean. As one can easily see, the filter has two parameters: σ_s governs the spatial extent of the kernel, and σ_r is what allows to define the minimum amplitude of an edge. This way the kernel shape depends on the image content, and no averaging is performed across the edges, revealing great performances. Despite this large potential, though, the bilateral filter is generally difficult to adequately configure. Every image has its own preferable sigma settings, depending on the features of the details we want to preserve and on the noise we want to sieve; but these settings are possibly very hard to find. Durand[24] suggests a shortcut letting σ_r be a sort of difference between the dynamic range extrema, divided by a factor of 10; as for σ_s, it is said to be somehow proportional to the dimensions of the image, so that the minimum between height and width is divided by a factor of 16, or the 2% of the image diagonal.

3.1.2 Total Variation Denoising: TVD

The TVD-ROF algorithm is funded on a general solution to denoising, called Total Variation Denoising (TVD), pioneered by L. I. Rudin et al.[25] The basic idea is that a signal with a high level of details, some of which could eventually be noise, has to show a high level of total variation; in other words, the integral of its absolute gradient is expected to be high. The algorithm reduces denoising to a minimization problem. The total variation of a generic signal f is defined in the mono-dimensional domain as:

$$V(f) = \sum_n |f_{n+1} - f_n| \tag{2}$$

The aim of the total variation algorithm is then to minimize the difference between the input signal, say s, and its approximation $\widehat{s}$, such that the total variation of the latter is lower than the formers, but the two signals shapes are still close to each other. Closeness can be calculated as the sum of square errors:

$$E(s, \widehat{s}) = \frac{1}{2} \sum_n (s - \widehat{s})^2 \tag{3}$$

and the whole problem can be formulated as:

$$min_{\widehat{s}} [E(s, \widehat{s}) + \lambda V(\widehat{s})] \tag{4}$$

The extension to the 2-D domain is trivial. It is easy to see that the choice of λ is key to a proper denoising. With $\lambda \to 0$ denoising tends to no action, and the result is the same as the input signal; when $\lambda \to \infty$, the total variation term plays a very strong role in the linear combination, and forces the result to have a smaller total variation. These dynamics give the idea of how important an optimization can be in this context, choosing λ so that just the right amount of details is eliminated. With Meyers work,[21] the application of the TVD-ROF algorithm has been extended to the purpose of detexturing. Basically, the signal gets split into a structural part (eventually containing all of the main scenic characteristics) and a texture part, revealing patterns and other details. For this reason, it was our idea to use such a method to recover blocky images, letting the algorithm approximate the blocky signal to the texture part, thus retrieving a clean but still detailed scene (structure part). The implementation we used was essentially based on multiple reiterations of the TVD-ROF. Thus, the optimization revolved around two parameters: λ and N that is the number of iterations. Thanks to a grid sampling test, we were able to reveal the operative domain of the TVD-ROF detexturer, and found the values $\lambda = \frac{1}{8}$ and $N = 100$, as indicated by the authors, pretty worth to be considered the standard setting.

3.2 Quality metrics

3.2.1 Full Reference Metric: FRI

As to our framework, we took inspiration from the full-reference metric by Ginesu et al.[22] The metric replicates some psychovisual conditions by means of a frequency-domain Contrast Sensitivity Function and the implementations of some well-modeled effects such as contrast and luminance masking. The perceptibility of blockiness is analyzed through a variance test on every block border region, and it is supported by a second test revealing which part of the extracted orthogonal information is due to a coincidental square texture. A total of three indexes is computed, one for blockiness, one for edge errors and one for the visual impairments, further condensed in a single metric value, formally the FRI, Full Reference Index. In our case, this metric has been partially cut for time save issues, and reduced to the blockiness index only. The lower is the metric value the higher the image quality.

3.2.2 No reference Metric: BIQI

As no-reference metric to evaluate blockiness, we took inspiration from the Blind Image Quality Index (BIQI).[23] The algorithm relies on a Support Vector Machine (SVM) to classify the image into one of five distortion categories (JPEG, JPEG2000, white noise, blur and fast fading), producing a probability match for each one. These information relate to the amount of each distortion present in the image. Then the algorithm computes an overall quality index of the image thanks to Support Vector Regression. Through the use of a series of $\nu - SVMs$, each one trained for a specific distortion category, the processing of the previously computed feature vector leads to a quality quantification. Coupled with the above distortion classification, this quantification yields a BIQI quality score. We just used the JPEG quality assessment submodule of the metric instead of the whole system. The lower is the metric value the higher the image quality.

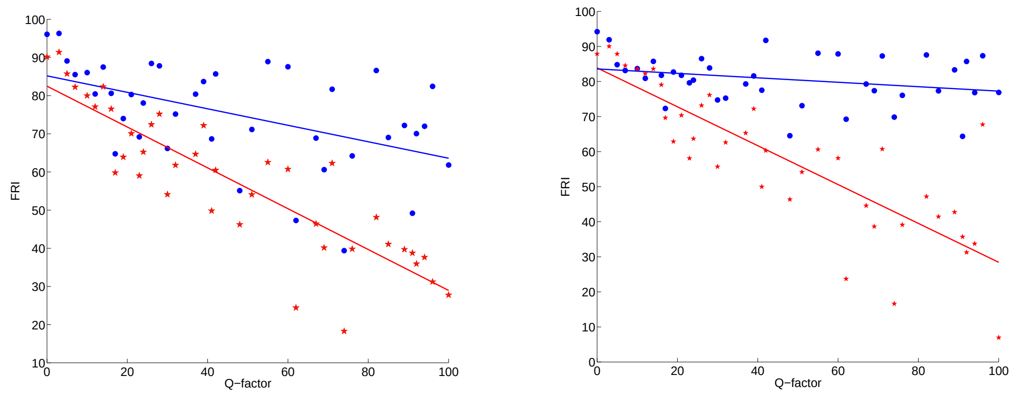
4. EXPERIMENTAL RESULTS

As for the database, we relied on the LIVE Image Quality Assessment Database release 2.[26] We chose, on the basis of image content, three training-sets consisting of three images each. The complementary of every training-set formed a test-set. We run genetic optimization sessions over each training-set and merged the final results into a single OQP table using the best quality gain criterion. We treated the test-sets with both the restoration algorithms using either the default and the genetically-optimized parameters fetched from the final OQP table. We evaluated the results through the full-reference metric FRI to have a more accurate quality gain estimation. The outcomes show that a better visual quality is achieved through the optimized parameters over the entire range of compression. This is true either for the Bilateral Filter and for the TVD-ROF detexturer, as shown on Figure 3, where the FRI values obtained on test set images, processed respectively by the BF (a) and the TVD-ROF (b), are reported, in case of parameters optimized by our framework (stars) and standard parameters (diamonds) and with respect to the Q-factors.

As a further experiment, we randomly picked some degraded images from the Internet. Each image was then evaluated by the BIQI no-reference quality metric and applied the restoring algorithm using the BIQI-related entry of a genetic-born OQP-table. Although the visual outcome counts as the capital proof of a successful processing, we also quantify the quality gain in terms of the same BIQI metric. In Figure 4 top row is shown an example of compressed image, with BIQI score equal to 65.1. This image was processed with the Bilateral Filter and the parameters obtained by our OQP-table. Perceptually speaking, we obtained a remarkable result in deblocking and denoising the image, as shown in Figure 4 bottom. This visually improvement is also confirmed buy the BIQI score that for the so processed image has become 60.5.

5. CONCLUSIONS

In this paper we have shown how parameters of restoring algorithms can be set with an off-line module based on a genetic optimization, so that they produce a better visual quality when compared with results obtained with default parameters. In particular we have proven the feasibility of our framework in the case of JPEG distortion, using as restoration algorithms two general purpose methods, widely used in a variety of applications. In the future we want to investigate the relation between the semantic content of an image and its frequency content, hypothesizing that semantically-driven approaches could produce better restoration results.



(a) BF (b) TVD-ROF

Figure 3. FRI scores for the test set images, processed by the BF (a) and TVD-ROF (b) with the optimized parameters (stars) and the standard parameters (diamonds).

REFERENCES

[1] Z. Wang, A. C. Bovik, H. R. Sheikn, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing* **13**(4), pp. 600–612, 2004.

[2] G. H. Golub, M. Heath, and G. Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter," *Technometrics* **21**(2), pp. 215–223, 1979.

[3] D. A. Girard, "The fast monte-carlo cross-validation and c_l procedures: comments, new results and application to image recovery problems," *Computational statistics* **10**, pp. 205–231, 1995.

[4] P. C. Hansen, "Analysis of discrete ill-posed problems by means of the l-curve," *SIAM Review* **34**(4), pp. 561–580, 1992.

[5] P. C. Hansen and D. P. OLeary, "The use of the l-curve in the regularization of discrete ill-posed problems," *SIAM Journal on Scientific Computing* **14**(6), pp. 1487–1503, 1993.

[6] C. Stein, "Estimation of the mean of a multivariate normal distribution," *Annals of Statistics* **9**, pp. 1135–1151, 1981.

[7] S. Ramani, T. Blu, and M. Unser, "Monte-carlo sure: A black-box optimization of regularization parameters for general denoising algorithms," *IEEE Transactions on Image Processing* **17**(9), pp. 1540–1554, 2008.

[8] X. Zhui and P. Milanfar, "Automatic parameter selection for denoising algorithms using a no-reference measure of image content," *IEEE Transactions on Image Processing* **19**(12), pp. 3116–3132, 2010.

[9] D. Pearson and M. Whybray, "Transform coding of images using interleaved blocks," *IEE Proceedings Radar and Signal Processing* **131**, pp. 466–472, 1984.

[10] P. M. Farrelle, *Recursive Block Coding for Image Data Compression*. Springer-Verlag New York, Inc.

[11] B. Hinman, J. Bernstein, and D. Staelin, "Short-space fourier transform image processing," *IEEE Acoustics, Speech and Signal Processing* **1**, pp. 481–484, 1984.

[12] B. Hinman and D. H. Staelin, "The lot: Transform coding without blocking effects," *IEEE Transaction Acoustics Speech and Signal Processing* **37**, pp. 553–559, 1989.

[13] W. Gao, C. Mermer, and Y. Kim, "A de-blocking algorithm and a blockiness metric for highly compressed images," *IEEE Transaction on Circuits and Systems for Video Technology* **12**(12), pp. 1150–1159, 2002.

[14] A. Averbuch, A. Z.and Schclar and D. L. Donoho, "Deblocking of block-transform compressed images using weighted sums of symmetrically aligned pixels," *IEEE Transaction on Image Processing* **14**(2), pp. 200–212, 2005.

[15] E. V. Tolstaya, M. N. Rychagov, S. H. Kim, and D. C. Choi, "Removal of blocking and ringing artifacts in jpeg-coded images," in *Proc. SPIE*, **7537**, pp. 75370O–1–75370O–12, SPIE, 2010.

[16] X. Gan, A. W. Liew, and H. Yan, "Blocking artifact reduction in compressed images based on edge-adaptive quadrangle meshes," *Journal of Visual Communication and Image Representation* **14**(4), pp. 492–507, 2003.

[17] T. C. Hsung, D. P. Lun, and W. Siu, "A deblocking technique for block-transform compressed images using wavelet transform modulus maxima," *IEEE Transactions on Image Processing* **7**(10), pp. 1488–1496, 1998.

[18] N. C. Kim, I. H. Jang, D. H. Kim, and W. H. Hong, "Reduction of blocking artifact in block-coded images using wavelet transform," *IEEE Transaction on Circuits and Systems for Video Technology* **8**(3), pp. 253–257, 1998.

[19] A. Foi, V. Katkovnik, and K. Egiazarian, "Pointwise shape-adaptive dct for high-quality denoising and deblocking of grayscale and color images," *IEEE Transactions on Image Processing* **16**(5), pp. 1395–1411, 2007.

[20] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," *Proc. IEEE Int. Conf. Computer Vision* , pp. 839–846, 1998.

[21] Y. Meyer, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations: The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures.* American Mathematical Society, Boston, MA, USA, 2001.

[22] G. Ginesu, F. Massidda, and D. Giusto, "A multi-factors approach for image quality assessment based on a human visual system model," *Signal Processing: Image Communication* **20**(4), pp. 316–333, 2006.

[23] A. K. Moorthy and A. C. Bovik, "A two-step framework for constructing blind image quality indices," *IEEE Signal Processing Letters* **17**(5), pp. 513–516, 2010.

[24] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," *ACM Transactions on Graphics (TOG)* , pp. 257–266, 2002.

[25] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena* **60**(1-4), pp. 259–268, 1992.

[26] H. Sheik, Z. Wang, L. Cormakc, and A. Bovik, *LIVE Image Quality Assessment Database Release 2.* http://live.ece.utexas.edu/research/quality.

Figure 4. Top: original image, heavily affected by blockiness, with a zoomed region; Bottom: image restored applying a bilateral filter with optimal parameters and the corresponding zoomed region.