OPTICAL REVIEW Vol. 20, No. 6 (2013) 491–495

Image Restoration with Dual-Prior Constraint Models Based on Split Bregman

Li Ao, Li Yibing*, Yang Xiaodong, and Liu Yue

Institute of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China

(Received September 2, 2013; Accepted September 29, 2013)

In order to utilizing the local and non-local information in the image, we proposed a novel sparse scheme for image restoration in this paper. The new scheme includes two important contributions. The first one is that we extended the image prior model in conventional total variation to the dual-prior models for combining the local smoothness and non-local sparsity under regularization framework. The second one is we developed a modified iterative Split Bregman majorization method to solve the objective function with dual-prior models. The experimental results show that the proposed scheme achieved the state-of-the-art performance compared to the current restoration algorithms.

Keywords: image restoration, non-local sparsity, Split Bregman, total variation, regularization

1. Introduction

As is known to all, image restoration is a significant field of image processing, which wants to reconstruct the original signal from its degradation observation. Generally, the observation can be written as the following form:

$$y = Hx + n \tag{1}$$

where y is the degradation observation, x and n is the original signal and additive white noise, and the matrix H is usually the degradation matrix. With the model in Eq. (1), the restoration is implemented by minimizing the following objective function:

$$\min_x\{\|y - Hx\|_2^2 + \lambda prior(x)\}. \tag{2}$$

Here, the $prior(x)$ represents the prior knowledge of original signal x and λ is called regularization parameter. By the Eq. (2), we can get that the prior model of signal plays a great role in image restoration. It means that the different prior model decided the different restoration performance. In the past decades, for this inverse problem, there are many creative prior expressions to make the performance more and more perfect. Its development looks like the biological evolution, initially it adopted the energy form $\|x\|_2$, and then the smoothness form $\|Lx\|_2^2$ (L is a smooth operator), next robust statistical distribution form for $p(Lx)$ and so on.[1] Although the results are improved but we still did not find a reasonable form until the total variation (TV) produced.[2–4] The TV method is though to be most effective among the classic regularization models, because the natural image satisfies the smoothness constrain commonly. It means that the performance is guaranteed by the piecewise constant assumption, but which is only satisfied in the local area of image practically. And then, for improving the prior model, some methods based sparsity regularization emerged. These methods understand that the original signal x is consist of linear combination of base function. Meanwhile, the representation coefficients satisfy the sparsity property under the base framework. So we can explore some robust prior model for the coefficient, which is widely used to solve the inverse problem.[5–7] In this paper, what we want is to combine the TV and sparsity prior model. On one hand, for utilizing their advantages respectively, we take the TV prior model as the local smoothness constrain item and the sparsity prior model as the non-local sparsity constrain item. On the other hand, for solving the novel objective function, we develop a modified iterative Split Bregman majorization.

The paper is organized as follows. We began in Sect. 2 with a description of the TV and sparsity prior model, and how to establish the novel objective function with them. In Sect. 3, we will give a modified iterative Split Bregman majorization and its application to the objective function in Sect. 2. Section 4 shows the simulation results compared to the current algorithms. We summarize and conclude in Sect. 5.

2. Dual-Prior Constraint Models

As to the minimization problem in Eq. (2), many publications adopted the TV regularization which is widely used in the image inverse problem to solve it in recent years.[8,9] In the TV method, the prior model is defined as follows:

$$prior(x) = TV(x) = \sqrt{(\nabla_{i,j}^h x)^2 + (\nabla_{i,j}^v x)^2}, \tag{3}$$

where the $\nabla_{i,j}^*$ is the first-order partial differential operator of the patch whose center is $x(i,j)$, h, and v represent the horizontal direction and vertical direction respectively. That is, the TV regularization can be seen as

$$\min_x\{\|y - Hx\|_2^2 + \lambda TV(x)\}. \tag{4}$$

And it often shows good performance with the local smoothness assumption. Yet, Bioucas-Dias talked in Ref. 3 that though both smooth and sharp edge have the similar $TV(x)$, but it did not mean the TV regularization favor the sharp edge relatively to the smooth one. In addition, it only constrained the local information, but ignored the non-local geometrical similarity.

*E-mail address: liyibing0920@sina.cn

Unlike the classic regularization whose prior model is established for the original signal itself, the sparse regularization established the prior model for its coding with certain dictionary.[10] Take the Laplace distribution as the prior model, it is usually written by the Bayesian framework as follows:

$$\min_{\alpha}\{\|y - HD\alpha\|_2^2 + \lambda\|\alpha\|_1\}, \quad x = D\alpha. \tag{5}$$

Here, α is the coefficient of x and D is the base function set. The optimization in Eq. (5) is also called sparse coding, which is effective in many signal processing applications.[11,12] The advantage of prior model in Eq. (5) is that it can constrain the sparsity. Specially, if D is a non-singular matrix, the formularization in Eq. (5) can be rewritten as:

$$\min_{x}\{\|y - Hx\|_2^2 + \lambda\|T(x)\|_1\}. \tag{6}$$

Here, $T(x)$ can be seen as the invertible transformation for x.

In order to develop the advantages both of the TV and sparse model, we try to combine the two above prior models, and establish a dual-prior optimization problem. Motivated by the non-local mean method, we want to utilize some non-local geometrical similarity in image. First, we divided the image into many patches whose size is 8×8. Then, suppose $x_{i,j}$ is a patch of them and (i,j) can be decided by the center pixel of each patch. $X_{i,j}$ is the matrix that includes the patches similar to $x_{i,j}$. We construct the following novel objective function:

$$\begin{aligned} &\min_{x}\{\|y - Hx\|_2^2 + \lambda_1 TV(x) + \lambda_2\|T_{\text{sparse}}(X_{i,j})\|_1\} \\ &X_{i,j} = [column_1, column_2, \cdots] \end{aligned} \tag{7}$$

where $X_{i,j}$ can be seen a set of patches similar to $x_{i,j}$. We suppose the $x_{i,j}^1$, $x_{i,j}^2$, $x_{i,j}^3$, ... are patches similar to $x_{i,j}$, and we construct the $X_{i,j}$ as follows:

$$\begin{aligned} x_{i,j} &\in R^{8\times 8} \to column_1 \in R^{64\times 1} \\ x_{i,j}^1 &\in R^{8\times 8} \to column_2 \in R^{64\times 1} \\ x_{i,j}^2 &\in R^{8\times 8} \to column_3 \in R^{64\times 1} \\ &\vdots \end{aligned}$$

And the $\|\cdot\|_1$ is a sparsity measure. Equation (7) realizes our idea to establish an objective function with dual-prior model, which can constrain not only the local smooth but also the non-local sparsity among the similar patches. As to the $T_{\text{sparse}}(\cdot)$, we can select any 2D transformation, such as wavelet, contourlet and so on.

3. A Scheme for Dual-Prior Constraint Models Optimization

In this section, a new scheme for solving the optimization problem in Eq. (7) will be introduced. Recently, a so-called Split-Bregman method proposed by Goldstein and Osher[13] shows its advantage for solving the l_1 minimization.[14] Based on the Split-Bregman, we develop a modified iterative Split-Bregman algorithm to solve the dual-prior constraint model proposed in Sect. 2.

With the splitting technique, the optimization can be changed to be the following constraint formulation:

Table 1. Synthetic algorithm.

Algorithm I
1. Initialization: $k = 0$, $p^0 = q^0 = 0$, $\hat{\omega}^0 = \hat{s}^0 = 0$, $\lambda_1 = \lambda_2 = 0.2$, $\eta_1 = \eta_2 = 0.1$;
2. For $k = 0$ to MaxIterNum do
3. $\hat{x}^{k+1} = \arg\min\left\{\begin{matrix} \|y - Hx\|_2^2 + \eta_1\|x - p^k - \hat{\omega}^k\|_2^2 + \\ \eta_2\|x - q^k - \hat{s}^k\|_2^2 \end{matrix}\right\}$; ①
4. $\tilde{\omega}^k = \hat{x}^{k+1} - p^k$, $\mu = 2\lambda_1/\eta_1$;
5. $\hat{\omega}^{k+1} = \arg\min\{\mu\|\omega - \tilde{\omega}^k\|_2^2 + TV(\omega)\}$; ②
6. $\tilde{s}^k = \hat{x}^{k+1} - q^k$, $\gamma = 2\lambda_2/\eta_2$;
7. $\hat{s}^{k+1} = \arg\min\{\gamma\|s - \tilde{s}^k\|_2^2 + \|T_{\text{sparse}}(S)\|\}$; ③
8. $p^{k+1} = p^k - (\hat{x}^{k+1} - \hat{\omega}^{k+1})$;
9. $q^{k+1} = q^k - (\hat{x}^{k+1} - \hat{s}^{k+1})$;
End

Table 2. Algorithm for optimization ③.

Algorithm II
1. Initialization: $\tilde{s}^k$, $\gamma = 2\lambda_2/\eta_2$
2. Search the similar patches of $\tilde{s}^k$ in the image. The measurement is the Euclidean distance between the two patches: $S = \{s \mid \|s - \tilde{s}^k\|_2 \le \tau\}$ (τ is the threshold to control the precision);
3. Implement the 2D Sparse Transformation to the S: $C = T_{\text{sparse}}^{2D}(S)$
4. Shrinkage the coefficients: $C^{\text{new}} = shrink(C)$;
5. Implement the inverse transformation to the C^{new}: $S^{\text{new}} = T^{-1}(C^{\text{new}})$;
6. Output: the first column of S^{new} as the $\hat{s}^{k+1}$;

$$\begin{aligned} &\min_{x,\omega,s}\{\|y - Hx\|_2^2 + \lambda_1 TV(\omega) + \lambda_2\|T_{\text{sparse}}(S)\|_1\} \\ &s.t.\omega = x \text{ and } S = X \end{aligned} \tag{8}$$

Next, by the Bregman, we can change Eq. (8) into the following unconstraint optimization problem.

$$\begin{aligned} &(\hat{x}^{k+1}, \hat{\omega}^{k+1}, \hat{s}^{k+1}) \\ &= \arg\min\left\{\begin{matrix} \|y - Hx\|_2^2 + \lambda_1 TV(\omega) + \lambda_2\|T_{\text{sparse}}(S)\|_1 \\ + \eta_1\|x - (\omega + p^k)\|_2^2 + \eta_2\|X - (S + Q^k)\|_2^2 \end{matrix}\right\} \end{aligned} \tag{9}$$

Then, p^k and Q^k are updated by:

$$\begin{aligned} p^{k+1} &= p^k - (\hat{x}^{k+1} - \hat{\omega}^{k+1}) \\ Q^{k+1} &= Q^k - (\hat{X}^{k+1} - \hat{S}^{k+1}) \end{aligned} \tag{10}$$

With the Eqs. (9) and (10), a modified Bregman method is proposed to divide the optimization into three minimization problems and get the $\hat{x}^{k+1}$, $\hat{\omega}^{k+1}$, and $\hat{S}^{k+1}$ separately. The pseudo code for algorithm is shown in Algorithm I. Specially, we set $\hat{s}^k$ and $\hat{q}^k$ are the first column of the $\hat{S}^k$ and $\hat{Q}^k$ respectively.

In Algorithm I, we should notice that the optimization ① is a quadratic convex programming, so the solution can be expressed as:

$$\begin{aligned} \hat{x}^{k+1} &= (H^T H + \mu I)^{-1} A, \\ A &= H^T y + \eta_1(p^k + \hat{\omega}^k) + \eta_2(q^k + \hat{s}^k) \end{aligned} \tag{11}$$

Here, $\mu = \eta_1 + \eta_2$ and I is the identity matrix.

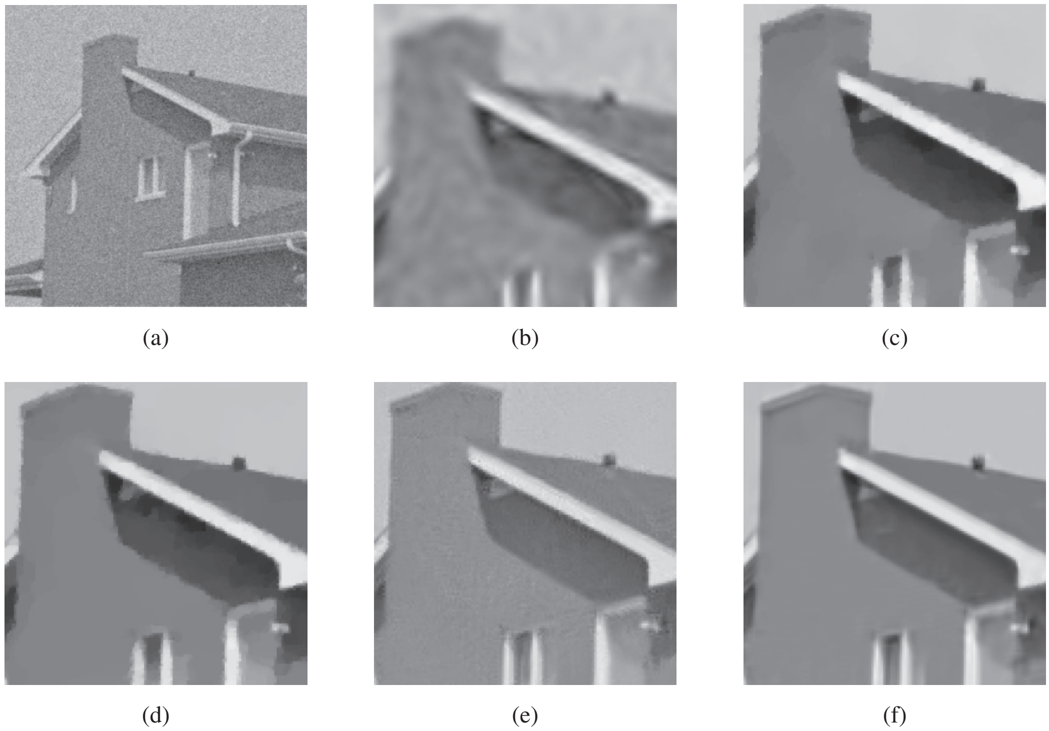

Fig. 1. The results comparison of House with the noise deviation 20. (a) Noisy; (b) C-LLD; (c) TV; (d) SRD; (e) NL-mean; (f) proposed algorithm.

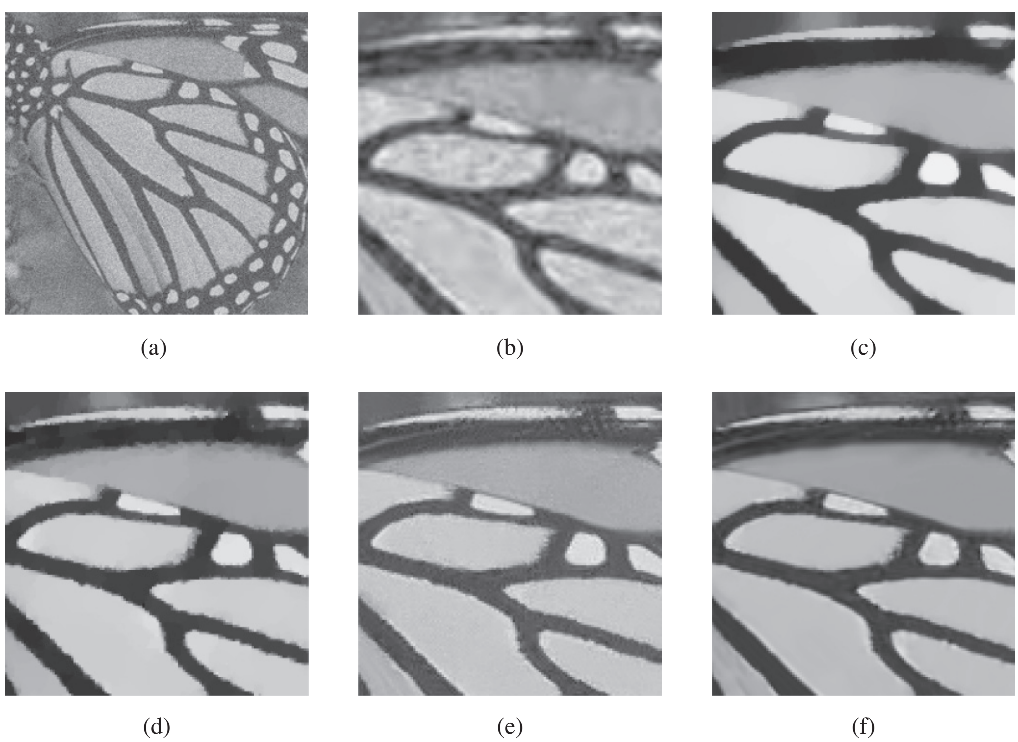

Fig. 2. The results comparison of Butterfly with the noise deviation 20. (a) Noisy; (b) C-LLD; (c) TV; (d) SRD; (e) NL-mean; (f) proposed algorithm.

About the optimization ②, we take the method called FISTA in Ref. 8 to solve our total variation problem. As to the optimization ③, due to its huge computation cost for searching the similar patches, we adopt the shrinkage scheme[15,16] whose pseudo code is outlined in Algorithm II to get the approximate solution.

The transformation in Algorithm II can select any 2D sparse transformation. Here, we adopt the non-subsampling

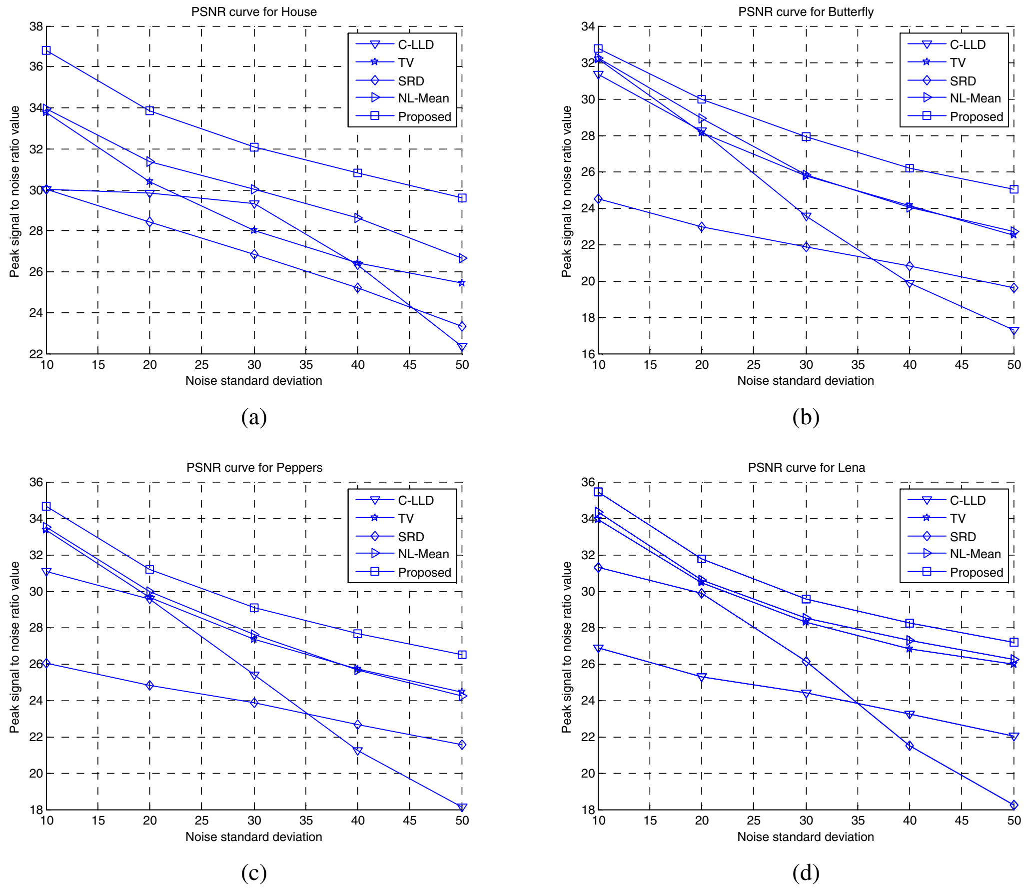

Fig. 3. (Color online) PSNR curves with different deviations. (a) House; (b) Butterfly; (c) Peppers; (d) Lena.

contourlet transformation (NSCT),[17)] which shows so many advantages in representing the image signal.

4. Experimental Results

To evaluate the proposed algorithm, we carried out several experiments on the natural image compared to the current restoration algorithms. Here, we take the denoising task as the example. That is, we set the degradation matrix $H = I$, only leaving the noise matrix n. The test image are respectively Lena (512×512), Peppers (256×256), House (256×256), and Butterfly (256×256). The comparison is conducted between the C-LLD (clustering-based locally learned dictionary),[18)] TV (total variation regularization) in Ref. 9, SRD (sparse restoration by l_1 minimization with over-completed DCT) by Ref. 12, NL-Mean (non-local mean filter) in Ref. 19 and the proposed algorithm. Due to the limited space, we only present the local results of House and Butterfly in Figs. 1 and 2, respectively. Meanwhile, the peak signal-to-noise ratio (PSNR) curves with different noise deviations are depicted in Fig. 3.

Among the visual results, the C-LLD gets the worst effect and generates many scratches. But specially, as to the Lena (it is not shown here), C-LLD shows some advantages in presenting the hair. The TV and SRD are good at recovering the smooth regions, but fail in the texture regions. The result of NL-mean is a little mottled, but shows some advantage on PSNR. The proposed algorithm shows not only best visual quality but also the satisfactory PSNR value among the comparison methods. Beside, compared to the other curves, our algorithm exhibits largo decline with the increasing noise deviation. In addition, we test the performance with two specific cases, whose parameters are $\lambda_2 = \eta_2 = 0$, $\lambda_1 = 0.2$, $\eta_1 = 0.1$ and $\lambda_1 = \eta_1 = 0$, $\lambda_2 = 0.2$, $\eta_2 = 0.1$ separately. We do so is to test the performance constrained by only TV or non-local sparsity regularizer. Compared to the second group of parameters, the first group (only with TV regularizer) generates better performance on local smoothness, but worse result on the texture that can be guaranteed by the non-local geometrical similarity. So, the proposed algorithm with dual-prior constraint models can show advantage on both of the smoothness and texture.

5. Summary and Future Work

In this paper, we proposed a novel method for image restoration. Firstly, in consideration of the advantage of the TV model and sparse model, we combined the two models in one optimization problem and construct an objective function with dual-prior constraint model, which took the TV model to constrain local smoothness and the sparse model to constrain non-local sparsity. Secondly, we developed a scheme to solve the novel optimization, which can be seen as a modified iterative Split-Bregman method. The results showed the proposed algorithm can be a valuable tool for the image restoration task. Future work can be extended to: i) applications to image super-resolution or image inpainting and so on; ii) a study of how to select the 2D Transformation with different task.

Acknowledgement

This work is funded by the Defense advanced research foundation of state shipbuilding (10J3.1.6), the Nation Nature Science Foundation of China (No. 61301095 and No. 61201237), and the Fundamental Research Funds for the Central Universities (No. HEUCFZ1129, No. HEUCF130810, and No. HEUCF130817).

References

1) M. Elad: *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing* (Springer, New York, 2010).
2) L. I. Rudin, S. Osher, and E. Fatemi: Physica D **60** (1992) 259.
3) J. M. Bioucas-Dias, M. A. T. Figueiredo, and J. P. Oliveira: IEEE Int. Conf. Acoustic, Speech and Signal Processing, 2006, p. 861.
4) M. Zhu, S. J. Wright, and T. F. Chan: Comput. Appl. Math. Rep. **47** (2010) 377.
5) M. Sun, N. Feng, Y. Shen, J. Li, L. Ma, and Z. Wu: Chin. Opt. Lett. **9** (2011) 061002.
6) M. Elad and M. Aharon: IEEE Trans. Image Process. **15** (2006) 3736.
7) J. A. Tropp: IEEE Trans. Inf. Theory **50** (2004) 2231.
8) A. Beck and M. Teboulle: IEEE Trans. Image Process. **18** (2009) 2419.
9) J. Dahl, P. C. Hansen, S. H. Jensen, and T. L. Jensen: Numer. Algorithms **53** (2010) 67.
10) S. S. Chen, D. L. Donoho, and M. A. Saunders: SIAM J. Sci. Comput. **20** (1998) 33.
11) J. K. Romberg: Annu. Conf. Information Science and Systems, 2006, p. 213.
12) Y. Nesterov: ECORE Discussion Paper **76** (2007).
13) T. Goldstein and S. Osher: SIAM J. Image Sci. **2** (2009) 323.
14) J. Zhang, Z.-H. Wei, and L. Xiao: Intell. Comput. Theories Appl. **7390** (2012) 189.
15) D. L. Donoho and I. M. Johnstone: Ann. Stat. **26** (1998) 879.
16) J. Friedman, T. Hastie, H. Höfling, and R. Tibshirani: Ann. Appl. Stat. **1** (2007) 302.
17) A. L. da Cunha, J. Zhou, and M. N. Do: IEEE Trans. Image Process. **15** (2006) 3089.
18) P. Chatterjee and P. Milanfar: IEEE Trans. Image Process. **18** (2009) 1438.
19) A. Buades, B. Coll, and J.-M. Morel: IEEE Conf. Computer Vision and Pattern Recognition, 2005, p. 60.