THE NOISE CLINIC : A UNIVERSAL BLIND DENOISING ALGORITHM

M. Lebrun, J.M. Morel

CMLA, Ecole Normale Supérieure de Cachan
61 avenue du Président Wilson,
94235 Cachan CEDEX, France

M. Colom

Universitat de les Illes Balears
Crta de Valldemossa, km 7.5
07122 Palma de Mallorca, Spain

ABSTRACT

Denoising is probably the operation with the highest impact to improve image quality. Indeed the presence of noise hides image details and limits other image improvements such as contrast enhancement, color balance and gamma-correction. An immense effort has been dedicated to improve denoising methods, but most papers assume a fixed noise model, mainly white Gaussian. Yet in most images handled by the public or even by scientific users, the noise model is imperfectly known or even unknown. Very little has been written on actual image denoising methods including the noise estimation step. In this paper we propose to estimate a rather general noise model, both signal and frequency dependent, coupled with a general denoising multiscale method able to cope with general noise distortion, including JPEG compression. On noisy images coming from various sources (JPEG, scans of old photographs,...) we show perceptually convincing results, for which of course no ground truth is available.

***Index Terms*—** Blind Denoising, multiscale algorithm, noise estimation, denoising

Note to referees: a demo of the Noise Clinic algorithm is available on line at `http://dev.ipol.im/~colom/ipol_demo/noise_clinic_exper_v6/`

1. INTRODUCTION

Our goal here is to provide a "blind" image denoising method. Its denoising part is preceded by an accurate noise estimate made from the image itself. Our minimal assumption on noise will be that it is signal and frequency dependent. This assumption is for example compatible with noise in JPEG images, which is the result of a (signal dependent) Poisson noise which has undergone a quantization of its DCT coefficients. We shall see that a recent state of the art method, Nonlocal Bayes (NL-Bayes), can be adapted to that purpose, but most other denoising methods can be adapted as well. Our main motivation comes from the fact that most image users in science and technology do not actually dispose of both the raw image and the noise model, but only often of an "end product" which is not necessarily fully denoised and has anyway undergone several frequency and signal alterations. Thus, blind denoising must be able to cope with both raw and preprocessed images of all sorts.

This helplessness of image users can be observed in the journal IPOL archives of the online executions of six state of the art or emblematic denoising methods (NL-means [1], DCT denoising [2], TV denoising [3], K-SVD [4], BM3D [5] and NL-Bayes [6]). Whereas it allows users to upload noise-free images, to add the noise and denoise them, it appears that most of submitted images are not noise-free images, nor even white noise images, which leads to inefficient and even misleading results. Only from this fact comes clear that the demand for image denoising exceeds widely the white noise case. "Blind" methods are required for a good diffusion of state of the art image processing methods among other scientific disciplines.

This blind denoising approach was studied by Javier Portilla [7], [8], by Tamer Rabie [9] and by Liu, Freeman, Szeliski and Kang [10]. For this purpose, Portilla modified his state of the art denoising algorithm BLS-GSM and adapted it to deal with homogeneous, Gaussian or mesokurtotic noise, which provides the only state of the art blind denoising algorithm to our knowledge. Liu, Freeman, Szeliski and Kang proposed a unified framework for JPEG image for two tasks : 1) automatic estimation and 2), removal of color noise from a single image. The paper proposed by Rabie seems less effective and works only for Gaussian noise, where the blind denoising filter is based on the theory of robust statistics.

Our plan follows from the above considerations. Section 2 explains how the original NL-Bayes algorithm may be adapted to the current general noise framework. The noise estimation procedure is described in Section 3. Section 4 details the computation of the noise covariance matrix. Section 5 gives the multiscale denoising procedure. The final synthetic description of the whole blind denoising method can be found in Section 6. Some results on real noisy images with unknown preprocessing and comparison with the state of the art algorithm of [7] are presented in Section 7.

2. A QUICK REMINDER OF THE NL-BAYES ALGORITHM

Like most current state of the art methods, NL-Bayes denoises all noisy square patches extracted from the image and then obtains the final denoised image $\hat{u}$ by replacing every image pixel value by an average of the denoised values obtained for this pixel in all denoised patches containing it. We shall denote by $\tilde{P}$ a reference patch extracted from the image, and by $\mathcal{P}(\tilde{P})$ the set of patches $\tilde{Q}$ similar to the reference patch $\tilde{P}$. Assuming that the patches similar to a given patch follow a Gaussian model, a first basic estimation of the denoised patch P can be obtained by

$$P^{\textbf{basic}} = \overline{\tilde{P}} + \left[\mathbf{C}_{\tilde{P}} - \mathbf{C}_n\right] \mathbf{C}_{\tilde{P}}^{-1} \left(\tilde{P} - \overline{\tilde{P}}\right) \tag{1}$$

where $\overline{\tilde{P}}$ is the average of patches similar to $\tilde{P}$, $\mathbf{C}_n$ is the covariance matrix of the noise and $\mathbf{C}_{\tilde{P}}$ is the covariance matrix of the patches similar to $\tilde{P}$. For pure Gaussian signal-independent noise, we simply have $\mathbf{C}_n = \sigma^2 \mathbf{I}$.

Seeing the basic estimation as an "oracle", equation (1) becomes:

$$P^{\textbf{final}} = \overline{\tilde{P}}^{\textbf{basic}} + \mathbf{C}_{\tilde{P}}^{\textbf{basic}} \left[\mathbf{C}_{\tilde{P}}^{\textbf{basic}} - \mathbf{C}_n\right]^{-1} \left(\tilde{P} - \overline{\tilde{P}}^{\textbf{basic}}\right). \tag{2}$$

Then to adapt the original NL-Bayes algorithm to signal dependent noise, one has to provide an estimated covariance matrix of the noise $\mathbf{C}_{n\mathbf{i}}$ for every group of similar patches $\mathcal{P}(\tilde{P})$ according to its average intensity value $\mathbf{i}$.

The way to obtain $\mathbf{C}_{n\mathbf{i}}$ is explained in section 4.

3. NOISE ESTIMATION

Most noise estimation algorithms capable of estimating the noise variance according to the frequency can be easily adapted to measure signal-dependent noise [11]. For the Noise Clinic we adapted an existing method by Ponamarenko et al. [12] to estimate the noise variance at each frequency.

The algorithm can be summarized as follows:

1. Extract from the input image (of size $N_x \times N_y$ pixels) all possible $M = (N_x - w + 1)(N_y - w + 1)$ overlapping $w \times w$ blocks (with $w = 4$) and compute its 2D orthonormal DCT-II.
2. Set $\mathbf{L} = \emptyset$ (the empty set).
3. For each DCT block m_1,
   - (a) Look for the block m_2 that minimizes $\text{PMSE}_{m_1,m2}$ (Eq. (3)). Consider only those blocks whose horizontal and vertical distance with respect to m_1 is belongs to the interval $[r_1, r_2] = [4, 14]$.
   - (b) Add block m_1 and its PMSE, $[m_1, \text{PMSE}_{m_1,m2}]$, to list $\mathbf{L}$.
4. Compute the mean of each block[1].
5. Classify the elements of list $\mathbf{L}$ into disjoint bins according to the intensity of the blocks [11]. Each bin contains (with the exception of the last) 42000 elements.

Then, for each bin,

1. Consider the set $\mathcal{S}_p$ made by the blocks inside the current bin whose PMSE is below the p-quantile, with $p = 0.005$.
2. Assign to the current bin the intensity I as the median of the means of the blocks that belong to the bin[2].
3. For each frequency $[i, j]$ with $[i, j] \in [0, w-1]^2$, $[i, j] \neq [0, 0]$,
   - (a) Compute the (biased[3]) variance of the noise at the current bin and frequency $[i, j]$ using the MAD estimator (Eq. (4)).
   - (b) Correct the biased variance and obtain the final estimate $\tilde{\sigma}[I][i, j] = 1.967\hat{\sigma}[I][i, j] - 0.2777$.

$$\text{PMSE}_{m_1,m_2} := \frac{1}{w^2} \sum_{i=0}^{w} \sum_{j=0}^{w} (D_{m_1}[i,j] - D_{m_2}[i,j])^2 (w^2 + 1 - i - j)^2. \tag{3}$$

$$\hat{\sigma}[i, j] = \text{MAD}(\mathcal{S}_p) = \text{median}_k \left[|\mathcal{S}_p[k][i, j]) - \text{median}_l (\mathcal{S}_p[l][i, j])| \right]. \tag{4}$$

Note that with this approach it is not possible to estimate the STD of the noise frequency $[0, 0]$ (DC). However, since the complete algorithm is multiscale (see Sec. 5), the missed information of the STD of the noise at the DC in a given scale is later recoved when the noise is estimated at the next scale.

4. HOW TO OBTAIN THE COVARIANCE MATRIX

At this point, we will suppose that for any given intensity $\mathbf{i}$, the multi-frequency noise estimate has provided us with $k^2 \times k^2$ matrices.

$$\mathbf{M}_\mathbf{i} = \mathbb{E}\left(\mathcal{D}N_\mathbf{i} \left(\mathcal{D}N_\mathbf{i}\right)^t\right) \tag{5}$$

where:

- $\mathcal{D}$ is the matrix of the discrete cosine transform (DCT) of size $k^2 \times k^2$;
- $N_\mathbf{i}$ denotes the $k \times k$ stochastic noise patch model at intensity $\mathbf{i}$.

[1] This operation is fast since the mean of the block can be obtained by dividing into w the value of the coefficient at frequency $[0, 0]$.

[2] The means of the blocks have been already computed in step 4.

[3] The estimate is biased because of the MAD estimator and because the variance is measured using a finite number of samples from $\mathbf{L}$.

From equation (5) and the definition of the covariance matrix of the noise, it comes for a given intensity $\mathbf{i}$ that

$$\mathbf{C}_{n\mathbf{i}} = Cov(N_{\mathbf{i}}) = \mathbb{E}\left(N_{\mathbf{i}} N_{\mathbf{i}}^t\right) = \mathcal{D}^t \mathbf{M}_{\mathbf{i}} \mathcal{D}.$$

5. THE MULTISCALE ALGORITHM

The state-of-the-art denoising algorithms such that DDID (Knaus et al. [13]), BM3D (Dabov et al. [14]), NL-means (Buades et al. [15]), K-SVD (Mairal et al. [16], [17]), Wiener filters applied on DCT (Yaroslavsky et al. [18], [19]) or on wavelet transform (Donoho et al. [20]) or even the total variation minimization (Rudin et al. [21]) achieve very good results for moderate noise,for large noise many artifacts inherent to each method start appearing, in particular low frequency noise. A natural idea to deal with is to involve a multiscale procedure, which promises three improvements: 1) it favors a better patch comparison, 2) at larger scales the noise decreases, 3) subsampling the image before denoising amounts to enlarge the real size of the neighborhood.

5.1. Down and Up Sampling

The sub-sampling is done by averaging four samples in the higher scale without any overlapping. As there is four ways to do it (depending on the starting pixel), the four sub-sampled images are kept to avoid aliasing. Then the noise estimation may work with the same amount of pixels at every scale. As the four sub-images are shifted by $\pm\frac{1}{2}$ in both coordinate directions, the up-sampling of higher scale pixels is done by averaging their four neighbors, each one belonging to each sub-image.

Figure 1 shows at each scale the result of the denoising.

5.2. Noise Estimation

If the input noisy image had pure Gaussian noise, then after each sub-sampling the noise should be divided by two and remain white. However, the proposed algorithm has been developed to deal with all kinds of noisy image, as shown in Figure 2 anything can happen to the noise curves, then the noise covariance matrices must be estimated at each scale. As the noise estimation is applied on the set of all sub-images for a given scale, then the estimation have the same accuracy whatever the current scale is. Then for one scale, all sub-images are denoised with the same set of noise covariance matrices.

Figure 2 shows an example of average noise curves over high and low frequencies for a three scales noise estimation.

6. THE FINAL NOISE CLINIC ALGORITHM

Now that all parts of the Noise Clinic algorithm have been detailed in the previous sections, we can summarize the whole algorithm in Algorithm 1.

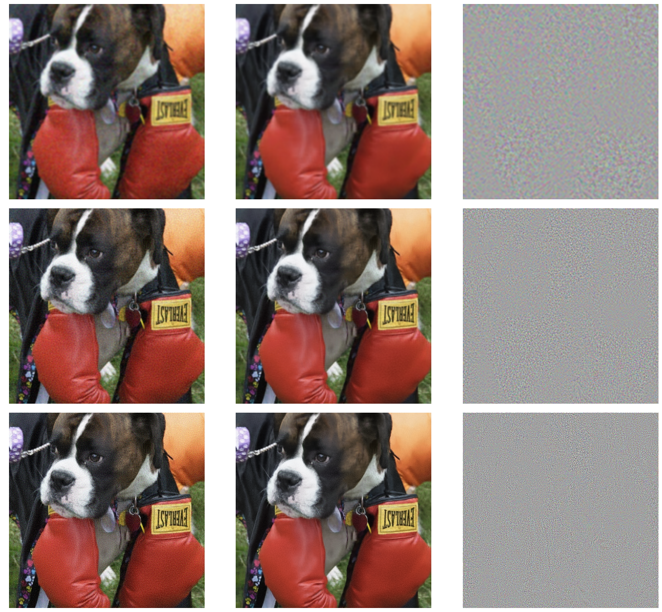

Fig. 1. Result of the Noise Clinic at each scale. From top to bottom : scale 2, scale 1 and scale 0. From left to right : noisy image, denoised image, difference image.

7. RESULTS AND COMPARISONS

A comparison with blind BLS-GSM introduced in [7] and [8] is shown in Figure 3 on some images with different kinds and values for the noise, extracted from [8]. Whereas for the left image the Noise Clinic better succeed to remove all the low frequency noise than blind BLS-GSM while preserving details, it re-enforces the strong structured noise in the right image, while blind BLS-GSM remarkably removes it. However one can argue that this structured noise may be seen and treated as detail belonging to the image.

Results over an old photography and a JPEG image are given in Figure 4. Both noisy images present a huge amount of noise with artifacts, but the Noise Clinic manages to remove a lot of it, while well preserving details and structure of the image.

8. CONCLUSION

The presented algorithm brings together state-of-the-art methods of both world of denoising and noise estimation to create a simple and effective blind denoising algorithm. The link is done via the covariance matrix of the noise. The power of the proposed method lie in the fact that very few assumptions on the nature of the noise are done, which allows the Noise Clinic to give really impressive and efficient results on almost any natural image, even if it has been modified by destructive applications such as JPEG compression. This power

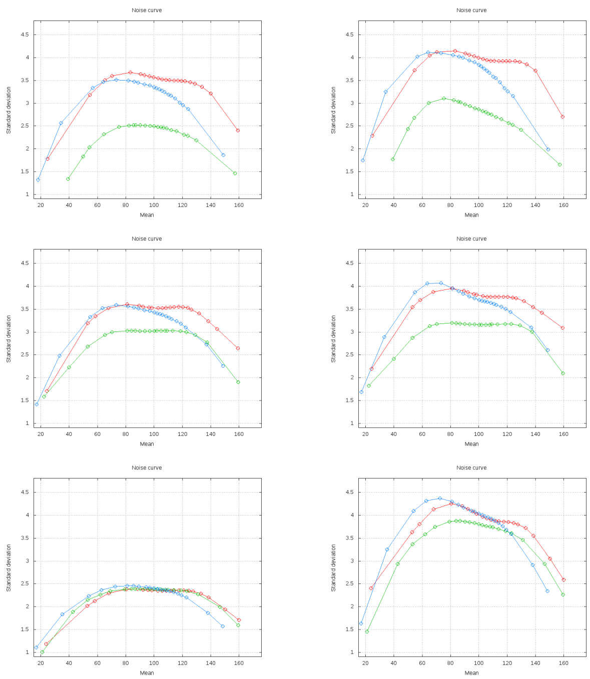

Fig. 2. Average noise curves. From left to right : low frequencies, high frequencies. From top to bottom : scale 2, scale 1, scale 0.

is strengthened by the multi-scale approach which allows to efficiently remove low-frequency noise while preserving tiny details.

However this method may still be improved by better dealing with structured noise, or adapting the NL-Bayes algorithm to have a better/more generic model than the current Gaussian one for patches.

This algorithm may be tested on line at this address : `http://dev.ipol.im/~colom/ipol_demo/noise_clinic_exper_v6/`

Algorithm 1 Noise Clinic Algorithm

Input : Noisy image $\tilde{u}_0$, Number of scales N
Output : Denoised image $\hat{u}_0$

Part 1 : Getting the image scale pyramid
for each scale $s = 1$ to $N-1$ **do**
  Let $\{\tilde{u}^k_{s-1}\}$ be the previously set of noisy sub-images
  for $k = 1$ to 4^{s-1} **do**
    Down-sample $\tilde{u}^k_{s-1}$ into 4 sub-images and save details

$$\tilde{d}^k_{s-1} = \tilde{u}^k_{s-1} - \mathcal{U}\left(\{\tilde{u}^{4(k-1)+i}_s\}_{i\in[\![1,4]\!]}\right)^k$$

  end for
  if $s = N-1$ **then**
    Set $\{\tilde{v}^k_{N-1}\}_k = \{\tilde{u}^k_{N-1}\}_k$
  end if
end for

Part 2 : Estimating the noise and denoising the pyramid
for $s = N-1$ to 0 **do**
  Estimate the noise covariance matrices over $\{\tilde{v}^k_s\}_k$.
  Denoise $\{\tilde{v}^k_s\}_k$ with the NL-Bayes algorithm by using noise covariance matrices $\{\mathcal{D}^t \mathbf{M}_\mathbf{i} \mathcal{D}\}_\mathbf{i}$ to get $\{\hat{u}^k_s\}_k$.
  if $s > 0$ **then**
    for $k = 1$ to 4^{s-1} **do**
      Up-sample $\{\hat{u}^{4(k-1)+i}_s\}_{i\in[\![1,4]\!]}$ and add the saved details $\tilde{d}^k_{s-1}$ to get $\tilde{v}^k_{s-1}$.
    end for
  else
    $\hat{u}_0 = \hat{u}^1_0$
  end if
end for

9. REFERENCES

[1] A. Buades, B. Coll, and J.M. Morel, “Non-Local Means Denoising,” *Image Processing On Line*, vol. 2011, 2011.

[2] G. Yu and G. Sapiro, “DCT image denoising: a simple and effective image denoising algorithm,” *Image Processing On Line*, vol. 2011, 2011.

[3] P. Getreuer, “Rudin-Osher-Fatemi Total Variation Denoising using Split Bregman,” *Image Processing On Line*, vol. 2012, 2012.

[4] M. Lebrun and A. Leclaire, “An Implementation and Detailed Analysis of the K-SVD Image Denoising Algorithm,” *Image Processing On Line*, vol. 2012, 2012.

[5] M. Lebrun, “An Analysis and Implementation of the BM3D Image Denoising Method,” *Image Processing On Line*, vol. 2012, 2012.

[6] M. Lebrun, A. Buades, and J.M. Morel, “Implementation of the ”Non-Local Bayes” (NL-Bayes) Image Denoising Algorithm,” *Image Processing On Line*, vol. 2013, pp. 1–42, 2013.

[7] J. Portilla, “Full blind denoising through noise covariance estimation using gaussian scale mixtures in the wavelet domain,” *Image Processing, 2004. ICIP’04. 2004 International Conference on*, vol. 2, pp. 1217–1220, 2004.

[8] J. Portilla, “Blind non-white noise removal in images using gaussian scale mixtures in the wavelet domain,” *Benelux Signal Processing Symposium*, 2004.

[9] T. Rabie, “Robust estimation approach for blind denoising,” *IEEE Transactions on Image Processing*, vol. 14, no. 11, pp. 1755–1765, November 2005.

[10] Liu, W. Freeman, R. Szeliski, and S. Kang, “Automatic estimation and removal of noise from a single image,”

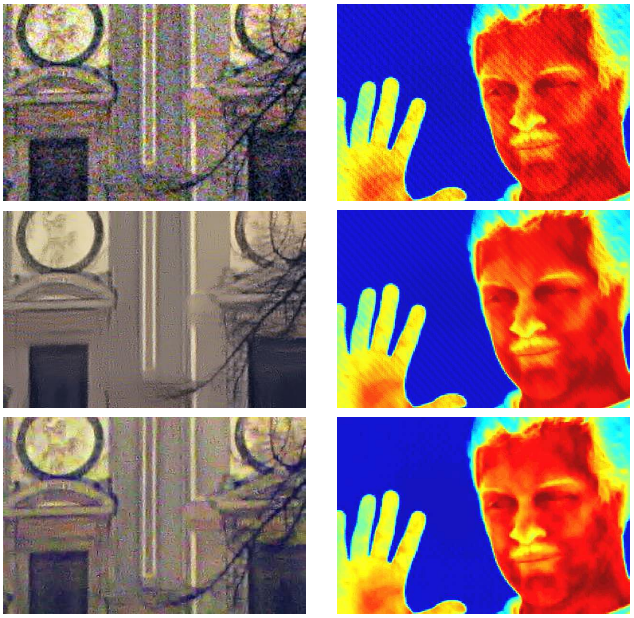

Fig. 3. Results of our blind denoising and of Blind BLS-GSM on different images from [8]. From Top to bottom : noisy image, Noise Clinic result, Blind BSL-GSM result.

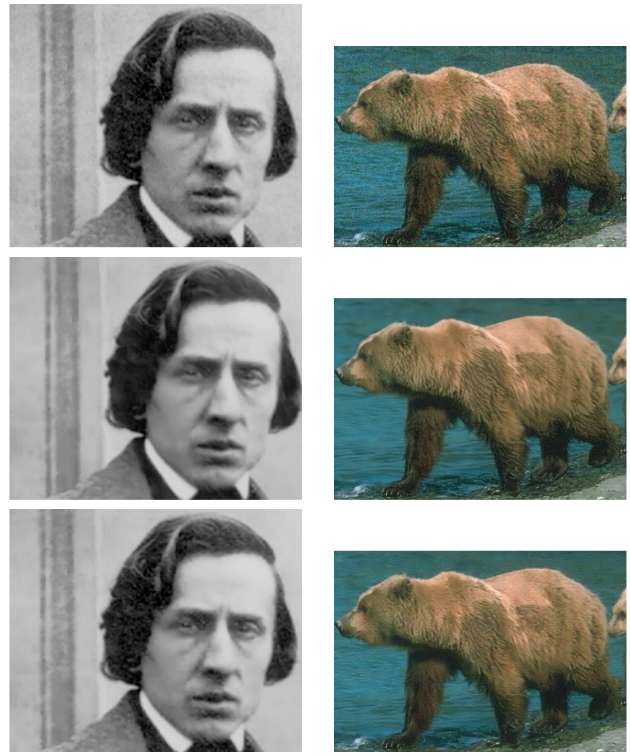

Fig. 4. Results of our blind denoising algorithm over noisy images. From Top to bottom : noisy image, Noise Clinic result, Blind BSL-GSM result.

IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, no. 2, pp. 299–314, February 2008.

[11] M. Lebrun, M. Colom, A. Buades, and J.M. Morel, "Secrets of image denoising cuisine," *Acta Numerica*, vol. 21, pp. 475–576, 2012.

[12] Nikolay N. Ponomarenko, Vladimir V. Lukin, Karen O. Egiazarian, and Jaakko T. Astola, "A method for blind estimation of spatially correlated noise characteristics," in *IS&T/SPIE Electronic Imaging*. International Society for Optics and Photonics, 2010, pp. 753208–753208.

[13] C. Knaus and M. Zwicker, "Dual-Domain Image Denoising," *ICIP*, pp. 440–444, 2013.

[14] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3d transform-domain collaborative filtering," *IEEE Transactions on image processing*, vol. 16, no. 82, pp. 3736–3745, 2007.

[15] A. Buades, B. Coll, and J.M. Morel, "A non local algorithm for image denoising," *IEEE Computer Vision and Pattern Recognition*, vol. 2, pp. 60–65, 2005.

[16] J. Mairal, M. Elad, and G. Sapiro, "Sparse representation for color image restoration," *IEEE Transactions on image processing*, vol. 17, no. 1, pp. 53–69, 2008.

[17] J. Mairal, G. Sapiro, and M. Elad, "Learning multiscale sparse representations for image and video restoration," *SIAM Multiscale Modeling and Simulation*, vol. 7, no. 1, pp. 214–241, 2008.

[18] L.P. Yaroslavsky, K.O. Egiazarian, and J.T. Astola, "Transform domain image restoration methods: review, comparison, and interpretation," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, May 2001, vol. 4304, pp. 155–169.

[19] L.P. Yaroslavsky, "Local adaptive image restoration and enhancement with the use of DFT and DCT in a running window," in *Proceedings of SPIE*, 1996, vol. 2825, pp. 2–13.

[20] J.L. Starck, E.J. Candès, and D.L. Donoho, "The curvelet transform for image denoising," *IEEE Transactions on image processing*, vol. 11, pp. 670–684, 2002.

[21] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, pp. 259–268, 1992.