Statistical Modeling of the Patches DC Component for Low-Frequency Noise Reduction

Antoine Houdard
CNRS, Univ. Bordeaux, IMB,
F-33400 Talence, France
antoine.houdard@math.u-bordeaux.fr

***Abstract*—This work is devoted to patch-based image denoising. Assuming an additive white Gaussian noise (AWGN) on patches, we derive corresponding models on centered patches and on their DC components. Then we propose a strategy for improving a given path-based denoiser. Finally, we provide experiments with the recent denoising method HDMI that shows improvement of the denoising quality, particularly for residual low frequency noise.**

***Index Terms*—patch-based image denoising, noise modeling, low frenquency noise reduction, patch centering**

I. Introduction

a) Context: In this paper, we focus on image denoising, which aims to estimate an image $\widehat{u}$ from its noisy observation

$$v = u + e \in \mathbf{R}^n, \tag{1}$$

where $e \sim \mathcal{N}(0, \sigma^2 \mathrm{I}_n)$ is an additive white Gaussian noise (AWGN) and u the underlying clean image. A popular way to deal with this problem, is to represent the image with the set of all its patches. This patch approach has, for instance, lead to the well-known denoising methods Non-Local means [1] and BM3D [2]. In this context, each patch $i \in \{1, \dots, n\}$ is seen as vector of size $p = s \times s$, and the model (1) yields the following model on the patch-space

$$Y_i = x_i + N_i, \tag{2}$$

where $Y_i \in \mathbf{R}^p$ is the observed random vector modeling the i-th patch, $x_i \in \mathbf{R}^p$ is the underlying clean patch and $N_i \sim \mathcal{N}(0, \sigma^2 \mathrm{I}_p)$ is a Gaussian white noise. Using this, some of the recent denoising methods are based on a statistical modeling of the image patches [3], [4], [5], [6], [7]. The idea behind these methods is to set a prior model on the clean patch x_i seen as a realization of a random vector X_i. The model therefore rewrites

$$Y_i = X_i + N_i, \tag{3}$$

and Bayes' theorem yields the posterior $X_i|Y_i$. Finally, each clean patch can be estimated with the conditional expectation

$$\widehat{x}_i = \mathbb{E}[X_i|Y_i = y_i]. \tag{4}$$

Convenient priors for computing this estimator (4) are Gaussian priors [3] or Gaussian mixture models (GMM) [4], [6], [7]. The use of these priors has been widely studied and it appears that the covariance matrix of these models can encode local structures up to some contrast change [8]. This permits to regroup more patches under the same Gaussian model and then allows for a better estimate of its parameters. There is however a drawback: grouping patches in this way makes the mean of the model less informative. This yields an estimate for each patch that has some bias. This produces the low frequency residual noise that appears in the result of model-based patch-based denoising methods. Figure 1 (b) illustrates this phenomenon in the case of the HDMI method [7], with strong noise and small patches. A large part of this low frequency noise seems to come from a poor estimation of the mean of each patch. Indeed, the image (d) from figure 1 made up of the mean of each patch of the noisy image (a) has the same patterns than the image (c) which is made up of the mean of each patch of the denoised image (b). Moreover, replacing the mean of each denoised patches from (b) with the true mean of the oracle (f), yields a denoised image (c) that is way better, both in terms of Peak Signal to Noise Ratio (PSNR) and visual quality, than the image (b). In addition, some methods from the literature [6], [5] also seem to suggest that removing the mean – also called the DC component – of the patches may improve the denoising quality.

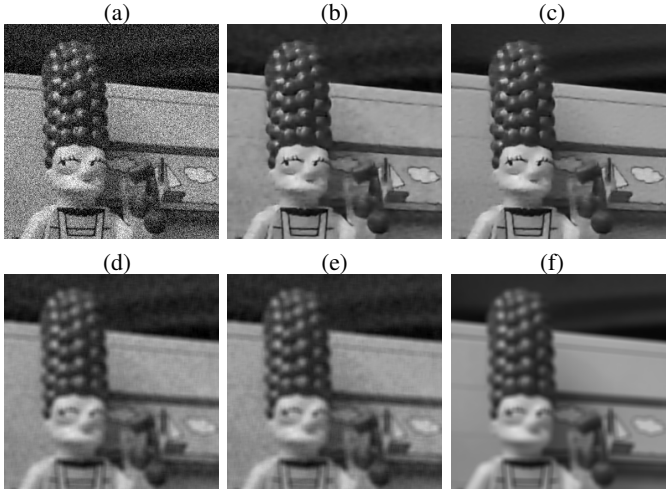

Fig. 1. **(a)** noisy image (standard deviation $30/255$). **(b)** image denoised with HDMI [7] (patches 7×7), PSNR = 31.92 dB. **(c)** image from the denoised patches of (b) with DC component corrected with oracle value (f), PSNR = 33.78 dB. **(d)** DC component of the patches of (a). **(e)** DC component of the patches of (b). **(f)** DC component of the oracle image (ground truth).

978-9-0827-9703-9/19/$31.00

b) Proposed work: In this work, we propose to study the decomposition of the patches into the DC component and the centered component for denoising purposes. To do so, we define the centered observed random variable $Y_i^c = Y_i - \bar{Y}_i 1_p$, where

$$\bar{Y}_i = \frac{1}{p}\sum_{j=1}^{p} Y_i(j), \tag{5}$$

is the mean of Y_i and $1_p = (1, \dots, 1) \in \mathbf{R}^p$. The model (3) can then be divided into the two following problems

$$\bar{Y}_i = \bar{X}_i + \bar{N}_i \in \mathbf{R}, \tag{6}$$

and

$$Y_i^c = X_i^c + N_i^c \in \mathbf{R}^p. \tag{7}$$

We propose to model the noise components N_i^c (section II) and $\bar{N}_i$ (section III) of these two problems. Then we propose in section IV to use this decomposition for enhancing the denoising results of existing patch-based denoising methods such as, for instance, [3], [6], [7]. Finally, we provide numerical experiments that shows improvement of the denoising quality in the case of the HDMI method [7].

II. Modeling the centered noise

The centered noise is defined by $N_i^c = N_i - \overline{N_i} 1_p$ and then the j-th entry of N_i^c is

$$N_i^c(j) = N_i(j) - \frac{1}{p}\sum_{k=1}^{p} N_i(k), \tag{8}$$

since N_i is a Gaussian random vector, N_i^c is also Gaussian. Then, we can compute its mean and its covariance matrix coordinate by coordinate. That gives for the mean

$$\mathbb{E}\left[N_i^c(j)\right] = \mathbb{E}\left[N_i(j)\right] - \frac{1}{p}\sum_{k=1}^{p}\mathbb{E}\left[N_j(k)\right] = 0. \tag{9}$$

And for the covariance matrix, we have $\forall k, l \in \{1, \dots, p\}$

$$\begin{aligned}
\mathbb{E}\left[N_i^c(k)N_i^c(l)\right] &= \mathbb{E}\left[N_i(k)N_i(l)\right] \\
&\quad - \frac{1}{p}\sum_{m=1}^{p}\mathbb{E}\left[N_i(m)\left(N_i(l) + N_i(k)\right)\right] \\
&\quad + \frac{1}{p^2}\sum_{m=1}^{p}\sum_{n=1}^{p}\mathbb{E}\left[N_i(m)N_i(n)\right] \\
&= \sigma^2\left(\delta_{kl} - \frac{2}{p} + \frac{p}{p^2}\right) \\
&= \sigma^2\left(\delta_{kl} - \frac{1}{p}\right).
\end{aligned}$$

Finally, we have $N_i^c \sim \mathcal{N}(0, \Sigma_{N_i^c})$ with

$$\Sigma_{N_i^c} = \frac{\sigma^2}{p}\begin{pmatrix} p-1 & -1 & \cdots & -1 \\ -1 & p-1 & & \vdots \\ \vdots & & \ddots & -1 \\ -1 & \cdots & -1 & p-1 \end{pmatrix}. \tag{10}$$

Since $\Sigma_{N_i^c}$ is a real symmetric matrix, there exists an orthonormal basis that diagonalizes it. Given that its eigenvalues are p (of multiplicity $p-1$) and 0, we can build an orthogonal matrix Q such that

$$\Sigma_{N_i^c} = Q\begin{pmatrix} \sigma^2 \mathrm{I}_{p-1} & 0 \\ 0 & 0 \end{pmatrix} Q^T. \tag{11}$$

It is worth noticing that the unit eigenvector corresponding to the eigenvalue 0 is $\frac{1}{\sqrt{p}}(1, \dots, 1)$. The change of basis Q^T applied to the centered noise then yields

$$Q^T N_i^c \sim \mathcal{N}\left(0, \begin{pmatrix} \sigma^2 \mathrm{I}_{p-1} & 0 \\ 0 & 0 \end{pmatrix}\right), \tag{12}$$

with the last line traducing that the mean of a centered vector is zero. Finally, centering the noise implies a dimension reduction and the total variance of the centered noise defined as

$$\begin{aligned}
\mathrm{TVar}(N_i^c) &= \mathbb{E}[\|N_i^c\|_2^2] && (13)\\
&= \mathrm{Tr}(\Sigma_{N_i^c}) = (p-1)\sigma^2 && (14)\\
&= \frac{p-1}{p}\mathrm{TVar}(N_i) && (15)
\end{aligned}$$

is also reduced by a factor $(p-1)/p$.

III. Modeling of the DC component

Since the initial noise model on a patch N_i is a Gaussian vector, its mean is a Gaussian random variable and we have $\bar{N}_i \sim \mathcal{N}(0, \frac{\sigma^2}{p})$. Then, reshaping problem (6) as an image yields the new image denoising problem with additive Gaussian noise

$$\bar{Y} = \bar{X} + \bar{N}, \tag{16}$$

where $\bar{Y}$, $\bar{X}$ and $\bar{N}$ are the images of $\mathbf{R}^n$ whose values at pixel i are $\bar{Y}_i$, $\bar{X}_i$ and $\bar{N}_i$. The major difference is that the noise is now colored. Indeed, if we consider two random variables $\bar{N}_i$ and $\bar{N}_j$ within a same area of $s \times s$ pixels, they are issued from two overlapping patches and thus are not independent.

However, we can still perform patch-based image denoising on this problem: let us consider patches of the same size $p = s \times s$ from this new image. We define the patches $Z_i = \pi_i(\bar{Y})$, $W_i = \pi_i(\bar{X})$ and $M_i = \pi_i(\bar{N})$, where π_i is the i-th patch extraction operator. We consider the patch noise model

$$Z_i = W_i + M_i, \tag{17}$$

with M_i modeling the noise. Since $M_i = (\bar{N}_{i_1}, \dots \bar{N}_{i_p})$, all its entries are linear combinations of the noise components of the problem (1) that are i.i.d following $\mathcal{N}(0, \sigma^2)$. Therefore, all linear combinations of entries of M_i are also Gaussian. This shows that M_i is a Gaussian vector. We can now compute its mean and its covariance matrix.

The mean of M_i is obviously 0_p the coefficients of the covariance matrix Σ_{M_i} are given by

$$(\Sigma_{M_i})_{kl} = \mathbb{E}\left[\bar{N}_{i_k}, \bar{N}_{i_l}\right] = \frac{\sigma^2}{p^2}C_{kl}, \tag{18}$$

where C_{kl} is the number of common pixels between the two patches of the original image from which $\bar{N}_{i_k}$ and $\bar{N}_{i_l}$ are derived. This yields after counting

$$\Sigma_{M_i} = \frac{\sigma^2}{p^2} B \otimes B \tag{19}$$

where

$$B = \begin{pmatrix} s & (s-1) & \cdots & 1 \\ (s-1) & s & \ddots & \vdots \\ \vdots & \ddots & \ddots & (s-1) \\ 1 & \cdots & (s-1) & s \end{pmatrix}, \tag{20}$$

and $\otimes$ is the Kronecker product. In order to use a denoising method that has been designed for AWGN, we want to perform a change of basis for the data. To do so, we study the structure of Σ_{M_i}. First, we show that B is symmetric positive-definite. Using the Sylvester's criterion, it is sufficient to show that all of the leading principal minors of B are positive. These minors of size $d \in \{1, \dots, s\}$ are given by

$$m_d = \begin{vmatrix} s & (s-1) & \cdots & (s-d+1) \\ (s-1) & s & \ddots & \vdots \\ \vdots & \ddots & \ddots & (s-1) \\ (s-d+1) & \cdots & (s-1) & s \end{vmatrix}, \tag{21}$$

Adding the first column in the last one yields

$$m_d = (2s-d+1) \begin{vmatrix} s & (s-1) & \cdots & 1 \\ (s-1) & s & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ (s-d+1) & \cdots & (s-1) & 1 \end{vmatrix}, \tag{22}$$

then subtracting the second and the last column to the first one gives

$$m_d = (2s-d+1) \begin{vmatrix} 2 & (s-1) & \cdots & 1 \\ 0 & s & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & (s-1) & 1 \end{vmatrix}, \tag{23}$$

Finally, developing the determinant with respect to the first column and repeating the two last steps yields

$$m_d = (2s-d+1)2^{d-2} > 0. \tag{24}$$

This shows that B is a positive-definite matrix. Then $B \otimes B$ is also symmetric positive-definite as the Kronecker product of two positive-definite matrices and the Cholesky decomposition yields an invertible matrix L such that $B \otimes B = LL^T$. Therefore, the problem

$$L^{-1} Z_i = L^{-1} W_i + L^{-1} M_i, \tag{25}$$

is an AWGN problem with noise variance $\sigma^2/p^2 \mathrm{I}_p$ and a denoising method such as HDMI can be used to find an estimate $\widehat{L^{-1}W_i}$ that gives an estimate of W_i by $\widehat{W}_i = L\widehat{L^{-1}W_i}$.

Fig. 2. Left: noisy DC image ($\sigma = 50/255$). Middle: denoised DC image *PSNR* 38.58 *dB*. Right: ground-truth DC image.

IV. Experiments

In this section, we take advantage of the previous modeling in order to improve the denoising result of model-based patch-based denoising methods. That is, methods that denoise each patch with a denoising operator $f_{denoise}$. For this testing part, we propose to use $f_{denoise}$ = HDMI [7] which has the advantage of using only statistical tools. The principle of this method is rather simple:

1) It uses a GMM with intrinsic lower dimensions to model clean patches;
2) This model is inferred on the noisy patches with an expectation–maximization (EM) algorithm;
3) The clean patches are estimated with the conditional expectation (4), which has a closed-form and is numerically stable (proposition 1 of [7]).

A. Denoising the DC component

In section III, we have proposed to reshape the DC component of all patches into an image and to extract patches of this image in order to perform patch-based denoising on it. We showed that the noise model on these patches is an AWGN in a given basis L. We can therefore apply the HDMI method directly on (25). Figure 2 shows the result of the denoising of the DC images from the images *simpson* and *lena* with a noise of standard deviation $\sigma = 30/255$. Note that the results are quite good since the problem (25) is an easier problem than the original one (3). Indeed, the dynamic of the DC image is quite the same as the one of the original image whereas the dynamic of the noise is reduced by a factor p. Therefore, the signal-to-noise ratio of the problem (17) is about p times larger than the one of the original problem (3).

Finally, with this step, we obtain for each patch $i \in \{1, \dots, n\}$ of size $s \times s$ from an image $u \in \mathbf{R}^n$, an accurate estimate $\widehat{\bar{X}_i}$ of its DC component as follow :

- construct an image U such that each pixel i is the mean $\bar{Y}_i$ of the patch i;
- extract the patches Z_i from U of size $s \times s$;
- denoise each patch in the new basis $L^{-1}Z_i$ with the denoising operator $f_{denoise}$
- estimate $\widehat{W}_i = L\, f_{denoise}\left(L^{-1}Z_i\right)$ then reconstruct the denoised image $\widehat{U}$;
- an estimate of the DC component of each patch of the original image is then given by $\widehat{\bar{X}_i} = \widehat{U}_i$.

Fig. 3. Images used for the experiments. Grey-scale: *Simpson*, *Lena* and *Barbara*. Color: *Traffic*, *Dice* and *Flower*.

B. Denoising the patches

In order to perform the final patch denoising, we explored two strategies :

(S1) denoise the centered patches from (7) then add the DC estimate from section IV-A;

(S2) denoise the original patches from (3), remove afterwards their DC component, then add the DC estimate from section IV-A.

The results of these two strategies are shown in table I for the three grey-scale images *simpson*, *lena* and *barbara* with different level of noise. The second strategy (S2) always performs better than the first one (S1). Although the first strategy performs well in the constant areas, it fails in the complex parts and makes texture more blurry. This can be explained, since the signal-to-noise ratio of the problem on centered patches (7) is lower than on original patches (3) (the centering of the patches has reduced their dynamics). In this section, we then propose to study the denoising improvement of the strategy (S2) that correct the DC component of the patches afterwards.

TABLE I
RESULTS IN PSNR (DB) OF TWO STRATEGIES (S1) AND (S2) USED ON THE HDMI METHOD WITH PATCHES OF SIZE 10×10 FOR DIFFERENT NOISE LEVELS. IMAGES *simpson*, *lena* AND *barbara* FROM FIGURE 3.

Image	σ	HDMI [7]	(S2)	diff.	(S1)	diff.
Simpson	10	38.89	38.93	+0.04	38.93	+0.04
	20	34.88	34.97	+0.09	34.96	+0.08
	30	32.46	32.58	+0.12	32.58	+0.12
Lena	10	35.81	35.82	+0.01	35.81	+0.00
	20	32.86	32.91	+0.05	32.89	+0.03
	30	31.09	31.19	+0.10	31.14	+0.05
Barbara	10	34.81	34.82	+0.01	34.79	-0.02
	20	31.42	31.45	+0.03	31.42	+0.00
	30	29.38	29.42	+0.04	29.36	-0.02

1) Correction of the DC component afterwards: Hereafter, we propose a method for improving the denoising result of any denoising method that uses an operator $f_{denoise}$ that remove AWGN. The main steps are as follows:

1) compute an estimate $\widehat{X}_i = f_{denoise}(Y_i)$ of each patch of the image;

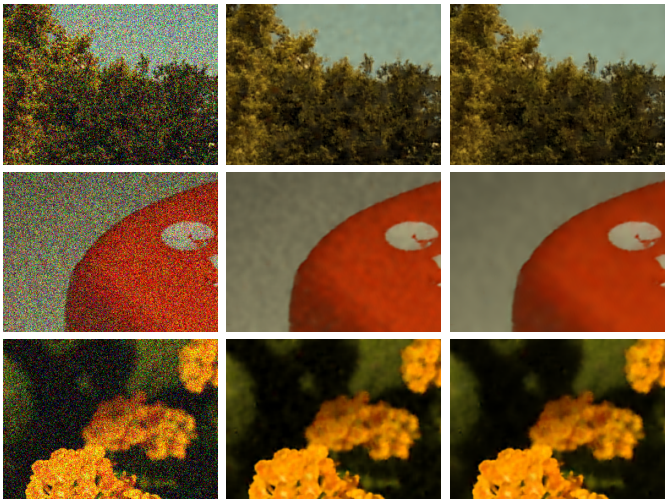

Fig. 4. From left to right. Noisy images (standard deviation $50/255$). HDMI resulting groups. Images denoised by HDMI with patches 7×7 (top 26.97dB, middle 34.61dB, bottom 31.76dB). Images denoised with DC correction method (top 27.07dB, middle 36.06dB, bottom 32.16dB).

2) compute an estimate $\widehat{\widehat{X}}_i$ of the DC component of each patch as show in section IV-A;
3) correct each patch with the new estimate of its DC component:

$$\text{Est}\,(X_i) = \widehat{X}_i - \widehat{\widehat{X}}_i 1_p + \widehat{\widehat{X}}_i 1_p. \tag{26}$$

This strategy improves the quality of denoising compared to the original method. The figure 4 illustrates this improvement with $f_{denoise}$ = HDMI for two images. The low frequency noise is reduced visually and the results are also better in terms of PSNR, $+0.10$dB for the upper image, that has fine textures, and $+1.45$dB for the middle image, which is smoother, and $+0.40$db for the bottom image, which has texture and constant parts.

C. Discussion

Table II shows the results obtained for different patch sizes, different noise levels the three images presented in figure 3. Our approach systematically improves the quality of denoising, with a much more significant improvement for the least textured images. This improvement allows the HDMI method to be close to the performance of the state-of-the-art FFDnet [9] method which uses deep learning. Even the HDMI method was already above the BM3D method [2], the proposed approach allows HDMI to really outperforms this method that used to be the state-of-the-art for path-based image denoising. Another trend that appears is that the improvement is significantly better in the case of high variance than in the case of low variance. The residual low-frequency noise being proportional to the variance, this trend shows that this residual noise is actually reduced. Finally, we can see that the smaller are the patches, the greater is the improvement. This behavior can be explained by the fact that the variance of the noise of the DC component is σ^2/p where p is the

TABLE II

RESULTS IN PSNR (dB) FOR THE THREE IMAGES FROM FIGURE 3 AND VARIOUS NOISE STANDARD-DEVIATION σ. HDMI$_s$ INDICATES THAT WE USED PATCHES OF SIZE $s \times s$, *basic* CORRESPONDS TO HDMI [7], *improved* IS OUR PROPOSED IMPROVEMENT, *diff.* IS THE DIFFERENCE. RESULTS OF THE STATE-OF-THE ART ALGORITHMS BM3D [2] AND FFDNET [9]ARE SHOWN FOR SAKE OF COMPARISON.

Images	σ	HDMI$_5$			HDMI$_7$			HDMI$_{10}$			BM3D [2]	FFDnet [9]
		basic	improved	diff.	basic	improved	diff.	basic	improved	diff.		
Traffic	20	31.42	31.53	+0.11	31.48	31.53	+0.05	31.29	31.31	+0.02	30.81	31.74
	30	29.24	29.40	+0.16	29.39	29.46	+0.07	29.25	29.28	+0.03	28.83	29.79
	40	27.74	27.95	+0.21	27.99	28.09	+0.10	27.94	27.98	+0.04	27.45	28.48
	50	26.60	26.86	+0.26	26.97	27.07	+0.10	26.97	27.02	+0.05	26.43	27.52
Dice	20	38.51	40.00	+1.49	39.76	40.74	+0.98	40.49	41.01	+0.52	39.98	41.06
	30	36.08	37.79	+1.71	37.48	38.76	+1.28	38.55	39.30	+0.75	38.01	39.36
	40	34.33	36.14	+1.81	35.90	37.21	+1.31	37.01	37.95	+0.94	36.52	38.01
	50	32.98	34.88	+1.90	34.61	36.06	+1.45	35.89	36.85	+0.96	35.19	36.72
Flower	20	36.25	36.71	+0.46	36.76	36.95	+0.19	36.81	36.85	+0.04	35.89	37.19
	30	33.91	34.53	+0.62	34.58	34.86	+0.28	34.73	34.81	+0.08	33.74	35.18
	40	32.22	32.94	+0.72	33.00	33.36	+0.36	33.26	33.37	+0.11	32.13	33.73
	50	30.90	31.74	+0.84	31.76	32.16	+0.40	32.11	32.27	+0.16	30.94	32.51
Mean	20	35.39	36.08	+0.69	36.00	36.41	+0.41	36.20	36.39	+0.19	35.56	36.66
	30	33.08	33.91	+0.83	33.82	34.36	+0.54	34.18	34.46	+0.29	33.53	34.78
	40	31.43	32.34	+0.91	32.30	32.89	+0.59	32.74	33.10	+0.36	32.03	33.41
	50	30.16	31.16	+1.00	31.11	31.76	+0.65	31.66	32.05	+0.39	30.85	32.25

number of pixels of the patch. Thus, for large patches, the noise for the DC component is already low and correcting it does not improve the final result that much.

V. CONCLUSION

In this work, we have studied the effect of patch centering for model-based patch-based denoising methods. For this purpose, we have proposed a modeling of the centered noisy patches and a modeling of the DC component of the noise. These modeling have led us to a strategy for improving the quality of denoising when we have a denoiser $f_{denoise}$ for the patches. The proposed strategy shows improvement, both visually and in term of PSNR, when used with the HDMI method, especially for the reduction of the low-frequency residual noise.

In future work, we would like to study the links between this approach and multiscale frameworks.

REFERENCES

[1] A. Buades, B. Coll, and J.M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Modeling and Simulation*, vol. 4, no. 2, pp. 490–530, 2006.

[2] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian, "Image denoising by sparse 3-d transform-domain collaborative filtering," *Image Processing, IEEE Transactions*, 2007.

[3] M. Lebrun, A. Buades, and J. M. Morel, "A Nonlocal Bayesian Image Denoising Algorithm," *SIAM J. Imaging Sci.*, vol. 6, no. 3, pp. 1665–1688, Sept. 2013.

[4] Yi-Qing Wang and Jean-Michel Morel, "SURE Guided Gaussian Mixture Image Denoising," *SIAM J. Imaging Sci.*, vol. 6, no. 2, pp. 999–1034, May 2013.

[5] Daniel Zoran and Yair Weiss, "From learning models of natural image patches to whole image restoration," in *2011 Int. Conf. Comput. Vis.* Nov. 2011, pp. 479–486, IEEE.

[6] Afonso M Teodoro, Mariana SC Almeida, and Mário AT Figueiredo, "Single-frame image denoising and inpainting using gaussian mixtures.," in *ICPRAM (2)*, 2015, pp. 283–288.

[7] Antoine Houdard, Charles Bouveyron, and Julie Delon, "High-dimensional mixture models for unsupervised image denoising (hdmi)," *SIAM Journal on Imaging Sciences*, vol. 11, no. 4, pp. 2815–2846, 2018.

[8] Julie Delon and Antoine Houdard, *Gaussian Priors for Image Denoising*, chapter , pp. 125–149, Springer International Publishing, Cham, 2018.

[9] Kai Zhang, Wangmeng Zuo, and Lei Zhang, "Ffdnet: Toward a fast and flexible solution for cnn based image denoising," *IEEE Transactions on Image Processing*, 2018.