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

Antoine Houdard

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

Antoine Houdard

MAP5, Université Paris Descartes

houdard.wp.imt.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 to treat the two components separately. Finally, we provide experiments with the recent denoising method HDMI [1] and we show that our approach improves the denoising quality, particularly for residual low frequency noise.

1 Introduction

Context   In this paper, we focus on image denoising, which aims to estimate an image $\hat{u}$ 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. Some of the recent denoising methods are based on a statistical modeling of the image patches [2, 3, 4, 5, 1]. For each patch $i \in \{1, \ldots, n\}$ seen as vector of size $p = s \times s$, model (1) yields

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

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. The idea is then 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

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

1

Convenient priors for computing this estimator (4) are Gaussian priors [2] or Gaussian mixture models (GMM) [3, 5, 1]. 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 [6]. 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 [1], 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 [5, 4] also seem to suggest that removing the mean – also called the DC component – of the patches may improve the denoising quality.

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 $Y_i^c = Y_i - \bar{Y}_i \mathbf{1}_p$, where

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

is the mean of Y_i and $\mathbf{1}_p = (1, \ldots, 1) \in \mathbf{R}^p$. 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 2) and $\bar{N}_i$ (section 3) of these two problems. Then we suggest in section 4 various solutions that can enhance the denoising results of the patch-based denoising method HDMI [1], and we provide numerical experiments for these solutions.

2   Modeling the centered noise

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

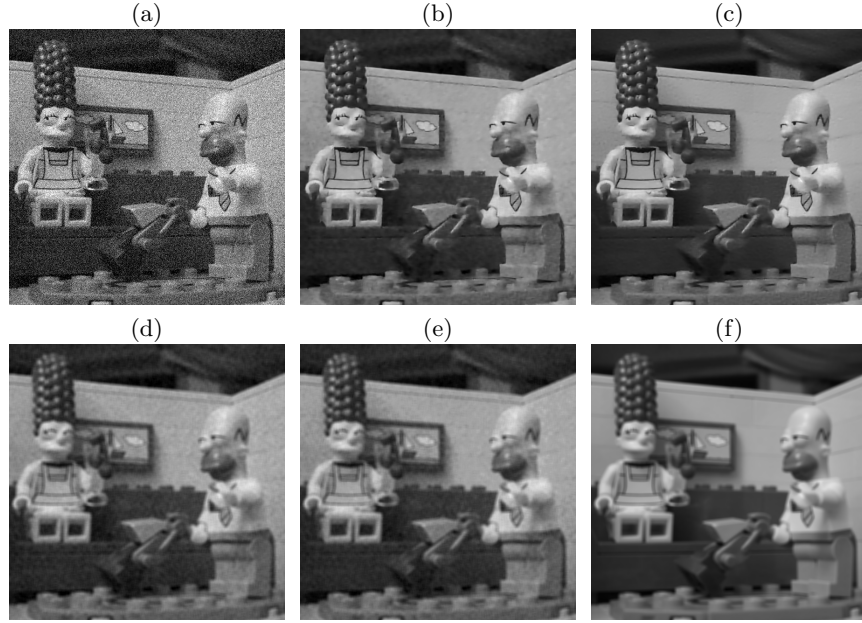$$N_i^c(j) = N_i(j) - \frac{1}{p} \sum_{k=1}^{p} N_i(k), \tag{8}$$

Figure 1: **(a)** noisy image (standard deviation 30/255). **(b)** image denoised with HDMI [1] (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).

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 $\forall k, l \in \{1, \ldots, p\}$

$$\mathbb{E}\left[N_i^c(k)N_i^c(l)\right] = \mathbb{E}\left[N_i(k)N_i(l)\right]$$

$$-\frac{1}{p}\sum_{m=1}^{p}\mathbb{E}\left[N_i(m)\left(N_i^c(l)+N_i^c(k)\right)\right]$$

$$+\frac{1}{p^2}\sum_{m=1}^{p}\sum_{n=1}^{p}\mathbb{E}\left[N_i(m)N_i^c(n)\right]$$

$$=\sigma^2\left(\delta_{kl}-\frac{2}{p}+\frac{p}{p^2}\right)$$

$$=\sigma^2\left(\delta_{kl}-\frac{1}{p}\right).$$

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 $\frac{1}{\sqrt{p}}(1,\ldots,1)$. The change of basis Q^T applied to the centered noise then yields $Q^T N_i^c \sim \mathcal{N}(0,\mathrm{diag}(\sigma^2\mathrm{I}_{p-1},0))$. Finally, centering the noise implies a dimension reduction. The total variance of the centered noise defined as

$$\mathrm{TVar}(N_i^c) = \mathbb{E}[\|N_i^c\|_2^2] \tag{12}$$

$$= \mathrm{Tr}(\Sigma_{N_i^c}) = (p-1)\sigma^2 \tag{13}$$

$$= \frac{p-1}{p}\mathrm{TVar}(N_i) \tag{14}$$

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

3 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 $\bar{N}_i \sim \mathcal{N}(0,\frac{\sigma^2}{p})$. Then, reshaping problem (6) as an image yields the new image denoising problem with additive Gaussian noise

$$\bar{Y} = \bar{X} + \bar{N}, \tag{15}$$

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

$$Z_i = W_i + M_i, \tag{16}$$

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 covariance matrix.

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

$$(\Sigma_{M_i})_{kl} = \mathbb{E}\left[\bar{N}_{i_k}, \bar{N}_{i_l}\right] = \frac{\sigma^2}{p^2}C_{kl}, \tag{17}$$

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

$$\Sigma_{M_i} = \frac{\sigma^2}{p^2}B \otimes B \tag{18}$$

where

$$B = \begin{pmatrix} s & (s-1) & \cdots & 1 \\ (s-1) & s & \ddots & \vdots \\ \vdots & \ddots & \ddots & (s-1) \\ 1 & \cdots & (s-1) & s \end{pmatrix}, \tag{19}$$

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

$$m_d = \begin{vmatrix} s & (s-1) & \cdots & (s-d+1) \\ (s-1) & s & \ddots & \vdots \\ \vdots & \ddots & \ddots & (s-1) \\ (s-d+1) & \cdots & (s-1) & s \end{vmatrix}, \tag{20}$$

Adding the first column in the last one yields

$$m_d = (2s - d + 1) \begin{vmatrix} s & (s-1) & \cdots & 1 \\ (s-1) & s & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ (s-d+1) & \cdots & (s-1) & 1 \end{vmatrix}, \qquad (21)$$

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

$$m_d = (2s - d + 1) \begin{vmatrix} 2 & (s-1) & \cdots & 1 \\ 0 & s & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & (s-1) & 1 \end{vmatrix}, \qquad (22)$$

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

$$m_d = (2s - d + 1)2^{d-2} > 0. \qquad (23)$$

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

$$L^{-1}Z_i = L^{-1}W_i + L^{-1}M_i, \qquad (24)$$

is an AWGN problem with noise variance $\sigma^2/p^2 I_p$ 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 $\widehat{W_i} = L\widehat{L^{-1}W_i}$.

4 Experiments

In this section, we take advantage of the previous modeling in order to propose various strategies for improving the denoising result of model-based patch-based denoising methods. For this testing part, we propose to use the HDMI method [1] 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 [1]).

4.1 Denoising the DC component

In section 3, 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 (24). 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 (24) 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 (16) is about p times larger than the one of the original problem (3). Finally, with this step, we obtain for each patch i an accurate estimate $\widehat{\overline{X}_i}$ of its DC component.

Figure 2: Left: noisy DC image. Middle: denoised DC image ('simpson' PSNR = 40.64 dB , 'lena' PSNR = 40.89 dB). Right: oracle DC image. Top line is the from the *simpson* image and bottom line from the *lena* image both with noise of standard deviation $\sigma = 30/255$.

4.2 Denoising the patches

In order to perform the final patch denoising, we expose hereafter two strategies. For each of these strategies, we propose numerical examples and we discuss the improvements they bring for the denoising task.
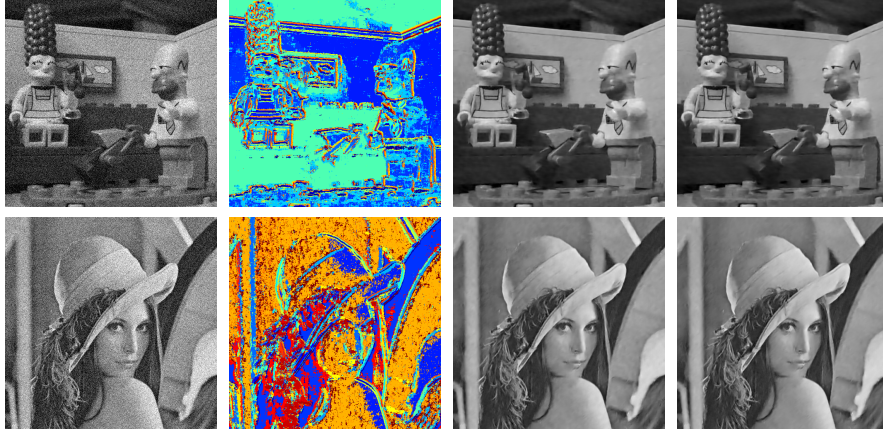
Figure 3: From left to right. Noisy images (standard deviation 30/255). HDMI resulting groups. Images denoised by HDMI (lena 31.09dB, simpson 32.46dB). Images denoised with method 4.2.1 (lena 31.19dB, simpson 32.58dB).

4.2.1 Correction of the DC component afterwards

The first idea is to run the HDMI method on the original patches y_i, yielding for each patch i an estimate $\widehat{X_i}$, and then to correct this estimate with the estimated DC component $\overline{\widehat{X_i}}$. That is, to consider the final estimate

$$\mathrm{Est}_1\left(X_i\right) = \widehat{X_i} - \overline{\widehat{X_i}}\mathbf{1}_p + \widehat{\overline{X_i}}\mathbf{1}_p. \tag{25}$$

This strategy improves the denoising quality compared to the original denoising, as illustrated in Figure 3. The low frequency noise is visually reduced and in term of PSNR, the results show an improvement of $\approx 0.1dB$ for both images. However, this does not use the modeling of the centered noise from section 2, and the grouping of the patches from the HDMI model remains the same.

4.2.2 Learn a model on the centered patches

The second strategy is to learn the model on the centered patches X^c in order to group together patches with various contrasts. To do so, we can apply the change of basis Q^T from section 2 on the problem (7). This yields

$$Q^T Y_i^c = Q^T X_i^c + Q^T N_i^c, \tag{26}$$

with the last line of this equality being $\overline{Y_i^c} = \overline{X_i^c}$ which is always true since $\overline{Y_i^c} = \overline{X_i^c} = 0$ by definition. Removing this last dimension then yields an AWGN problem in dimension $p-1$ that can be denoised with the HDMI algorithm. This gives us an estimate of the centered patch $\widehat{X_i^c}$ and therefore the final estimate

$$\mathrm{Est}_2\left(X_i\right) = \widehat{X_i^c} + \widehat{\overline{X_i}}\mathbf{1}_p. \tag{27}$$
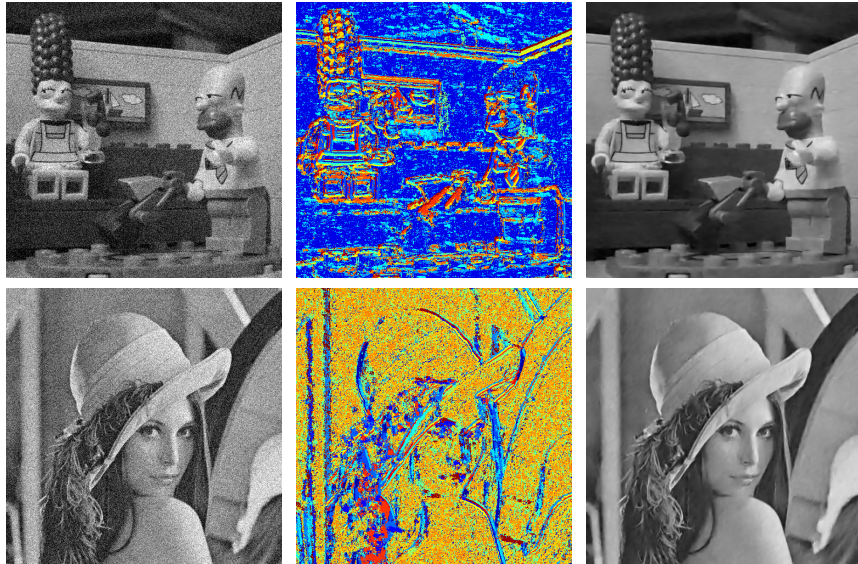
Figure 4: From left to right. Noisy images (standard deviation 30/255). Resulting groups of method 4.2.2. Denoising with method 4.2.2 (lena 31.14dB, simpson 32.58dB).

Figure 4 shows the result of this strategy. The advantage of learning the model on the centered patches is that the patches with same structure but different contrasts are grouped together (figure 4). This can make the estimate of the covariance matrix easier. However, this centered problem has a lower signal-to-noise ratio than the original problem since the centering of the patches has reduced the dynamic. Therefore, the algorithm is more likely to missclassify some fine textures. This strategy still leads to a visual and PSNR improvement but seems not to be better than the previous solution.

4.3 Discussion

Finally, we propose in table 1 the results of the two presented strategies for the three images *simpson*, *lena* and *barbara* with different level of noise. This shows that the first strategy always improve the denoising quality, with a larger improvement for images that have less textures like *simpson*. The second strategy seems also to work well in this precise case but fail to improve the result in the case of *barbara* which has a lot of fine texture. The other trend that appears is the better improvement for high variance noise than for low variance noise. This is due to the use of the same patch size for all experiments. Indeed, for a given patch size, the residual low frequency noise increases with the variance of the initial noise.

Table 1: Results in PSNR (dB) of the enhancement methods from section 4.2.1 and 4.2.2 for the images *simpson*, *lena* and *barbara* with different noise levels.

Image	σ	HDMI [1]	sec. 4.2.1	diff.	sec. 4.2.2	diff.
	10	38.89	38.93	+0.04	38.93	+0.04
Simpson	20	34.88	34.97	+0.09	34.96	+0.08
	30	32.46	32.58	+0.12	32.58	+0.12
	10	35.81	35.82	+0.01	35.81	+0.00
Lena	20	32.86	32.91	+0.05	32.89	+0.03
	30	31.09	31.19	+0.10	31.14	+0.05
	10	34.81	34.82	+0.01	34.79	-0.02
Barbara	20	31.42	31.45	+0.03	31.42	+0.00
	30	29.38	29.42	+0.04	29.36	-0.02

5 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 strategies for improving the result of existing denoising methods, especially to reduce low-frequency residual noise. The results obtained with the HDMI method on grayscale images show improvements, both visually and in term of PSNR.

In future work, we would like to study the links between this approach and multiscale frameworks. Then, in a second step, we would like to propose a discussion on the strategy to be adopted for color images.

References

[1] 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.

[2] 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.

[3] 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.

[4] 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.

[5] 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.

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