FROM PATCHES TO PIXELS IN NON-LOCAL METHODS: WEIGHTED-AVERAGE REPROJECTION

J.Salmon *

Université Paris 7- Diderot
Laboratoire de Probabilité et Modèles Aléatoires,
CNRS-UMR 7599,
175 rue du Chevaleret 75013 Paris

Y.Strozecki,

Université Paris 7- Diderot
Équipe de logique Mathématique,
FRE 3233
175 rue du Chevaleret 75013 Paris

ABSTRACT

Since their introduction in denoising, the family of non local methods, whose Non-Local Means (NL-Means) is the most famous member, has proved its ability to challenge other powerful methods such as wavelet based approaches or variational techniques. Though simple to implement and efficient in practice, the classical NL-Means suffers from ringing artifacts around edges. In this paper, we present an easy to implement and time efficient modification of the NL-means based on a better reprojection from the patches space to the original (image) pixel space. We illustrate the performance of our method on a toy example and on some classical images.

Index Terms— Image processing, Image restoration, Image edge analysis, Statistics

1. INTRODUCTION

In recent years, major progresses in image denoising has been recorded using patch-based methods. These approaches take advantage of the presence of a relatively large number of similar small sub-images, called patches, to remove noise by statistical proceedings. As far as we know, the first patch-based procedure has been introduced in image processing by Efros and Leung [1] for texture synthesis, and then extended to inpainting by Criminisi et al. [2]. The introduction of patch-based method in the context of image restoration is due to Buades et al. [3, 4]. The algorithm they developed, the Non-Local Means (NL-Means) and its variants [5] give some of the best results in denoising. Moreover two state-of-the-art methods in denoising [6, 7], though more sophisticated, are also patch oriented.

The principle of the NL-Means is the following. First, transform an image in a collection of patches. Then, estimate any patch by a weighted average on this entire collection. Eventually, reproject the information obtained from the patches to denoise the pixels themselves. The weights used are calculated thanks to the similarity between the patch to denoise and the other patches candidates. This method can be interpreted as a kernel smoothing approach in the space of patches. In statistical regression framework, the NL-Means could be seen as a Nadaraya-Watson estimator in this space. Originally [3], the kernel used is Gaussian but several authors proposed to enlarge the variety of relevant kernels [8], and here we consider the simple case of the flat kernel.

Then, since the space of patches is of larger dimension than the original image space, there is a large variety of estimators to recover pixels from patches estimators. In this paper we investigate the existing reprojection functions available to connect these two spaces. We also define a new one, that both increases PSNR and reduces ringing artifacts on standard images.

*The work of Joseph Salmon is supported by ANR Parcimonie.

2. CLASSICAL DEFINITION OF THE NL-MEANS

In this paper, we are concerned with the problem of restoration of noisy images. We assume that we are given a grayscale image I being a noisy version of an unobservable image $I^\star$. It is classical to model the noise as an additive Gaussian term:

$$I(\mathbf{x}) = I^\star(\mathbf{x}) + \varepsilon(\mathbf{x}) \,, \tag{1}$$

where $\mathbf{x} = (x, y)$ is any pixel in the image and ε is a centered Gaussian noise with known variance σ^2. From now on, we assume that the images are of size $N \times M$ and denote by Ω the set of pixels $\{1, \dots, N\} \times \{1, \dots, M\}$. Traditional neighborhood filters treat this model (1) as a problem of estimating a two-dimensional regression function. So, they apply various smoothing techniques, see e.g. [9] and the references therein.

The approach of [3] proposes to replace the spatial regression by the regression of the pixel value on the values of its neighbors. More precisely, for some odd integer $W = 2W_1+1 > 0$ and for some pixel $\mathbf{x} \in \Omega$, the patch with W^2 elements and center $\mathbf{x}$ is by definition the vector $\mathsf{P}_\mathbf{x} = (I(\mathbf{x}') : \|\mathbf{x} - \mathbf{x}'\|_\infty \leq W_1)$. Thus, for estimating the value $I^\star(\mathbf{x})$, a non-parametric regression estimation is carried out taking as covariate $\mathsf{P}_\mathbf{x}$ and as response $I(\mathbf{x})$. Under the stationarity assumption, using for instance kernel smoothing, this leads to the estimator

$$\hat{I}_{\text{NLM}}(\mathbf{x}) = \frac{\sum_{\mathbf{x}'} I(\mathbf{x}') \cdot K(\|\mathsf{P}_{\mathbf{x}} - \mathsf{P}_{\mathbf{x}'}\|/h)}{\sum_{\mathbf{x}''} K(\|\mathsf{P}_{\mathbf{x}} - \mathsf{P}_{\mathbf{x}''}\|/h)} , \qquad (2)$$

where $\mathbf{x}'$ runs in Ω, K is a kernel function and $h > 0$ is a bandwidth width. In practice, to avoid the violation of the stationarity assumption, the summation is restricted to a searching zone $\Omega_R(\mathbf{x})$, a moderately small neighborhood of $\mathbf{x}$ of size $R \times R$. Remark that substituting $\Omega_R(\mathbf{x})$ to Ω is also done to speed up the algorithm in practice as the complexity of the algorithm decreases from $O(N^2M^2W^2)$ to $O(R^2NMW^2)$.

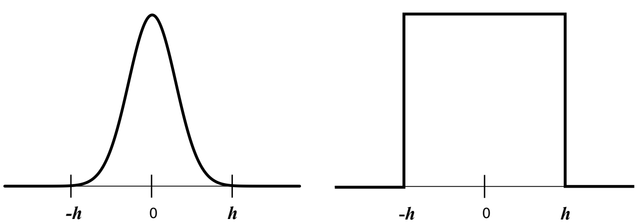
Now, let us define the weights by

$$w(\mathbf{x}, \mathbf{x}') = \frac{K(\|\mathsf{P}_{\mathbf{x}} - \mathsf{P}_{\mathbf{x}'}\|/h)}{\sum_{\mathbf{x}'' \in \Omega_R(\mathbf{x})} K(\|\mathsf{P}_{\mathbf{x}} - \mathsf{P}_{\mathbf{x}''}\|/h)} , \qquad (3)$$

and let us recast the NL-Means estimators as a weighted average over the searching zone

$$\hat{I}_{\text{NLM}}(\mathbf{x}) = \sum_{\mathbf{x}' \in \Omega_R(\mathbf{x})} w(\mathbf{x}, \mathbf{x}') \cdot I(\mathbf{x}') . \qquad (4)$$

Note that different variants of the NL-Means estimator are obtained by choosing different kernel functions K [8] and changing the norm that measures the distance between patches [10]. Default choices are the Gaussian kernel $K(x) = e^{-x^2/2}$ and the Euclidean norm but other alternatives using PCA on patches were also proposed [11].



(a) Gaussian Kernel (b) Flat Kernel

Fig. 1. Two types of Kernel considered and their bandwidth h

We want to emphasize that the choice of the kernel should not drastically alter the performance of the method. But, there are several good reasons for using a kernel with a compact support rather than a Gaussian kernel.

Firstly, the NL-Means estimator encounters a problem when the searching zone is too large [12]. As many coefficients $w(\mathbf{x}, \mathbf{x}')$ are almost zero without being null, they create perturbations that lower the impact of "the good candidates" i.e. those with largest $w(\mathbf{x}, \mathbf{x}')$. This phenomenon limits the performance of increasing R. Thus, hard thresholding the small coefficient with a compact support kernel leads to improvements for large R.

Secondly, except for regular functions, no statistical theoretical result favors any kernel, so we use the flat kernel $K(x) = \mathbb{1}_{[-1,1]}(x)$ to minimize the computation time.

Thirdly, for the Flat kernel, using an approach based on χ^2 behavior like in [4], it is possible to give a choice of the bandwidth h based on statistical argument. h is the most crucial tuning parameter in NL-Means methods, so it might be useful to have a good idea of its scale. To this purpose, one can test whether two patches $\mathsf{P}_{\mathbf{x}}$ and $\mathsf{P}'_{\mathbf{x}}$ are equal under the assumption of Gaussian noise (assuming that the patches can be considered independent). If two patches are equal, the renormalized square L^2-norm between them, $D(\mathbf{x}, \mathbf{x}') = \|\mathsf{P}'_{\mathbf{x}} - \mathsf{P}'_{\mathbf{x}}\|^2/2\sigma^2$ is $\chi^2(W^2)$ distributed. Thus, we can reject with probability α the hypothesis that the patches are equal if $D(\mathbf{x}, \mathbf{x}') > q_{1-\alpha}^{W^2}$, where $q_{1-\alpha}^{W^2}$ is the quantile of the $\chi^2(W^2)$ distribution. In practice, we used $\alpha = 0.01$ for natural images.

3. REPROJECTION FROM THE PATCHES SPACE

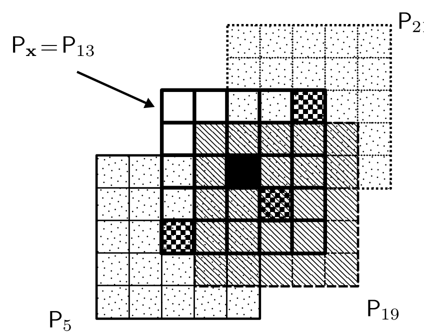
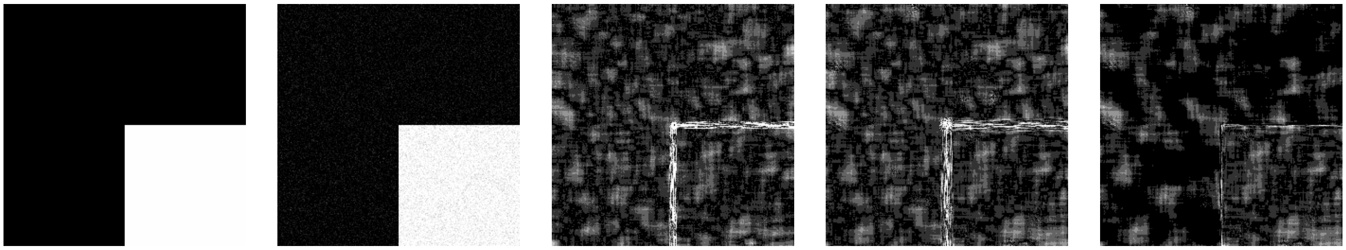


Fig. 2. Overlapping patches and their order for $W = 5$. The centers are highlighted.

Once patches have been introduced, one can figure out that the NL-Means procedure estimates every patch in the image before using this information to denoise pixelwise. Remind that the patches are overlapping and that every pixel belongs to W^2 patches. So we can rewrite the NL-Means estimator in two part. On the one hand, one defines a nonparametric estimator for every patch in the image. To this aim, one uses a weighted average with the same weights as before, meaning that an estimator $\widehat{\mathsf{P}_{\mathbf{x}}} = \sum_{\mathbf{x}' \in \Omega_R(\mathbf{x})} w(\mathbf{x}, \mathbf{x}') \cdot \mathsf{P}_{\mathbf{x}'}$ is find for any patch. On the other hand, one needs to define a reprojection function Ψ to get an estimate $\hat{I}(\mathbf{x})$ in the pixel domain. Ψ maps the space of W^2 patches to which $\mathbf{x}$ belongs into the pixel space. We name $\mathsf{P}_1, \ldots, \mathsf{P}_{W^2}$ those patches enumerated in column major order from the top left corner (cf Fig.2). The patch centered on $\mathbf{x}$ is $\mathsf{P}_{\mathbf{x}} = \mathsf{P}_{\frac{W^2+1}{2}}$ and the pixel $\mathbf{x}$ can be expressed for $i = 1, \ldots, W^2$ as $\mathbf{x} = \mathsf{P}_i(W^2 - i + 1)$. Formally we can now define the pixel estimate

$$\hat{I}(\mathbf{x}) = \Psi\left(\widehat{\mathsf{P}_1}, \ldots, \widehat{\mathsf{P}_{W^2}}\right) . \qquad (5)$$

It is interesting to note that as we use information on a bigger space (of dimension R^{W^2}) than the original one, there are



(a) (b) $PSNR = 22,08$ (c) $PSNR = 44,99$ (d) $PSNR = 45,72$ (e) $PSNR = 48,79$

Fig. 3. (a) original image $I^\star$, (b) noisy image I with $\sigma = 20$ ($PSNR = 22.08$), (c) absolute difference (the whiter the bigger) between I and the denoised image with the original NL-Means, (d) absolute difference for NL-Means with flat kernel and Av-reprojection and (e) with Wav-reprojection ($W = 9, R = 21$ for all cases).

many alternatives to reproject into the pixel domain. Moreover, any function Ψ define a variation of the NL-Means.

The classical method [3] proposes to only use the center of the denoised patch (there is a center as W is artificially constrained to be odd), called here $\hat{I}_{center}(\mathbf{x})$

$$\hat{I}_{center}(\mathbf{x}) = \widehat{\mathsf{P}_{\frac{W^2+1}{2}}}\left(\frac{W^2+1}{2}\right) . \tag{6}$$

It is obvious that such a method looses an important part of the information provided by the patches. This in particular leads to well known ringing artifacts of the NL-Means (cf: Fig.(3)-c for a typical example).

A first improvement [3, 13] is to average the different estimators of the pixel $\mathbf{x}$ based on the collection of W^2 patches $\mathbf{x}$ belongs to. Formally, we define the Av-reprojection estimator by

$$\hat{I}_{Av}(\mathbf{x}) = \frac{1}{W^2}\sum_{i=1}^{W^2}\widehat{\mathsf{P}}_i\left(W^2 - i + 1\right) . \tag{7}$$

It leads to a significant increase in PSNR in practice (cf. Tab.1) but does not remove the ringing artifacts (cf. Fig.3-d).

Let us now introduced a more refined method designed to reduce this problem. Sometimes, a pixel may belong to a patch with many repetitions, although the patch centered on this pixel may only have a few similar patches. To see this, consider the piece-wise constant image "Corner" in Fig.3-a, corrupted with a Gaussian white noise (Fig.3-b). This image is a good toy model for dealing with edges and corners. Using the original NL-Means with patches of width $W = 9$ and searching zone of width $R = 21$, we can observe a blurry phenomenon of width $W - 1$ centered on the edge (see Fig.3-c where we show the absolute difference between the original image and the denoised one). Note that this appears for any global or local choice of h.

Now, we want to build a better weighted average of the W^2 estimators of $I(\mathbf{x})$ than a simple uniform average. Define the pixel weighted average estimator (Wav-reprojection) as

$$\hat{I}_{Wav}(\mathbf{x}) = \sum_{i=1}^{W^2}\beta_i\widehat{\mathsf{P}}_i\left(W^2 - i + 1\right) . \tag{8}$$

Assume the patches corresponding selected that lead to those estimators satisfy the test "they are independent noisy observations of the same original patch". This means that the patches selected can be considered of zero bias, so is their combination. Thus, to minimize the expectation of the quadratic error of our weighted average estimator, we only need to minimize the variance, under the constraint that $\sum_{i=1}^{W^2}\beta_i = 1$. Using a Lagrangian, this leads for all $i = 1, \ldots, W^2$ to choose

$$\beta_i = \frac{\left[\mathrm{Var}(\widehat{\mathsf{P}}_i\left(W^2 - i + 1\right))\right]^{-1}}{\sum_{i=1}^{W^2}\left[\mathrm{Var}(\widehat{\mathsf{P}}_i\left(W^2 - i + 1\right))\right]^{-1}}.$$

For the flat kernel case, the weight β_i is proportional to the number of selected patches used to denoise P_i.

Reprojections considering every patch a pixel belongs allows to denoise efficiently with a smaller searching zone, as the influence of each pixel is increased by moving the patches positions. It reduces the computation time as R can be chosen smaller. Moreover, W is no longer constraint to odd values.

Let us return to the " Corner" image (size: 256×256). Using our reprojection method also with $R = 21, W = 9$, we have almost no ringing artifact (cf Fig 3-e). Expressed in term of PSNR, improvements are also important: for the classical NL-Means $PSNR = 45.51$, whereas for the Av-reprojection $PSNR = 48,65$ for the Wav-reprojection $PSNR = 49,56$. We used $h^2 = 2\sigma^2 q_{0.99}^{W^2}$ for the last two methods and for the original NL-Means we have chosen $h = 1.35\sigma$.

On natural images, the mean reprojection is already quite an improvement. On all images tested, our reprojection procedure is of the same order of performance (up to 0.1dB loss in PSNR) except for Bridge, Cameraman, Fingerprint and Flinstones, where the gain is of 0.5dB in our favor. The improvement is mostly around edges and thus barely visible through PSNR, but visually, the ringing artifacts have almost disappeared. We present the example of Cameran Fig. 4 (size: 256×256), with $\sigma = 20$ and $W = 9, R = 11$. We get $PSNR = 28,18$ for the original NL-Means ($h = 4.5\sigma$, and Gaussian kernel). For the flat kernel and center-reprojection $PSNR = 27,70$, for Av-reprojection $PSNR = 28.68$ and for Wav-reprojection $PSNR = 29.13$ (all with $h^2 = 2\sigma^2 q_{0.99}^{W^2}$).

Matlab code is available on-line on the author's webpage.

(a) $PSNR = 27,70$ (b) $PSNR = 28.19$ (c) $PSNR = 28,68$ (d) $PSNR = 29.13$

Fig. 4. Denoising with $R = 11, W = 9$, $\sigma = 20$ ($PSNR = 22.14$) (a) original NL-Means $h = 4.5\sigma$,Gaussian Kernel, (b) Center-reprojection, (c) Av-reprojection and (d) Wav-reprojection : the last three with $h^2 = 2\sigma^2 q_{0.99}^{W^2}$

Image	Flat	Gaussian	Flat Av.	Flat Wav.
Barbara	28.67	29.00	29.99	30.15
Boats	28.47	28.80	29.47	29.53
Bridge	24.33	25.68	25.42	25.78
Camera.	27.62	28.17	28.68	29.14
Couple	28.14	28.52	29.17	29.28
Fingerp.	25.13	25.87	26.66	27.21
Flinst.	26.04	26.45	27.31	27.93
Hill	26.39	27.45	27.62	27.77
House	31.15	30.98	32.36	32.36
Lena	31.11	31.18	32.08	32.13
Man	28.52	29.13	29.60	29.61
Peppers	28.89	29.06	30.25	30.38

Table 1. PSNR of NL-Means variants ($\sigma = 20, W = 9, R = 9$), in order: original NL-Means algorithm with Flat and Gaussian kernel for center-reprojection, then Av-reprojection and Wav-reprojection with flat kernel. $h = 4.5\sigma$ for the Gaussian case and $h^2 = 2\sigma^2 q_{0.99}^{W^2}$ for the three other methods.

4. CONCLUSION

We have presented a new way to handle the information obtained in considering patch-based approaches in the context of denoising. Ongoing work is to automatically select the h without assuming σ is known, and to speed up the NL-Means as in [5], computing the weights only on a sub-sampled grid. This decreases the performance in general, but with our Wav-reprojection, the way we handle the information of the patches allows to do so without degrading performances as much as with the center-reprojection.

5. REFERENCES

[1] A. A. Efros and T.K. Leung, "Texture synthesis by non-parametric sampling," in *ICCV*, 1999, pp. 1033–1038.

[2] A. Criminisi, P. Pérez, and K. Toyama, "Object removal by exemplar-based inpainting," in *CVPR*, 2003, vol. 2, pp. 721–728.

[3] A. Buades, B. Coll, and J-M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 490–530, 2005.

[4] A. Buades, B. Coll, and J-M. Morel, "Nonlocal image and movie denoising," *Int. J. Comput. Vision*, vol. 76, no. 2, pp. 123–139, 2008.

[5] Ch. Kervrann and J. Boulanger, "Optimal spatial adaptation for patch-based image denoising," *IEEE Trans. Image Process.*, vol. 15, no. 10, 2006.

[6] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, 2007.

[7] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-local sparse models for image restoration," *ICCV*, 2009.

[8] B. Goossens, Q. Luong, A. Pizurica, and W. Philips, "An improved non-local denoising algorithm," in *LNLA*, 2008.

[9] J. Polzehl and V. Spokoiny, "Image denoising: pointwise adaptive approach," *Ann. Statist.*, vol. 31, no. 1, pp. 30–57, 2003.

[10] D. van de Ville and M. Kocher, "SURE-based Non-Local Means," *IEEE Signal Process. Lett.*, vol. 16, pp. 973–976, 2009.

[11] N. Azzabou, N. Paragios, and F. Guichard, "Image denoising based on adapted dictionary computation," in *ICIP*, 2007, pp. 109–112.

[12] J. Salmon, "On two parameters for denoising with Non-Local Means," *IEEE Signal Process. Lett.*, vol. 17, pp. 269–272, 2010.

[13] Ch. Kervrann, J. Boulanger, and P. Coupé, "Bayesian non-local means filter, image redundancy and adaptive dictionaries for noise removal," in *SSVM*, 2007, vol. 4485, pp. 520–532.