IMAGE DENOISING BASED ON ADAPTED DICTIONARY COMPUTATION

Noura Azzabou (1,2) *Nikos Paragios* (1) *Frédéric Guichard* (2)

(1) Ecole Centrale de Paris, Laboratoire MAS , France
(2) DxOLabs, 3 Rue Nationale Boulogne Billancourt, France

ABSTRACT

This paper introduces a new denoising technique that consists in recovering the image using a filtering function adapted to the image content. The definition of such a function relies on the computation of similarity between pixels of a given neighborhood. Our contribution consists in the definition of a new similarity criterion which is more robust to noise. This measure is computed from a dictionary that is adapted to image content. The projection of the image content to this subspace are used then to define a metric between a pixel and the neighborhood ones. Very promising experimental results show the potential of our approach.

***Index Terms*—** Neighborhood filtering, Principle Component Analysis, similarity measure, Non Local Means

1. INTRODUCTION

Natural image denoising is still a challenging and open problem in image processing in particular when dealing with rich content like texture. State of the art techniques in image enhancement refer to global methods and local ones. Global approaches represent images through a set of invertible transformations [1, 2] or a specific designed dictionary. Noise reduction is achieved through the modification of the coefficients with limited importance in the reconstruction process. Despite their performance, these techniques are strongly dependent on the choice of decomposition basis. In the most general case, finding an optimal one for all natural images is not trivial.

Approaches based on total variation minimization [3], or partial differential equations [4] are efficient tools in image regularization field. Nevertheless, these approaches are based on a local smoothness hypothesis and quite often fail to preserve texture. In order to address this problem, separating structure from texture is the most prominent technique to deal with such limitation and has gained significant attention in the past years [5]. However, these methods fail to separate noise from texture because like noise, texture is an oscillatory pattern. Further more, these models are complex and rely on data fidelity term that cannot be computed directly and can be only approximated.

Filtering approaches that do not make specific assumption on noise model rely on a weighted mean based estimation of the noise free image. The weights definition is dependent on the similarity between pixels. This similarity is defined according to the spatial and photometric distance between pixel in the case of bilateral filter [6], or a distance between local image patches [7, 8, 9, 10]. One has to point out that these algorithms are based on a distance computation between noisy observations which reduces their robustness.

In the present paper, we propose a novel filtering method that exploits a new definition of similarity between pixels. A pixel will be represented with a set of coefficients that corresponds to the projection of its local neighborhood on an appropriate dictionary adapted to the image content. To this end we build a compact dictionary of local image content using principal component analysis.

The paper is organized in the following fashion: in section 2 we present our denoising technique, while the next section will be devoted to the validation of the method and the comparison results. Finally, we will conclude in section 4.

2. IMAGE DENOISING BASED ON NEIGHBORHOOD FILTERING

Image denoising based on neighborhood filter is a quite standard technique. It refers to restoring a pixel taking a weighted average of the neighboring pixels. Such a filtering can be expressed as:

$$\hat{U}(x) = \frac{1}{N_h(x)} \iint_{R_x} h(x,y)U(y)dy \qquad (1)$$

where U is the noisy image, $N_h(x)$ is the normalization constant defined as $N_h(x) = \iint_{R_x} h(x,y)dy$, $\hat{U}$ is the reconstructed image and R_x is a local neighborhood associated to x. The filtering function h is a monotonically decreasing and depends on the photometric distance between the pixel x and its neighbor. In fact, samples that have similar content to x will have a strong contribution in the gray level estimation.
The main focus of our work, is to propose an adaptive filtering function h using a more appropriate distance to account for the image content while being more robust to noise.

2.1. Image Dictionary Computation

To evaluate a similarity between pixels one has to define a set of features (e.g. image intensities), and an appropriate metric in the space of these features. The most common feature space is the image itself, while the L-1 and the L-2 norms have been frequently considered. In [7], authors suggest as a feature vector the intensity values within a local patch around the pixel. The corresponding definition of the filtering function is

$$h(x, y) = \exp - \frac{\|\mathbf{u}_x - \mathbf{u}_y\|_{L_2}}{2\sigma^2}$$

where $\mathbf{u}_x$ is a vector of dimension p and related to a definition of a local neighborhood $\mathcal{N}_x$ (of size p) and expressed as $\mathbf{u}_x = \{U(y) \text{ such that } y \in \mathcal{N}_x\}$. σ is a parameter that is fixed according to the noise level.

Such a definition of a set of features is not able to preserve image texture. In fact, this pixel characterization is very simple and thus not robust to noise. Furthermore, it does not incorporate image structure at local neighborhood. To overcome this limitation, one needs a feature vector definition that better describes the data structure on one hand and is not sensitive to noise on the other hand.

Image decomposition in subspaces, like wavelets, Fourier, etc. reduce the dimensionality of the problem and often associate noise to the least important components. The central idea of our approach is to decompose image content into a dynamic dictionary, or a subspace and then use this subspace (in particular the projection of the image patch to the base) to define a metric. One can claim that such an image representation will remove certain amount of the noise, that is critical when determining similarities between observations of neighborhood pixels.

Several image filters can be used to perform local features extraction of the image. For simplicity, we consider principle component analysis (PCA) which is a powerful tool that provides a compact representation of the most prominent element in the image. Such a representation can be considered as a set of filters adapted to image content. This technique performs image decomposition on an orthonormal basis where each component is relative to a group of image features. Thus, the most prominent vectors allows the extraction of smooth image content as well as texture and edges. The remaining components which are related to a small variation in the image emphasize the noise component. In this paper, we will take advantage of the PCA decomposition to compute a new feature vector. In the following we will explain in details how we perform the PCA to obtain such a decomposition.

If we consider the set of vectors $\mathbf{u}_x$ that correspond to the set of observations, the new orthonormal basis obtained using a PCA corresponds to the eigenvectors of the correlation matrix defined as

$$Cr(i, j) = corr(V_i, V_j)$$

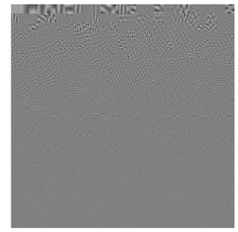


Fig. 1. Eigenvectors obtained through PCA decomposition of barbara image corrupted by Gaussian noise ($\sigma_n = 10$) (patch size 25×25)

(a) (b) (c)

Fig. 2. (a) Projection of the observations set on the first eigenvector of barbara image (b) Projection of the observation set on the 17^{th} eigenvector (c) Projection of the observations set on the Last eigenvector (49^{th})

$$V_i = [\mathbf{u}_1(i), \mathbf{u}_2(i) \dots \mathbf{u}_N(i)] \quad 1 \leq i \leq p$$

The eigenvalue associated to a given eigenvector can be interpreted as the variance of the projection of the observation set on this vector. In the remainder of the paper, we will note $(e_1, e_2, \dots e_p)$ the new orthonormal basis of the eigenvectors associated to the eigenvalues ($\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$) of the matrix Cr. The eigenvectors can be interpreted as a dictionary where the elements correspond to image characteristics (cf figure [Fig.(1)]). The projection of the set of observations $(\mathbf{u}_x)_{1 \leq x \leq N}$ on each vector on the new basis is equivalent to a correlation between an image patch around x and an element of the dictionary learned using PCA. The projection on some principle components is illustrated in figure [Fig.(2-a),Fig.(2-b),Fig.(2-c)]. We can notice that the first principle component acts as a low pass filter that captures all smooth contents of the image. The projection on other components extracts the texture and small details in the image. It's important also to point out that vectors associated to small eigenvalues that are close to the noise variance emphasize the random component introduced by noise.

2.2. The image filtering

We define in this section the new filtering function using the projection of the vectors $(\mathbf{u}_x)_{x=1 \dots N}$ on the PCA basis. To this end, we compute the distance between observations in a sub-space of the space induced by the eigenvector basis. We specify that after projection, the noise variance (σ_n^2) remains constant while the variance of observation changes according to projection direction. In our algorithm, we will consider only the vector that consists of observations that have more important variance than the noise one. Taking into consideration that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p$, we will restrain our selves to the subspace (noted E_q) generated by the first q vector $(e_l)_{1 \leq l \leq q}$ such that $\lambda_l \geq \sigma_n$. In fact, as shown in figure [Fig.(2-c)], vectors with small eigenvalues correspond to the noisy component of the image and do not contain information about image structure. Under these considerations, we will measure the similarity between pixels in the new subspace (E_q) of dimension $q \leq p$. Thus, if we note $\mathbf{v}_x$ the projected vector of $\mathbf{u}_x$ on E_q (inner product between the eigen vector and the image patch), the new definition of h is

$$h(x, y) = \exp - \frac{\|\mathbf{v}_x - \mathbf{v}_y\|_{L_2}}{2\sigma^2}$$

In the following section, we propose an experimental validation of our approach and we give a comparison with state of the art methods.

3. EXPERIMENTAL RESULTS

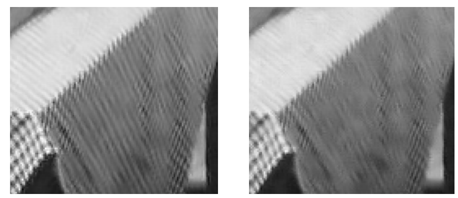
Toward objective validation of our method, we have used natural images corrupted by an additive white Gaussian noise (σ_n=20 and σ_n=10) as well as digital images corrupted by real camera noise. We compared our approach to well known filtering techniques such as the bilateral filter [6], the Non Local Mean approach [7], the total variation [3] and the anisotropic diffusion [4] using an edge stopping function of the type $(1 + |\nabla I|^2 / K^2)^{-1}$. The parameters of the considered methods were tuned to get a good balance between texture preservation and noise suppression as well as the highest possible PSNR value. We specify that for our method we considered the following parameters for all experiments $p = 5 \times 5$, $\sigma = 15$ for σ_n=20 ($\sigma = 9$ for σ_n=10) and a rectangular window of size 7×7 for R_x.

As far as subjective criteria are concerned, we adopt the whole aspect of the image in term of noise suppression and small detail preservation. Visual comparison results of denoising [Fig.(3),Fig.(4)] show that our denoising method outperforms the other ones. For comparison, we also considered the "noise image" which is the difference between the noisy image and the restored one. Figure [Fig.(5)] shows that our "noise image" does not contain structures and details contrarily to the other ones. In [Fig.4] a zoom on a textured region in barbara image as well as the result obtained with our method and the NL-means algorithm are shown. It is clear that our method ensures a better reconstruction of texture than the NLmean algorithm.

When considering real digital camera noise with an unknown noise model, we conclude that, as shown figure [Fig.(6)], better reconstruction is obtained when similarity between pixels is computed using the new vector basis introduced by the PCA. The NLmean algorithm deteriorates the skin texture in the image.

(a) (b) (c)

Fig. 3. (a) Noisy Image (b) Image filtered with our method , (c) Image filtering using NLmean



(a) (b)

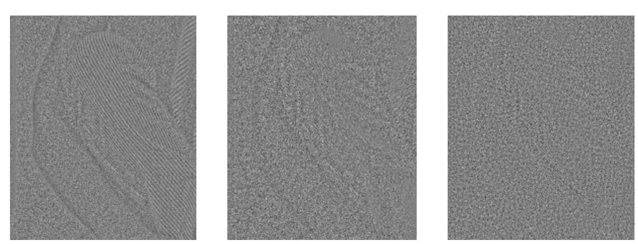
Fig. 4. (a) Zoom on a resulting image obtained with our method (d) Zoom on a resulting image obtained with NLmean

As far as a quantitative validation is concerned we used the Peak Signal to Noise Ratio criterion defined by

$$PSNR = 10log_{10}\frac{255^2}{MSE} \quad MSE = \frac{1}{\|\Omega\|}\sum_{x \in \Omega}(U(x) - \hat{U}(x))^2$$

Table (1) confirms the subjective results and show that our method has better performances then the other state of the art methods. In addition, we have to point out that our approach is faster then the classical NLmean because we use only a smaller number of features then the NLmean to compute the distance between patches.

In our validation experiments we studied also the impact of the choice of the parameter q which corresponds to the number of the principle components retained while computing pixel similarity. The curve in figure [Fig.(7)] shows that



(a) (b) (c)

Fig. 5. Zoom on the "noise image" obtained with the different methods: (a) bilateral filter (b) Non local mean (c) Our approach

	barbara		Boat		FingerPrint		House		Lena		baboon	
σ_n	20	10	20	10	20	10	20	10	20	10	20	10
TV	26.18	29.60	27.72	32.17	26.08	30.65	28.43	33.86	28.45	33.83	25.18	-
AD	26.45	30.85	28.06	31.92	24.81	29.02	29.41	33.72	29.27	33.36	23.68	-
Bilateral	26.75	31.05	27.82	31.52	24.12	28.81	29.18	33.40	29.28	33.01	24.95	29.31
NLmean	28.78	32.96	28.92	32.49	26.45	30.60	30.86	34.66	31.13	34.65	25.18	29.54
ACP+Convol	**29.99**	**33.33**	**29.72**	**32.61**	**27.14**	**30.67**	**32.04**	**34.87**	**31.97**	**34.84**	**26.11**	**29.77**

Table 1. PSNR values for denoised image corrupted by additive Gaussian noise (The PSNR of the original noisy image is 22.15 for σ_n=20 and 28.11 for σ_n=10)

(a)

(b)

(c)

Fig. 6. Results for real digital camera noise (a) Noisy image (b) Image restored using our algorithm (c) Image restored using NLmean

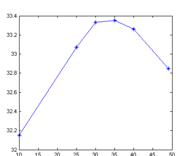


Fig. 7. Evolution of the SNR corresponding to noisy barbara image ($\sigma_n = 10$) according to the parameter q where $1 \leq q \leq 49$

the performance of denoising depends on the number of components involved in feature vector computation. One can claim that considering few principal directions is not sufficient to incorporate local image information in pixel characteristics vector. On the other hand, considering an important number of projection direction will introduce noise content on the observations. Consequently, the accuracy of the similarity computation will be decreased. The optimal value was obtained for $q = 35$ which corresponds to the eigenvalue λ_q that verifies $\lambda_l \geq \sigma_n$ for $l \leq q$.

4. CONCLUSION

In this paper we have presented a novel, simple, efficient and robust approach toward image denoising. Our main contribution consists of defining a similarity metric between image observations where noise has been eliminated leading to a more appropriate selection of pixels and weights contributing to the reconstruction process. To this end, an adaptive image basis is built (local content), and the projection of the observation space to this basis is used to determine content similarities. Experimental results show the potential of our method mainly in terms of texture and small detail preserving. A future direction is to improve the algorithm is to adapt the spatial bandwidth of the kernel to image content. In other words, information on texture at local scale must be taken into consideration for better reconstruction.

5. REFERENCES

[1] S. Mallat, “A theory for multiscale signal decomposition: The wavelet representation,” *IEEE Trans on Pattern and Machine Intelligence*, pp. 674–693, 1989.

[2] E. Le Pennec and S. Mallat, “Sparse geometric image representations with bandelets,” *IEEE Trans on Image Processing*, pp. 423– 438, 2005.

[3] L. Rudin, S. Osher, and E. Fatemi, “Nonlinear Total Variation Based Noise Removal,” *Physica D*, vol. 60, pp. 259–268, 1992.

[4] P. Perona and J. Malik, “Scale space and edge detection using anisotropic diffusion,” *IEEE Trans on Pattern and Machine Intelligence*, vol. 12, pp. 629–639, 1990.

[5] L. Vese and S. Osher, “Modeling Textures with Total Variation Minimization and Oscillating Patterns in Image Processing,” *Journal of Scientific Computing*, vol. 19, pp. 553–572, 2003.

[6] C. Tomasi and R. Manduchi, “Bilateral filtering for gray and color images,” in *ICCV*, 1998, pp. 839–846.

[7] A. Buades, B. Coll, and J-M. Morel, “A non-local algorithm for image denoising,” in *CVPR*, 2005, pp. 60–65.

[8] N. Azzabou, N. Paragios, and F. Guichard, “Random walks, constrained multiple hypotheses testing and image enhancement,” in *ECCV*, 2006, pp. 379–390.

[9] Ch. Kervrann and J. Boulanger, “Unsupervised patch-based image regularization and representation,” in *ECCV*, 2006, pp. 555–567.

[10] N. Azzabou, N. Paragios, F; Guichard, and F. Cao, “Variable bandwidth image denoising using image-based noise models,” in *CVPR*, 2007, pp. –.