Combination of closest space and closest structure to ameliorate non-local means method

Quoc Bao DO, Azeddine BEGHDADI, Marie LUONG
L2TI, Université Paris 13
Villetaneuse, France

***Abstract*— Recently non-local means (NLM) has been known to be one of the most attractive denoising algorithms. It alters each pixel by a weighted average of pixels in the image. The weights express the level of similarity between two small patches defined for two involved pixels. There are many propositions to ameliorate the performance of this method. One of branches is to seek the whole image the most similar patches for a given one. In this paper, we investigate this approach and show that it is suitable for only highly textured images. Moreover, we show that combination of this approach and the original NLM yields better result for all image types.**

Keywords- denoising, filtering, image restoration, kNN

I. INTRODUCTION

Image denoising is one of the classical image pre-processing issues. Many methods share the same basic idea: denoising each pixel is carried out by averaging all pixels which are similar to it. These methods are based on the observation that any image often contains self-similarity and some spatial redundancy. If the noise is considered as an independent and identically distributed (i.i.d.) random signal, it could be smoothed out by averaging similar pixels. The main difference between these methods is the way to determine similar pixels. For example, in [1], Gaussian filter simply considers neighborhoods as similar pixels. In [2], Yaroslavsky proposes another approach which incorporates the pixel intensity similarity in the design of the kernel. It has been shown that this approach outperforms the classical one. However, Yaroslavsky's filter is based on a pixel-based similarity which is not effective enough to determine similar pixels. To cope with this limitation, non-local means filter [3-5] uses block-based similarity approach. Two pixels are considered similar if two small blocks around them are similar. To deal with computational burden, instead of searching in the entire image as in Yaroslavsky's filter, the authors restrain search-zone in a small window around the target-patch. However, this limited search-zone makes the method no longer non-local in strict sense. In the literature, there are several propositions [6-9] to ameliorate this weakness. Their search-zone is entire image and each one proposes a different scheme to fast prune useless patches. For example, in [6] and [7], the authors choose the average intensity and the local orientation as criteria to select the best candidates. In [8], Orchard *et al.* exploit singular value decomposition (SVD) to project blocks into several orthonormal subspaces and use only the most significant ones to eliminate dissimilar blocks. In [9], Liu and Freeman propose to use a k-nearest neighbor (kNN) algorithm to quickly seek the most similar patches for a given one. From now, we call these methods "closest structure" in the sense that they take into account only similar blocks for a given one and we call the original NLM method "closest space" in the sense that it uses only closest blocks in a small window.

In this paper, the performance of the closest structure approach is investigated. We will show that it is only suitable for highly textured image. Since the closest space and the closest structure methods are complementary, a combination of these methods is also proposed to ameliorate the denoising performance.

The paper is organized as follows. A review of original NLM method and closest structure ones are presented in section II and section III, respectively. The performance analysis of the closest structure approach is reported in section IV. Our proposed method and the experimental results are shown in section V. Section VI is devoted to the conclusions.

II. NON-LOCAL MEANS METHOD

First, let us define some notations: I_p, I_q are intensities of pixel p and q respectively; f_p, f_q are two small patches defined for pixel p and q respectively; $\|a\|_2^2$ is the square of Euclidean norm of a. In NLM method, the estimated value for pixel p is written as follows:

$$I_p^{NLM} = \sum_{q \in \Omega} w(p,q) I_q \tag{1}$$

where $w(p,q)$ is weight which expresses the amount of similarity between patches f_p and f_q:

$$w(p,q) = \frac{1}{Z(p)} \exp\left(\frac{-\|f_p - f_q\|_2^2}{h} \right) \tag{2}$$

where $Z(p) = \sum_{q \in \Omega} w(p,q)$ is a normalizing factor, Ω is the whole image and h is the filter parameter that controls the decay of the exponential expression. A very large h will oversmooth the image whereas noise still remains if h is very small. As can be

978-1-4244-9915-1/11/$26.00 ©2011 IEEE

seen, the above estimation is required for all pixel pairs in the image. Therefore, if image size is MxN, patch size is dxd, the computational complexity of the algorithm is $O(M^2N^2d^2)$. To cope with this computational burden, Buades *et al.* propose to restrain the search-zone in a small window Ω_s of size sxs around pixel p. In this case, the computational complexity is reduced to $O(MNs^2d^2)$. If sxs is much smaller than MxN, the computational time is significantly reduced. However, this proposition makes the method no longer non-local in the strict sense. Consequently, there are many suitable candidates which are not taken into account. For example, as can be seen in Fig. 1, due to small search-zone (in red square), the patch b_2 which is very similar to the target-patch b is not taken whereas the patch b_1 which is completely different from b is used. We can see easily that there is a trade-off between the computational time and good-candidate selection. If search-zone is small, the time saving is considerable but many suitable patches could not be used. In the other hand, if search-zone is large, there are more chances to select suitable patches but at the cost of a significant increase of computational burden. In the literature, there are many approaches [6-9] which deal with this trade-off. Indeed, their search-zone is always the whole image but each one proposes a different algorithm to fast find out the most adequate patches b_i for each target-patch b. Hence only these suitable patches are used to restore the central pixel of the target-patch b. In the next section, we do a short review these methods.

Figure 1. Weakness of limited search-zone in NLM method: for target-patch b, due to small search-zone, the patch b_2 which is very similar to b is not taken whereas the patch b_1 which is completely different from b is used.

III. CLOSEST STRUCTURE METHODS

In [6], Mahmoudi and Sapiro propose to prune dissimilar blocks by using the average intensity and average gradient orientation. Firstly, they pick only patches with average intensity closest to that of the target-one. This premature list is further reduced by excluding patches whose average gradient vectors differs from that of the target-patch an angle that exceeds a threshold. In brief, the weight $w(p,q)$ of this method is defined as follows:

$$w(p,q)=\begin{cases}\dfrac{1}{Z(q)}\exp\left(\dfrac{\left\|f_p-f_q\right\|_2^2}{h}\right), & if \quad \left[\left\|\nabla f_p\right\|<\sigma_\nabla \; or \; \left\|\nabla f_q\right\|<\sigma_\nabla \; or \; \left|\theta(p,q)\right|<\sigma_\theta\right] \; and \left(\eta_1<\dfrac{\overline{f_p}}{\overline{f_q}}<\eta_2\right)\\ 0, & otherwise\end{cases} \tag{3}$$

where σ_∇, σ_θ, η_1 and η_2 are thresholds; $\overline{f}$ and $\left\|\nabla f\right\|$ stand for the average and norm of gradient of patch f, respectively; $\theta(p,q)$ is the angle between two average gradient vectors of block f_p and f_q:

$$\theta(p,q)=\angle\left(\overline{\nabla f_p},\overline{\nabla f_q}\right) \tag{4}$$

where average gradient vector of block f is define as follows:

$$\overline{\nabla f}=\left(\overline{f^x},\overline{f^y}\right) \tag{5}$$

($\overline{f^x}$ and $\overline{f^y}$ are the average horizontal and vertical derivatives of block f).

In [7], Bilcu and Vehvilainen use the same principle as Mahmoudi and Sapiro except instead of exploiting block gradient, they estimate block orientation by estimating mean and variance in the 4 directions at 0^0, 45^0, 90^0, 135^0.

In [8], each patch of size $D=dxd$ is presented as a point in the space $\mathbb{R}^D$. It is well known that most of patches of a natural image are placed in a relatively modest lower-dimensional manifold. Therefore, by using SVD, the authors project the points in the space $\mathbb{R}^D$ onto several subspaces and only coordinates of the most significant subspaces are used for comparison. More precisely, approximately 90% of dissimilar patches are eliminated thanks to comparison of only the first coordinate and 20% are eliminated for each subsequent coordinates after that. By this way, they need MxN norm terms for comparison of the first coordinate but only 0.02xMxN terms for the next. Using this idea, they progressively narrow the set of the best candidates for a target-patch.

In [9], Liu and Freeman propose to use a kNN algorithm namely PatchMatch [10] to rapidly find out the most similar patches for a given one. Given a point in the space $\mathbb{R}^D$, kNN algorithm will seek k nearest points in term of Euclidean distance. It contains three phases: *initialization*, *propagation* and *random search*. More details about this algorithm can be found in [10].

In this paper, the computational and the complexity issues will not be addressed. The main intent of this paper is to evaluate the performance of the closest structure approach. From that, we propose a new method to improve the denoising performance.

IV. Performance Analysis of Closest Structure Approach

In the NLM method [3-6], the authors suggest to set patch size *d*x*d*=7x7 and search-zone *s*x*s*=21x21. We adopt these parameters for our tests. An image of size 100x441 (see Fig.2a) with two intensity levels (dark on top haft and light on bottom half) is perturbed by additive Gaussian noise. The noisy image is filtered by two approaches: closest space (i.e. original NLM) and closest structure. For the second one, we use a kNN algorithm which is based on kd-trees structure [11, 12] and priority search [13]. The C++ code source of this algorithm is available on in [14]. This algorithm is slower than PatchMatch [10] but yields similar results. In the original NLM approach, each patch uses 21x21 = 441 closest ones. Therefore to be fair, in the closest structure approach, we seek in the whole image the best 441 blocks for each one. The results of the first approach and the second one are shown in Fig. 2c, d respectively (***Please use your monitor to view the images and zoom-in for more details***). As can be seen, the contours are clearly better preserved in the second case than the first one. However, in the homogeneous region, the noise is better eliminated in the first case than the second one. This is due to the fact that the closest structure approach searches in the entire image for the best matches, then in flat regions, the noise pattern of the target-patch will match with that of the best candidates. Averaging these similar noise patterns can not effectively remove the noise.

The test is also carried out on several natural images such as Lena, Barbara, Peppers and Fingerprint. They are contaminated by additive Gaussian noise with standard deviation 10. Fingerprint is a typical example for highly textured image while Peppers contains homogeneous regions. On the other hand, Lena and Barbara have all elements (homogeneous, contour, texture). The PSNR comparison of the closest space and the closest structure approaches is reported in Table I. As can be seen, the latter yields only better result for Fingerprint image and is runner-up for the remaining cases.

TABLE I. PSNR comparison of closest space and closest structure approaches

Method	**Image**			
	Barbara	***Lena***	***Peppers***	***Fingerprint***
Closest space	31.8108	32.6141	32.9041	28.1550
Closest structure	30.7513	31.8532	32.6208	29.1180

The restored images of two approaches are shown in Fig. 3, 4, 5, 6. As can be seen, noise still remains in uniform regions of Lena, Barbara and Peppers images in the case of the closest structure method while it is effectively removed in the case of the closest space approach. In contrary, contour of the restored Fingerprint image of the closest structure method is sharper than that of the closest space method.

From above analyses, we can conclude that the closest structure approach is only suitable for highly texture images.

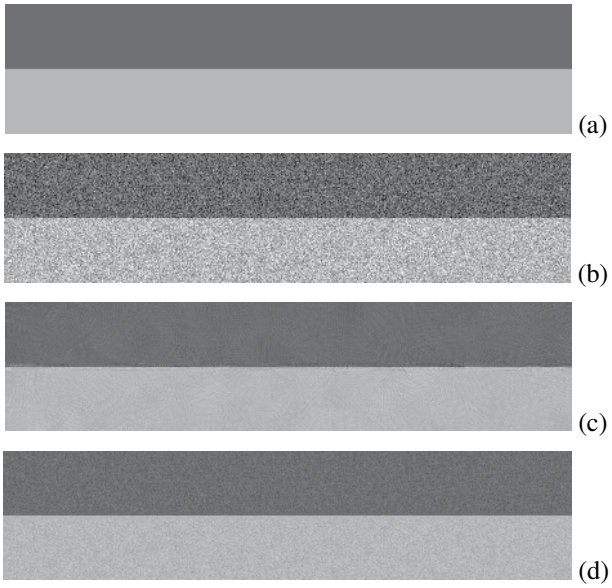

Figure 2. Performance comparison of closest space and closest structure approaches on simulated image.

V. Proposed Method And Experimental Results

Based on the above analysis, NLM method could be improved by using a block classification before filtering. The blocks are classified into two groups: homogeneous blocks and structured blocks (texture, contour). For the first class, the closest space strategy will be used whereas for the second case, the closest structure will be applied. The main difficulty of this idea is how to do the classification task in noisy context.

In this paper, we do not follow this idea. Indeed, we found that the closest space and the closest structure approaches are complementary. Therefore, for each patch, among its 441 candidates, we select *n* patches from the closest structure approach and 441-*n* patches from the closest space one. By this way, all of patches are treated in the same manner and we do not need classification step. *n* is a parameter of the method and the optimal value is estimated from experimental tests.

The experiments are carried out on 4 natural images Lena, Barbara, Peppers and Fingerprint as shown section IV. The parameter *n* varies from 0 to 441. For each *n*, 5 values of the decay parameter *h* are tested, i.e. 40, 50, 60, 70, 80. For objective evaluation, beside PSNR, we use also two other metrics namely MAD [15] and $PSNR_W$ [16] which are based on the human visual system (HVS). Note that the smaller the value of MAD, the better the image quality. However, the higher the value of $PSNR_W$ and PSNR, the better the image quality. The quality measures against *n* are shown in Fig.7, 8, 9, 10. In the case of Fingerprint image, as can be seen in Fig. 7, for all values of the decay parameter *h*, all metrics confirm that the closest structure method outperforms the closest space one (In PSNR-$PSNR_W$/MAD curves, the right-most data points (where *n*=441) are always higher/lower than the left-most ones (where *n*=0)). For the remaining cases, i.e. Lena, Barbara and Peppers images, they all show that the closest

space approach is better than the closest structure one. (In PSNR-$PSNR_W$/MAD curves, the right-most data points are always lower/ higher than the left-most ones). The results of two additional quality metrics ($PSNR_W$ and MAD) reinforce our remark in the above section. Moreover, the best point is neither the left-most points nor the right-most ones. It is always in the middle of the curves. This happens for all images and does not depend on the decay parameter *h*. It means that the idea of combining the two approaches outperforms both closest space and closest structure methods. It could be observed that for each metric an optimal value for *n*, is obtained. But this value is always in the interval ranging from 10 to 20.

VI. CONCLUSIONS

In this paper, the denoising performance of NLM (closest space) and its amelioration (closest structure) is investigated. It has been shown that the closest structure approach is only suitable for highly textured images. The experimental tests have proven that by combining these two approaches better results could be obtained.

REFERENCES

[1] R.C. Gonzalez and R.E. Woods, "Digital image processing", 2nd edition, Prentice Hall International Editions, 2002.

[2] L.P. Yaroslavsky, "Digital Picture Processing - An Introduction", Springer Verlag, 1985.

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

[4] A. Buades, B. Coll and J. M. Morel, "Image and movie denoising by nonlocal means," International Journal of Computer Vision, vol. 76, 2008.

[5] A. Buades, B. Coll and J.M. Morel, "A non-local algorithm for image denoising," IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), vol.2, 2005.

[6] M. Mahmoudi and G.Sapiro, "Fast image and video denoising via nonlocal means of similar neighborhoods," IEEE Signal Processing Letter, vol. 12, no. 12, pp. 839-842, 2005.

[7] R.C. Bilcu and M. Vehvilainen, "Fast nonlocal means for image denoising," The International Society for Optical Engineering (SPIE), vol. 6502, 2007.

[8] J. Orchard, M. Ebrahimi and A. Wong, "Efficient nonlocal-means denoising using the SVD," IEEE International Conference on Image Processing, pp. 1732 - 1735, 2008

[9] C. Liu and William T. Freeman. "A High-Quality Video Denoising Algorithm based on Reliable Motion Estimation," European Conference on Computer Vision, 2010.

[10] C. Barnes, E. Shechtman, A. Finkelstein, and D.B. Goldman, "PatchMatch: A Randomized Correspondence Algorithm for Structural Image Editing," ACM Transactions on Graphics (Proc. SIGGRAPH) 28(3), August 2009.

[11] J. L. Bentley, "K-d trees for semidynamic point sets," In Proc. 6th Ann. ACM Sympos. Comput. Geom., pp.187-197, 1990.

[12] J. H. Friedman, J. L. Bentley, and R. A. Finkel, "An algorithm for finding best matches in logarithmic expected time." ACM Transactions on Mathematical Software, pp. 209-226, 1977.

[13] S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman, and A. Wu, "An optimal algorithm for approximate nearest neighbor searching," J. ACM, pp. 891-923, 1998.

[14] http://www.cs.umd.edu/~mount/ANN/

[15] E.C. Larson, D.M. Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy," Journal of Electronic Imaging, 19 (1), 2010.

[16] A. Beghdadi, B. Pesquet-Popescu, "A New Image Distortion Measure Based Wavelet Decomposition," In Proc. ISSPA, pp.485-488, 2003.

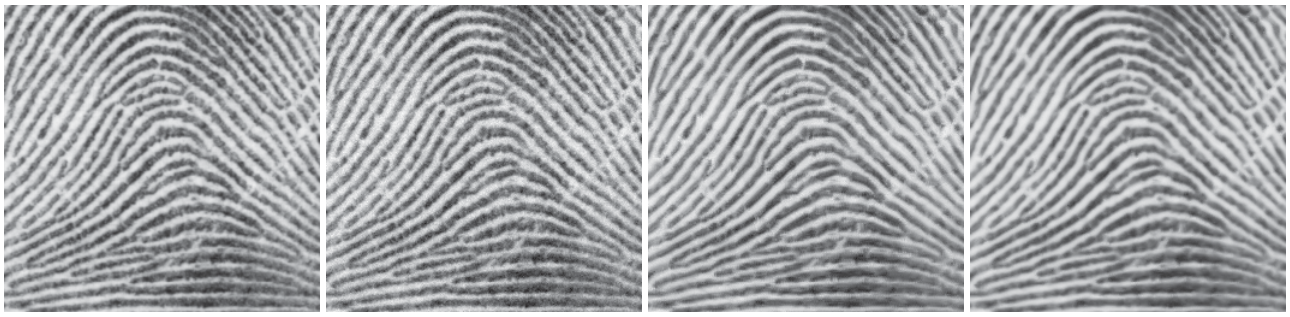

Figure 3. Zooms of Fingerprint image (from left to right): the original image, noisy image, closest space's result, closest structure's result

Figure 4. Zooms of Peppers image (from left to right): the original image, noisy image, closest space's result, closest structure's result

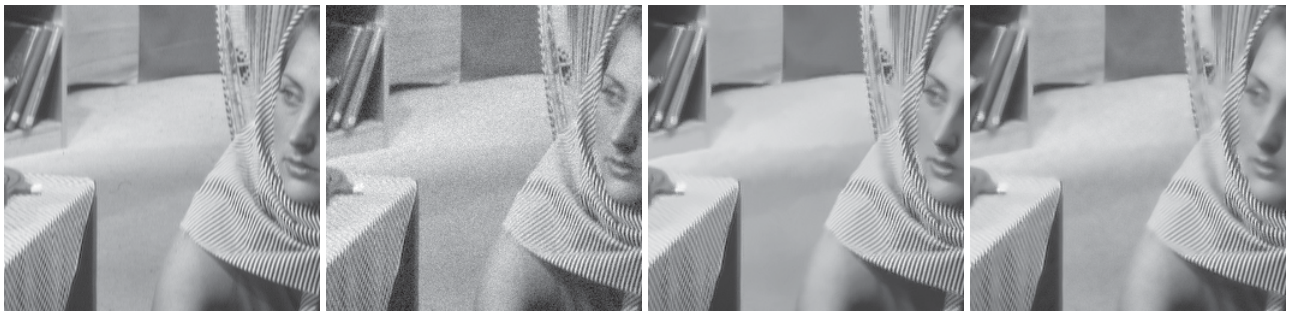

Figure 5. Zooms of Barbara image (from left to right): the original image, noisy image, closest space's result, closest structure's result

Figure 6. Zooms of Lena image (from left to right): the original image, noisy image, closest space's result, closest structure's result

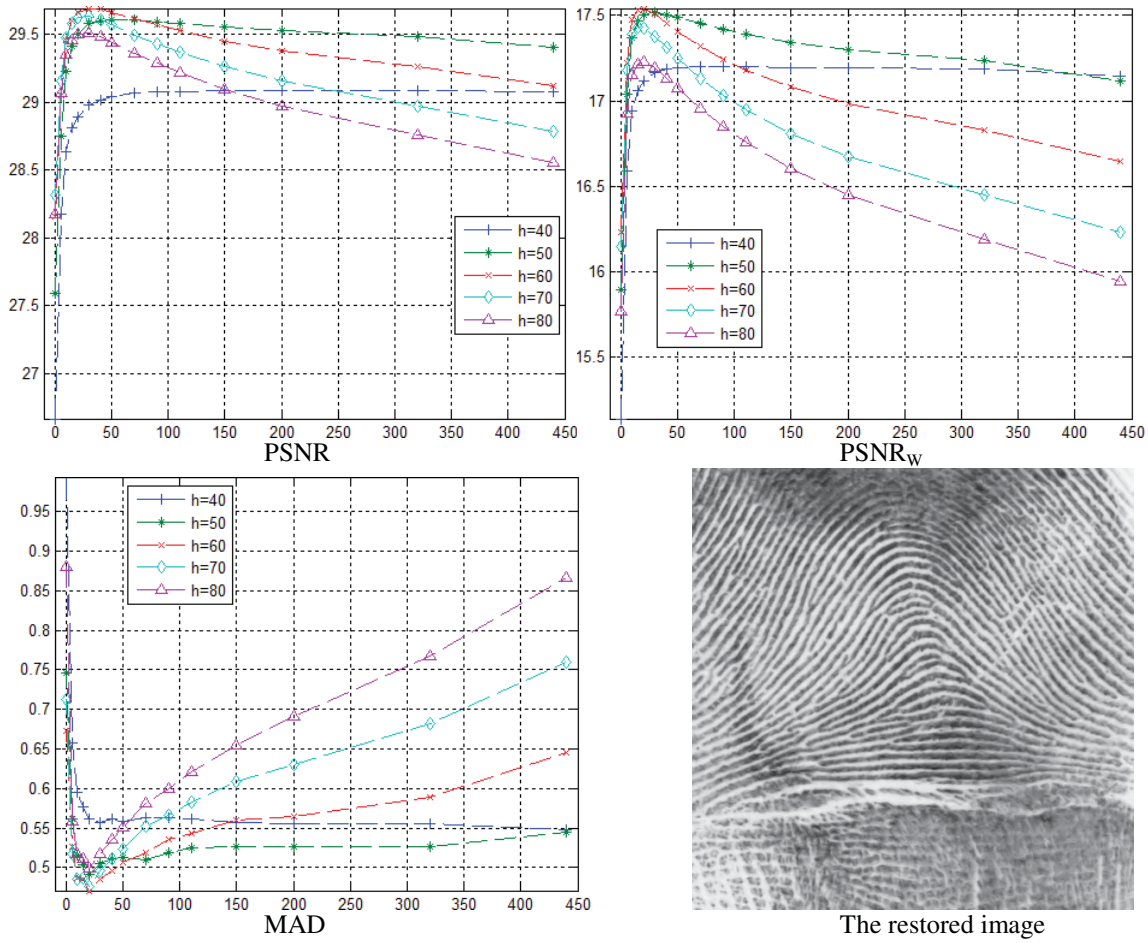

Figure 7. Fingerprint image (from left to right, top to bottom): PSNR, PSNR$_W$, MAD measurements and the restored image with h=60, n=20

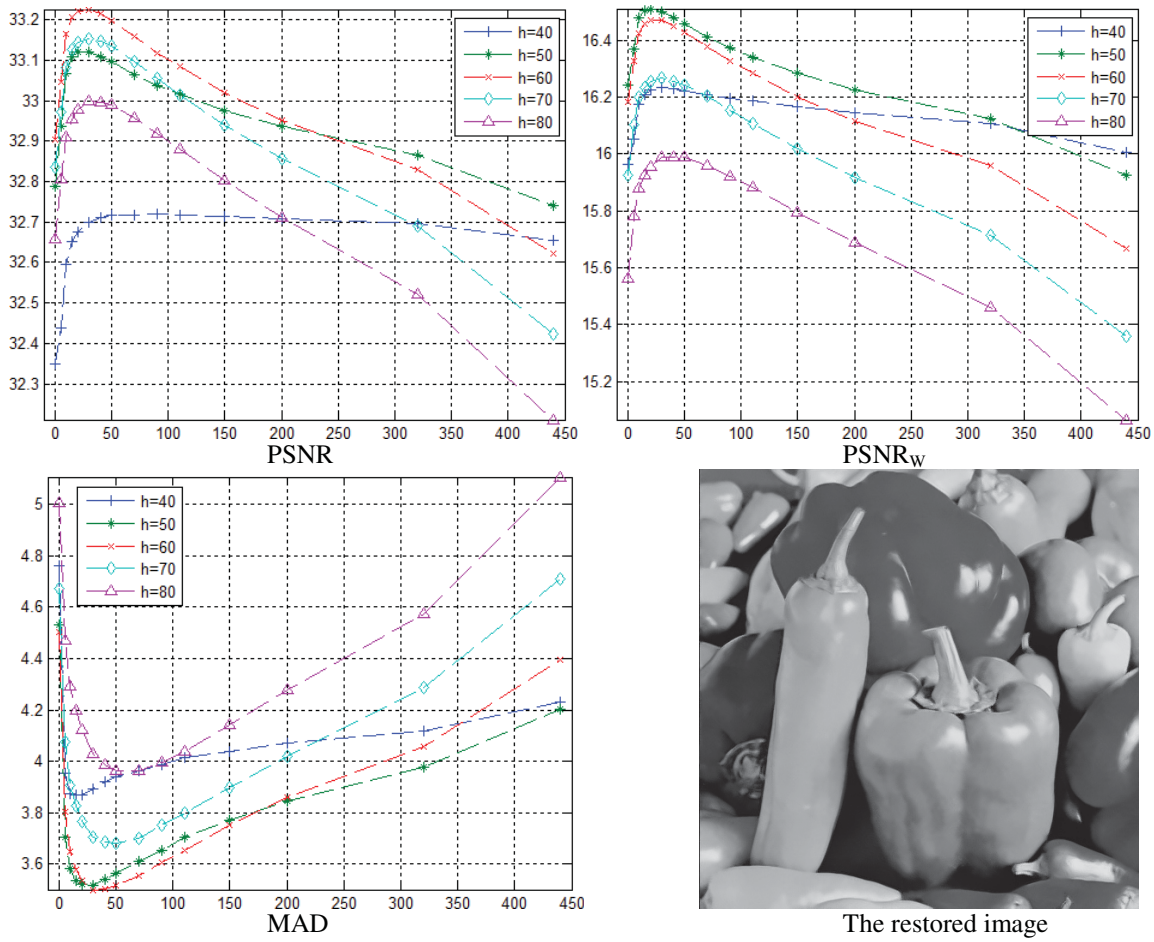

Figure 8. Peppers image (from left to right, top to bottom): PSNR, $PSNR_W$, MAD measurements and the restored image with h=60, n=10

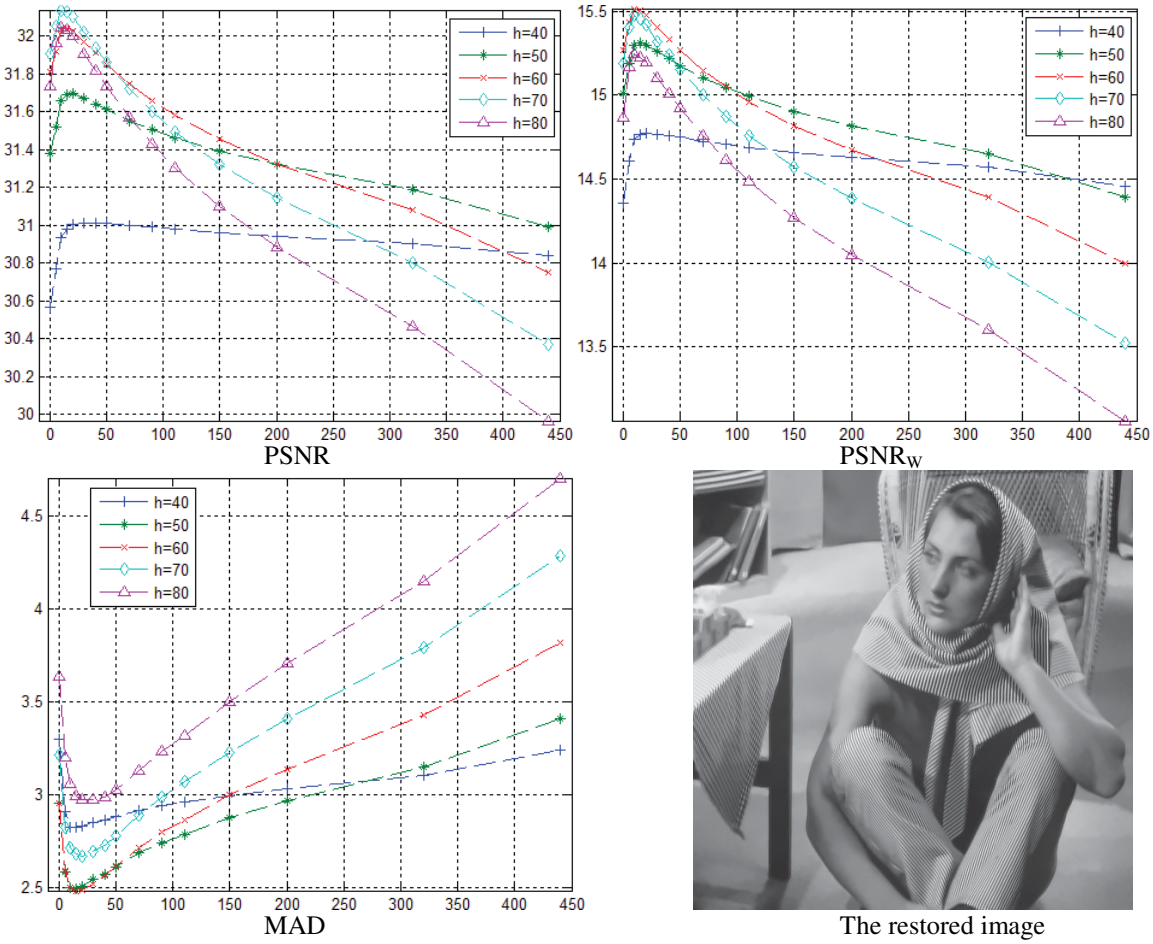

Figure 9. Barbara image (from left to right, top to bottom): PSNR, $PSNR_W$, MAD measurements and the restored image with h=60, n=15

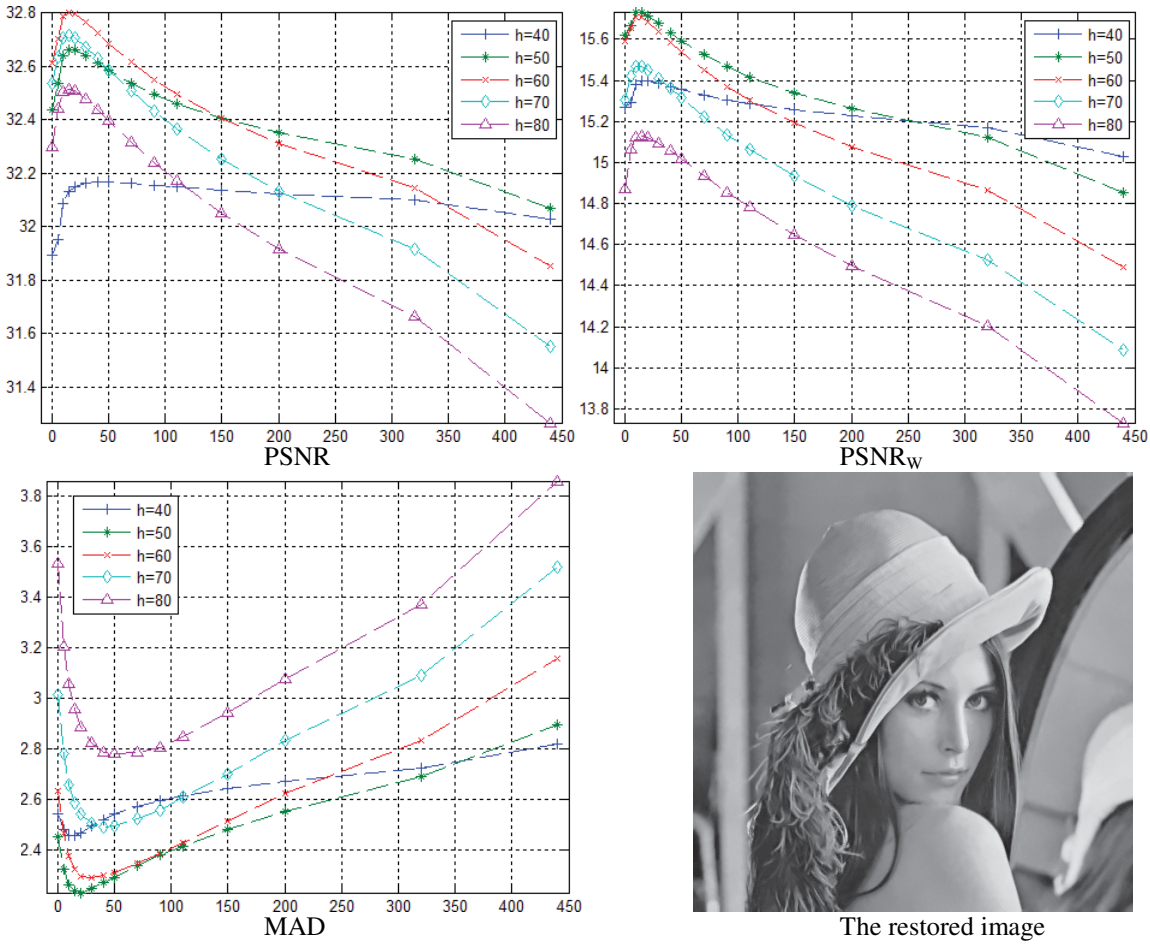

Figure 10. Lena image (from left to right, top to bottom): PSNR, PSNR$_W$, MAD measurements and the restored image with h=50, n=20