

# Improved Non-Local Means Algorithm Based on Dimensionality Reduction

Golam M. Maruf and Mahmoud R. El-Sakka<sup>(✉)</sup>

Department of Computer Science, University of Western Ontario, London,  
Ontario, Canada  
{gmaruf, melsakka}@uwo.ca

**Abstract.** Non-Local Means is an image denoising algorithm based on patch similarity. It compares a reference patch with the neighboring patches to find similar patches. Such similar patches participate in the weighted averaging process. Most of the computational time for Non-Local Means scheme is consumed to measure patch similarities. In this paper, we have proposed an improvement where the image patches are projected into a global feature space. Then we have performed a statistical t-test to reduce the dimensionality of this feature space. Denoising is achieved based on this reduced feature space. The proposed modification exploits an improvement in terms of denoising performance and computational time.

**Keywords:** Non-Local Means algorithm · Image denoising · Image smoothing · Image enhancement · Additive white Gaussian noise · Spatial domain filtering

## 1 Introduction

An image may be numerically represented as a two dimensional discrete function  $u$ , in the spatial coordinates  $x$  and  $y$ . Intensity or gray level is the amplitude of  $u$  at any pair of coordinates. A digital image is composed of finite number of elements called pixels. An image may be contaminated with noise during acquisition, transmission or transformation. Noise is a variation of pixel intensity. Such noise can be additive or multiplicative. Additive noise is generally independent of image data whereas multiplicative noise is dependent on image data. Additive noise can be formularized as,

$$v(i) = u(i) + n(i), \quad (1)$$

whereas, multiplicative noise is formularized as,

$$v(i) = u(i) \times n(i). \quad (2)$$

Here,  $u(i)$  is the original value,  $n(i)$  is the noise value and  $v(i)$  is the observed value at pixel  $i$ . Despite the good quality of acquisition devices, an image denoising

method is always required to reduce unwanted noise signals. An image denoising scheme is used to find the best estimate of the original image from its noisy version.

Some of the basic filtering such as Gaussian and average filtering have a drawback of over-smoothing on edges and losing image details. Wavelet based denoising method [1], anisotropic diffusion [2], and bilateral filtering [3] try to overcome this drawback and preserve the image quality by preserving edges. But they may introduce a staircase effect or false edges. Recently, Buades et al. [4] proposed a denoising algorithm called Non-Local Means (NLM) which allows neighboring patches in the search window to participate in the denoising process for a certain reference patch in the noisy image. Most of the computational time for NLM is allocated to the similarity assessments between patches. In a general case, NLM needs to search the entire image for similar patches and performs weighted average based on the similarities. However, searching in a fixed area around the pixel of interest (POI) can reduce this computational time. Our main focus is to further reduce this computational time and improve denoising performance over the original Non-Local Means algorithm.

Many improvements have been suggested on the Non-Local Means algorithm in recent years. Bhujle et al. [5] proposed a dictionary based denoising in which patches with similar photometric structures are clustered together to create groups. Mahmoudi et al. [6] accelerate the NLM algorithm by pre-classifying neighborhood patches based on average gray values, gradient orientation, or both. Chaudhury et al. [7] claimed that the denoising performance of the Non-Local Means algorithm can be improved by replacing the mean operation by a median operation. Vignesh et al. [8] proposed a speed up technique for the Non-Local Means algorithm based on a probabilistic early termination (PET). Tasdizen et al. [9] proposed principal component based Non-Local Means algorithm where a global feature space is created to select important features. Brox et al. [10] proposed a technique to improve the performance of the NLM method using a clustering tree.

In this paper, we have proposed to create feature vectors for the noisy image. Then we have implemented a statistical *t-test* on these feature vectors to reduce their dimensionality. Our proposed method reduces the computational time and improves the overall performance of the original NLM algorithm.

## 2 Methods

The Non-Local Means algorithm searches neighboring patches to match with the reference patch. The original algorithm requires an extensive amount of time to select patches similar to the reference patch. These similar patches contribute to the weighted averaging process to denoise the center pixel of the reference patch. The computation time for the NLM algorithm can be reduced by improving this searching process.

We have formalized our proposed work into three steps. In the first step (pre-processing), we have created a global feature space and stored all possible

neighborhood pixels. Then in the second step, we have performed a statistical *t-test* to reduce the dimensionality of the feature space on the previous feature points. Finally, we implement the non-local means algorithm, where we have used the selected feature points for calculating similarity measures of image neighborhood.

## 2.1 Preprocessing

In this step, we have created a feature vector space for the noisy image. An image patch is linearized and represented as a row vector of size  $j$ . Thus the dimension of this feature vector space will be  $j \times N$ , where  $N$  is the total number of image pixels. Feature vectors can be represented as matrix  $C$ ,

$$C = \begin{bmatrix} c(1,1) & \cdots & c(1,j) \\ \vdots & \ddots & \vdots \\ c(N,1) & \cdots & c(N,j) \end{bmatrix} \quad (3)$$

Here, for example if we have a patch size of  $7 \times 7$  then  $j$  will be equal to 49. This matrix will be used during the dimensionality reduction process.

## 2.2 t-Testing

We have implemented a paired *t-test* of the null hypothesis. This test is performed on the matrix  $C$ . For each test case (i.e., each column in the matrix  $C$ ), once the  $t$  value is determined, the students  $t$ -distribution lookup table is used to find the value of  $p$ . When the calculated  $p$  value is below a given threshold value, then the null hypothesis is rejected. In our denoising problem, we have considered each patch as a feature vector. The hypothesis tries to accept or reject a feature (i.e. an entire column in the matrix  $C$ ). Here, the null hypothesis is whether a feature is significant or not. In calculating the null hypothesis, one uses the following normalization equation

$$T = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \quad (4)$$

Where,  $\bar{x}$  is the sample mean,  $\mu_0$  is the population mean,  $s$  is the sample standard deviation and  $n$  is the sample size. When the null hypothesis is accepted, it concludes that the feature is significant. Otherwise, this feature is not significant. Thus the entire column is deleted and hence reduces the size of matrix  $C$ .

### 2.3 Non-Local Means Algorithm

In the Non-Local Means algorithm a discrete noisy image  $v = \{v(j) \mid j \in I\}$ , where  $I$  is the input image, can be denoised by the estimated value  $NL[v](i)$  for a pixel  $i$ . It is computed as a weighted average for all of the pixels in the image,

$$NL[v](i) = \sum_{j \in I} w(i, j)v(j) \quad (5)$$

where, the weight  $w(i, j)$  depends on the similarity between the pixel  $i$  and the pixel  $j$  of the intensity gray level vectors  $v(N_i)$  and  $v(N_j)$ . Here,  $N_k$  is the square patch around the center pixel  $k$ . The weight is then assigned to value  $v(j)$  to denoise pixel  $i$ . The summation of all weight is equal 1 and each weight value  $w(i, j)$  has a range between  $[0, 1]$ . To measure similarity between patches, the Euclidean distance between patches is calculated,

$$\|v(N_i) - v(N_j)\|_{2, \sigma}^2 \quad (6)$$

where,  $\sigma > 0$  is the standard deviation of the Gaussian kernel. The weight  $w(i, j)$  are computed as follows,

$$w(i, j) = \frac{1}{z(i)} e^{-\frac{\|v(N_i) - v(N_j)\|_{2, \sigma}^2}{h^2}} \quad (7)$$

where,  $Z(i)$  is a normalization constant such that,

$$Z(i) = \sum_j e^{-\frac{\|v(N_i) - v(N_j)\|_{2, \sigma}^2}{h^2}} \quad (8)$$

Here,  $h$  is a smoothing kernel width which controls decay of the exponential function and therefore controls the decay of the weights as a function of the Euclidean distances. In our proposed method  $N_k$  is replaced by  $f_k$ , where  $f_k$  is the reduced feature vector. Then we have selected similar patches and calculated weights based on this reduced feature vector.

Our proposed algorithm is summarized as follows.

---

**Algorithm** Improved Non-Local Means

---

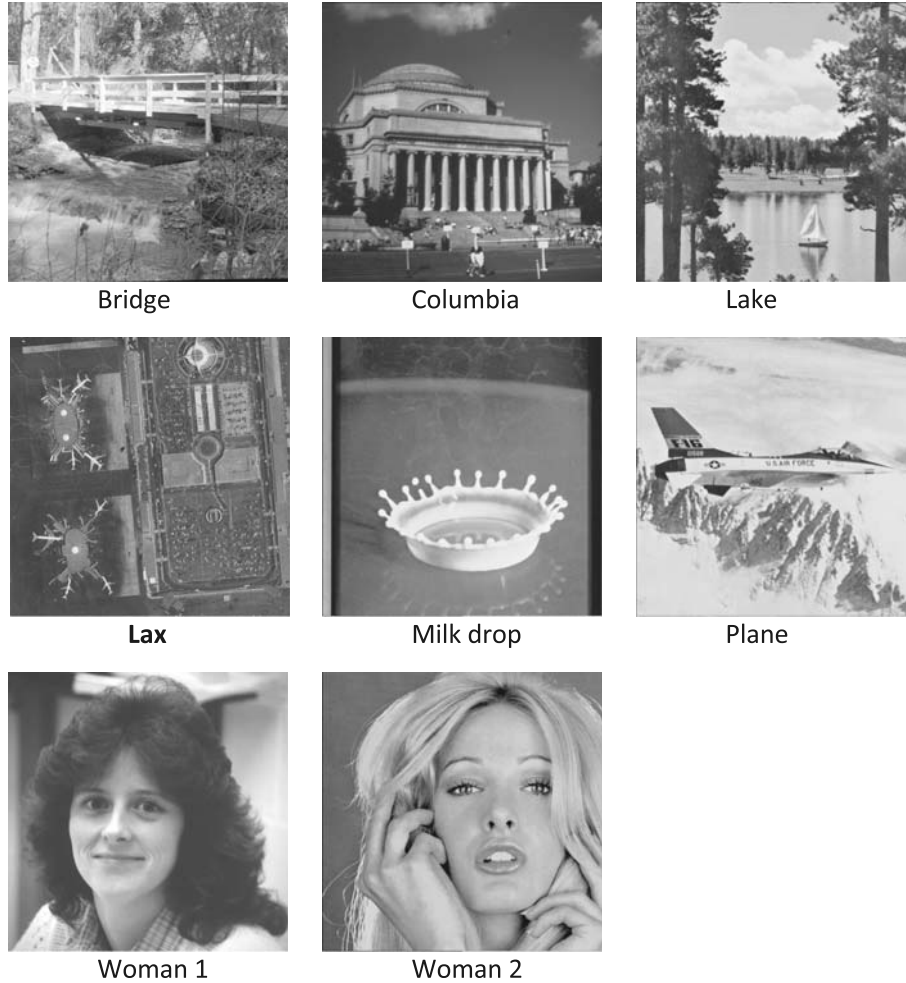
**Input I:** Image with additive white Gaussian noise**Output NL(I):** Denoised image

1. Create a global feature matrix  $C$  (as shown in Equation 3).
  2. Perform the  $t$ -test on matrix  $C$  to produce the reduced row matrix  $f_k$ .
  3. For each pixel  $i$ , where  $i \in [1, N]$ ,
  4. Do
    - 4.1. For each pixel in  $N_k$ , i.e., the square patches around the center pixel  $k$
    - 4.2. Do
      - 4.2.1. Evaluate the normalization constant  $Z(i) \leftarrow \sum_j e^{-\frac{\|v(f_i) - v(f_j)\|_2^2}{h^2}}$ , where  $j$  refers to the  $N_k$  patches.
      - 4.2.2. Calculate the weight matrix  $W(i, j) \leftarrow \frac{1}{Z(i)} e^{-\frac{\|v(f_i) - v(f_j)\|_2^2}{h^2}}$
      - 4.2.3. Done
    - 4.3. Denoise pixel  $i$ :  $NL[v](i) \leftarrow \sum_{j \in 1} w(i, j)v(j)$
    - 4.4. Done
- 

### 3 Results

The performance of our proposed method is compared in terms of PSNR with other denoising schemes, namely the original NLM method, the principal component analysis based NLM method (PCA-NLM), the patch regression based NLM method (NLM-Patch) and the BM3D method. Eight  $512 \times 512$  test images (Bridge, Columbia, Lake, Lax, Milk drop, Plane, Woman1 and Woman2) are utilized to assess the performance of these schemes. See Fig. 1.

Tables 1 and 2 show the average PSNR and SSIM comparative performance, respectively, for all test images. The bolded values in Tables 1 and 2 represent the highest PSNR and SSIM values, respectively, among all of the algorithms for a given noise level.



**Fig. 1.** The set of the test images ( $512 \times 512$ ) for performance analysis.

For noise level  $\sigma < 50$ , the proposed method performs better than any other denoising scheme, including the BM3D method. Yet, for noise level  $\sigma > 50$  the proposed method performs better than the original NLM and its variants. Yet, the BM3D method performs better at higher noise levels.

Table 3 compares the average running time performance for all test images for the proposed method and the other denoising schemes. It has been found that our proposed method outperforms the NLM method, variants of the NLM method and the BM3D method at all noise levels, as it requires fewer features to compare and calculate weights. Thus the computational time is dramatically reduced while keeping the denoising performance in an acceptable range.

**Table 1.** Average PSNR(dB) comparison for all test images among the proposed method, the NLM method, variants of the NLM method and the BM3D method for various noise levels.

Noise level	NLM	PCA-NLM	NLM-patch	Proposed method	BM3D
10	32.52	32.94	31.47	<b>33.94</b>	33.84
20	29.87	29.95	29.04	<b>31.0</b>	30.50
30	28.13	28.26	27.45	<b>28.96</b>	28.38
40	26.69	26.43	25.87	<b>27.72</b>	27.70
50	25.49	25.38	24.61	26.49	<b>26.86</b>
60	23.85	23.87	22.75	24.30	<b>25.94</b>
70	22.90	22.81	22.31	23.22	<b>25.29</b>
80	22.32	22.32	21.92	22.60	<b>24.75</b>
90	21.73	21.57	20.89	21.86	<b>24.18</b>
100	21.13	20.94	20.14	21.19	<b>23.68</b>
Average	25.46	25.45	24.64	26.15	<b>27.11</b>

**Table 2.** Average SSIM comparison for all test images among the proposed method, the NLM method, variants of the NLM method and the BM3D method for various noise levels.

Noise level	NLM	PCA-NLM	NLM-patch	Proposed method	BM3D
10	0.9078	0.9015	0.9051	0.9201	0.9124
20	0.8625	0.8605	0.8610	0.8785	0.8711
30	0.8389	0.8341	0.8291	0.8469	0.8415
40	0.8071	0.8065	0.8017	0.8202	0.8201
50	0.7689	0.7597	0.7659	0.7810	0.7841
60	0.7487	0.7491	0.7412	0.7524	0.7617
70	0.7059	0.7032	0.7015	0.7195	0.7217
80	0.6925	0.6912	0.6907	0.7079	0.7138
90	0.6857	0.6815	0.6851	0.6992	0.7051
100	0.6711	0.6504	0.6522	0.6975	0.7004
Average	0.7878	0.7819	0.7823	0.8022	0.8099

**Table 3.** Running time (in milliseconds) for Lena image among the proposed method, the NLM method, variants of the NLM method and the BM3D method for different noise levels.

Noise level	NLM	PCA-NLM	NLM-patch	Proposed method	BM3D
10	209.5	195.1	208.1	<b>161.2</b>	223.2
20	210.7	196.7	210.6	<b>164.5</b>	224.2
30	212.3	197.4	210.0	<b>165.7</b>	225.1
40	212.6	198.8	211.5	<b>169.9</b>	229.3
50	212.4	200.3	211.0	<b>173.9</b>	230.2
60	213.0	204.4	212.8	<b>181.2</b>	230.8
70	214.5	207.9	213.1	<b>182.5</b>	231.3
80	214.9	208.6	213.9	<b>184.0</b>	231.6
90	216.0	209.9	214.0	<b>185.2</b>	232.9
100	217.1	210.1	216.4	<b>185.9</b>	233.1
Average	213.3	202.9	212.1	<b>175.4</b>	229.8

## 4 Conclusions

Non-Local Means is a popular image denoising algorithm implemented in the spatial domain. In this research, we have proposed a statistics based improvement for the Non-Local Means algorithm. The key of this improvement is to reduce the size of the feature space, which reduces the patch similarity measurement time and increases the overall denoising performance. We have utilized a statistical *t-test* to reduce the dimensionality of the feature space. Experimental results show that our proposed method provides the best running time among all other algorithms in all test cases at various noise levels. It also provides a good denoising improvement in terms of the PSNR and the SSIM values. In addition, it performs better than the NLM method and its variants at all noise levels and perform better than the BM3D method for lower noise levels.

## References

1. Portilla, G., Strela, J., Wainwright, V., Simoncelli, M.: Image denoising using scale mixtures of gaussians in the wavelet domain. *IEEE Trans. Image Process.* **12**(11), 1338–1351 (2003)
2. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.* **12**(7), 629–639 (1990)
3. Tomasi C., Manduchi, R.: Bilateral filtering for gray and color images. In: *IEEE Sixth International Conference on Computer Vision*, pp. 839–846 (1998)
4. Buades, A., Coll, B., Morel, J-M.: A non-local algorithm for image denoising. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 60–65 (2005)
5. Bhujle, H., Chaudhuri, S.: Novel speed-up strategies for non-local means denoising with patch and edge patch based dictionaries. *IEEE Trans. Image Process.* **23**(1), 356–365 (2014)
6. Mahmoudi, M., Sapiro, G.: Fast image and video denoising via nonlocal means of similar neighborhoods. *IEEE Signal Process. Lett.* **12**(12), 839–842 (2005)
7. Chaudhury, K., Singer, A.: Non-local Euclidean medians. *IEEE Signal Process. Lett.* **19**(11), 745–748 (2012)
8. Vignesh, R., Oh, B., Kuo, C.-C.: Fast non-local means (NLM) computation with probabilistic early termination. *IEEE Signal Process. Lett.* **17**(3), 277–280 (2010)
9. Tasdizen, T.: Principal components for non-local means image denoising. In: *15th IEEE International Conference on Image Processing*, pp.1728–1731 (2008)
10. Brox, T., Kleinschmidt, O., Cremers, D.: Efficient nonlocal means for denoising of textural patterns. *IEEE Trans. Image Process.* **17**(7), 1083–1092 (2008)