LFAD- Locally and Feature-Adaptive Diffusion Based Image Denoising

Ajay K. Mandava, Emma E. Regentova, George Bebis*

Department of Electrical and Computer Engineering
University of Nevada, Las Vegas, NV 89154, USA
*Department of Computer Science & Engineering
University of Nevada, Reno, NV 89557, USA
*Email: mandavaa@unlv.nevada.edu, emma.regent@unlv.edu,
bebis@cse.unr.edu*

Abstract

LFAD is a novel locally and feature adaptive diffusion method for denoising an additive Gaussian noise in images. The method approaches each image region individually and uses different number of diffusion iterations per region for attaining best objective quality according to PSNR. Unlike block-transform based methods which perform with a pre-determined optimum block size and clustering-based denoising methods which use a fixed optimum number of classes, our method searches for an optimum patch size through iterative diffusion starting with a small patch size and proceeds with aggregating patches until a best PSNR is attained. The diffusion model has a substitution of the gradient value with the inverse difference moment (IDM) which is a robust feature in determining the amount of local intensity variation in the presence of noise. Experiments with benchmark images and various noise levels show that the designed LFAD outperforms advanced diffusion based denoising methods, and it is competitive with the state-of-the-art block-transformed techniques by yielded PSNR levels, producing however lesser visible blocking or ringing artifacts.

Keywords: Diffusion, Patch, Region, Over-segmentation.

1. Introduction

Nonlinear anisotropic diffusion has drawn considerable attention over the past decade and has experienced significant developments as it gracefully diffuses the noise in the intra-regions while inhibiting inter-region smoothing. Introduced first by Perona and Malik (PM diffusion) [1] the diffusion process is mathematically described by the following equation:

$$\frac{\partial}{\partial t} I(x, y, t) = \nabla \bullet (c(x, y, t)\nabla I)$$

(1),

where $I(x,y,t)$ is the image, t is the iteration step and $c(x,y,t)$ is the diffusion function monotonically decreasing of the magnitude of the image gradient. Two diffusivity functions proposed are:

$$c_1(x, y, t) = \exp\left(-\left(\frac{|\nabla I(x, y, t)|}{k}\right)^2\right)$$

(2)

and

$$c_2(x, y, t) = \frac{1}{1 + \left(\frac{|\nabla I(x, y, t)|}{k}\right)^2}$$

(3),

where k is referred to as a diffusion constant. Depending on the choice of the diffusivity function, equation (1) covers a variety of filters. The discrete diffusion structure is translated into the following form:

$$I_{i,j}^{n+1} = I_{i,j}^{n} + (\nabla t) \bullet \left[\begin{array}{l} c_N\left(\nabla_N I_{i,j}^n\right) \bullet \nabla_N I_{i,j}^n + c_S\left(\nabla_S I_{i,j}^n\right) \bullet \nabla_S I_{i,j}^n + \\ c_E\left(\nabla_E I_{i,j}^n\right) \bullet \nabla_E I_{i,j}^n + c_W\left(\nabla_W I_{i,j}^n\right) \bullet \nabla_W I_{i,j}^n \end{array} \right]$$

(4).

Subscripts N, S, E and W (North, South, East and West) describe the direction of the local gradient, and the local gradient is calculated using nearest-neighbor differences as

$$\nabla_N I_{i,j} = I_{i-1,j} - I_{i,j}; \quad \nabla_S I_{i,j} = I_{i+1,j} - I_{i,j} \quad \nabla_E I_{i,j} = I_{i,j+1} - I_{i,j}; \nabla_W I_{i,j} = I_{i,j-1} - I_{i,j} \quad (5).$$

Generally, the effectiveness of the anisotropic diffusion is determined by (a) the efficiency of the edge detection operator to distinguish between noise and edges; (b) the accuracy of an "edge-stopping" function to promote or inhibit diffusion; and (c) the adaptability of a convergence condition to terminate the diffusion process automatically. The model in [1] has several practical and theoretical limits. It needs a reliable estimate of image gradients because with the increase of the noise level, the effectiveness of the gradient calculation degrades and thus deteriorates the performance of the method. Secondly, the equal number of iterations in the diffusion of all the pixels in the image leads to blurring of textures and fine edges while the smooth regions benefit.

Let us apply the PM diffusion to two different image patches; each representing a certain structural content, for example, a texture and a smooth region. Fig.1 indicates significant differences in PSNR levels/iterations for provided examples. Two examples in Fig.2 show how the image quality varies among two iterations (22 and 30); in the left image pixels are corrupted in a smooth region and, in the right image details are severely blurred. This happens because diffusion stops when the PSNR which is calculated per entire image reaches its maximum.

Several authors have independently addressed this problem. Catte et al. [2] used a smoothed gradient of the image, rather than the true gradient. Let G_σ be a smoothing kernel, then

$$\frac{\partial}{\partial t} I(x, y, t) = \nabla \bullet (c(\|\nabla G_\sigma * I\|)\nabla I)$$

.

The smoothing operator removes some of the noise which might have deceived the original PM filter. In this case, the scale parameter σ is fixed. In [3], authors have proposed the inhomogeneous anisotropic diffusion which includes a separate mutiscale edge detection.
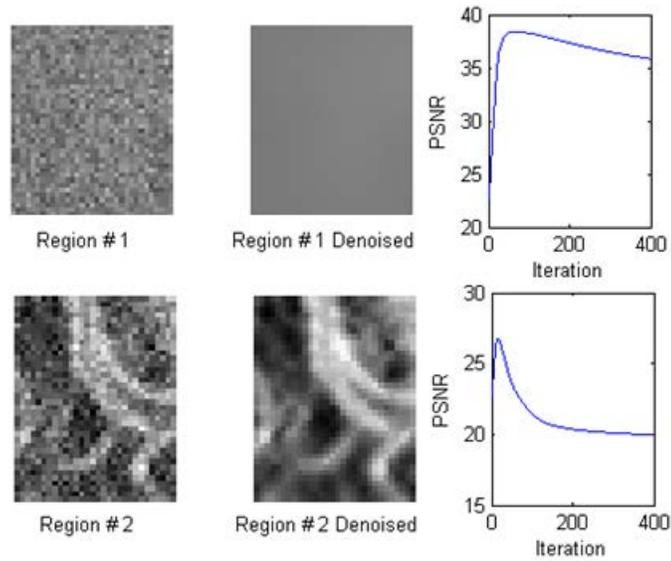
Fig.1. Denoising results for two different structural contents.

Fig.2. First row: PM denoised "Lena" image for two different iterations (left = 22 iterations, PSNR = 29.37 dB; right = 30 iterations, PSNR = 28.52 dB) for additive white Gaussian noise level σ =20;

Yu et al. [4] have incorporated the SUSAN edge detector into the model:

$$\frac{\partial}{\partial t} I(x, y, t) = \nabla \bullet (SUSAN(c(\|\nabla G_\sigma * I\|))\nabla I)$$

SUSAN is capable to guide the diffusion process in an effective manner due to the noise suppression. Li et al. [5] proposed a context adaptive anisotropic diffusion via weighted diffusivity function

$$\frac{\partial}{\partial t} I(x, y, t) = \nabla \bullet (w(x, y, t)c(x, y, t)\nabla I),$$

where the combined term $w(x,y,t)c(x,y,t)$ is referred to as the weighted diffusivity function and $w(x,y,t)$ is a pixel-wise feature dependent weight function.

Chao and Tsai [6] proposed a diffusion model which incorporates both the local gradient and the gray-level variance. When the level of noise is high; noisy pixels in the image generally involve larger magnitudes of gray level variance and gradients than those of actual edges and fine details. Thus, the method is becoming inefficient quite soon. Wang et al. [7] proposed a local variance controlled scheme wherein spatial gradient and contextual discontinuity of a pixel are jointly employed to control the evolution. However, a solution to estimating the contextual discontinuity leads to an exhaustive search procedure, which causes algorithm to be too computationally expensive. Yu and Acton [8] proposed speckle-reducing anisotropic diffusion (SRAD), which integrated spatially adaptive filters into the diffusion and provided considerable improvement in speckle suppression compared to other conventional diffusion methods. Adb-Elmoniem et al. [9] devised a coherence-enhancing nonlinear coherent diffusion (CENCD) model for speckle reduction. This method combines isotropic diffusion, anisotropic coherent diffusion and mean curvature motion. The pursuit is

to maximally filter those regions which correspond to fully developed speckle while preserving information associated with object structures. Zhang et al. [10] presented a Laplacian pyramid-based nonlinear diffusion (LPND) method where Laplacian pyramid was utilized as a multiscale analysis tool to decompose an image into subbands, and then anisotropic diffusion with different diffusion flux was used to suppress noise in each subband. LPND tries to introduce sparsity and multiresolution properties of multiscale analysis into anisotropic diffusion. Another approach to context-based diffusion was researched in [11]. The multi-scale stationary wavelet analysis of the local neighborhood across the scales provides the edge information partially free of noise and thus makes possible the tunable diffusion. As a result, and due to the shift invariance property of stationary wavelet transform the PSNR has been improved compared to Shih's diffusion [12].

Generally, the performance of anisotropic diffusion is influenced by a) effectiveness of the edge detection operator in the presence of noise; b) accuracy of the "edge-stopping" function to promote or impede diffusion; c) adaptability of the convergence condition to the automatic termination of the diffusion process. And the research in this area targets one or more from the above factors. State-of-the art denoising techniques all rely on patches, either for dictionary learning [13,14], collaborative denoising of blocks of similar patches [15] or for non-local sparse models [16]. Regularization with non-local patch-based weights has shown improvements on classical regularization involving only local neighborhoods [17, 18, 19]. The shape and size of patches should adapt to anisotropic behaviour of natural images [20, 21]. In spite of the high performance of the patch-based denoising methods they generally produce artifacts even at a comparatively moderate noise levels. Examples of such visual artifacts are presented in Fig.3 for two state-of-the-art methods such as KLLD [14] and

BM3D [15]. The size of the patch has a significant impact on the PSNR even for the similar or identical contents.

KLLD [14] result for σ = 25   BM3D [15] result for σ = 60

Fig.3. Results of patch based denoising methods.

Fig.4 shows that the equal size regions of the same structural content from different parts of the image could be diffused differently.

Thus, it would be feasible to incorporate adaptation to the image local structure within optimally sized patches. Unlike block-transform based methods such as BM3D which perform with a pre-determined optimum block size and clustering-based denoising methods such as KLLD which uses a predetermined optimum number of classes, our method searches for an optimum patch size through iterative diffusion starting with a small patch size, that is a large number of patches and proceeds with aggregating patches until a best PSNR is attained. To initialize the algorithm we use superpixel segmentation [22]. In our pursuit of determining the amount of diffusion we use the inverse difference moment (IDM) feature [28]. We demonstrate that the feature is robust in estimating local intensity variation in the presence of noise. Overall, the diffusion process converges to PSNR levels comparable to those by the

state-of-the-art methods with minimum visible blocking/patching artifacts. The method is called locally- and feature- adaptive diffusion (LFAD) method.
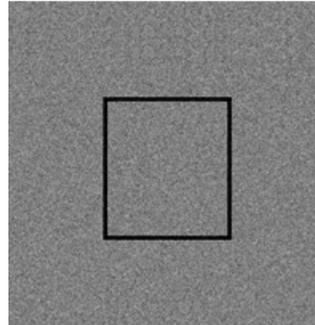
Noisy Image, σ = 20
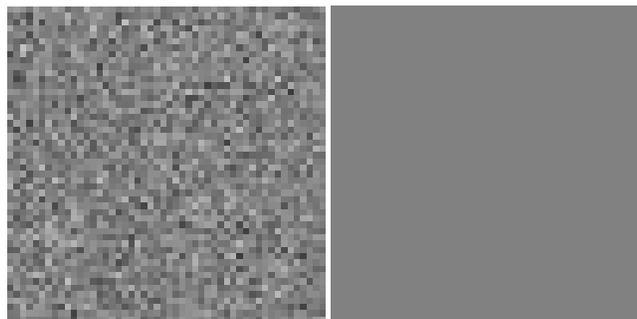
Inside the square   Diffusion outcome (PSNR = 73.54 dB)
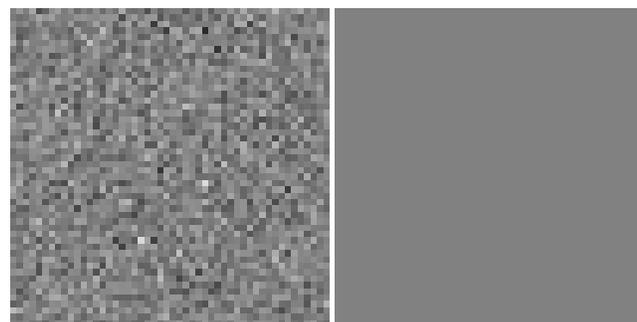
Outside the square   Diffusion outcome (PSNR = 65.71 dB)

Fig.4. Diffusion result of structurally identical patches.

The rest of the paper is organized as follows: Section 2 provides a theoretical background and introduces the method and implementation details. Section 3 presents results of the experiment; thereafter we conclude.

2. LFAD- Locally- and Feature- Adaptive Diffusion

2.1 Superpixel Segmentation

As it was pointed out earlier in this paper, we need the image to be over-segmented first. For this purpose we use superpixel segmentation. A single parameter of the method is k which is a desired number of approximately equally-sized superpixels. The procedure begins with an initialization step where k initial cluster centers C_i are sampled on regular grid space S pixels apart. To produce roughly equally sized superpixels, the grid interval is $S = \sqrt{N/k}$. The centers are moved to seed locations corresponding to the lowest gradient position in a 3x3 neighborhood, and thus avoid centering a superpixel on an edge. This reduces the chance of seeding a superpixel with a noisy pixel. Next, in the assignment step, each pixel i is associated with the nearest cluster center whose search region overlaps its location. A distance measure D, determines the nearest cluster center for each pixel. Since the expected spatial extent of a superpixel is a region of an approximate size SxS, the search for similar pixels is carried in a region of size $2Sx2S$ around the superpixel center. Once each pixel has been associated to the nearest cluster center, an update step adjusts the cluster centers to be the mean vector of all the pixels belonging to the cluster. The $L2$ norm is used to compute a residual error E between center locations of the new and the previous clusters. The assignment and update steps can be repeated iteratively until convergence. Experimentally, twenty iterations are sufficient for most images, and therefore throughout the rest of the paper we use this value.

2.2. Region Merging

Partition an image I in to sub-regions R_1, R_2,…, R_n. The following properties must hold true.

1. $R_1 \cup R_2 \cup … \cup R_n = I$

2. R_i is connected

3. $R_i \cap R_j$ is empty.

Algorithm:

1. From initial regions in the Image using superpixel segmentation.

2. For each region do:

    a. Consider its adjacent region and test to see if they are similar.

    b. For regions that are similar merge them i.e. merge R_i and $R_j \leq \alpha * \sigma^2$

3. Repeat step 2 with increasing α until all the regions are merged.

2.3. Modified Diffusion

The normalized inverse difference moment (IDM) feature captures texture details in both coarse and fine structures. IDM will get small contributions from non-homogenous region and larger values in homogenous regions. Ranging between 0 and 1; the value being 0 has an indication of a pixel being a part of a homogenous neighborhood. The value being 1 indicates that the pixel is a part of texture or an object boundary. One example of the visualized IDM feature is provided in Fig.5, wherein it is contrasted to the gradient image. Fig. 6 shows the line profile plots for both IDM and gradient values for Lena image with additive white Gaussian noise level $\sigma =40$ across the hat area region. From the results it indicates that IDM is more robust indicator compare to the gradient.

The diffusivity function of Eq.2 is modified to the following:

$$c_p = \exp\left(-\left(\frac{IDM(I)}{\lambda}\right)^2\right), p = N, S, W, E$$

,

where

$$IDM = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{1}{1+(i-j)^2} P(i, j)$$
.

Given an MxN neighborhood containing G gray levels, let f(m,n) be the intensity at sample

m, line n of the neighborhood.

Then

$$P(i, j \mid \Delta x, \Delta y) = W \cdot Q(i, j \mid \Delta x, \Delta y),$$

where

$$W = \frac{1}{(M - \Delta x)(N - \Delta y)};$$

$$Q(i, j \mid \Delta x, \Delta y) = \sum_{n=1}^{N-\Delta y} \sum_{m=1}^{M-\Delta x} A$$

and

$$A = \begin{cases} 1, iff\,(m,n) = i & and & f(m+\Delta x, n+\Delta y) = j \\ 0, elsewhere \end{cases}$$
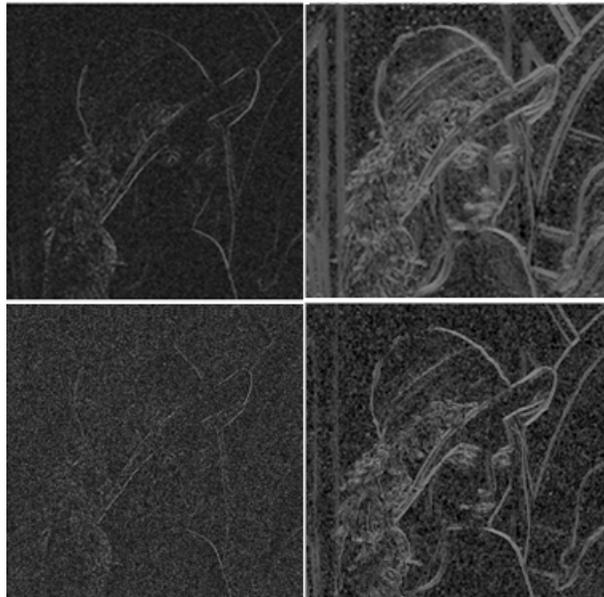
Fig.5 1st column: Gradient image for additive white Gaussian noise level σ =20, 40 for "Lena";
2nd column: Inverse difference moment (IDM) image for additive white Gaussian noise level σ=20,40. IDM is
calculated in 9x9 windows centered at pixel (i,j)
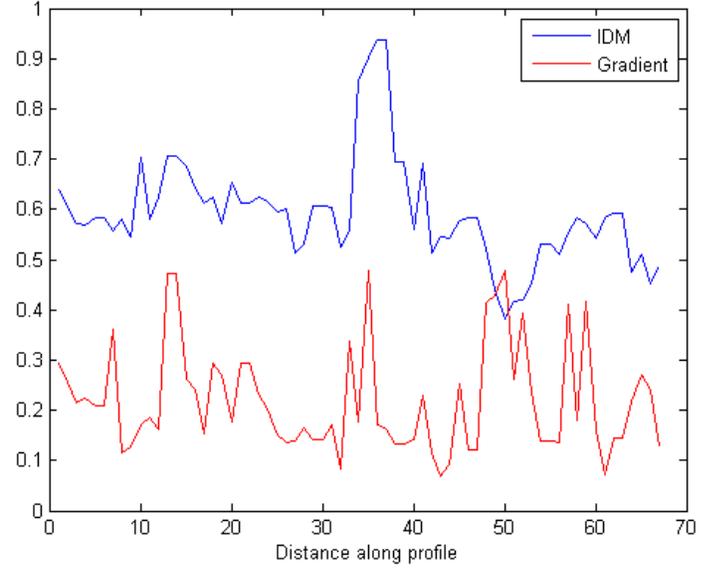
Fig.6 Left: Lena with additive white Gaussian noise level σ =40; Right: IDM and Gradient values along a line (red) segment in "Lena" Image.

2.4 LFAD Algorithm

The method performs according to the following steps:

1.   *Initialize the number of merging steps, k=0. Segment image into m (m≠1) regions using superpixel segmentation method. Calculate PSNR of the noisy image, $PSNR_1$ [$PSNR_m$ $^{(0)}$]$_{0 =}$ $PSNR_1$.*

2.   *Initialize n=0;*

3.   *Pixels of each region are diffused at an iteration step as*

$$I_{i,j}^{n+1} = I_{i,j}^{n} + (\nabla t) \bullet \left[\begin{array}{l} c_N \left(\nabla_N I_{i,j}^{n} \right) \bullet \nabla_N I_{i,j}^{n} + c_S \left(\nabla_S I_{i,j}^{n} \right) \bullet \nabla_S I_{i,j}^{n} + \\ c_E \left(\nabla_E I_{i,j}^{n} \right) \bullet \nabla_E I_{i,j}^{n} + c_W \left(\nabla_W I_{i,j}^{n} \right) \bullet \nabla_W I_{i,j}^{n} \end{array} \right] \quad (14),$$

where

$$c_p = \exp\left(-\left(\frac{IDM(I)}{\lambda} \right)^2 \right), p = N, S, W, E$$

4.   *Per region: if $[PSNR_m^{(n+1)}]_k > [PSNR_m^{(n)}]_k$, goto Step 3; elseif*

   $[PSNR_{m\text{-}}^{(0)}]_{k+1} < [PSNR_m^{(n+1)}]_k$ goto Step 7

5. *For $\forall$ pair of adjacent regions Ri and Rj, if variance of $R_i \cup R_j \leq \alpha * \sigma^2$ merge regions based on variance of neighbouring regions and threshold α=1.1; k=k+1; Update m. Goto Step 2 else Goto Step 6*

6. *α = α+0.1, Goto Step2*

7. *Stop*

3. Experimental Results

In order to verify the performance of LFAD we have tested it on a number of benchmark images degraded by the additive white Gaussian noise of zero mean μ=0 and σ = 10, 20, 30, 50 and 100. The comparison is made to other diffusion models such as PM [1], Catte [2], Li [5], LVCFAB [7], GSZFAB [26], RAAD [27]and the state of art method BM3D[15]. The evaluation is performed first based on PSNR calculated as below:

$$PSNR = 10\log\frac{I^2_{max}}{MSE},$$

where MSE is a mean square error.

Additionally we evaluate the method using the universal image quality index (UIQI) given by

$$Q = \frac{4\sigma_{xy}\,\overline{x}\,\overline{y}}{\left(\sigma_x^2 + \sigma_y^2\right)\left[\left(\overline{x}\right)^2 + \left(\overline{y}\right)^2\right]}$$

where $\overline{x}, \overline{y}$ are the means and σ_x, σ_y are the standard deviations and σ_{xy} represents the covariance. As it mentioned in [26], the average quality index UIQI coincides with the mean subjective ranks of observers.

Initial number of superpixel segments is set to 'k' = MxN/patch size; where MxN is the size of the image and the patch size is usually set globally (between 5x5 and 19x19). Levin and Nadler [23] derive bounds on how well any denoising algorithm can perform. The

bounds are dependent on the patch size, where larger patches lead to better results. For large patches and low noise, tight bounds cannot be estimated. However, Levin et al. [24] suggest that patch-based denoising can be improved mostly in smooth areas and less in textures. Chatterjee [25] studied that a smaller patch size can lead to the performance degradation from the lack of information captured by each patch, and a large patch size might capture regions of widely varying information in a single patch and also result in a fewer similar patches being present in the image. It was shown also that clusters with more patches are denoised better than clusters with fewer patches and the bound on the predicted MSE increases at different rates as the patch size grows from 5 x5 to 19 x 19 for the images, so it was concluded that a patch size of 11x11 can capture the underlying patch geometry while offering sufficient robustness in the search for similar patches. BM3D [15] uses a patch size of 8x8 for low noise levels i.e. $\sigma \leq 40$ and 11x11 for Wiener filtering and 12x12 for hard thresholding for high noise levels, i.e. $\sigma > 40$. In our work, we calculate the bounds with the patch/area size of 64 pixels for low noise levels i.e. $\sigma \leq 40$ and a larger patch/area size of 128 pixels for high noise levels i.e. $\sigma > 40$. To select the patch size among two above automatically, one can use one of available methods for estimation of the noise standard deviation. For example, one can suppress the image structure using the Laplacian mask such that the remaining part of the image is noise [29].

The diffusion equation needs the value of the diffusion constant, λ. Fig.7 displays PSNR of the outcomes of IDM based diffusion for a fixed noise level ($\sigma=50$) with different values of $\lambda= 5, 10, 15, 25$ and 50 for 1000 iterations for "Lena" image. The plot provides the indication that $\lambda=10$ is a best choice.

The above parameters were used to obtain Table I which shows PSNR values by the LFAD for benchmark images. Next, in Table II, the LFAD is compared to six diffusion based

methods which are considered the state-of- the-art techniques in diffusion based denoising. They are FAB, GSZ FAB [26], LVCFAB [7], and RAAD [27]. The improvement by LFAD for the given noise levels is ranging from 1.3 dB for low noise to 1.59dB for noise level with σ=100. It is interesting to note that the use of IDM feature helped with improving PSNR compared to the reference PM method by 0.86dB on average, from 0.65db for low noise to 1.03 dB for high noise. The proposed method outperforms all other diffusion models. The comparison to BM3D shows that the performance of LFAD is 0.35 dB lower compared to that of the BM3D for noise level σ=10 and 0.39 dB lower for noise level σ=100. Results for BM3D are publicly available at http://www.cs.tut.fi/~foi/GCF-BM3D/index.html#ref_results and therefore are not reproduced here. Table III provides UIQI values by the LFAD and BM3D and Table IV provides UIQI values by the proposed method and other diffusion models for same benchmark images. Fig.8 shows that lesser or no blocking/ringing artifacts are introduced by LFAD compared to those in BM3D denoised images. The denoising performance of the LFAD is further illustrated in Fig. 9 and Fig. 10, where we show fragments of a few noisy (σ=10, 20, 30 and 50) test images and fragments of the corresponding denoised ones. The denoised images show high visual quality in the areas of smooth intensity transition and lesser or no ringing around contours of extended objects.

4. Conclusion

We have proposed a new diffusion-based method of image denoising. The high performance of the method is attained due to the following properties: a) patch-based optimization of PSNR through iterative diffusion; b) agglomeration of patches and repetitive iteration of the process; c) modification of the diffusion function with IDM feature. The method has attained a highest performance in the class of advanced diffusion based methods. Being slightly inferior to the state-of-the-art BM3D method by yielded PSNR levels it however outperforms

the counterpart by reducing visible blocking and ringing artifacts generally inherent to block- and transform-based methods.

Acknowledgement
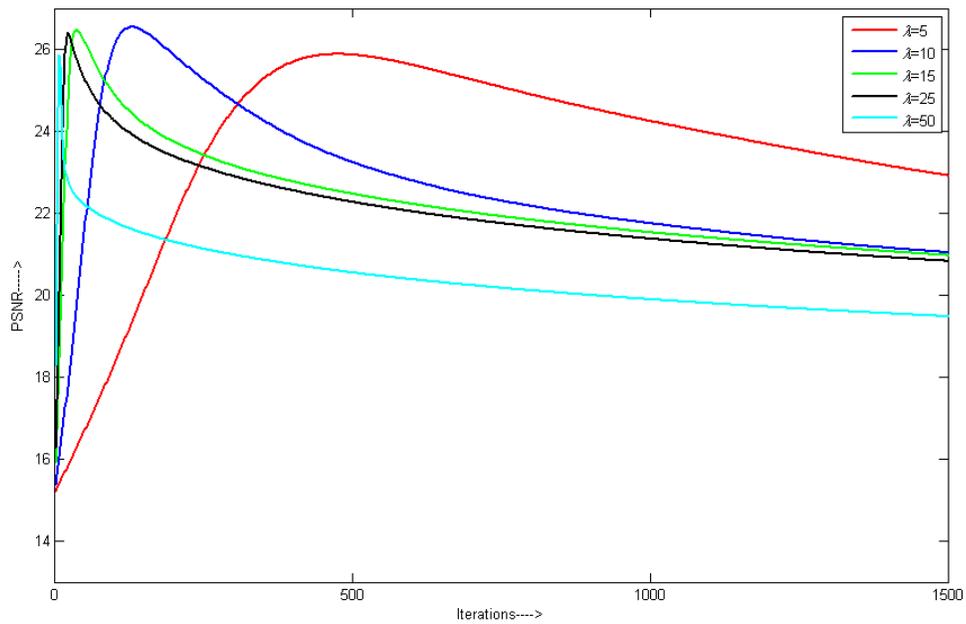
Fig.7. PSNR obtained using IDM with λ = 5, 10, 15, 25 and 50 with a noise level σ=50 for Lena image

Table I. PSNR of the proposed method

Image/Noise, σ	LFAD				
	10	20	30	50	100
Lena	35.56	32.61	30.85	28.59	25.56
House	35.94	32.93	31.11	28.68	25.12
Peppers	34.48	31.05	29.03	26.56	23.18
Cameraman	33.99	30.18	28.24	25.89	23.08

Table II. PSNR comparison of different anisotropic diffusion methods for Lena

Method/ σ	10	15	20
Noisy	28.15	24.62	22.14
PM [1]	32.70	30.71	29.37
Catte [2]	33.27	31.39	30.09
Li [5]	34.28	32.41	31.15
GSZ FAB [26]	32.49	29.86	28.29
LVCFAB [7]	31.90	28.21	26.67
RAAD [27]	34.33	32.53	31.24
LFAD	**35.56**	**33.86**	**32.61**

Table III. UIQI comparison of BM3D and LAFD methods.

	10		20		30		50		100	
	BM3D	LFAD	BM3D	LFAD	BM3D	LFAD	BM3D	LFAD	BM3D	LFAD
Lena	0.6976	0.6903	0.6107	0.5991	0.5489	0.5391	0.4661	0.4566	0.3440	0.3427
House	0.5894	0.5640	0.4505	0.4296	0.4001	0.3810	0.3443	0.3224	0.2691	0.2411
Peppers	0.8182	0.8148	0.7443	0.7361	0.6863	0.6777	0.6016	0.5931	0.4664	**0.4682**
Cameraman	0.5975	0.5908	0.4914	0.4908	0.4319	0.4275	0.3567	0.3496	0.2600	0.2383

Table IV. UIQI comparison of different anisotropic diffusion methods

Scheme	Image	10	15	20
GSZ FAB	Lena	0.6294	0.5391	0.4833
	Peppers	0.592	0.5237	0.4682
	Cameraman	0.539	0.4333	0.3789
LVCFAB	Lena	0.6309	0.4955	0.4337
	Peppers	0.5883	0.4819	0.4236
	Cameraman	0.5441	0.3967	0.3413
RAAD	Lena	0.6819	0.6232	0.5749
	Peppers	0.6325	0.5764	0.5395
	Cameraman	**0.5994**	0.5199	0.4622
LFAD	Lena	**0.6903**	**0.6396**	**0.5991**
	Peppers	**0.8148**	**0.7708**	**0.7361**
	Cameraman	0.5908	**0.5314**	**0.4908**

Fig.8. First row: "Lena" image and that with additive white Gaussian noise level σ =100;

Second row: results by BM3D and LFAD. Arrows show areas where LFAD performs comparatively better than BM3D

Fig.9. First Column: "Lena" image with additive white Gaussian noise level σ =10, 20, 30 and 50;

Second Column: corresponding results by LFAD.

Fig.10. First Column: "Peppers" image with additive white Gaussian noise level σ =10, 20, 30 and 50;
Second Column: corresponding results by LFAD

References:

1. P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, no. 7, pp. 629–639, 1990.

2. F. Catté, P.-L. Lions, J.-M. Morel, and T. Coll, "Image selective smoothing and edge detection by nonlinear diffusion," SIAM Journal on Numerical Analysis, vol. 29, no. 1, pp. 182–193, 1992.

3. V. B. Surya Prasath and Arindama Singh, "Well-Posed Inhomogeneous Nonlinear Diffusion Scheme for Digital Image Denoising," Journal of Applied Mathematics, vol. 2010, Article ID 763847, 14 pages, 2010. doi:10.1155/2010/763847

4. Jinhua Yu, Jinglu Tan, Yuanyuan Wang. Ultrasound speckle reduction by a SUSAN-controlled anisotropic diffusion method. Pattern Recognition, 2010: 3083-3092.

5. H. C. Li, P. Z. Fan, and M. K. Khan, "Context-Adaptive Anisotropic Diffusion for Image Denoising," IET Electronics Letters, vol.48, no.14, pp.827-829, 2012.

6. Shin-Min Chao, Du-Ming Tsai: An improved anisotropic diffusion model for detail- and edge-preserving smoothing. Pattern Recognition Letters 31(13): 2012-2023 (2010).

7. Y. Wang, L. Zhang, P. Li. Local Variance-Controlled Forward-and-Backward Diffusion for Image Enhancement and Noise Reduction. IEEE Transactions on Image Processing, 2007: pp.1854-1864.

8. Y. Yu and S.T. Acton, "Speckle reducing anisotropic diffusion," IEEE Transactions on Image Processing, vol. 11, pp. 1260-1270, 2002.

9. Zhang Fan, Mo Yoo Yang, Mong Koh Liang, and Kim Yongmin, "Nonlinear Diffusion in Laplacian Pyramid Domain for Ultrasonic Speckle Reduction," IEEE Trans. Med. Imaging. 26, 200-211 (2007).

10. Adb-Elmoniem K. Z., **,** Youssef A. M., **, and** Kadah Y. M.**,** " Real-time speckle reduction and coherence enhancement in ultrasound imaging via nonlinear anisotropic diffusion. ," IEEE Trans. Biomed. Eng.. 0018-9294 49, , 997–1014 ((2002)).

11. Ajay K. Mandava and Emma E. Regentova, "Image denoising based on adaptive nonlinear diffusion in wavelet domain", J. Electron. Imaging 20, 033016 (Sep 14, 2011)

12. Shih, A.C.-C., Liao, H.-Y.M., Lu, C.-S.: A New Iterated Two-Band Diffusion Equation: Theory and Its Applications. IEEE Transactions on Image Processing (2003) DOI: 10.1109/TIP. 2003.809017.

13. M. Aharon, M. Elad and A. M. Bruckstein, "The K-SVD: An Algorithm for Designing of Overcomplete Dictionaries for Sparse Representation," IEEE Transactions on Signal Processing, Vol. 54, No. 11, 2006.

14. P. Chatterjee and P. Milanfar, Clustering-based Denoising with Locally Learned Dictionaries (K-LLD), IEEE Transactions on Image Processing, vol. 18, num. 7, July 2009, pp. 1438-1451

15. K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080-2095, August 2007.

16. J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-Local Sparse Models for Image Restoration," *Proc. of ICCV*, September-October 2009.

17. Gilboa, G., Osher, S.: Nonlocal operators with applications to image processing. Multiscale Model. Simul. **7**(3), 1005-1028 (2008).

18. Peyre*, G., "*Image Processing with Non-Local Spectral Bases*" SIAM *Journal on Multiscale. Modeling and Simulation 7, 2 (2008) 703-730.*

19. X. Zhang, M. Burger, X. Bresson, and S. Osher, "Bregmanized Nonlocal Regularization for Deconvolution and Sparse Reconstruction", SIAM Journal on Imaging Sciences, 3(3), 253-276, 2010

20. K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "A nonlocal and shape-adaptive transform-domain collaborative filtering", Proc. Int. Workshop on Local and Non-Local Approx. in Image Process., LNLA 2008, Lausanne, Switzerland, August 2008.

21. Charles-Alban Deledalle, Vincent Duval and Joseph Salmon, *Non-Local Methods with Shape-Adaptive Patches (NLM-SAP)*, Journal of Mathematical Imaging and Vision, pp. 1-18, 2011

22. Achanta, R.; Shaji, A.; Smith, K.; Lucchi, A.; Fua, P.; Süsstrunk, S.; , "SLIC Superpixels Compared to State-of-the-Art Superpixel Methods," *Pattern Analysis and Machine Intelligence, IEEE Transactions on* , vol.34, no.11, pp.2274-2282, Nov. 2012

23. A. Levin and B. Nadler. Natural Image Denoising: Optimality and Inherent Bounds. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011.

24. A. Levin, B. Nadler, F. Durand, and W.T. Freeman. Patch complexity, finite pixel correlations and optimal denoising. In European Conference on Computer Vision (ECCV), 2012.

25. P. Chatterjee and P. Milanfar, "*Is Denoising Dead?*", IEEE Trans. on Img. Proc., vol 19, num. 4, Apr. 2010, pp 895-911.

26. G. Gilboa, N. Sochen, and Y. Y. Zeevi, "Forward-and-backward diffusion processes for adaptive image enhancement and denoising," *IEEE Trans. Image Process.* **11**(7), 689–703 (2002).

27. Y. Wang , R. Niu , L. Zhang and H. Shen "Region-based adaptive anisotropic diffusion for image enhancement and denoising", *Opt. Eng*. 49(11), 117007 (2010).

28. Cooper, G.R.J., "The Antialiased Textural Analysis of Aeromagnetic Data", Computers & Geosciences v.35, p.586-591 (2009).

29. J. Immerkær, "Fast Noise Variance Estimation", Computer Vision and Image Understanding, Vol. 64, No. 2, pp. 300-302, Sep. 1996