



ELSEVIER

Contents lists available at [SciVerse ScienceDirect](http://SciVerse.ScienceDirect.com)

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

On the choice of the parameters for anisotropic diffusion in image processing

Chourmouzos Tsotsios^{a,*}, Maria Petrou^{b,1}^a Department of Electrical and Electronic Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, UK^b Informatics and Telematics Institute, CERTH, Thessaloniki, Greece

ARTICLE INFO

Article history:

Received 19 March 2012

Received in revised form

10 September 2012

Accepted 5 November 2012

Keywords:

Anisotropic diffusion

Image denoising

Stopping criterion

Image quality measures

ABSTRACT

Anisotropic diffusion filtering is highly dependent on some crucial parameters, such as the conductance function, the gradient threshold parameter and the stopping time of the iterative process. The various alternative options at each stage of the algorithm are examined and evaluated and the best choice is selected. An automatic stopping criterion is proposed, that takes into consideration the quality of the preserved edges as opposed to just the level of smoothing achieved. The proposed scheme is evaluated with the help of real and simulated images, and compared with other state of the art schemes using objective criteria.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Based on the importance of the scale-space representation of images, which was introduced by Witkin [1], Perona and Malik suggested a new definition of scale-space through Anisotropic Diffusion (AD), a non-linear partial differential equation-based diffusion process [2]. Overcoming the undesirable effects of linear smoothing filtering, such as blurring or dislocating the semantically meaningful edges of the image, AD has become a very useful tool in image smoothing, edge detection, image segmentation and image enhancement. AD filtering can successfully smooth noise while respecting the region boundaries and small structures within the image, as long as some of its crucial parameters are determined or estimated correctly. The conductance function, the gradient threshold parameter and the stopping parameter form a set of parameters which define the behavior and the extent of the diffusion. Overestimating one of the parameters may lead to an oversmoothed blurry result, while underestimating it may leave the noise in the image unfiltered. Therefore, it is crucial that all parameters are determined in an optimal and automatic way in every step of the iterative process, by evaluating both the denoising needs and the quality of the edges of a given image.

Over the last years, a great amount of work has been done with respect to both the continuous and discrete form of AD filtering as stated by Perona and Malik [2]. As for the behavior of the continuous form of AD, there has been a considerable amount of

research proving the ill-posedness of the diffusion equation and developing new well-posed equations or regularizing methods [3–11]. In [12,13] a semi-implicit scheme was presented, while in [14] a different discrete implementation was proposed in order to obtain better isotropy. A multigrid approach leading to a well-posed, steady-state solution was proposed in [11]. In [15] a time-dependent numerical scheme was proposed. In [16] a modified method that considers also the variance of the brightness levels in a local neighborhood around each pixel was presented. However, the problem of the automatic estimation of the crucial parameters was not addressed. A modified diffusion scheme, suitable for images with low-contrast and uneven illumination, was described in [17]. In [18,19] the attention was drawn mostly to the discrete implementation of the scheme and the experimental results of the new conductance functions that were proposed. The automatic estimation of the method's parameters was also studied in these works. Since several conductance functions can be used, differentiating considerably the filtering results as shown in [18], it is necessary to define the appropriate one and scale it in a way that the edges remain the sharpest possible. The importance of the scaling of the conductance function is emphasized in the current work, leading to a comparison and selection of the most edge-preserving function.

In [20] the necessity of the gradient threshold parameter to be a decreasing function of time was first shown. In this way, the parameter adapts itself to the denoising needs of the filtered image after every iteration, preserving all the edges above a decreasing threshold. Various methods estimating this parameter were proposed using statistical characteristics of the image [2,18,19] and morphological operators [21]. These methods are compared in the current work, along with a proposed statistical

* Corresponding author. Tel.: +44 20 7594 6267.

E-mail addresses: c.tsotsios@imperial.ac.uk (C. Tsotsios),petrou@iti.gr (M. Petrou).¹ Tel.: +30 2311257700.

method which estimates two gradient threshold parameters and yields robust filtering results.

Since AD is an iterative process, the problem of choosing the optimal time to stop the iterations and prevent an oversmoothed result is crucial. Adding a fidelity term that keeps the resultant image close to the original image has been proposed in [8,22,23] but the noise in that case had not been sufficiently removed. In [24], a stabilized, insensitive to the number of iterations process was introduced. A multigrid algorithm was presented and evaluated in [11], introducing a Brent-NCP (Normalized Cumulative Periodogram) automatic stopping parameter method. A frequency approach of the problem was presented in [25]. Some criteria estimating a stopping parameter have been introduced in [26–29] based solely on the extend of the noise smoothing of the filter in every iteration. Spatially varying stopping methods that increase significantly the computational cost were presented in [30,31]. The quality of the preserved edges was not considered in any of the above methods. As it is shown in the present work, the evaluation of the image's edges is strongly related to the stopping time estimation problem. A novel automatic stopping criterion based on this approach is described and evaluated.

Anisotropic diffusion has been widely used in biomedical imaging [14,32–35]. Plenty of applications where nonlinear diffusion filtering has also been used can be found in [9].

The goal of the current work is to investigate fully the role of the parameters on the quality of the results of the AD discrete scheme and propose novel methodology for their automatic adaptation, as well as a novel termination criterion for the whole iterative process, so that the final denoised result is optimal.

This paper is organized as follows: in Section 2 we present the critical steps of the method and perform a comparative evaluation of the various proposed options. The right choice and scaling of the conductance function and the methods for estimating the gradient threshold parameter are considered in order to come up with the optimal automatic discrete scheme. In Section 3 we present a novel stopping criterion based on the quality of the image's edges and in Section 4 we evaluate the scheme using a set of natural images. The discussion and concluding remarks are presented in Section 5.

2. Anisotropic diffusion

2.1. Overview of the process

The basic equation of anisotropic diffusion equation as presented in [2] is

$$\frac{\partial I(x,y,t)}{\partial t} = \text{div}[g(\|\nabla I(x,y,t)\|)\nabla I(x,y,t)] \quad (1)$$

where t is the time parameter, $I(x,y,0)$ is the original image, $\nabla I(x,y,t)$ is the gradient of the version of the image at time t and $g(\cdot)$ is the so-called conductance function. This function is chosen to satisfy $\lim_{x \rightarrow 0} g(x) = 1$, so that the diffusion is maximal within uniform regions, and $\lim_{x \rightarrow \infty} g(x) = 0$, so that the diffusion is stopped across edges. Two such functions proposed by Perona and Malik were

$$g_1(x) = \exp\left[-\left(\frac{x}{K}\right)^2\right] \quad (2)$$

and

$$g_2(x) = \frac{1}{1 + \left(\frac{x}{K}\right)^2} \quad (3)$$

where K is the gradient magnitude threshold parameter that controls the rate of the diffusion and serves as a soft threshold

between the image gradients that are attributed to noise and those attributed to edges. Black et al. in [18], through an interpretation of AD in terms of robust statistics, defined a different conductance function, called Tukey's biweight function

$$g_3(x) = \begin{cases} \frac{1}{2} \left[1 - \left(\frac{x}{S} \right)^2 \right]^2, & x \leq S \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

where $S = K\sqrt{2}$.

The flow function ϕ defined as

$$\phi(x) \equiv g(x)x \quad (5)$$

represents the sum of the brightness flow that is generated. The maximum flow is generated at locations where $|\nabla I| = K$.

Perona and Malik discretized their anisotropic diffusion equation to

$$I_{t+1}(s) = I_t(s) + \frac{\lambda}{|\eta_s|} \sum_{p \in \eta_s} g_k(|\nabla I_{s,p}|) \nabla I_{s,p} \quad (6)$$

where I is a discretely sampled image, s denotes the pixel position in the discrete 2-D grid, t denotes the iteration step, g is the conductance function and K is the gradient threshold parameter. Constant $\lambda \in (0,1]$ determines the rate of diffusion and η_s represents the spatial 4-pixel neighborhood of pixel s : $\eta_s = \{N, S, E, W\}$, where N, S, E and W are the North, South, East and West neighbors of pixel s , respectively. Consequently, $|\eta_s|$ is equal to 4 (except for the image borders). The symbol ∇ which in the continuous form is used for the gradient operator, now represents a scalar defined as the difference between neighbouring pixels in each direction:

$$\nabla I_{s,p} = I_t(p) - I_t(s), \quad p \in \eta_s = \{N, S, E, W\}. \quad (7)$$

As mentioned by Perona and Malik, this scheme is not the exact discretization of the continuous equation, with more numerically consistent methods having been proposed in the literature. However, it is favoured due to its low computational complexity, preserving most of the properties of the continuous form [2].

2.2. Choosing the conductance function

According to Perona and Malik [2], the g_1 conductance function favours high-contrast edges over low-contrast ones, while the g_2 function favours wide regions over smaller ones. The g_3 function, according to Black et al. [18], yields sharper edges improving considerably the experimental results of the filtering, since the diffusion process converges faster. We will focus on the scaling and comparison done by Black et al. in order to examine the different experimental behavior of the conductance functions.

In order to be able to compare their efficiency, Black et al. scaled the conductance functions g_1 , g_2 and g_3 , so that their respective flow functions, given by (5), reach the same maximum value (producing the same amount of brightness flow) at the same point $x=K$ (Fig. 1a). This leads us to the following scaled conductance functions:

$$g_1(x) = \exp\left[-\left(\frac{x}{K\sqrt{2}}\right)^2\right] \quad (8)$$

$$g_2(x) = \frac{1}{1 + \left(\frac{x}{K}\right)^2} \quad (9)$$

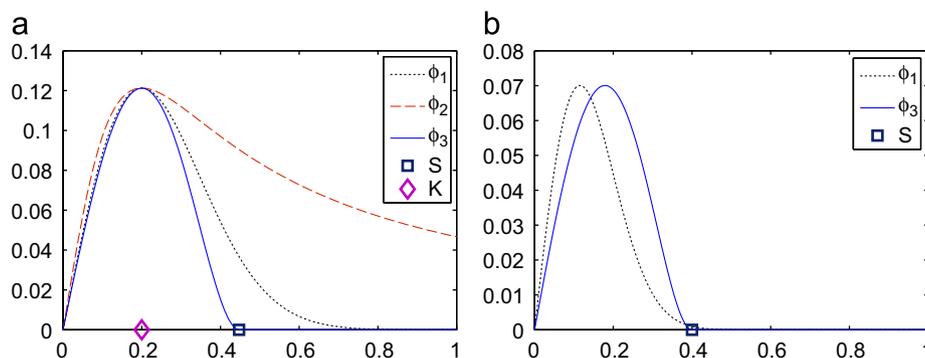


Fig. 1. The flow functions (a) ϕ_1 , ϕ_2 and ϕ_3 aligned and scaled so that the flow is maximum at the same point K . (b) ϕ_1 and ϕ_3 aligned and scaled so that the flow reaches zero near the same point S .



Fig. 2. (a) A zoomed part of the original Lena image. (b) The AD filtering result after 20 iterations using g_1 . (c) The AD filtering result after 20 iterations using g_3 .

$$g_3(x) = \begin{cases} 0.67 \left[1 - \left(\frac{x}{K\sqrt{5}} \right)^2 \right]^2, & x \leq K\sqrt{5} \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

As it can be easily noticed in Fig. 1a, using the Perona and Malik functions g_1 and g_2 , the flow continues to take place and smooth the image, while using Tukey's biweight g_3 function, the flow descends more rapidly and stops the diffusion, protecting this way the edges from becoming oversmoothed and blurred. Assuming, in this case, that an image is characterized only by strong edges above a certain threshold (the point $x=S$ where g_3 reaches zero), g_3 function will leave the edges untouched while the other functions will not.

This function scaling and comparison favours g_3 function which descends faster, preventing an amount of edges above a certain threshold S from becoming smoothed away. S treated as the boundary between noise and edges in practice means that the local gradients below S will be smoothed and those above S , treated as outliers in [18], will be preserved since the diffusion is stopped. Consequently, in order to compare the behavior of the different conductance functions, it is suggested that they are scaled so that the respective conductance and flow functions tend to zero at the same point S . Since the g_2 function descends very slowly to zero, it is regarded to be more noise smoothing-efficient than edge-preserving. This leads us to a comparison between the two other functions. The g_1 function does not reach zero until ∞ either, hence it is scaled so that it reaches a very small value at point S , where the g_3 function reaches zero, while both functions are aligned so that their flow functions generate the same maximum diffusion value (Fig. 1b). The small but finite value g_1 has to reach at S is important for the shape of the function. As we are working in the digital domain, where brightness is quantised into 256 levels, we may say that the implied digital 0 is equal to $0.5/256 \sim 0.002$. However, image enhancement is the subjective

improvement of an image, and so one has to take into consideration also the way humans perceive grey tone differences. It is known that humans on average cannot distinguish grey levels that differ for less than $\sim 2-3$ levels. So, we may consider that, for practical purposes, $g_1(S)$ is 0 when it takes the value $\sim 0.006 \sim \exp(-5)$. This corresponds to $g_1(S) < 0.01 \max\{g_1\}$. The conductance functions are now set to be

$$g_1(x) = \exp \left[- \left(\frac{x\sqrt{5}}{S} \right)^2 \right] \quad (11)$$

$$g_3(x) = \begin{cases} 0.67 \left[1 - \left(\frac{x}{S} \right)^2 \right]^2, & x \leq S \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

As it can be seen from Fig. 1b, ϕ_1 now descends faster and is expected to result in sharper discontinuities. The experimental results of diffusing the Lena image using the two functions can be seen in Fig. 2. The value of S was kept fixed to $S=0.1$ and $T=20$ iterations were conducted. It can be seen that the g_1 function produces sharper boundaries in some parts of the image.

Using the synthetic image of Fig. 3a and adding Gaussian noise of zero mean and different standard deviations $N(0, \sigma^2)$, we were able to measure the PSNR value of the resultant denoised image after diffusing with g_1 and g_3 . The PSNR is defined as

$$PSNR = 10 \log_{10} \frac{\max(I_0)^2}{MSE} \quad (13)$$

where $MSE = (1/rc) \sum_{i=1}^r \sum_{j=1}^c (I_t - I_0)^2$. I_t is the filtered image in iteration t , I_0 is the original noise-free synthetic image and r, c indicate the number of rows and columns of the image, respectively. The results show the advantage of g_1 over Tukey's biweight g_3 .

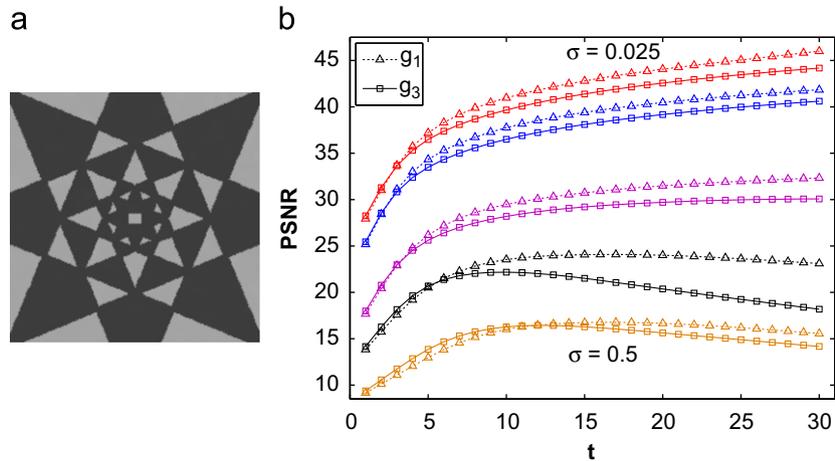


Fig. 3. (a) An original synthetic image. (b) Noise with different values of σ was added to the synthetic image and the PSNR of the diffused resultant images was measured in every iteration, using g_1 and g_3 . $\sigma = 0.025, 0.05, 0.075, 0.1, 0.5$, from top to bottom, respectively.

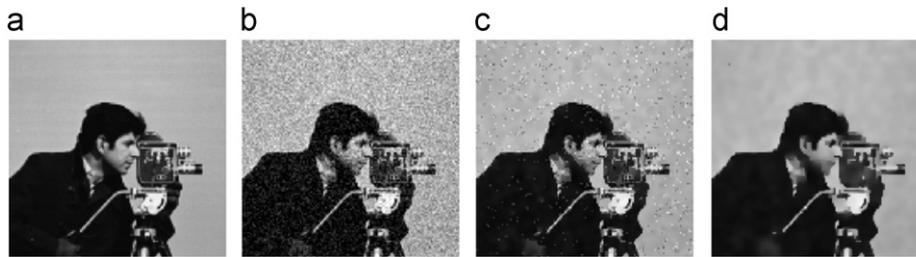


Fig. 4. (a) A zoomed part of the original Cameraman image. (b) A noisy version of it (Gaussian noise with $\sigma = 0.1$). (c) The basic AD filtering result after 15 iterations. (d) The AD filtering result after 15 iterations, by estimating the local gradients from a smoothed version of the image.

2.3. Calculating the gradient

The basic AD scheme has a good edge-preserving behavior after choosing the right conductance function, but is incapable of denoising efficiently images with high levels of noise. This problem lies in the fact that the image gradient is not a reliable measure since it is susceptible to noise [4,20,36]. A response to this problem is to replace term $g(\|\nabla I(x,y,t)\|)$ in (1) with $g(\|\nabla(G_\sigma * I(x,y,t))\|)$, where G_σ is a Gaussian filter of scale σ . This means that the local gradients that are the argument of the conductance function are now computed using a smoothed version of the image in every iteration.

In order to estimate automatically scale σ of the Gaussian filter according to the level of the Gaussian noise within an image, a sliding window of size between 25×25 and 64×64 pixels is used (in order to have enough pixels for reliable statistical calculation), so that the most uniform block of pixels within the image is detected. The uniformity measure used is the standard deviation of the pixels within each block. Finally, the standard deviation of the most uniform block is considered to be the scale σ of the Gaussian filter. From the value of σ the size of the smoothing Gaussian filter is then determined as described in [37]. As it can be seen from Fig. 4, computing the local gradients on a smoothed version of the noisy image can eliminate successfully the artifacts that are preserved in the resultant image when using the basic AD scheme.

2.4. Estimating the gradient threshold parameter

The estimation of the gradient threshold parameter plays a major role in the diffusion process, since it defines the threshold between the image gradients that are attributed to noise and

those attributed to true edges. Although in the literature the gradient threshold parameter is mentioned as parameter K , in the current work it will be mentioned as parameter S . As it was described in Section 2.2, S is the point where the conductance function reaches zero and the diffusion is stopped, preserving true edges. Some methods that have been proposed in order to estimate the gradient threshold parameter are the following:

1. Perona and Malik in [2] suggested the use of the “noise estimator” described by Canny [38], where a histogram of the absolute values of the gradient throughout the image is computed and S is set equal to the 90% value of its integral (cumulative sum) in every iteration.
2. Black et al. in [18] defined $S = \sigma_e \sqrt{5} = 1.4826 \text{MAD}(\nabla I)$, where MAD denotes the median absolute variation and is defined as $\text{MAD} = \text{median}(\|\nabla I - \text{median}(\|\nabla I\|)\|)$.
3. Voci et al. in [21] used the p-norm of the image and defined $S = \sigma \|I\|_p / rc$. Here σ is a constant, proportional to the image average intensity, r and c are the number of rows and columns of the image respectively, and $\|I\|_p$ is the p-norm of the image defined as $\|I\|_p = \left(\sum_{i,j \in I} |I|^p\right)^{1/p}$.
4. Voci et al. in the same work, used a morphological approach to estimate S as follows: $S = \sum_{i,j \in I} I \circ st / rc - \sum_{i,j \in I} I \bullet st / rc$, where st is a structuring element (usually 3×3 or 5×5) and the symbols “ \circ ” and “ \bullet ” represent the opening and closing operators, respectively.

The first two of the above methods are based on the use of the absolute value of the image gradient in order to estimate the gradient threshold parameter. We notice in (14) that in the discrete AD scheme, for every pixel s in the image there are four

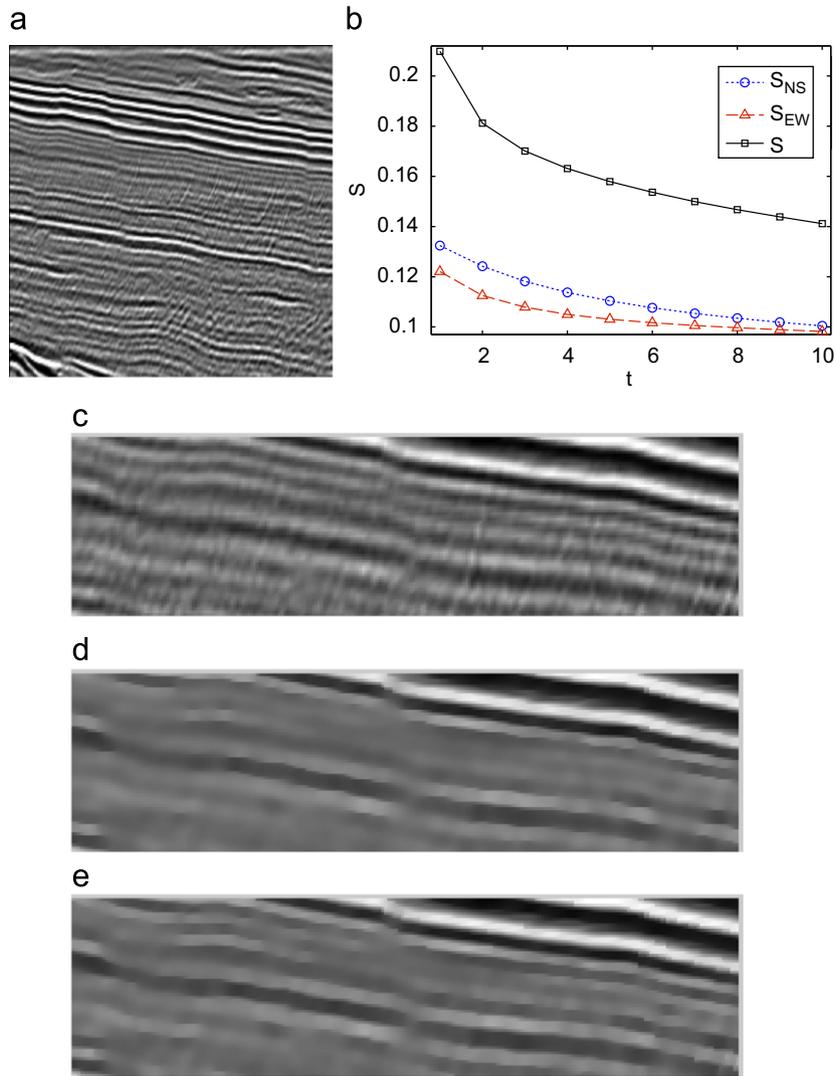


Fig. 5. (a) A seismic image. (b) The estimated gradient threshold parameter S and the estimated gradient threshold parameters S_{NS} and S_{EW} , for the seismic image in every iteration, using the knee algorithm. (c) A zoomed version of the seismic image (d) The filtered version after 10 iterations, estimating one gradient threshold parameter S . (e) The filtered version after 10 iterations, estimating two gradient threshold parameters S_{NS} and S_{EW} .

difference values, instead of one, which are defined in (7) as the difference between the brightness value of the pixel s and the brightness values of each one of its four neighbors in the 4-pixel neighborhood η_s . This happens because the ∇ operator indicates a scalar defined as the difference between the neighboring pixels in the discrete implementation, rather than a single gradient vector as in the continuous form (Section 2.1). This leads us to the idea that four difference threshold parameters should be used, each one estimated using the respective difference along the four directions. However, given the region of the entire image, *i.e.* in the statistical sense, the absolute values of the North and South differences are almost equal, while the same happens with the East and West ones. Thus, we propose the estimation of two gradient threshold parameters. These will be an S_{NS} parameter, which refers to the vertical (North–South) direction, and an S_{EW} parameter, which refers to the horizontal (East–West) direction. This changes the discrete anisotropic diffusion equation (14) to:

$$I_{t+1}(s) = I_t(s) + \frac{\lambda}{|\eta_s|} \left[\sum_{p \in N,S} \overbrace{g(\nabla I_{s,p})}^{S_{NS}} \nabla I_{s,p} + \sum_{p \in E,W} \overbrace{g(\nabla I_{s,p})}^{S_{EW}} \nabla I_{s,p} \right] \quad (14)$$

Estimating two gradient threshold parameters forms the discrete AD equation in a more precise way, given that there is not only one single gradient vector as in the continuous form. Furthermore, it is expected to lead to better experimental results, since the extent of the smoothing that takes place (defined by parameter S) is not the same for both directions. It differs according to the strength of the differences in each direction. Especially in images where the edges are orientated more towards one of the two directions, the two gradient threshold parameters may differ considerably, since the stronger differences in one direction would lead to an estimation of a higher S parameter in that direction.

In order to estimate the two gradient threshold parameters S_{NS} and S_{EW} , we use the corresponding histograms of the absolute values of the gradient (difference) component in each direction and we employ the so-called knee algorithm. The knee algorithm, which is used to estimate the threshold between two populations in histograms with one peak and a long tail, fits with straight lines, in the least square error sense, each population leading to the estimation of the threshold after an iterative process. A detailed description of the algorithm can be found in [37]. In our case, the population that has a roughly flat distribution and creates the long tail is that of the



Fig. 6. (a) A zoomed part of the original Lena image. (b) A noisy version of it (Gaussian noise with $\sigma = 0.07$). The AD filtering result after 15 iterations, using (c) the proposed method, (d) Perona–Malik method, (e) Black et al. method, (f) Voci et al., 2-norm method and (g) Voci et al., morphological method.

differences attributed to true edges, while the population with the steeper distribution is that of the differences attributed to noise. The S parameter is the estimated threshold between the two populations.

As it can be noticed in Fig. 5, the S_{NS} estimated in every iteration for the seismic image is higher than the S_{EW} , since most of the edges in the image are orientated more towards the horizontal direction. Fig. 5c–5e show the resultant diffused images after 10 iterations, estimating only one gradient threshold parameter and two gradient threshold parameters, respectively. The edges are preserved better in the second case. It can be seen that estimating one value for S seems to overestimate the threshold between noise and true edges. The threshold value was not overestimated only in images with structured content. The overestimation was evident in all natural images that were evaluated, leading to a quicker degradation of the image's edges.

Fig. 6 shows the resultant images, after applying AD filtering to a noisy version of the Lena image. The gradient threshold parameters were estimated in every iteration, using one of the above described methods in each case. The methods of Black et al. and the morphological approach of Voci et al. seem to underestimate the gradient threshold parameter since the noise in the resultant images is not smoothed properly. The Voci et al. p-norm and the Perona–Malik methods, on the contrary, oversmooth the image since some of the edges are smoothed away. The proposed method leads to the sharpest discontinuities, while the noise is sufficiently smoothed.

Using the synthetic image of Fig. 3a and adding Gaussian noise of zero mean and different standard deviations, we measured the PSNR value of the resultant image using (13), after using the above described methods. The PSNR was measured after 10 iterations, in every case. The result is shown in Fig. 7. The proposed method outperforms the other methods and turns out to be the most robust scheme, since it yields a better PSNR for all the different levels of noise added. The results were similar after 20 and 30 iterations.

3. Stopping criterion

Since AD filtering is an iterative process, it is highly sensitive to the number of iterations. The choice of the stopping time T is

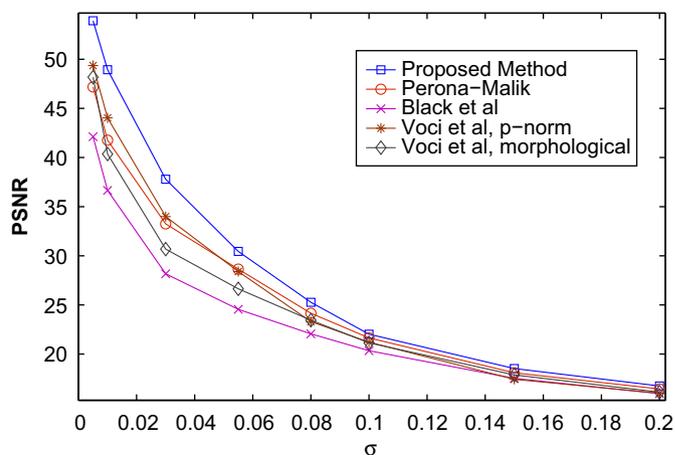


Fig. 7. Noise with different values of σ was added to the synthetic image of Fig. 3a and the PSNR of the diffused resultant images was measured after 10 iterations, using the described gradient threshold estimating methods.

crucial, since overestimating it may result in blurring the true edges, while underestimating it may leave unfiltered noise artifacts. Fig. 8 shows the measured PSNR in every iteration for an airplane image, using the scheme that was proposed in the previous sections. Choosing the optimal conductance function and gradient thresholding parameters leads to higher PSNR values as was previously proven and in a way ensures that due to the adaptability of the method, the slope of the decrease of PSNR will be low, keeping the diffused image close to the original one. However, PSNR is always maximized in a specific iteration, which is the time when the process should ideally be terminated ($T=26$ for the airplane image in Fig. 8). Obviously, in typical image processing problems the original noise-free version of the image is not known a priori. Therefore, the optimal stopping time T should be estimated using only the statistics of each filtered version of the noisy image.

A comprehensive study of the stopping time problem was done in [26,27], where two global stopping criteria were proposed and tested

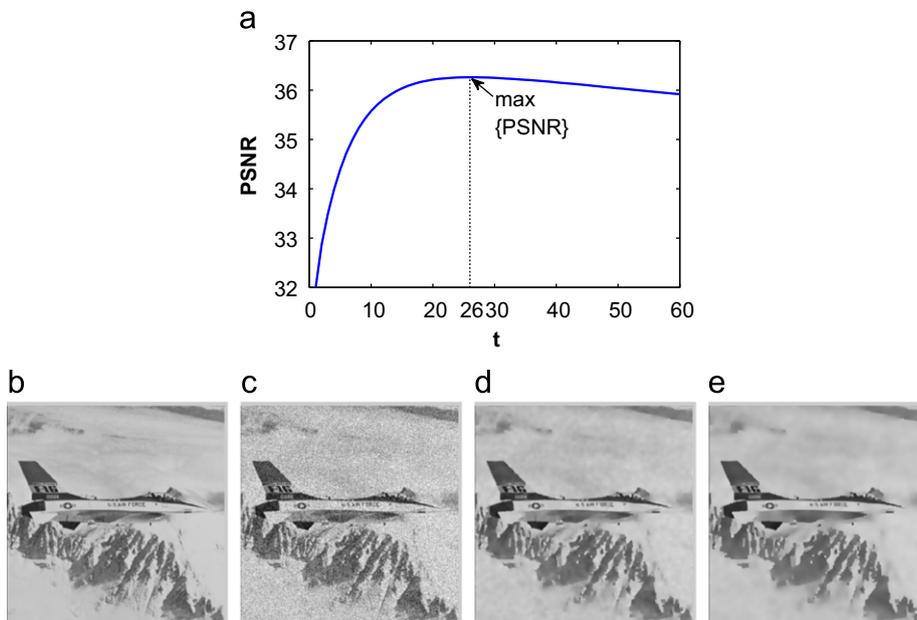


Fig. 8. (a) The measured *PSNR* value in every iteration, between the original and the filtered versions of a noisy airplane image using the proposed AD scheme. (b) The original noise-free image. (c) A noisy version of it (Gaussian noise with $\sigma = 0.1$). (d) The filtered version after $T=26$ iterations when the *PSNR* is maximized. (e) The filtered version after $T=60$ iterations when the result is oversmoothed.

through various denoising schemes. First, we give the definition of these two criteria and then we present our proposed method.

The input noisy image I_0 is considered to be the sum of the ideal noise-free image \bar{I} and additive uncorrelated noise \bar{N} of variance σ_{n0}^2 :

$$I_0 = \bar{I} + \bar{N} \quad (15)$$

Mrázek and Navara in [27] proposed a decorrelation criterion, selecting as time T the time that minimizes the correlation between $I_0 - I_t$ and I_t

$$T \equiv \arg \min_t \frac{\text{cov}(I_0 - I_t, I_t)}{\sqrt{\text{var}(I_0 - I_t)\text{var}(I_t)}} \quad (16)$$

considering $I_0 - I_T$ as an artificial substitute of noise \bar{N} and I_T as an artificial substitute of noise-free image \bar{I} .

Gilboa et al. in [26] stated a condition in order to achieve maximal SNR of the filtered image I_t :

$$\text{SNR} \equiv 10 \log \frac{\text{var}(\bar{I})}{\text{var}(I_t - \bar{I})} = 10 \log \frac{\text{var}(\bar{I})}{\text{var}(\bar{N} - (I_0 - I_t))}$$

Given the condition $\partial \text{SNR} / \partial \text{var}(I_0 - I_t) = 0$ for the maximum and rewriting $\text{var}(\bar{N} - (I_0 - I_t))$ as $\text{var}(\bar{N}) + \text{var}(I_0 - I_t) - 2 \text{cov}(\bar{N}, (I_0 - I_t))$ yields the condition for selecting the value of parameter T :

$$T = \arg \min_t \frac{\partial_t \text{cov}(\bar{N}, (I_0 - I_t))}{\partial_t \text{var}(I_0 - I_t)} \quad (17)$$

The variance of noise \bar{N} of the original image is considered a priori known, while term $\partial_t \text{cov}(\bar{N}, (I_0 - I_t))$ is calculated in every iteration using an image N_t instead of I_t , which contains only pure noise with $\text{var}(N_0) = \text{var}(\bar{N})$.

Both the above methods place the emphasis on reducing noise and they do not take into consideration at all the quality of the preserved edges. In [36] was shown that the edges using AD undergo a slow decay after a certain time, leading to the degradation of the image. It is reasonable to expect that the decay of the edges would be associated with the decrease in the *PSNR* of the resultant image (as in Fig. 8). In this paper, we propose a method that leads to the stopping time T examining directly the quality of the edges of the final image.

3.1. Proposed method

We deal with the stopping time problem, by evaluating the quality of the true edges in every iteration. Under the assumption that the optimal discrete AD scheme controls successfully the rate of the diffusion by discriminating true edges from noise, the initial contrast of an edge should be preserved while the brightness variations around it should be gradually smoothed away. As the effect of the noise increases in an image, the discrimination between edges and noise becomes a more difficult task, resulting in an unavoidable decrease of the strength of the edges that takes place along with the denoising effect of the filter. The proposed method evaluates, in every iteration, the quality of a percentage of the true edges of the image, taking into consideration the contrast and the noise brightness fluctuations around them, and leads to a judicious choice of the stopping time T that corresponds to the maximum overall quality of the edges. An estimate of the standard deviation σ_{n0} of the Gaussian noise of the original image is also required, which is considered to be a priori known. The steps of the proposed method are the following:

1. The original image I_0 is convolved with the kernels of the Sobel operator [37] and N edgels with the highest gradient magnitudes are chosen. In order to choose edgels that belong to different edges within the image, a minimum Euclidean distance D between any pair of edgels considered is required.
2. For each one of the N edgels, a local area that characterizes the edge is defined: given the coordinates (x_k, y_k) of an edgel k and the direction Θ (calculated from the Sobel operator) of the gradient for the specific edgel, the coordinates of 12 interpixel locations around k are defined:

$$x_{m,n} = x_k + m \cos \Theta - n \sin \Theta \quad \text{for } m = \{-2, -1, 1, 2\}, n = \{-1, 0, 1\}$$

$$y_{m,n} = y_k - m \sin \Theta - n \cos \Theta \quad \text{for } m = \{-2, -1, 1, 2\}, n = \{-1, 0, 1\}$$

3. Since the brightness value at interpixel positions is not known, bilinear interpolation is used in order to calculate it [37].

Each edgel is now characterized by the brightness values of two sets of interpixel positions around it, Ω_1 and Ω_2 :

$$\begin{aligned}\Omega_1 &= \{f(x_{m,n}, y_{m,n}) \mid m \in \{-2, -1\}, n \in \{-1, 0, 1\}\} \\ \Omega_2 &= \{f(x_{m,n}, y_{m,n}) \mid m \in \{1, 2\}, n \in \{-1, 0, 1\}\}\end{aligned}$$

Here $f(x_{m,n}, y_{m,n})$ is the estimated brightness value of interpixel location $(x_{m,n}, y_{m,n})$, using bilinear interpolation.

- A quantity Q is assigned to each edgel in every iteration, in order to evaluate the quality of the edge it belongs to. We define it as

$$Q \equiv |\mu_1 - \mu_2| - \alpha(\sigma_1 + \sigma_2) \quad (18)$$

where μ_1 and μ_2 are the mean values of Ω_1 and Ω_2 , respectively, and σ_1 and σ_2 are the standard deviations of Ω_1 and Ω_2 , respectively. The quantity $|\mu_1 - \mu_2|$ is an estimate of the strength of the edge, while the quantity $\sigma_1 + \sigma_2$ is an estimate of the noise fluctuations around an edge. α is a constant defined as $\alpha = 10\sigma_{n0}/\bar{\mu}_0$, where $\bar{\mu}_0 = (1/N)\sum_{i=1}^N |\mu_1(i) - \mu_2(i)|$ is the mean contrast of the N edgels calculated in the original image I_0 , and σ_{n0} is an estimate of the standard deviation of the noise in the original image I_0 .

- The average $\bar{Q}(t) = (1/N)\sum_{i=1}^N Q_i(t)$ quantity of all N edgels reflects the quality of the edges within the image, in every iteration. With this understanding, we select the stopping time T , so that it maximizes $\bar{Q}(t)$:

$$T = \arg \max_t \frac{1}{N} \sum_{i=1}^N Q_i(t) \quad (19)$$

The image of the two sets of interpixel positions around an edgel can be seen in Fig. 9. In order to ensure that the N edgels correspond to true edges, the selection (step 1 of the algorithm) is performed on a smoothed version of the original image, as in Section 2.3. Furthermore, constant N should be high enough to provide a representative enough subset of the true edges within the image. Typical values of N between 50 and 300 caused no significant difference in the experimental results for the images tested (sized between 256×256 and 880×600 pixels), with the lower limit being preferable for finely textured images, as we explain in the next section. The Euclidean distance D between the edgels differs according to the size of the input image, since it should be larger for bigger images so that the edgels belong to different edges across the image. It could be equal to a fraction of the image size, $(1/x)\min\{r, c\}$, where r and c are the width and height of the image in pixels, respectively.

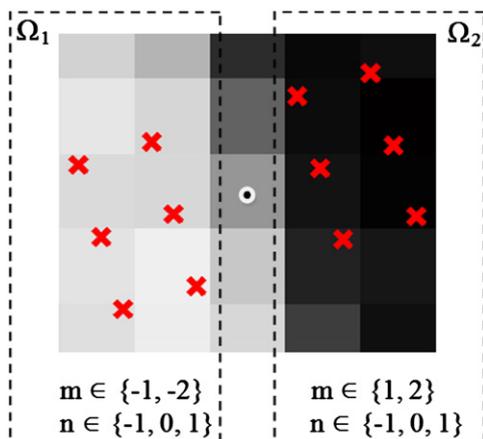


Fig. 9. The 12 interpixel locations around an edgel.

Constant α in (18) serves as an estimate of the effect of the noise to the edges of the original image and has to satisfy the following reasonable conditions. Given the case of very low noise, $\lim_{\sigma_{n0}/\bar{\mu}_0 \rightarrow 0} \alpha = 0$, so that the quality of the edges is expressed only by their contrast $|\mu_1 - \mu_2|$ in (18). Given the case of high noise level, $\lim_{\sigma_{n0}/\bar{\mu}_0 \rightarrow 1} \alpha = 10$, so that the product between α and the estimate of the noise $\sigma_1 + \sigma_2$ in (18) is one order of magnitude higher than the estimate of the contrast $|\mu_1 - \mu_2|$ of the edge, and priority is given to the denoising of the edge. As the noise increases, $\lim_{\sigma_{n0}/\bar{\mu}_0 \rightarrow \infty} \alpha = \infty$, the quality of the edges is expressed only by the estimate of the noise around them, since $|\mu_1 - \mu_2| \ll \alpha(\sigma_1 + \sigma_2)$.

4. Experimental results

In the previous sections we used partial evaluations of the various alternatives in order to come up with an edge-preserving scenario of AD, which terminates automatically evaluating the quality of its edges. Fig. 10 summarizes the final algorithm. In this section we evaluate the proposed stopping criterion and compare its results with those of the results obtained with the criteria proposed by Mrázek and Navara (MN method) and Gilboa et al. (GSZ method), described earlier, using in all cases the same algorithm. To do that, we used the set of images shown in Fig. 11 and we added to them Gaussian noise of zero mean and different standard deviations σ_{n0} . The MN method requires no prior estimation of the noise statistics, while GSZ and our method require an estimate of the standard deviation of the noise σ_{n0} of the input noisy image I_0 , which is considered to be a priori known for the presented experiments. The simple case of noise estimator described in Section 2.3 can be used instead, providing an estimate of the standard deviation of the noise in the initial image. None of the three methods requires a priori knowledge of the characteristics or the structure of the original, noise-free image. For all experiments, using the proposed method, $N=200$ edgels were selected for all the image sizes used (ranging between 256×256 and 880×600 pixels).

In order to evaluate the denoised results, we used three different quality measures between the filtered image I_T and the reference noise-free image \bar{I} . The Peak Signal-to-Noise Ratio (PSNR) is simple to calculate and has a clear physical meaning, but is not always in accord with the human judgement of quality, so the Visual Information Fidelity (VIF) [39] and Structural Similarity (SSIM) [40] criteria, that are closer to the human vision system, were used as well. A general brief description of the two criteria will be given since their analytical expressions are beyond the scope of this work. A detailed study of various quality measures can be found in [41].

Using the VIF criterion the signal fidelity is related to the total information that two signals share. This shared information is quantified in terms of mutual information, a term widely used in information theory. The signal fidelity measurement is based on the modeling of the original image by a wavelet-domain Gaussian scale mixture (GSM) which is associated with the non-Gaussian distributions of the wavelet coefficients of natural images. All distortions between the reference and distorted image are modeled by a simple image distortion model as the sum of a uniform wavelet-domain energy attenuation and an independent additive noise. The mutual information can be calculated in every spatial location in each wavelet subband of the image, with the VIF index being defined as the ratio of the summed mutual information of all local coefficient patches, of all subbands included between the original and the distorted image.

The SSIM criterion is based on the fact that natural images have a strong structural information which is extracted by the HVS.

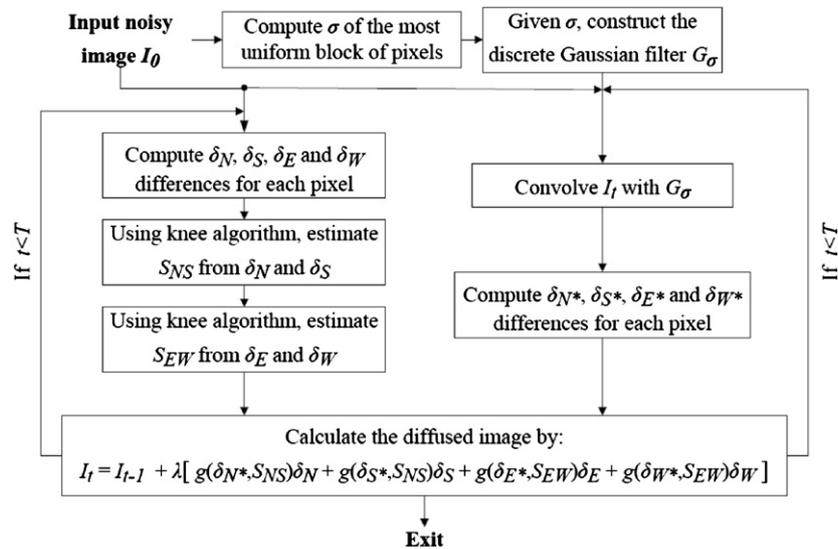


Fig. 10. Block diagram of the denoising algorithm.

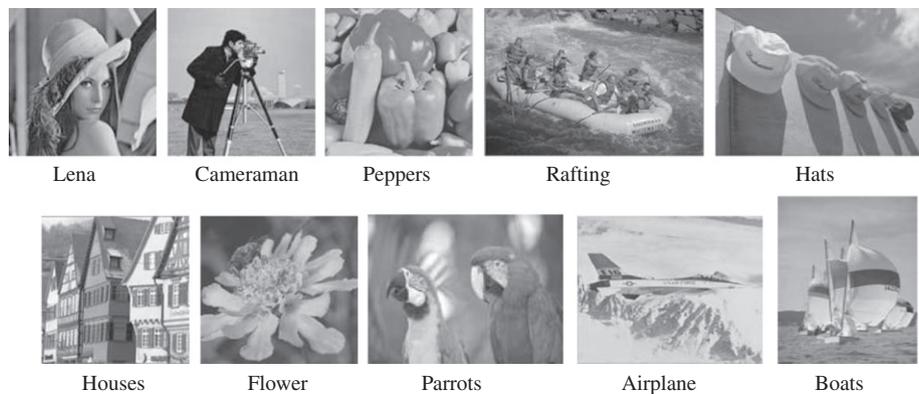


Fig. 11. Natural images used in the experiments.

Table 1

Comparison between different stopping criteria.

Image	PSNR			VIF			SSIM		
	MN	GSZ	Prop	MN	GSZ	Prop	MN	GSZ	Prop
Lena	33.557	33.758	33.760	7.642	7.672	7.684	8.303	8.394	8.407
Cameraman	33.367	33.585	33.679	6.954	7.188	7.196	8.032	8.154	8.205
Boats	34.792	35.132	35.227	7.031	7.290	7.354	8.467	8.581	8.660
Peppers	35.324	35.253	35.272	7.720	7.696	7.697	8.769	8.723	8.737
Houses	30.598	31.929	31.719	6.692	7.648	7.679	7.116	8.347	8.174
Flower	35.370	35.452	35.467	7.542	7.623	7.621	8.696	8.734	8.757
Parrots	36.104	35.965	36.027	7.289	7.276	7.267	8.698	8.623	8.666
Rafting	32.167	33.178	33.121	6.541	7.170	7.176	8.255	8.908	8.888
Airplane	34.047	34.872	34.964	6.783	7.281	7.294	8.104	8.630	8.707
Hats	36.092	36.070	36.107	7.209	7.280	7.288	8.592	8.546	8.566
Average	34.142	34.519	34.534	7.140	7.412	7.425	8.303	8.564	8.577

Distortions such as the additive noise or blur in an image cause a significant alteration of the structures of the objects within the image. The SSIM index is a multiplication of three terms indicating different characteristics of patches within the original and the distorted image: the luminance of the patch, the similarity of the local patch contrasts and the similarity of the local patch structures.

In our experiments, the PSNR was calculated as in (13). The VIF and SSIM implementations used can be found in [42]. For all three quality criteria, a higher measure suggests that the filtered image I_t is closer to the reference noise-free image \bar{I} .

To each image of Fig. 11, Gaussian noise with zero mean and five different standard deviations was added ($\sigma_{n0} = 0.025, 0.05, 0.075, 0.1, 0.125$) and the different stopping times T , using the three compared stopping criteria, were estimated. The filtered image I_t , in every case, was evaluated using the PSNR, VIF and SSIM quality criteria. For each image, the results for all five levels of noise were averaged, calculating the average for each quality criterion separately. The results of the measurements are shown in Table 1. The proposed method, achieved better results for most of the images, in terms of all three quality measures. The performance degraded notably in more

detailed images (Houses, Rafting), where the estimated time T was slightly overestimated. The degradation of the performance was mainly depicted through the $PSNR$ and $SSIM$ criteria, while the average VIF score was still higher using the proposed method. The overestimation of T , in more detailed images, is due to the presence of fine-scale details near the edges, which were erroneously considered as fluctuations due to noise, in the local interpixel area near each edgel. In this case, a smaller number of edgels could be selected, in the initial step of the proposed method, so that only the strongest edgels are taken into consideration.

In Fig. 12, the average quality \bar{Q} of the $N=200$ edgels selected is shown for different images corrupted by Gaussian noise. In general, the estimation of the stopping time T in the proposed method lies on the assumption that the calculated $\bar{Q}(t)$ is unimodal with a unique maximum. Although the uniqueness of the maximum is not ensured mathematically, in practice we have not encountered cases that yielded more than a single maximum for the average quality $\bar{Q}(t)$. As it was mentioned earlier, it is reasonable to assume that the quality of the edges is directly associated to the image quality measures. The similarity between the form of the estimated $\bar{Q}(t)$ in Fig. 12b and the form of the $PSNR$ in Fig. 8a for the airplane image is apparent.

In the case of Fig. 12a, where Gaussian noise with $\sigma_{n0} = 0.075$ was added to Lena image, the proposed stopping time T coincided with the maximum $PSNR$ and $SSIM$ measurements and was closer to the VIF measurement than the other two compared methods. Therefore, according to all three criteria, the proposed method was closer to the optimal stopping time, while GSZ method yielded an underestimated and MN method yielded an overestimated stopping time. The optical results for each method can be seen in Fig. 13.

Similarly, in Fig. 12b, where Gaussian noise with $\sigma_{n0} = 0.1$ was added to the airplane image, the proposed T was optimal according to VIF measurement and was closer to optimal $PSNR$ and $SSIM$ measurements than the other two methods. The optical results can be seen in Fig. 14. A case where the proposed method failed is shown in Fig. 12c, where Gaussian

noise with $\sigma_{n0} = 0.05$ was added to the Houses image. The proposed time T was overestimated in comparison with the maximum quality measurements. The overestimation of the proposed method is attributed to the presence of many texture edges and fine-scale details, as described earlier. The optical results can be seen in Fig. 15.

5. Discussion and conclusions

In this paper we studied carefully all steps of the Anisotropic Diffusion algorithm and came up with the best choice among the various options at each step, describing a complete image-adapted denoising tool. We also proposed a stopping criterion which does not only take into consideration the level of the removed noise, but also the quality of the preserved edges. The scheme was evaluated using several images with different levels of noise. The most difficult images to improve with AD are those with many details and texture. This is not surprising, as highly textured images have significant levels of energy in the high frequencies too, where noise is supposed to dominate. It is expected, therefore, that any noise suppression scheme will have difficulty with such images.

The estimation of two gradient threshold parameters using the knee algorithm improves the adaptability of the filter, yielding stronger edges as the estimation of one parameter tends to oversmooth the image as it was shown in Section 2.4. This property also introduces a sense of directionality which could lead to a comparison with a filter class that smoothes differently along distinct directions. Such anisotropic filters are usually modeled by diffusion tensors instead of scalar-valued diffusions [12,43] and are widely used in images with strong oriented structures, such as seismic [44,45] or medical [46–48] images. However, such a comparison would be beyond the scope of the current work which aims at coming up with the most edge-preserving generic AD scheme, making it available for denoising all types of natural images.

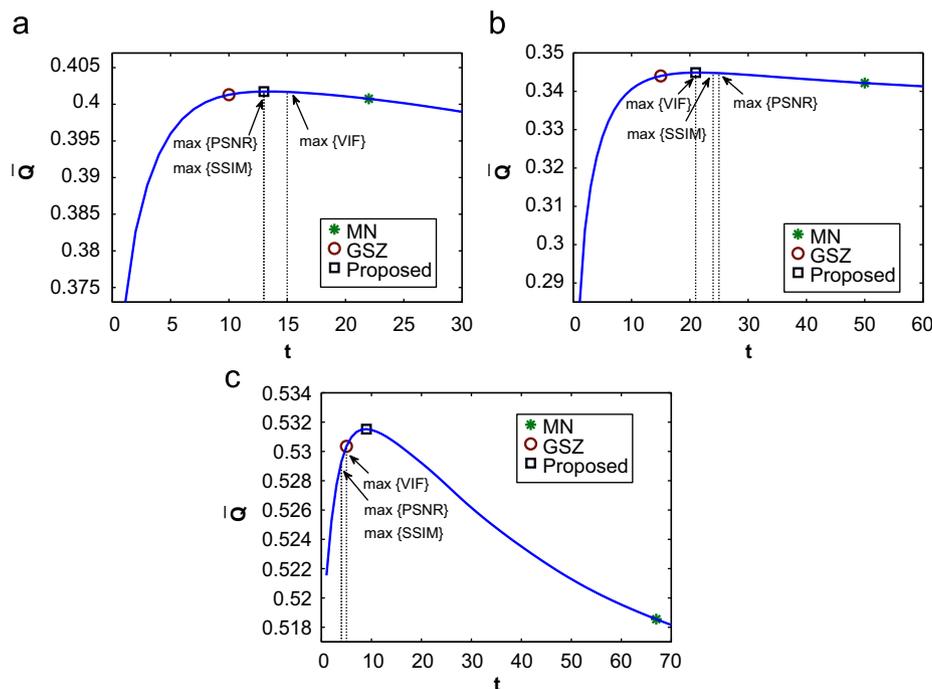


Fig. 12. The average quality $\bar{Q}(t)$ for $N=200$ edgels, the iterations that correspond to the maximum quality measures and the iterations that correspond to the estimated stopping time T using the described stopping criteria. (a) Lena image with Gaussian noise of $\sigma = 0.075$, (b) airplane image with Gaussian noise of $\sigma = 0.1$ and (c) houses image with Gaussian noise of $\sigma = 0.05$.

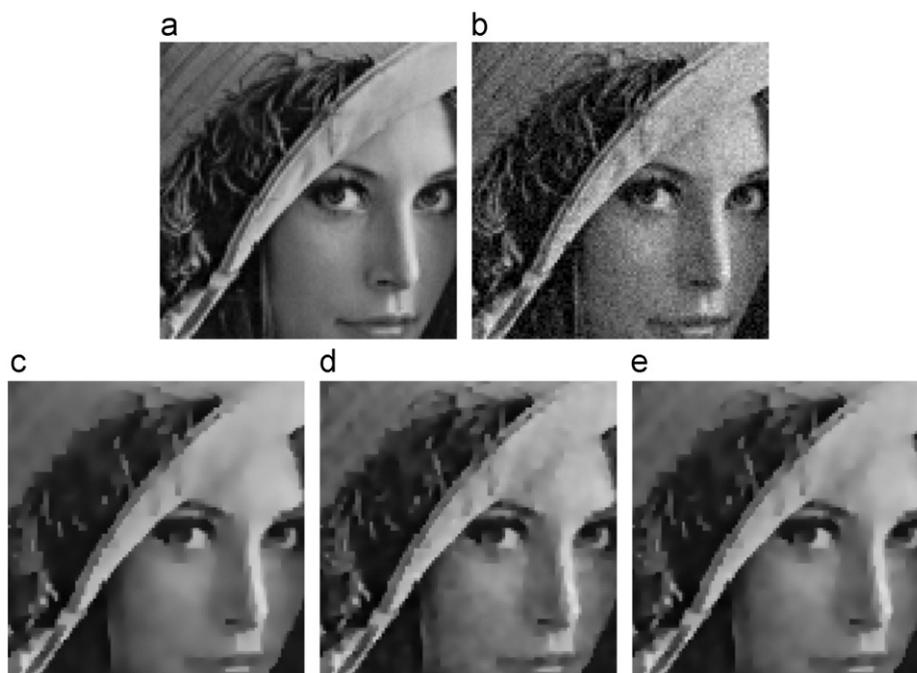


Fig. 13. (a) A zoomed part of the original Lena image. (b) A noisy version of it (Gaussian noise with $\sigma = 0.075$). The AD filtering result after estimating T with (c) MN method, (d) GSZ method and (e) the proposed method.

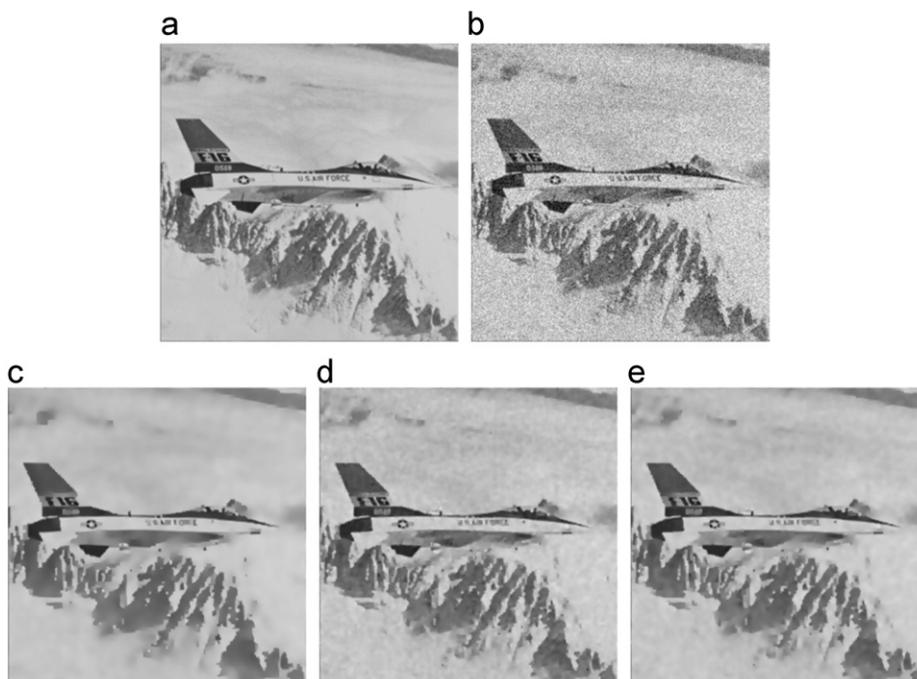


Fig. 14. (a) A zoomed part of the original Airplane image. (b) A noisy version of it (Gaussian noise with $\sigma = 0.1$). The AD filtering result after estimating T with (c) MN method, (d) GSZ method and (e) the proposed method.

The stopping criterion is based on a statistical representation of an image by its true edges. The numerical evaluation of Section 4 showed that this approach can provide a robust solution for the stopping time problem. The number N of the averaged examined edgels does not change with image size. It simply has to be high enough to allow the calculation of reliable statistics. A local approach of the proposed method could be considered in the future, where the diffusion would stop separately for different edges according to their quality measurement. However, such a method would increase significantly the computational cost of

the algorithm. It should also be mentioned that the proposed method fails at corners since there are not two half-planes separating one edge side from the other.

In the testing cases of the stopping criteria the estimate of variance of the noise was considered a priori known, so that a fair comparison with the Gilboa et al. criterion (GSZ), which uses an estimate of the variance as well, was done. However, the simple case of noise estimator that is used for the calculation of the gradients may be employed, as described in Section 2.3. This estimator yielded reliable estimates for the noise variance

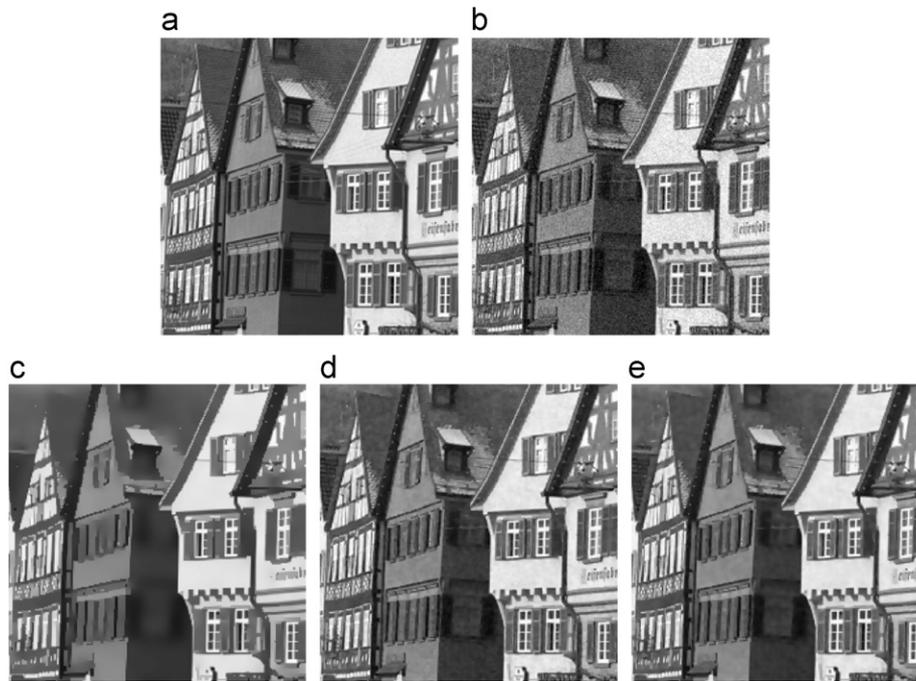


Fig. 15. (a) A zoomed part of the original Houses image. (b) A noisy version of it (Gaussian noise with $\sigma = 0.05$). The AD filtering result after estimating T with (c) MN method, (d) GSZ method and (e) the proposed method.

(estimates within 10% deviation from the real value). Using such a noise estimator, the whole presented scheme comprises a completely automatic tool for image denoising, adapted to the denoising needs of the image and the condition of its edges.

References

- [1] A. Witkin, Scale-space filtering, in: International Joint Conference on Artificial Intelligence, Karlsruhe, West Germany, 1983, pp. 1019–1021.
- [2] P. Perona, J. Malik, Scale space and edge detection using anisotropic diffusion, *IEEE Transactions on Image Processing* 12 (8) (1990) 629–639.
- [3] L. Alvarez, P.L. Lions, J.M. Morel, Image selective smoothing and edge detection by nonlinear diffusion ii, *SIAM Journal on Numerical Analysis* 29 (3) (1992) 845–867.
- [4] F. Catté, P.L. Lions, J.M. Morel, T. Coll, Image selective smoothing and edge detection by nonlinear diffusion, *SIAM Journal on Numerical Analysis* 29 (1) (1992) 182–193.
- [5] Y. Chen, C. Barcelos, B. Mair, Smoothing and edge detection by time-varying coupled nonlinear diffusion equations, *Computer Vision and Image Understanding* 82 (2) (2001) 85–100.
- [6] S. Kichenassamy, The Perona–Malik paradox, *SIAM Journal on Applied Mathematics* 57 (1997) 1328–1342.
- [7] M. Nitzberg, T. Shiota, Nonlinear image filtering with edge and corner enhancement, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14 (8) (1992) 826–833.
- [8] L.I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D* 60 (1992) 259–268.
- [9] J. Weickert, A review of nonlinear diffusion filtering, *Scale-space theory in computer vision*, Lecture Notes in Computer Science 1252 (1997) 3–28.
- [10] Y.L. You, W. Xu, A. Tannenbaum, M. Kaveh, Behavioral analysis of anisotropic diffusion in image processing, *IEEE Transactions on Image Processing* 5 (11) (1996) 1539–1553.
- [11] D. Chen, S. MacLachlan, M. Kilmer, Iterative parameter-choice and multigrid methods for anisotropic diffusion denoising, *SIAM Journal on Scientific Computing* 33 (October (5)) (2011) 2972–2994.
- [12] J. Weickert, *Anisotropic Diffusion in Image Processing*, Ser. ECMI, Teubner, Stuttgart, Germany, 1998.
- [13] J. Weickert, B.M.T.H. Romeny, M.A. Viergever, Efficient and reliable schemes for nonlinear diffusion filtering, *IEEE Transactions on Image Processing* 7 (March (3)) (1998) 398–410.
- [14] G. Gerig, O. Kubler, R. Kikinis, F. Jolesz, Nonlinear anisotropic filtering of MRI data, *IEEE Transactions on Medical Imaging* 11 (2) (1992) 221–232.
- [15] X. Yu, C. Wu, T. Jia, S. Chen, A time-dependent anisotropic diffusion image smoothing method, in: 2nd International Conference on Intelligent Control and Information Processing, vol. 2, July 2011, pp. 859–862.
- [16] S.M. Chao, D.M. Tsai, An improved anisotropic diffusion model for detail and edge-preserving smoothing, *Pattern Recognition Letters* 31 (October) (2010) 2012–2023.
- [17] S.M. Chao, D. Tsai, Anisotropic diffusion with generalized diffusion coefficient function for defect detection in low-contrast surface images, *Pattern Recognition* 43 (5) (2010) 1917–1931.
- [18] M.J. Black, G. Sapiro, D. Marimont, D. Heeger, Robust anisotropic diffusion, *IEEE Transactions on Image Processing* 7 (3) (1998) 421–432.
- [19] J. Monteil, A. Beghdadi, New interpretation and improvement of the nonlinear anisotropic diffusion for image enhancement, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21 (9) (1999) 940–946.
- [20] X. Li, T. Chen, Nonlinear diffusion with multiple edginess thresholds, *Pattern Recognition* 27 (8) (1994) 1029–1037.
- [21] F. Voci, S. Eiho, N. Sugimoto, H. Sekiguchi, Estimating the gradient threshold in the Perona–Malik equation, *IEEE Signal Processing Magazine* 21 (3) (2004) 39–46.
- [22] G. Gilboa, N. Sochen, Y.Y. Zeevi, Forward-and-backward diffusion processes for adaptive image enhancement and denoising, *IEEE Transactions on Image Processing* 11 (7) (2002) 689–703.
- [23] J. Weickert, Applications of nonlinear diffusion in image processing and computer vision, *Acta Mathematica Universitatis Comenianae* 70 (2001) 33–50.
- [24] A.K.W. Sum, P.Y.S. Cheung, Stabilized anisotropic diffusions, in: *IEEE International Conference on Acoustics Speech Signal Proceedings*, vol. 1, 2007, pp. 709–712.
- [25] A. Ilyevsky, E. Turkel, Stopping criteria for anisotropic PDEs in image processing, *Journal of Scientific Computing* 45 (2010) 333–347.
- [26] G. Gilboa, N. Sochen, Y.Y. Zeevi, Estimation of optimal PDE-based denoising in the SNR sense, *IEEE Transactions on Image Processing* 15 (8) (2006) 2269–2280.
- [27] P. Mrázek, M. Navara, Selection of optimal stopping time for nonlinear diffusion filtering, *International Journal of Computer Vision* 52 (2/3) (2003) 189–203.
- [28] G. Papandreou, P. Maragos, A cross-validatory statistical approach to scale selection for image denoising by nonlinear diffusion, in: *IEEE Conference on Computer Vision and Pattern Recognition*, 2005, pp. 625–630.
- [29] V. Solo, Automatic stopping criterion for anisotropic diffusion, in: *International Conference on Acoustics, Speech, and Signal Processing*, vol. 6, 2001, pp. 3929–3932.
- [30] G. Gilboa, Nonlinear scale space with spatially varying stopping time, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30 (December (12)) (2008) 2175–2187.
- [31] Jian Sun, Zongben Xu, Scale selection for anisotropic diffusion filter by Markov random field model, *Pattern Recognition* 43 (8) (2010) 2630–2645.
- [32] K.Z. Abd-Elmoniem, A.B.M. Youssef, Y.M. Kadah, Real-time speckle reduction and coherence enhancement in ultrasound imaging via nonlinear anisotropic diffusion, *IEEE Transactions on Biomedical Engineering* 49 (September (9)) (2002) 997–1014.
- [33] K. Krissian, S. Aja-Fernandez, Noise-driven anisotropic diffusion filtering of MRI, *IEEE Transactions on Image Processing* 18 (October (10)) (2009) 2265–2274.

- [34] K. Krissian, C.F. Westin, R. Kikinis, K.G. Vosburgh, Oriented speckle reducing anisotropic diffusion, *IEEE Transactions on Image Processing* 16 (May (5)) (2007) 1412–1424.
- [35] Deepti Mittal, Vinod Kumar, Suresh Saxena, Niranjana Khandelwal, Naveen Kalra, Enhancement of the ultrasound images by modified anisotropic diffusion method, *Medical and Biological Engineering and Computing* 48 (2010) 1281–1291.
- [36] R.T. Whitaker, S.M. Pizer, A multi-scale approach to nonuniform diffusion, *CVGIP: Image Understanding* 57 (1) (1993) 99–110.
- [37] M. Petrou, C. Petrou, *Image Processing the Fundamentals*, John Wiley, 2010.
- [38] J. Canny, A computational approach to edge detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8 (November (6)) (1986) 679–698.
- [39] H.R. Sheikh, A.C. Bovik, Image information and visual quality, *IEEE Transactions on Image Processing* 15 (2) (2006) 430–444.
- [40] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, *IEEE Transactions on Image Processing* 13 (4) (2004) 600–612.
- [41] H.R. Sheikh, M.F. Sabir, A.C. Bovik, A statistical evaluation of recent full reference image quality assessment algorithms, *IEEE Transactions on Image Processing* 15 (11) (2006) 3449–3451.
- [42] H.R. Sheikh, Z. Wang, L. Cormack, A.C. Bovik, Live image quality assessment database release 2 (2005).
- [43] J. Weickert, Coherence-enhancing diffusion filtering, *International Journal on Computer Vision* 31 (1999) 111–127.
- [44] J. Won-Ki, R. Whitaker, M. Dobin, *Interactive 3d Seismic Fault Detection on the Graphics Hardware*, 2006.
- [45] O. Laviaille, S. Pop, C. Germain, M. Donias, S. Guillon, N. Keskes, Y. Berthoumieu, Seismic fault preserving diffusion, *Journal of Applied Geophysics* 61 (2) (2007) 132–141.
- [46] D.C. Fernández, H.M. Salinas, C.A. Puliafito, Automated detection of retinal layer structures on optical coherence tomography images, *Optics Express* 13 (December (25)) (2005) 10200–10216.
- [47] R. Manniesing, M.A. Viergeever, W.J. Niessen, Vessel enhancing diffusion: a scale space representation of vessel structures, *Medical Image Analysis* 10 (6) (2006) 815–825.
- [48] A.S. Frangakis, R. Hegerl, Noise reduction in electron tomographic reconstructions using nonlinear anisotropic diffusion, *Journal of Structural Biology* 135 (3) (2001) 239–250.

Chourmouzi Tsotsios received his ME degree in Electrical and Computer Engineering in 2010, from Aristotle University of Thessaloniki, Greece. He completed his Master's thesis in collaboration with Informatics and Telematics Institute of Thessaloniki and he was a visiting student in the Department of Electronics, University of Barcelona in 2008. He is currently working towards a PhD degree in the Department of Electrical and Electronic Engineering in Imperial College London. His research interests include computer vision, underwater imaging and photometric stereo.

Maria Petrou studied Physics at the Aristotle University of Thessaloniki, Greece, and obtained her Master degree in Applied Mathematics, her PhD in Astronomy and her DSc in Engineering all from Cambridge University, UK. She is currently the Chair of Signal Processing at Imperial College London and the Director of the Informatics and Telematics Institute of CERTH, Greece. She has published two books on Image processing (*Image Processing, the Fundamentals* (1999 and 2nd edition 2010) and *Image Processing, dealing with Texture*, 2006), both with John Wiley. She has published more than 140 journal papers, more than 300 papers in conference proceedings and more than 15 book chapters. She has supervised to successful completion 45 PhD students. She is a Fellow of the Royal Academy of Engineering (FREng), Fellow of the International Association for Pattern Recognition (IAPR), Distinguished Fellow of the British Machine Vision Association (BMVA) and Fellow of the Academia Europaea. Among her many professional activities, she has been the Chairman of BMVA (2000–2003), the Treasurer of IAPR (2003–2007) and a Trustee of IET (2007–2009).