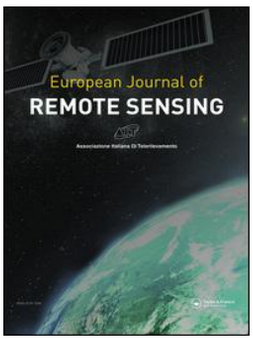

European Journal of Remote Sensing

ISSN: (Print) 2279-7254 (Online) Journal homepage: https://www.tandfonline.com/loi/tejr20

Adaptive Restoration of Multispectral Datasets used for SVM classification

Amin Zehtabian, Avishan Nazari, Hassan Ghassemian & Marco Gribaudo

To cite this article: Amin Zehtabian, Avishan Nazari, Hassan Ghassemian & Marco Gribaudo (2015) Adaptive Restoration of Multispectral Datasets used for SVM classification, European Journal of Remote Sensing, 48:1, 183-200, DOI: 10.5721/EuJRS20154811

To link to this article: https://doi.org/10.5721/EuJRS20154811

© 2015 The Author(s). Published by Taylor & Francis.

Published online: 17 Feb 2017.

Submit your article to this journal

Article views: 67

View Crossmark data

Citing articles: 1 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tejr20

European Journal of Remote Sensing - 2015, 48: 183-200
doi: 10.5721/EuJRS20154811
Received 27/12/2014, accepted 10/04/2015

European Journal of Remote Sensing
An official journal of the Italian Society of Remote Sensing

www.aitjournal.com

Adaptive Restoration of Multispectral Datasets used for SVM classification

Amin Zehtabian[1], Avishan Nazari[2], Hassan Ghassemian[1]* and Marco Gribaudo[2]

[1]Faculty of Electrical and Computer Engineering, Tarbiat Modares University, Tehran, Iran
[2]Department of Information Technology, Politecnico di Milano, Milan, Italy
*Corresponding Author, e-mail address: ghassemi@modares.ac.ir

Abstract
Removing noise from images while keeping its important details unchanged is a challenging issue in image restoration. In this paper, we propose a novel approach based on partial differential equations (PDE) in order to mitigate three well-known types of noises from remote sensing data while important features such as edges are preserved. In the presented method, after performing the Watershed-based segmentation as a preprocessing step, optimum values of PDE parameters are adaptively found based on the noise type and the image texture. In order to evaluate the performance of the proposed algorithm, Peak Signal-to-Noise Ratio (PSNR) criterion is applied. Moreover, feeding the original/noisy/denoised images into SVM classifier and exploring the classification ratios are suggested as an application-based assessment. The gained results prove a considerable enhancement both in quantitative metrics (PSNR and MSE) and SVM classification ratios (from 71.71% to 95.07%).

Keywords: Remote sensing, multispectral image restoration, partial differential equations, watershed partitioning, support vector machines.

Introduction
Remotely Sensed (RS) images (which may be single-layer or multiband) are often contaminated by additive or multiplicative noises while they are taken by sensors or during transformation through media or communication channels. In presence of noise, data interpretability can be considerably reduced and extracting required information from images can be a challenging task. “Image restoration” is the processing of infected images in order to approximate the initial image and earning required features from the denoised image. Traditional filters such as Mean, Median and Gaussian filters are some examples of these methods [Gonzales and Woods, 2004]. Moreover, Wavelet transformation [Jensen, 2001] as well as Wiener filtering [Song and Choi, 2005] is frequently used among researchers for denoising purposes. The subspace-based approaches are also applicable for multispectral RS image noise reduction, as we proposed in Zehtabian and Ghassemian [2013].
Retaining image structure and important features untouched, while performing an acceptable

CC BY

level of denoising, is a bottleneck in image restoration techniques. Being more clear, a given restoration method may guarantee a considerably high level of noise reduction, but it may also provide some blurring effects. Dislocation or smoothing important image details such as edges is another probable consequence of such approaches. Due to this problem and other defects, scientists have been inspired to perform denoising processes using Partial Differential Equation (PDE) based approaches [Perona and Malik, 1990]. In PDE-based techniques, changes in image pixels has been considered like temperature varying model in which the air moves from warmer environments to cooler spaces until two places reach equilibrium. In the process of applying PDE approaches, the model of changing in gray level is frequently considered as a Gaussian function. Based on this assumption, any sudden change in image will be recognized as noise. PDE-based methods, different from other kinds of noise removal techniques, makes the interior regions of an image smoothed while leaves the edges and intra regions relatively unaffected [Zehtabian and Ghassemian, 2015].

In PDE based methods, there is an important parameter called "diffusion coefficient" which has an important role in the denoising process. If this coefficient is a constant value, then the diffusion process will be equilateral in each direction. This isotropic process is usually called as "linear diffusion". Considering a nonlinear function for this parameter will result in anisotropic diffusion or "nonlinear diffusion". This idea was first presented by Perona and Malik [1990]. They introduced a function which has a reverse relation with image gradients around each pixel. By this idea, they could control the smoothing level regarding the amount of details in each region of the image. Applying this strategy, the image details have been preserved.

One of the most significant parameters in the Perona-Malik (P-M) equation is the "scale constant" (also named as "gradient threshold") which is usually denoted by letter "k" in literatures. Smoothing level of images as well as the sensitivity of denoising procedure to edges has been controlled by this parameter. Assigning low values to this parameter causes lower level of denoising, while high values of k leads to stronger denoising as well as probable mitigation of edges in image. Unlike the traditional PDE-based noise removal methods (where this parameter has been determined experimentally), in this paper (which is an extended version of our recently work in Nazari et al. [2014]), we propose an innovative method to adaptively estimate the parameter k. The adaptively defined parameter is able to deal with different kinds of image textures and noises. In other words, by proposing some metrics which estimate the level of noise in images, we let the image restoration system to adaptively set the parameter k and consequently tune the noise removal level for each case. Meanwhile, in order to obtain more accurate results, we propose to localize the adaptation procedure. To this end, each band of data is partitioned into several disjoint segments and then the proposed adaptation algorithm is applied separately to each segment.

The rest of the paper is ordered as follows: the next section provides a brief review on the most well-known methods dealing with the restoration of RS images. Then we present the theory of PDE in the third section. The fourth section introduces our proposed method in order to adaptively adjust the parameter k. The proposed Watershed-based restoration method (called as WshSeg-APDE) is explained in the fifth section. The sixth section comprises the outcomes of an extensive set of experiments which has been performed on satellite images. As a final point, conclusion and future works are presented in the last section.

Background and Literature Review

There have been valuable researches carried out in the past decades to favorably restore the RS datasets. Authors in Hird and McDermid [2009] review two novel logistic and asymmetric Gaussian function-fitting methods as well as four alternative filtering approaches for noise reduction from MODIS (Moderate Resolution Imaging Spectro-radiometer) remote sensing dataset. In Chui et al. [2012] an image denoising technique is introduced based on a nonlinear thresholding function with adaptive Bayes threshold in Non-Subsampled Contourlet Transform (NSCT) domain. In this method, first the noise deviation of each sub-band is estimated and then the thresholds of every sub-band are estimated by Bayesian threshold estimation technique. Although this method is able to considerably remove additive white Gaussian noise (AWGN) as well as keeping the image texture relatively unchanged, it has been only tested on single-band images.

Authors in Huang et al. [2009] develop a technique to overcome the problem of speckle noises which infect the synthetic aperture radar (SAR) images. Their main goal is to improve the detection ratio for ship targets from SAR imaging mechanism. In Di Martino et al. [2013] a framework is developed in order to simulate SAR images with speckle statistics. A so-called 'BiShrink' threshold as well as a novel remote sensing image denoising method is also presented in Li et al. [2011]. The method is based on the shift-invariance and better directionality of the Dual-Tree Complex Wavelet Transform (DTCWT), and incorporating the neighboring wavelet coefficients with BiShrink. Applying the singular value decomposition (SVD) based approaches in a wide variety of image processing applications has been considerably attracted by researchers. Authors in Herries et al. [1996] comprehensively study the applications of singular value decomposition in applied remote sensing. As stated in his article, SVD had been previously applied to the problems of the agricultural land classification and proved its accuracy and efficiency. They also mention some other applications of SVD in remote sensing such as key vector analysis, dimensional reduction, robustness, and finally the noise removal.

Another study which has been carried out to show the applications of SVD in RS data denoising can be referred in Su et al. [2012]. This technique (called as "K-SVD") leads to relatively considerable denoising at lower noise levels, but as reported in the article, the efficiency of the technique degrades in the environments with relatively high noise levels. Authors in Ragged et al. [2012] develop a novel spatial-spectral schema to derive the eigenvectors and separate those eigenvectors representing signal as opposed to noise. Their proposed approach is applied to the AVIRIS hyperspectral dataset.

We have recently developed a novel subspace based approach which employs SVD to separate the noise subspace and the signal subspace from noisy RS Multispectral images to overcome the AWGN processes [Zehtabian and Ghassemian, 2013].

As mentioned in the introduction section, simple denoising methods such as Mean, Median, and Gaussian filters are also commonly used in many researches but they suffer from some shortages. The idea of Mean filtering (also called as Averaging or Box filtering) is to replace the gray-level of each pixel in the image with the average value of its neighbors (often including the pixel itself). Mean filter is widely used because of its simplicity and easy implementation, but since it generates a new gray-level value in each step, some details may be lost and never be restored again. Therefore, it can be said that this method is a tradeoff between the denoising level and detail preservation. Moreover, this filter performs a better

denoising for images corrupted by Gaussian noise rather than salt and pepper noise. Indeed, the considerable difference between the noisy pixel intensities (i.e. value "0" for "pepper" and "255" for "salt") may lead to distort the pixel average values from their surroundings.
In Median filtering, instead of replacing the pixel value with the average of its neighbors, the median value of the surrounding pixels is applied. In this method, first the pixel intensities should be sorted and then the value in the middle will be selected. Compared to the Mean filter, the pixel values in Median filter will be substituted with a gray-level value which is already existed in the image. Therefore, it prevents the image from large changes in its intensity values. The efficiency of Median filter in removing the salt and pepper noise is better compared to the Mean filter. In case of Gaussian noise, however, Median filter behaves weaker. On the other hand, Median filter is more difficult to be implemented since sorting the neighbor values into numerical order takes more time compared to calculating the average values.
Referring to Equation [1], the Wiener filter attempts to perform filtering process by diminishing the Mean Square Error (MSE) between the original image ($I_{Original}$) and denoised image ($I_{Denoised}$):

$$\left|d\right|^2 = \left\|I_{Original} - I_{Denoised}\right\|^2 \quad [1]$$

In Wiener filtering, it is often assumed that both image and noise have a Gaussian distribution. Therefore, the Wiener filter works very well in case of Gaussian noise while it may not be suitable for other noises types.

Nonlinear Diffusion Filters

Researchers proposed PDE-based filtering techniques as a remedy to the shortages of the traditional noise removal methods stated in the aforementioned section. To be more precise, they bring the idea of applying nonlinear diffusion techniques. Diminishing the noise pollution from images where the important features still remain unaffected is an ability of PDE-based restoration techniques. In other words, PDE makes inter-regions smoothed while the intra-region parts (such as edges) remain relatively unsmoothed [Zehtabian and Ghassemian, 2015].
The basic idea of using PDE as a multi-scale smoothing filter goes back to 1983 [Witkin, 1983]. But first time in 1990, Perona and Malik came with a nonlinear diffusion equation for filtering purposes [Perona and Malik, 1990]. Later, the researchers continued their efforts to improve the PDEs and use them for image smoothing as well as restoration [Nadernejad et al., 2007]. In this paper, an innovative version of PDE has been used which includes some key progresses in its implementation. We also propose an adaptation method in order to assign an optimum value for PDE parameters.
The partial differential equation for image *I(x, y)* can be shown as below:

$$\frac{\partial I(x,y,t)}{t} = \nabla^2 I(x,y,t) = \frac{\partial^2 I(x,y,t)}{\partial x^2} \quad [2]$$

Where ∇ is the Gradient operator and $I(x, y, 0) = I_0(x, y)$ is the initial image. In heat diffusion equations, "t" represents the time, while here $t = \sigma^2/2$ where σ is the standard deviation of the corresponding Gaussian filter. There is another form for Equation [2] which can be written as:

$$\frac{\partial I(x,y,t)}{t} = \nabla.\left(c(x,y,t)\nabla I(x,y,t)\right) \quad [3]$$

Where $c(x, y, t)$ is the "diffusion coefficient" and (∇.) denotes divergence operator. Coefficient "c" has a vital impact on denoising process and defines the power of the smoothing/denoising procedure. Setting a constant value for this parameter will result in Isotropic diffusion (linear diffusion) which performs the homogenous smoothing in all directions. An anisotropic diffusion (nonlinear diffusion) will be achieved by embedding a nonlinear function which depends on the image local features. In other words, in "nonlinear diffusions", the diffusion coefficient is also varied according to the local changes and details in the image. Therefore, it controls the noise removal level in each region of the image, regarding the amount of details embedded in that region. Since diffusion coefficients must be dependent on the localized details in the image, they are usually nonlinear functions of the image gradients around each pixel. Indeed, the gradient operators are beneficial because of their capability to detect edges and boundaries.

There are a few number of equations which have been proposed by researchers in order to relate the diffusion coefficient with the image gradients. In this paper, we apply Equation [4] which is proposed by Perona and Malik to calculate the diffusion coefficient:

$$c(x,y,t) = \frac{1}{(1+\frac{|\nabla I|^2}{k^2})} \quad [4]$$

As previously introduced, k is the scale constant. Controlling the diffusion process can be done using this parameter. Equation [3] can be restructured as [5] which is implemented in an iterative form:

$$I(x,y,t+\Delta t) = I(x,y,t) + \Delta t(d_N.c_N + d_S.c_S + d_E.c_E + d_W.c_W) \quad [5]$$

where derivations along four cardinal directions have been implied by d_N, d_S, d_E and d_W. As two examples, we rewrite d_S and c_S in [6] and [7] which respectively indicate derivation along the south direction and its corresponding diffusion coefficient:

$$d_S = I(x,y+1,t) - I(x,y,t) \quad [6]$$

$$c_S = \frac{1}{1+(\frac{d_s}{k})^2} \quad [7]$$

As can be inferred from [5], the gradients have been calculated along four cardinal directions (i.e. north, south, west, and east). As a result, traditional PDE-based implementations are only capable of detecting and preserving the horizontal and vertical edges. In this study, we also benefit from derivations along the other four sidelong directions besides cardinal directions as suggested by Nadernejad et al. [2007]. By this, it is guaranteed to preserve the slant boundaries as well. The second-ordered neighbors will be also considered in calculating the derivations in the presented paper. Therefore, we update [5] as the following equation:

$$I(x,y,t+\Delta t) = I(x,y,t) + \Delta t(d_{N_1} c_{N_1} + d_{S_1} c_{S_1} + d_{E_1} c_{E_1} + d_{W_1} c_{W_1} + \alpha(d_{NE_1} c_{NE_1} + d_{SE_1} c_{SE_1} + d_{NW_1} c_{NW_1} + d_{SW_1} c_{SW} + d_{N_2} c_{N_2} + d_{S_2} c_{S_2} + d_{E_2} c_{E_2} + d_{W_2} c_{W_2}) + \beta(d_{NE_2} c_{NE_2} + d_{SE_2} c_{SE_2} + d_{NW_2} c_{NW_2} + d_{SW_2} c_{SW_2})) \quad [8]$$

The directions (where gradients are calculated) have been labeled by subscripts in [8].

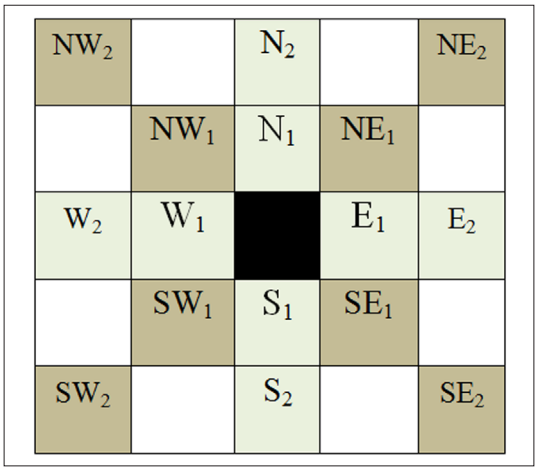

Figure 1 - Calculating the gradients along new directions according to a wider neighboring structure [Nadernejad et al., 2007].

Figure 1 illustrates the neighboring structure where these directions have been extracted from. Similar to [6], gradients along new directions can be calculated. Two examples are stated in [9] and [10]. Meanwhile in [8], the parameters α and β denote the weighting factors which define the priority level of the first and second-order slant neighbors. In this paper, these values are experimentally set to 0.30 and 0.10, respectively.

$$d_{NW_1} = I(x-1,y-1,t) - I(x,y,t) \quad [9]$$

$$d_{SE_2} = I(x+2,y+2,t) - I(x,y,t) \quad [10]$$

Proposed Adaptation Method for PDE-based Restoration

The scale constant k in [4] plays a key role in controlling the diffusion along denoising process. Various experiments on different datasets depicted that a small k will result in slow diffusion and less denoising while a large one leads to severe denoising and probable edge fading. Up to now, parameter k has been usually calculated experimentally. However, finding an appropriate value for scale constant k in an adaptive manner is one of the main goals of this paper. In order to define a suitable value for k, one idea is to approximate the amount of noise from the noisy dataset. To this end, we propose calculating the variations and changes of image gradients along different directions. As a result, if there is a large amount of variations in image gradients, it can be a sign of more pollution of noises or unwanted information in that image. Therefore, more smoothing should be performed for the image (i.e. parameter k should be larger). Furthermore, fewer changes in image pixels and lower variations in the gradients can be the results of less amount of noise. In these cases, smaller values for k should be chosen.

In order to evaluate the variation of gradients in a noisy image, one solution is calculating the gradient values of each pixel along four cardinal directions. However, our experiments prove that for a higher level of denoising (besides better preserving of slant edges), diagonal directions should be used as well (the applied neighboring structure can be recall from Fig. 1). Hence, first we calculate the image gradients around each pixel. This has been done for all the pixels from the whole image and along all eight cardinal and diagonal directions. Consequently, one matrix of gradients is achieved from the noisy image for each direction. The eight matrices are depicted by $(d_n, d_s, \ldots, d_{ne}, d_{nw}, \ldots)$. The differences of these gradients are then computed in horizontal, vertical and diagonal axes:

$$\begin{aligned} d_{ns} &= d_n - d_s \\ d_{ew} &= d_e - d_w \end{aligned} \quad [11]$$

and

$$\begin{aligned} d_{ne\,sw} &= d_{ne} - d_{sw} \\ d_{nw\,se} &= d_{nw} - d_{se} \end{aligned} \quad [12]$$

In order to produce a positive value, each element of these matrices is then powered by two. Since the resulted matrices are not necessarily identical to the initial matrices which are entirely powered by two, we use the notations of $d_{ns}^{\boxed{2}}, d_{ew}^{\boxed{2}}, d_{ne\,sw}^{\boxed{2}}, d_{nw\,se}^{\boxed{2}}$ instead of $d_{ns}^2, d_{ew}^2, d_{ne\,sw}^2, d_{nw\,se}^2$ to emphasize the difference. Now by applying equations [13] and [14], two matrices φ_c and φ_d are defined. These matrices are related to the gradient differences along the cardinal and diagonal directions respectively:

$$\varphi_c = \sqrt[2]{d_{ns}^{\boxed{2}} + d_{ew}^{\boxed{2}}} \quad [13]$$

$$\varphi_d = \sqrt[2]{d_{ne\,sw}^{\boxed{2}} + d_{nw\,se}^{\boxed{2}}} \quad [14]$$

After calculating the variances of elements in matrices φ_c and φ_d, the parameter F which shows the average value of the results, can be driven as below:

$$v_c = var(\varphi_c) \quad [15]$$

$$v_d = var(\varphi_d) \quad [16]$$

$$F = \frac{[v_c + v_d]}{2} \quad [17]$$

The adaptively set k can be then resulted by multiplying F by a scale factor:

$$K_{adapt} = F \times \gamma \quad [18]$$

Referring to extensive sets of experiments, there can be practically a linear relation between parameter k and the gained F. After performing some experiments on datasets, optimum values can be assigned to the scaling factor γ as well. Figure 2 illustrates the block-diagram of the proposed method which is called APDE, hereafter. As can be seen in this figure, each band of data is infected by a noisy process. The proposed adaptation method is then applied to each band and the PDE parameters are consequently tuned. The PDE-based restoration algorithm is then used in order to reduce the noise from each band. The restored data are then assessed using the quantitative metrics as well as an application-based approach which is based on SVM classification (Fig. 2).

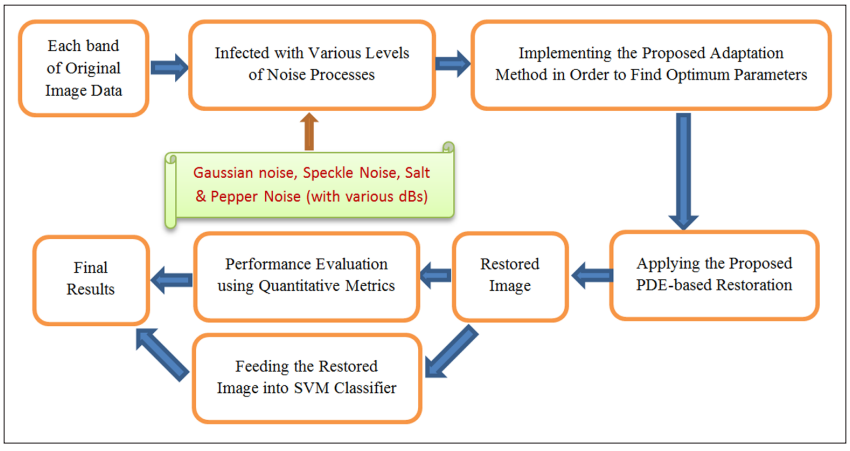

Figure 2 - Block-diagram of the proposed PDE-based image restoration algorithm (APDE method) which benefits from an innovative adaptation process.

Localizing the Adaptation Process Using Watershed Preprocessing (WShSeg-APDE)

The previous chapter reviewed the proposed PDE-based algorithm with adaptively set parameters (APDE). In this chapter, we improve the proposed method by performing a segmentation preprocessing step. The applied segmentation method is based on Watershed algorithm. Hereafter, this version of the proposed approach is called WShSeg-APDE. In WShSeg-APDE, each band of the given noisy data is partitioned into a set of relatively homogenous regions. By this, we localize the adaptation and consequently the denoising process to each segment. In other words, the APDE algorithms are individually applied to each segment instead of the whole image. Hence, a unique parameter k is found for each segment and the PDE algorithms (with different parameters) can be then applied distinctly to each segment.

Watershed-based segmentation methods have been widely used in various remote sensing applications [Tarabalka et al., 2010]. These techniques can be considered as the combination of edge-based and region-based approaches. In Watershed method, intensity of each pixel is interpreted as its elevation. Therefore, brighter pixels stand for higher altitudes while darker pixels represent the valleys. The Watershed transform then divides the image into catchment basins. Each basin is associated with one local minima (point surrounded by other points with higher gray-levels) of the gradient image.

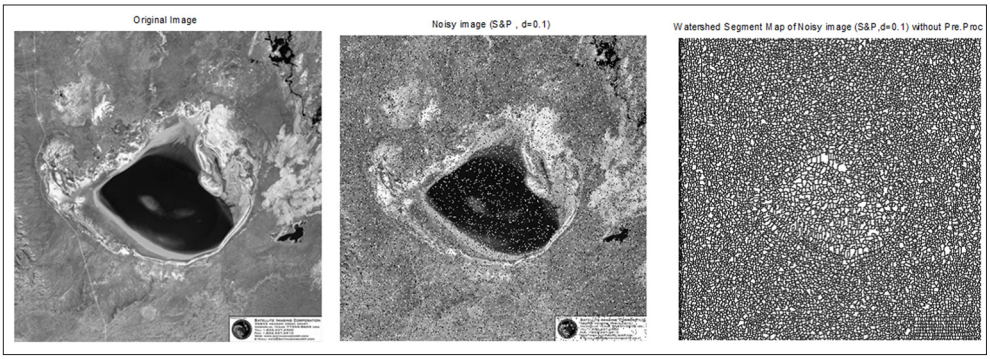

Figure 3 - The result of applying Watershed-based segmentation algorithm to a noisy single-layer dataset without using smoothing preprocessing step (noise is S&P with density equal to 0.1).

After applying the Watershed segmentation to the noisy image, the number of extracted segments may be considerably high (Fig. 3). In this situation, localizing the adaptation process for each segment is practically impossible or very time consuming. Therefore, we propose to apply a Gaussian filter to the noisy image before extracting the segments. As shown in Figure 4, after using the smoothing Gaussian filter, the number of segments is significantly reduced. The localized adaptation step can be now easily implemented to each segment. The segments maps in Figure 4 are colorized to provide better understanding of the segmentation results. From this figure, higher level of smoothing leads to less number of segments. In these experiments, a satellite image taken from Patagonia in Argentina has been used as the land-cover dataset. This image is gained using the Landsat-7 satellite which benefits from ETM+ (Enhanced Thematic Mapper Plus) sensors with spatial resolution of 30 meters. Our main reason for choosing Watershed technique among all

available segmentation methods is its ability to precisely find the image boundaries as well as its relatively straight-forward implementation. After adaptively finding the parameters for each segment and applying the PDE-based restoration algorithm, the denoised segments should be remerged together to form the whole restored image.

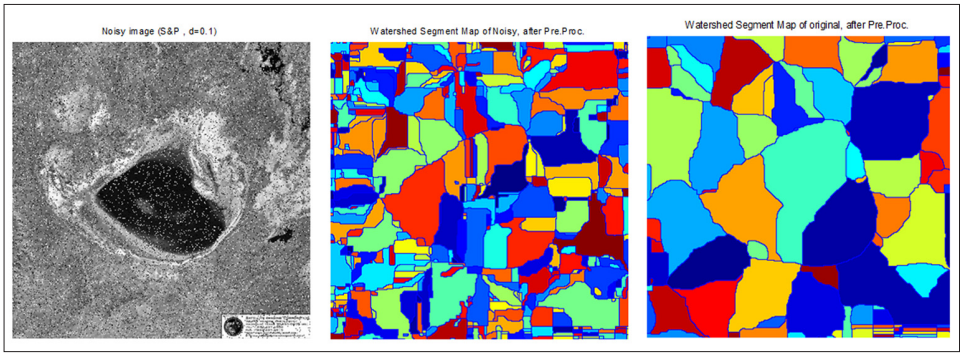

Figure 4 - The result of applying Watershed-based segmentation algorithm to the noisy Volcano after using two Gaussian filters with different smoothing levels: lower smoothing level (middle) and higher smoothing level (right).

Experimental Results

In this section, the outcomes of implementing the proposed approaches as well as the other restoration methods are reported in figures and tables.

Parameters Setting

The gained results show that for a special type of noise process, a unique γ can be achieved regardless of the level of noise. In other words, it is guaranteed that a predefined value for γ can be applied to different noisy images corrupted by the same noise type. In the experiments carried out in this paper, the value of γ is set to "0.01" for salt and pepper noise. Meanwhile, this value is set to "0.05" for the Gaussian as well as the Speckle noise.

The iteration number in numerical implementation of PDE is also set to 100 for all the experiments. Parameters α and β in the PDE process are equal to 0.30 and 0.10, respectively. For the experiments in which the SVM classification is applied, the number of training samples is set to 5% from each class. The other available samples are used for test. In other words, to fairly evaluate the gained results, the pixels from the training set are excluded from the test set and vice versa. The number of Monte-Carlo runs for the executed experiments is equal to 50 to obtain more robust results.

Datasets Description

Several single-layered remotely sensed images (with different textures) and multispectral datasets are used in our experiments. However, in this paper we report only the results of using one of the single-layered datasets as well as the F210 multispectral dataset in order to shorten the manuscript. Figure 5 shows the used single band data which is a high resolution panchromatic image with original spatial resolution of 10 meters per pixel. The resolution and size of the image are reduced to decrease the computational time. This data is obtained using the Advanced Land Observing Satellite (ALOS) via the Advanced Visible and Near

Infrared Radiometer-2 (AVNIR-2) sensor, in August 2007. The location of this urban image is the Fairbanks city in Alaska. To simplify the text, this single-layered dataset is called as "City", hereafter.

Figure 5 - One of the single-layered remotely sensed datasets utilized in our experiments:The Fairbanks city in Alaska.

The second data used in this paper is a multispectral dataset named as F210, comprising 12 bands, each in size of 220 by 140. The GTM (Ground Truth Map) of F210 covers 8 classes as well as a "non-farm" region. This dataset contains the agricultural data of the Indiana state in USA. The twelve bands of F210 are illustrated in Figure 6.

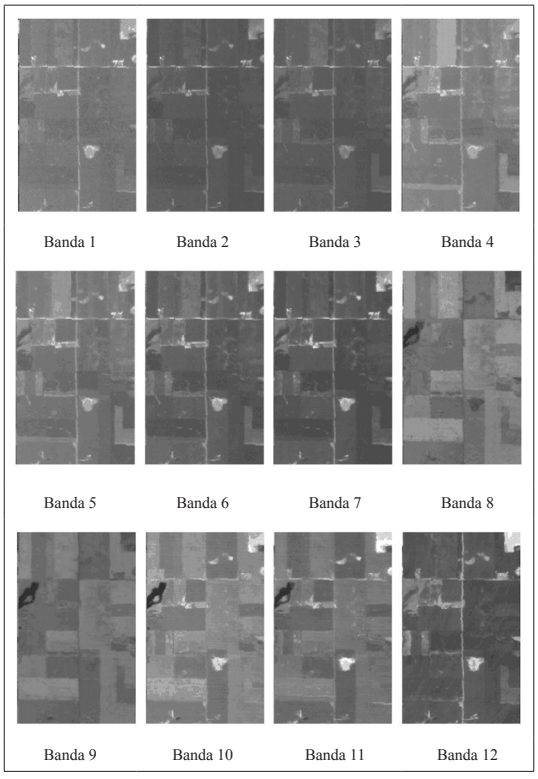

Figure 6 - Visual illustration of the twelve spectral bands of F210 dataset.

Quantitative Results of Applying the Proposed APDE Method

The gained outcomes of utilizing the proposed APDE method as well as other well-known restoration methods to the noisy single-layered City dataset have been depicted in Tables 1 and 2. In these tables, the PSNR of the restored methods as well as the Standard Deviations (SD) of the results (after 50 Mont-Carlo runs) are reported. In these experiments, we applied three common kinds of noise processes (Gaussian, Salt and Pepper, and Speckle) with different levels of variances and densities. In order to have a fair comparison with our proposed method, we report the best achieved results of the other restoration filters, in the upcoming tables.

Table 1 - Comparison of the gained results from applying different restoration methods on the noisy City dataset which is infected by various levels of Salt & Pepper noise and Gaussian noise (best results are emphasized by bold text).

		PSNR (Salt & Pepper noise case) (d= density of noise, P= PSNR of noisy data)				PSNR (Gaussian noise case) (v= variance of noise, P= PSNR of noisy data)			
		d=0.05	**d=0.08**	**d=0.1**	**d=0.2**	**v=0.01**	**v=0.02**	**v=0.03**	**v=0.04**
		P=23.22	**P=21.24**	**P=20.20**	**P=17.20**	**P=25.29**	**P=22.45**	**P=20.85**	**P=19.77**
Mean Filter (Optimal Kernel)	PSNR	27.47 [3x3]	26.59 [3x3]	26.036 [3x3]	23.89 [3x3]	28.16 [3x3]	27.27 [3x3]	26.58 [3x3]	25.98 [3x3]
	SD	0.0174	0.0280	0.0240	0.0299	0.0098	0.0085	0.0103	0.0142
Median Filter (Optimal Kernel)	PSNR	27.26 [3x3]	26.51 [3x3]	25.89 [3x3]	23.34 [5x5]	26.02 [3x3]	24.68 [3x3]	23.73 [3x3]	22.95 [3x3]
	SD	0.0290	0.04	0.0317	0.0106	0.0103	0.0112	0.0156	0.0193
Gaussian Filter (Optimal Kernel)	PSNR	28.14 [3x3] sig=1	27.05 [3x3] sig=1	26.36 [3x3] sig=1	23.96 [3x3] sig=2	29.07 [3x3] sig=1	**27.89** [3x3] sig=1	27.02 [3x3] sig=1	26.29 [3x3] sig=1
	SD	0.0389	0.0347	0.0197	0.0268	0.0160	0.0108	0.0116	0.0097
Wiener Filter (Optimal Kernel)	PSNR	25.83 [5x5]	25.08 [5x5]	24.67 [5x5]	23.25 [5x5]	**29.13** [3x3]	27.50 [3x3]	26.42 [3x3]	25.57 [3x3]
	SD	0.0268	0.0188	0.0237	0.0320	0.0103	0.0124	0.0122	0.0123
Mode Filter (Optimal Kernel)	PSNR	23.22	21.24	20.20	17.20	25.29	22.45	20.85	19.77
	SD	0.0284	0.0368	0.0271	0.0269	0.0105	0.0130	0.0145	0.007
P-M-PDE K=25	PSNR	28.97	27.60	26.47	22.26	26.71	24.69	23.40	22.42
	SD	0.0424	0.0632	0.0450	0.0541	0.0098	0.0139	0.0126	0.0170
P-M-PDE K=Optimal	PSNR	29.86 (k=46)	29.00 (k=49)	28.24 (k=55)	24.59 (k=67)	27.80 (k=60-100,>)	26.10 (k=60-100,>)	24.95 (k=80-100,>)	24.01 (k=80-100,>))
	SD	0.0196	0.0341	0.0503	0.0420	0.0097	0.0087	0.0149	0.0131
APDE K=Adaptive	PSNR	29.63 (k=31)	29.47 (k=39)	29.28 (k=44)	28.02 (k=62)	28.56 (k=82)	27.72 (k=100)	27.10 (k=119)	26.56 (k=136)
	SD	0.008	0.009	0.0216	0.0330	0.004	0.005	0.001	0.0112
APDE K = Optimal	PSNR	**29.67** (k=41)	**29.48** (k=43)	**29.28** (k=43)	**28.05** (k=54)	28.57 (k=4,70)	27.72 (k=60-100,>)	**27.11** (k=80-100,>130)	**26.57** (k=80-100,>150)
	SD	0.0073	0.0164	0.0235	0.0298	0.0068	0.0108	0.0067	0.0098

Table 2 - Comparison of the gained results from applying different restoration methods on the noisy City dataset which is infected by Speckle noise. (Results are averaged after 50 Monte-Carlo runs, v = variance of noise).

	v=0.04	v=0.05	v=0.06	v=0.07
	PSNR=25.9571	PSNR=25.0247	PSNR=24.2751	PSNR=23.6433
Mean Filter (Optimal Kernel)	28.22 [3x3]	27.97 [3x3]	27.76 [3x3]	27.54 [3x3]
Median Filter (Optimal Kernel)	25.97 [3x3]	25.59 [3x3]	25.25 [3x3]	24.92 [3x3]
Gaussian Filter (Optimal Kernel)	**29.16** [3x3] sig=1	**28.82** [3x3] sig=1	**28.55** [3x3] sig=1	**28.26** [3x3] sig=2
Wiener Filter (Optimal Kernel)	29.04 [3x3]	28.48 [3x3]	28.00 [3x3]	27.58 [3x3]
Mode Filter (Optimal Kernel)	25.95 [1x1]	25.02 [1x1]	24.27 [1x1]	23.64 [1x1]
P-M-PDE K=25	26.57	25.93	25.37	24.888
P-M-PDE K = Optimal	27.93 (k=95)	27.41 (k=98)	26.98 (k=113)	26.57 (k=124)
APDE K = Adaptive	28.48 (k=95)	28.22 (k=100)	27.92 (k=107)	27.51 (k=113)
APDE K = Optimal	28.48 (k=90)	28.22 (k=100)	27.93 (k=130)	27.53 (k=150)

Moreover, the results of applying the traditional Perona-Malik algorithm (P-M-PDE) to this dataset is reported for two different conditions listed as below:

1) when the parameter k is fixed to 25, as suggested by some of the other literatures;

2) when the value of parameter k is optimally found via testing all possible integer values for k (which are constrained in a reasonably defined range) and reporting the best output.

From these tables, the PSNR values after applying the proposed APDE approach are usually the best among the restoration methods, especially for Gaussian and Salt & Pepper noise cases. Meanwhile, the gained results using the adaptively set parameter k are very close to the optimal results obtained through exploring all possible values for k.

Another achievement from the tables is this fact that the optimum results obtained by the proposed method are generally better than the optimum outcomes achieved after using the P-M-PDE method. This fact is more obvious where the level of noise is considerably high. The main reason of this achievement is the more proper neighboring structure which is used in the numerical implementation of the proposed method. The visual results are also shown in Figure 7 for the City dataset infected by Salt and Pepper noise process.

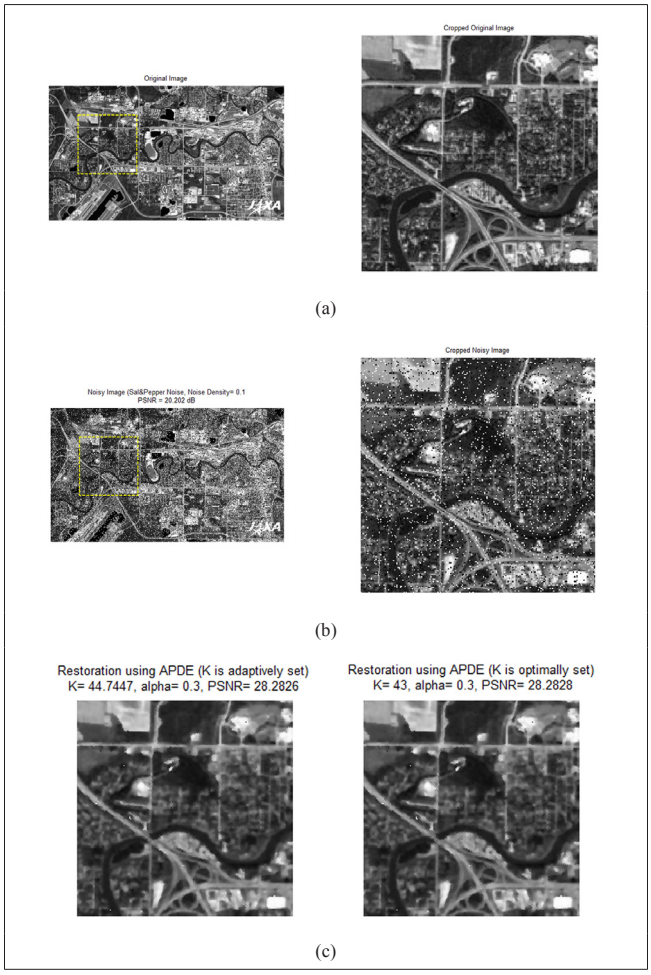

Figure 7 - (a) The original City image and a cropped area from the source image, (b) The noisy City image (Salt and Pepper noise with Density = 0.1) and a cropped area from this noisy image, (c) Two cropped areas from the restored versions of the noisy City image using proposed APDE method for two different conditions: (left) Parameter *k* is adaptively set, (right) Parameter *k* is optimally set after trying a range of integer values for *k*.

Classification Results of Applying the Proposed APDE Method

According to Table 1, the proposed APDE-based approach commonly expresses better results in terms of PSNR value, compared to other restoration techniques. Nevertheless, PSNR improvement is not assured for some probable conditions such as speckle noise cases (Tab. 2). Indeed, the main superiority of the proposed method is to keep the needed features in images while performing the denoising process. In cases involved in applications such as classification, this feature can be very helpful. Classification of the multispectral land-cover images is a common application in remote sensing which has been also addressed in Machala and Zejdova [2014]. Thus, first we utilize the proposed restoration method on the F210 multispectral dataset,

and then the original, noisy and denoised data will be fed into the SVM classifier. Note that, all the values reported as the results of the competing methods are optimally found. In other words, the reported values in the following figures are gained using the restoration filters with optimal features (i.e. kernel size and/or Sigma). For the P-M-PDE algorithm, the parameter k is also optimally set. Finally for the APDE method, this must be noted that the adaptive values for parameter k are individually calculated for each band of F210 data.

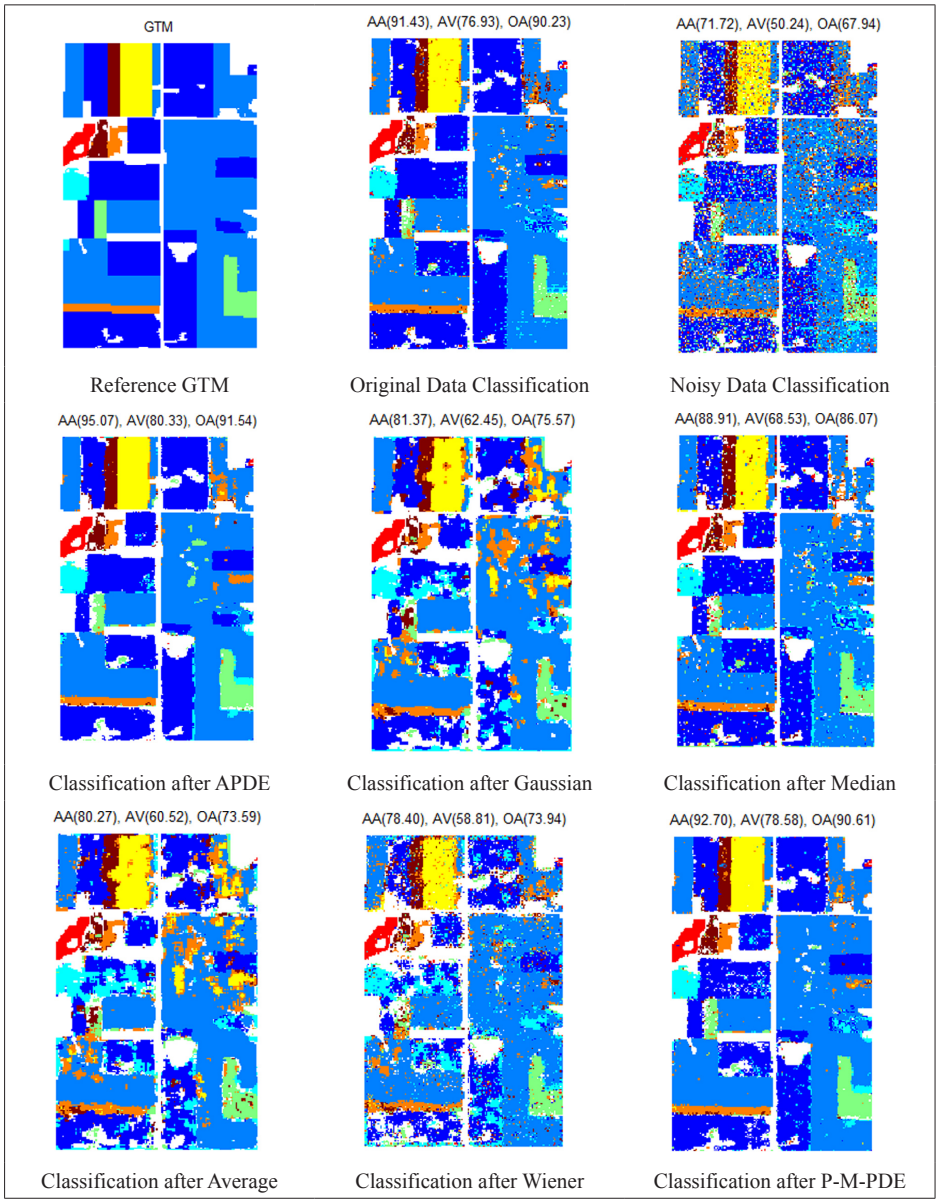

Figure 8 - The SVM classifier results after applying to F210 dataset, before and after using the restoration techniques, where the noise is Salt and Pepper with density equal to 0.05.

Figure 8 illustrates the results of applying the SVM classifier to the F210 dataset before and after using the restoration techniques. The reference Ground Truth Map (GTM) of the F210 data is shown in Figure 8a. Generally, more similarity between the reference GTM and the resulted class maps (after applying the restoration method) implies better accuracy of that given restoration method. The visual comparisons besides the gained classification ratios clearly prove the efficiency of our proposed approach. As can be seen in Figure 8, using the APDE restoration method enhances the averaged classification accuracy of SVM classifier from 71.71% (for the noisy data) to 95.07%. The results of the two PDE-based methods are even better than the outcomes which are gained by classifying the original data. However, this is not surprising since the smoothing effect of PDE leads to remove some unwanted details from the original data.

Results of Applying the Proposed WshSeg-APDE Method

Figure 9 illustrates the results of utilizing the proposed WshSeg-APDE method to a noisy single-layer dataset. After extracting the segments from noisy image and calculating the adaptive parameters, PDE algorithm is applied to each segment separately and then the restored segments are merged together to form Figure 9a. In order to compare the outcomes of applying the two proposed methods (WshSeg-APDE vs. APDE), Figure 9b is also provided. It concludes that by applying the proposed WShSeg-APDE method instead of the previously presented APDE technique, there exists a slight progress in the PSNR value (from 26.6 dB to 27.5 dB).

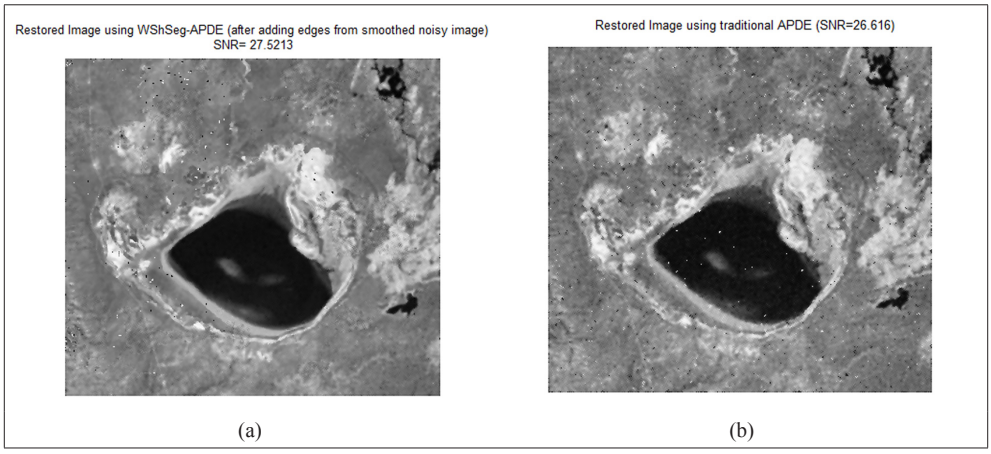

Figure 9 - (a) the result of the proposed WshSeg-APDE restoration method, after adding edges from the smoothed noisy image, and (b) the result of applying the traditional APDE.

Conclusion and Future Works

Performing the image restoration process in order to filter the probable noise from remote sensing datasets without dislocating or changing the important image features is the main goal of this paper. For this purpose, a restoration technique based on partial differential equations with adaptive parameters (APDE) is proposed by us. Moreover, a segmentation preprocessing step has been applied to each band of data before performing the restoration process (WShSeg-APDE). Using the Watershed segmentation technique, each band of the

noisy data is first partitioned into a set of disjointed segments and then the parameters are adaptively set in each segment. The proposed methods are implemented on several standard remote sensing datasets (including single-layer land-cover images as well as multispectral satellite images) and the obtained results are evaluated using a quantitative criterion (i.e. PSNR), and the application-based metrics which are the enhancement levels of the classification accuracy and validity after the denoising process. To have a fair comparison, we compare the performance of the proposed methods with those of the other restoration techniques such as Median, Mean (Average), Gaussian, Mode and Wiener filters as well as the traditional PDE algorithm presented by Perona and Malik. The achieved results represent a higher level of noise filtering along with better classification accuracies compared to the other restoration approaches. The adaptation method proposed in this paper, comparing to the other texture-based tools (such as the methods based on Gray-Level Co-occurrence Matrix (GLCM)) includes an outstanding priority which is its extensively low computational time. In the real-time applications, this counts as a key feature.

Although the traditional PDEs (which are implemented in real domain) are capable to denoise data while exempting edges from being smoothed, they cannot sharpen the edges. Therefore, our future works will emphasize on implementing some more state-of-the-art diffusion-based techniques (such as the higher order PDEs and non-linear complex PDEs) to gain better results. Moreover, we pursue our research to develop some robust statistical algorithms in order to automatically tune the other free parameters (such as Alpha and Beta) which have been used in the proposed approach. Applying some other modern classifiers instead of the traditional versions of SVM is also one of the main goals in our future works which may considerably result in better classification ratios.

References

Chui M., Feng Y., Wang W., Li Z., Xu X. (2012) - *Image Denoising Method with Adaptive Bayes Threshold in Nonsubsampled Contourlet Domain*. AASRI Procedia, 1: 512–518. doi: http://dx.doi.org/10.1016/j.aasri.2012.06.080.

Di Martino G., Iodice A., Riccio D., Ruello G. (2013) - *A Physical Approach for SAR Speckle Simulation: First Results*. European Journal of Remote Sensing, 46: 823-836. doi: http://dx.doi.org/10.5721/EuJRS20134649.

Gonzalez R.C., Woods R.E. (2004) - *Digital Image Processing*. Prentice Hall.

Herries G., Selige T., Danaher S. (1996) - *Singular value decomposition in applied remote sensing.* IEEE Colloquium on Image Processing for Remote Sensing.

Hird J.N., McDermid J. (2009) - *Noise reduction of NDVI time series: An empirical comparison of selected techniques*. Remote Sensing of Environment, 113 (1): 248-258. doi: http://dx.doi.org/10.1016/j.rse.2008.09.003.

Huang S., Liu D., Gao G., Guo X. (2009) - *A novel method for speckle noise reduction and ship target detection in SAR images*. Pattern Recognition, 42 (7): 1533-1542. doi: http://dx.doi.org/10.1016/j.patcog.2009.01.013.

Jansen M. (2001) - *Noise Reduction by Wavelet Thresholding*. Springer Verlag New York Incorporation. doi: http://dx.doi.org/10.1007/978-1-4613-0145-5.

Li M., Jia Z., Yang J., Hu Y., Li D. (2011) - *An Algorithm for Remote Sensing Image Denoising Based on the Combination of the Improved BiShrink and DTCWT*. Procedia Engineering, 24: 470-474. doi: http://dx.doi.org/10.1016/j.proeng.2011.11.2678.

Machala M., Zejdova L. (2014) - *Forest Mapping Through Object-based Image Analysis of Multispectral and LiDAR Aerial Data.* European Journal of Remote Sensing, 47: 117-131. doi: http://dx.doi.org/10.5721/EuJRS20144708.

Nadernejad E., Hassanpour H., MiarNaimi H. (2007) - *Image Restoration Using a PDE-based Approach*. IJE Transactions B: Applications, 20 (3).

Nazari A., Zehtabian A., Ghassemian H., Gribaudo M. (2014) - *Remotely Sensed Image Restoration Using Partial Differential Equations and Watershed Transformation.* Proceedings of SPIE, 9445, Seventh International Conference on Machine Vision (ICMV), Milan, Italy. doi: http://dx.doi.org/10.1117/12.2181817.

Perona P., Malik J. (1990) - *Scale-space and edge detection using anisotropic diffusion*. IEEE Transaction on Pattern Analysis and Machine Intelligence, 12 (7): 629-639. doi: http://dx.doi.org/10.1109/34.56205.

Ragged D., Bachmann M., Rivard B., Feng J. (2012) - *A spatial-spectral approach to deriving eigenvectors for remote sensing image transformations*. Geoscience and Remote Sensing Symposium (IGARSS): 4942-4945. doi: http://dx.doi.org/10.1109/IGARSS.2012.6352503.

Song Y.-C., Choi D.-H. (2005) - *Kernel Adjusted Wiener Filter for Image Enhancement*. Japanese Journal of Applied Physics, 44 (6): 3996. doi http://dx.doi.org/10.1143/JJAP.44.3996.

Su K., Fu H., Du B., Cheng H., Wang H., Zhang D. (2012) - *Image denoising based on learning over-complete dictionary.* 9th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD). doi: http://dx.doi.org/10.1109/FSKD.2012.6234041.

Tarabalka Y., Chanussot J., Benediktsson J.A. (2010) - *Segmentation and classification of hyperspectral images using watershed transformation*. Pattern Recognition, 43 (7): 2367-2379. doi: http://dx.doi.org/10.1016/j.patcog.2010.01.016.

Witkin A.P. (1983) - *Scale-Space Filtering: A New Approach To Multi-Scale Description*. Report of Fairchild Laboratory for Artificial Intelligence Research.

Zehtabian A., Ghassemian H. (2015) - *An Adaptive Pixon Extraction Technique for Multispectral /Hyperspectral Image Classification*. IEEE Geoscience and Remote Sensing Letters, 12 (4): 831-835. doi: http://dx.doi.org/10.1109/LGRS.2014.2363586.

Zehtabian A., Ghassemian H. (2013) - *A noise-robust SVD-ML based classification method for multi-spectral remote sensing images*. 21st Conference on Electrical Engineering (ICEE), pp. 1-6. doi: http://dx.doi.org/10.1109/IranianCEE.2013.6599592.

© 2015 by the authors; licensee Italian Society of Remote Sensing (AIT). This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).