



HAL
open science

Enhancement of Component Images of Multispectral Data by Denoising with Reference

Sergey Abramov, Mikhail Uss, Vladimir Lukin, Benoit Vozel, Kacem Chehdi,
Karen Egiazarian

► **To cite this version:**

Sergey Abramov, Mikhail Uss, Vladimir Lukin, Benoit Vozel, Kacem Chehdi, et al.. Enhancement of Component Images of Multispectral Data by Denoising with Reference. Remote Sensing, 2019, 11 (6), pp.611. 10.3390/rs11060611 . hal-02135627

HAL Id: hal-02135627

<https://univ-rennes.hal.science/hal-02135627>

Submitted on 7 Jul 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Article

Enhancement of Component Images of Multispectral Data by Denoising with Reference

Sergey Abramov ¹, Mikhail Uss ¹, Vladimir Lukin ^{1,*}, Benoit Vozel ², Kacem Chehdi ² and Karen Egiazarian ³

¹ Department of Information and Communication Technologies, National Aerospace University, 61070 Kharkov, Ukraine; s.abramov@khai.edu (S.A.); mykhail.uss@gmail.com (M.U.)

² Engineering School of Applied Sciences and Technology, University of Rennes 1, 22305 Lannion, France; benoit.vozel@univ-rennes1.fr (B.V.); kacem.chehdi@univ-rennes1.fr (K.C.)

³ Computational Imaging Group, Tampere University, Tampere 33720, Finland; karen.egiazarian@tuni.fi

* Correspondence: lukin@ai.kharkov.com

Received: 12 January 2019; Accepted: 10 March 2019; Published: 13 March 2019



Abstract: Multispectral remote sensing data may contain component images that are heavily corrupted by noise and the pre-filtering (denoising) procedure is often applied to enhance these component images. To do this, one can use reference images—component images having relatively high quality and that are similar to the image subject to pre-filtering. Here, we study the following problems: how to select component images that can be used as references (e.g., for the Sentinel multispectral remote sensing data) and how to perform the actual denoising. We demonstrate that component images of the same resolution as well as component images of a better resolution can be used as references. To provide high efficiency of denoising, reference images have to be transformed using linear or nonlinear transformations. This paper proposes a practical approach to doing this. Examples of denoising tests and real-life images demonstrate high efficiency of the proposed approach.

Keywords: remote sensing; multispectral imaging; DCT-filtering; vectorial (three-dimensional) filtering; BM3D-filtering; filtering with reference

1. Introduction

Remote sensing (RS) is widely used in many applications [1,2]. It provides high information content of images, fast data collection possibility for large territories, availability of different sensors both airborne and spaceborne, and so on. Modern remote sensing tends to improve the spatial resolution of sensors and to make them multichannel, for example, multi-polarization radar, hyperspectral, and multispectral [1–4]. Recently, a multispectral sensor Sentinel 2 has been launched and has already produced valuable and interesting data [5].

Multichannel data contain more information about a sensed terrain compared with single-channel data. However, there exists the following problem in multichannel sensing—images in one or a few components are corrupted by noise [4,6,7] (actually, noise is present in all images, but its influence in some components is negligible, as will be shown later). If a noise is intensive (input peak signal-to-noise ratio (PSNR) is low), it is worth applying pre-filtering in order to enhance RS data and to improve the performance of the next RS data processing, such as classification, segmentation, parameter estimation, and so on [4,8].

There are many approaches to filter multichannel images. They can be classified into component-wise, vectorial (three-dimensional, 3D), and hybrid. Component-wise denoising is the simplest among them, allowing parallel processing of component images [7–11]. However, similar to

filtering of color images [9–11], component-wise filtering is not able to exploit inter-channel correlation of component images inherent for practically all types of multichannel images [11–14]. Meanwhile, its exploiting often leads to considerably more efficient denoising [7–9]. A question is how large is the inter-channel correlation and how to exploit it properly and efficiently.

Among first filters exploiting inter-channel correlation are vector filters based on order statistics (see the works of [9–11] and references therein). Originally, they were oriented on removal of impulsive and mixed noises. However, impulse noise is rarely met in RS data produced by modern multispectral sensors.

Later, denoising methods based on orthogonal transforms appeared with the main application to color [15–17] images where components are processed jointly. Some of these methods have been modified to work with multichannel RS images [15,18]. The necessity of such modifications appears because a noise can be of different intensity, and even different type, in component images of multichannel RS data [3,4,6,15,19]. This either makes inapplicable filtering techniques designed to cope with identical characteristics of the noise in all components [10,19] or reduces their performance.

There are several approaches to deal with the aforementioned non-identical characteristics. The most typical ones are to carry out proper variance stabilizing transforms [8], normalize component images in channels [15], perform pre-filtering [15], modify the algorithm [18], and so on. One problem is that this makes filtering more complex and makes it necessary to have a priori information on noise characteristics or to estimate them accurately in a blind manner arises.

An important peculiarity and positive feature of this group of methods is that usually the largest positive effect due to filtering occurs for component images that are “the noisiest” [15,18,20]. The joint processing of more component images might provide more efficient denoising [20], but this does not happen necessarily. Meanwhile, the joint use of more component images leads to difficulties in processing dealing with more memory and time needed. Thus, the amount of jointly processed component images has to be either optimized or chosen in a reasonable way. Unfortunately, such an optimization has not been done yet.

A new group of methods of multichannel data filtering has recently appeared that can be treated as a hybrid. They exploit inter-channel correlation in different ways. The main idea is that in multichannel RS data, there can be the so-called low quality or “junk channels” (component images) and high quality component images in the sense of high input peak signal-to-noise ratio (PSNR) and absence of other distortions. There was a discussion concerning is it worth keeping junk channels for further processing and analysis [4,21]. Currently, many researchers consider that it is worth keeping them for further consideration under the condition that images in “junk channels” are pre-filtered with high efficiency [4,22–24]. A question is how such filtering can be done?

There are many proposed solutions that employ different principles. The method by Yuan et al. [25] uses the total variation algorithm applied in both spatial and spectral views. The problem is that the possible signal-dependent nature of the noise has not been taken into account. A method based on the parallel factor analysis (PARAFAC) approach to denoising has been proposed in the work of [26], but it also assumes an additive noise model. Anisotropic diffusion is applied for hyperspectral imagery enhancement in the work of [27], demonstrating also improvement of classification, but the noise model is not specified. Chen and Qian have proposed to filter hyperspectral data using principal component analysis and wavelet shrinkage [28], but again, the additive noise model was considered. Meanwhile, as will be shown in the next section, the signal-dependent component of the noise can be present.

Recently, the use of non-local based approaches to denoising multichannel images has become popular [2,24,29]. The main progress and benefits result from the fact that similar patches that can be used in collaborative denoising can be found not only in a given component image, but also in other component images. Other positive outcomes result from the fact that in multichannel RS data, there can be almost noise-free component images (called references) that are quite similar to a noisy component image that needs enhancement [22–24,30,31]. The main ideas are either to retrieve

and exploit some information from the reference (for example, about positions of edges [22]) or to incorporate reference image(s) into processing directly. Important items here are to find a proper reference and to make it as “close” to the noisy image as possible (e.g., by appropriate nonlinear transformation [31]). The approach [24,30,31] allows using both DCT [15] and BM3D [32] filters as well as to easily cope with signal-dependent noise in the component image to be denoised by applying a proper variance stabilizing transform (VST) to it before filtering.

These properties can be very useful in the denoising of junk components in multispectral data, for example, Sentinel-2 recently put into operation for which noise has been shown to be signal-dependent [33] and having quite different characteristics in different component images. One more specific property of multispectral data acquired by Sentinel-2 is that different component images are characterized by different spatial resolutions. There are three component images (##1, 9, and 10) that have a resolution of $60 \times 60 \text{ m}^2$; six component images (## 5, 6, 7, 8A, 11, and 12) that have a resolution of $20 \times 20 \text{ m}^2$, while the remaining four (##2, 3, 4, and 8) possess the best resolution of $10 \times 10 \text{ m}^2$. This feature distinguishes Sentinel-2 multispectral data from hyperspectral data partly discussed above, which have approximately the same spatial resolution in all sub-band images. This difference shows that methods of joint processing of two or more component images that have different resolution have to take this fact into consideration.

The aforementioned peculiarities (signal-dependent character, sufficiently different input PSNR and resolution) of Sentinel-2 multispectral data determine the novelty of the problem statement—to design methods for noise removal in component images that originally have low input PSNR. Recall that recent studies [34,35] show that it is difficult to expect high efficiency of any kind of image denoising if input PSNR is high and/or image is textural or contains a lot of fine details (these are just the cases for many RS images). So, we focus on noise removal in particular component images of Sentinel-2 data supposing that filtering of other component images is not needed (this allows for saving time and resources for data pre-processing).

The novelty of our proposed approach consists of the following two aspects. First, we show that component images with a resolution better than a component image to be denoised can be used. Second, by analyzing component image similarity, we propose a method to select component images that can be used as references.

2. Materials and Methods

2.1. Image/Noise Model and Basic Principles of Image Denoising with Reference

A general image/noise model considered below is as follows:

$$I_{ij}^n = I_{ij}^t + n_{ij}(I_{ij}^t), \quad i = 1, \dots, I_{Im}, \quad j = 1, \dots, J_{Im}, \quad (1)$$

where I_{ij}^t denotes the true image value in an ij -th pixel; $n_{ij}(I_{ij}^t)$ is the noise statistical properties, which are dependent on I_{ij}^t ; and I_{Im}, J_{Im} define the processed image size. If one deals with a multichannel image, index q can be added to all components in (1). Note that if a multispectral image is considered, even I_{Im} and J_{Im} should have index q because spatial resolution and the number of pixels in each component image, respectively, is individual.

Let us explain from the very beginning why we rely on the signal-dependent model of the noise (1). The model that assumes variance in an ij -th pixel is the following:

$$\sigma_{ij}^2 = \sigma_0^2 + kI_{ij}^t, \quad (2)$$

where σ_0^2 is the variance of the signal independent (SI) noise component and k is the parameter that determines the properties of the signal-dependent (SD) component tested in the work of [33] for Sentinel-2 multispectral images, provided after applying light compression by JPEG2000. Moreover, even more complicated models of signal dependence have been considered in the work of [33]

(where specific effects appear as a result of a lossy compression), but we will further accept the general model (2).

For this model, one can use the so-called equivalent noise variance that is equal to $\sigma_{\text{eq}}^2 = \sum_{i=1}^{I_{\text{Im}}} \sum_{j=1}^{J_{\text{Im}}} (I_{ij}^n - I_{ij}^t)^2 / (I_{\text{Im}} \times J_{\text{Im}})$ if the true image is available. Alternatively, it can be estimated as $\sigma_{\text{eq}}^2 = \sigma_0^2 + k \sum_{i=1}^{I_{\text{Im}}} \sum_{j=1}^{J_{\text{Im}}} I_{ij}^t / (I_{\text{Im}} \times J_{\text{Im}}) \approx \hat{\sigma}_0^2 + \hat{k} I_{\text{mean}}$ if the true image is not available, but quite accurate estimates $\hat{\sigma}_0^2$ and \hat{k} were obtained in a blind manner from images at hand [6,33,36], where I_{mean} is the image mean. This means that if an equivalent noise variance σ_{eq}^2 in an image is sufficiently larger than σ_0^2 , then noise should be considered signal-dependent and this feature has to be taken into account in image processing. Gaussianity tests carried out for manually chosen homogeneous regions show that noise is practically Gaussian.

Let us analyze multispectral data from Sentinel-2 using the estimates of σ_0^2 and k provided by the method [33]. The noise parameter estimates for two granules (sets of multispectral data) are presented in Table 1. As one can see, in practically all component images, the equivalent variance is considerably larger than σ_0^2 , although the contribution of the SD component is always smaller than that of the SI component. The only exception is the component image in channel #10, where σ_0^2 is practically the same as the corresponding equivalent variance. This means that the signal-dependent nature of the noise has to be taken into account.

Note that the equivalent variance of the noise is the smallest in component image #10. Thus, one might think that this image is the least noisy. However, this conclusion is not correct, as we have not taken into account the range of image representation. Let us also analyze peak signal-to-noise ratio. To avoid the possible presence of hot pixels in data and bright points, consider below the so-called robust estimate of input PSNR determined as $\text{PSNR}_{\text{inp}}^{\text{rob}} = 10 \log_{10} (D_{\text{rob}}^2 / \sigma_{\text{eq}}^2)$, where $D_{\text{rob}} = I^{(p)} - I^{(r)}$, $p = 0.99 I_{\text{Im}} J_{\text{Im}}$, and $r = 0.01 I_{\text{Im}} J_{\text{Im}}$; $I^{(p)}$ and $I^{(r)}$ are the p -th and r -th order statistics of image values, respectively. The obtained values of $\text{PSNR}_{\text{inp}}^{\text{rob}}$ are presented in Table 1. It is seen that the values of this metric are larger than 45 dB for 12 out of 13 component images. This means that these images are of high quality and noise cannot be noticed in visualized component images [37] (one example is shown in Figure 1a). Meanwhile, there is also an image in sub-band 10 for which $\text{PSNR}_{\text{inp}}^{\text{rob}}$ is only 11.6 dB and, therefore, noise is visible (one example is given in Figure 1b). As it is seen, noise is not white because specific diagonal structures are observed. Such artifacts can be, most probably, suppressed in frequency domain by special pre- or post-processing. However, their removal is out of the scope of this paper.

One more observation is that these component images are similar to each other and the cross-correlation factor $R_{\#10}$ for them is equal to 0.57.

The cross-correlation factors $R_{\#10}$ for the component image in channel #10 and other component images are given in Table 1. One can see that the correlation is low for component images (##1...4) that relate to the visible range, but it increases and exceeds 0.77 for the components number 11 and 12. If resolutions in channel 10 and another channel image are different, the corresponding downsampling is applied before calculation of the correlation factor.

One question concerns a stability of noise properties. To check this, we carried out estimation of noise and image parameters for another granule. They are given in the lower part of Table 1. As one can see, there are certain differences, but the main tendencies are the same. There is a comparable contribution of both signal independent and dependent components. The most “noisy” is the image in channel #10. The most similar images to the image in channel #10 are the images in channels ## 11 and 12.

We processed the image in Figure 1b by the 2D (component-wise) DCT based filter [38] with standard settings. The output is presented in Figure 2 and it is seen that the noise has been partly removed, but the image quality still remains poor (details and edges are smeared, strip-like interferences remain). This means that more efficient denoising is required.

We also analyzed other fragments and other granules of multispectral data produced by Sentinel-2. The image and noise properties are similar in the sense of noise nature and characteristics as well as values of $\text{PSNR}_{\text{inp}}^{\text{rob}}$ and inter-channel correlation.

Table 1. Noise parameters in component images of Sentinel-2. PSNR—peak signal-to-noise ratio.

Granule 1													
Channel Name	01	02	03	04	05	06	07	08	8A	09	10	11	12
σ_0^2	218.0	75.7	26.4	53.3	27.5	42.9	71.7	92.8	103.1	38.2	9.8	31.7	54.7
k	0.042	0.024	0.015	0.027	0.013	0.019	0.030	0.028	0.035	0.052	0.010	0.011	0.022
σ_{eq}^2	291.5	109.5	43.5	82.4	42.0	68.1	114.9	132.4	156.3	65.6	10.0	45.2	72.3
$\text{PSNR}_{\text{inp}}^{\text{rob}}$	45.7	51.6	56.0	54.7	57.9	56.2	54.1	53.1	52.9	46.9	11.6	54.7	50.3
$R_{\#10}$	0.18	0.21	0.29	0.36	0.42	0.49	0.52	0.53	0.56	0.570	1.00	0.782	0.772
Granule 2													
Channel Name	01	02	03	04	05	06	07	08	8A	09	10	11	12
σ_0^2	110.9	66.45	1.87	37.66	9.47	10.15	57.14	53.42	35.65	42.51	7.68	15.27	36.98
k	0.003	0.026	0.023	0.030	0.017	0.048	0.025	0.040	0.059	0.014	0.024	0.017	0.044
σ_{eq}^2	114.7	96.46	22.61	61.05	25.21	78.23	97.37	114.5	139.1	49.50	7.77	44.77	94.28
$\text{PSNR}_{\text{inp}}^{\text{rob}}$	29.36	32.72	41.55	41.31	46.90	51.58	52.40	51.38	51.47	44.11	11.91	52.48	47.89
$R_{\#10}$	0.366	0.384	0.412	0.467	0.490	0.243	0.229	0.243	0.261	0.308	1.00	0.651	0.618

For image denoising with a reference, it is assumed that a reference image or a set of reference images I_{ijs}^{ref} , $i = 1, \dots, I_{\text{Im}}$, $j = 1, \dots, J_{\text{Im}}$, $s = 1, \dots, S$ are available, where $S \geq 1$ defines a number of potential reference images. All candidate reference images are supposed to be noise-free, or at least such that input PSNRs for them are 10 dB or more larger than input PSNR for the image to be denoised. It is also supposed that downsampling is applied if the reference image has other resolution than the noisy one.

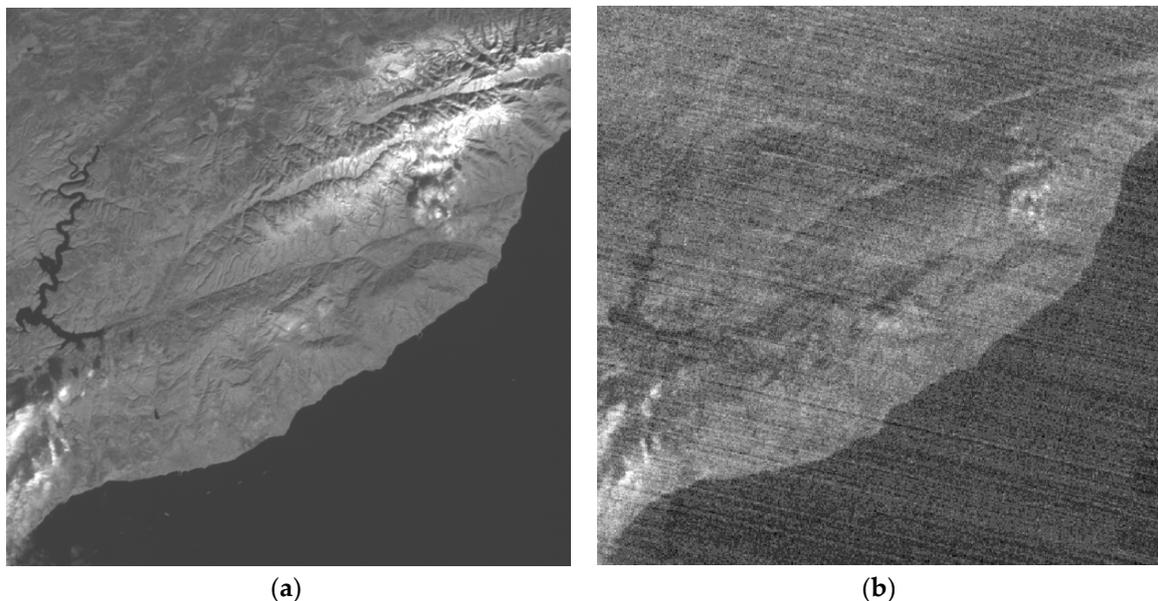


Figure 1. Visualized fragments of Sentinel-2 images of size 512×512 pixels in band #09 (a) and band #10 (b).

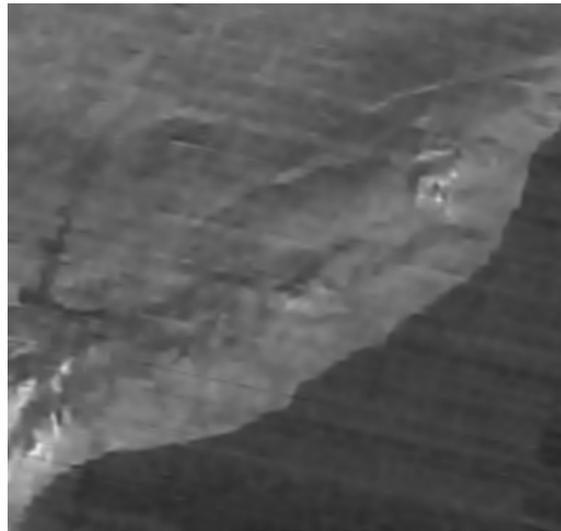


Figure 2. The output of the DCT-based filter.

Another assumption is that potential reference images are in some sense similar to I_{ij}^n , $i = 1, \dots, I_{Im}$, $j = 1, \dots, J_{Im}$. It is known that similarity of images can be measured differently—mean square errors (MSE) between images, cross-correlation factor, and so on. One can also apply a linear or nonlinear transform of reference image(s) before calculating measures of closeness. In this work, we assume that a linear or non-linear transform has been applied to the reference image in order to make it as close as possible in MSE sense to the noisy image subject to denoising.

There are several possible cases. Let us consider them more in detail with a discussion of when and why each of them takes place.

The first practical case is that noise in I_{ij}^n , $i = 1, \dots, I_{Im}$, $j = 1, \dots, J_{Im}$ is additive, and then the main metric that describes similarity is

$$\text{MSE}_{n \text{ rmod}} = \sum_{i=1}^{I_{Im}} \sum_{j=1}^{J_{Im}} \left(I_{ij}^n - I_{ij}^{\text{ref mod}} \right)^2 / (I_{Im} \times J_{Im}). \quad (3)$$

Here, $I_{ij}^{\text{ref mod}}$, $i = 1, \dots, I_{Im}$, $j = 1, \dots, J_{Im}$ defines the modified reference image, which can be either linearly transformed as

$$I_{ij}^{\text{ref mod}} = S_0 I_{ij}^{\text{ref}} + \Delta_0, \quad i = 1, \dots, I_{Im}, \quad j = 1, \dots, J_{Im}, \quad (4)$$

or nonlinearly transformed as

$$I_{ij}^{\text{ref mod}} = \Psi(I_{ij}^{\text{ref}}), \quad i = 1, \dots, I_{Im}, \quad j = 1, \dots, J_{Im}, \quad (5)$$

where S_0 , Δ_0 denote the parameters of linear least MSE regression (4) (case 1) and $\Psi(I_{ij}^{\text{ref}})$ defines nonlinear transformation (case 2) that leads to minimizing $\text{MSE}_{n \text{ rmod}}$.

Two other cases relate to the noise model described by (1) and (2). Then, if the noise is signal-dependent, it is usually recommended to apply a proper homomorphic or variance stabilizing transform (VST) to deal with an additive noise (although often non-Gaussian) in filtering [8,39]. An advantage of this approach is that the additive nature of the noise in an image to be denoised allows for applying a wider set of efficient filters [34]. As VST, the generalized Anscombe transform [8] or logarithmic transform [39] can be used, depending on the type of signal-dependent noise one deals in each particular case.

If VST is applied, one has I_{ij}^{nVST} , $i = 1, \dots, I_{Im}$, $j = 1, \dots, J_{Im}$ and may use either a linear transform (case 3) or nonlinear transform (case 4) and minimize either

$$MSE_{n \text{ rmod}} = \sum_{i=1}^{I_{Im}} \sum_{j=1}^{J_{Im}} \left(I_{ij}^{nVST} - S_0 I_{ij}^{ref} - \Delta_0 \right)^2 / (I_{Im} \times J_{Im}) \quad (6)$$

or

$$MSE_{n \text{ rmod}} = \sum_{i=1}^{I_{Im}} \sum_{j=1}^{J_{Im}} \left(I_{ij}^{nVST} - \Psi(I_{ij}^{ref}) \right)^2 / (I_{Im} \times J_{Im}), \quad (7)$$

respectively.

Let us now recall how denoising with a reference is carried out for the simplest case of having I_{ij}^n , $i = 1, \dots, I_{Im}$, $j = 1, \dots, J_{Im}$ and a properly chosen $I_{ij}^{ref \text{ mod}}$, $i = 1, \dots, I_{Im}$, $j = 1, \dots, J_{Im}$. Then, the noisy and the reference images are denoised jointly. A two-point DCT is applied first in the “vertical direction”, getting “sum” and “difference” images. The obtained images are filtered by the 2D DCT-based filter or by BM3D with properly selected hard thresholds. After this, inverse two-point DCT is applied and the obtained first component is considered as the filtered image.

If VST is used, then the same operations are applied to I_{ij}^{nVST} , $i = 1, \dots, I_{Im}$, $j = 1, \dots, J_{Im}$ and $I_{ij}^{ref \text{ mod}}$, $i = 1, \dots, I_{Im}$, $j = 1, \dots, J_{Im}$ obtained by minimizing (6) or (7). The only difference is that the denoised image has to be subject to inverse VST. The described operations are illustrated in Figure 3.

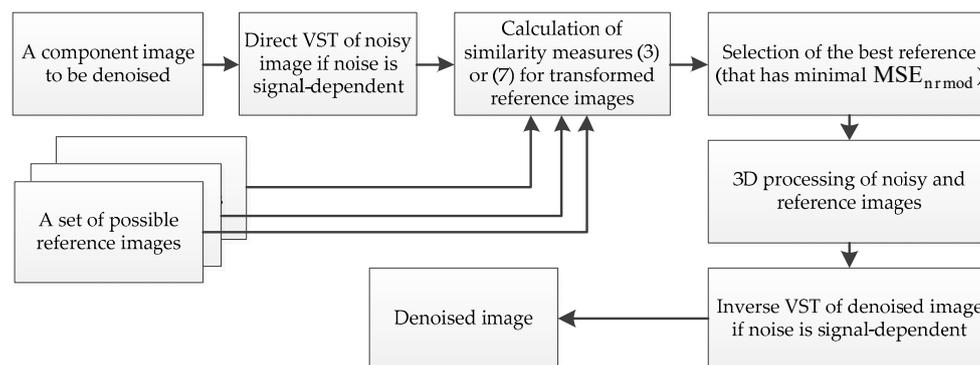


Figure 3. Block-diagram of the proposed processing approach. VST—variance stabilizing transform; MSE—mean square error.

Also, note that it is possible to have two or more modified reference images after a three-point or more DCT applied in the vertical direction to decorrelate data. We considered two reference images instead of one in the work of [31], and filtering has occurred to be more efficient in terms of standard metrics, such as output PSNR, and visual quality metrics, such as PSNR-HVS-M, which takes into account two important properties of human vision system (HVS), namely, less sensitivity to distortions in high frequency components and masking (M) effect of image texture and other heterogeneities [40]. Besides, after two- or three-point DCT, it is possible to apply component-wise different filters including standard DCT, BM3D, or others. Usually, if a given filter is more efficient in component-wise (single-channel) denoising, its use is also beneficial in the considered denoising with a reference [30]. It is also worth stressing that optimal (recommended) parameters of thresholds applied in DCT coefficient thresholding have been determined in the literature [24,30,31]. These thresholds differ from those usually recommended for the cases in which these filters are employed for noise removal in single channel images. Thus, in our further studies, we will use just optimal thresholds.

2.2. Performance Criteria

We start analyzing the performance of methods of image filtering with reference(s) for simulated data [24,30,31]. In our simulations, four test images typical for remote sensing, presented in Figure 4

and denoted as FR01, FR02, FR03, and FR04, and two high quality component images denoted RS1 and RS2 of AVIRIS hypercube of data were used. Additive White Gaussian Noise (AWGN) with variance σ^2 was artificially added to these images (note that noise in original component images of hyperspectral images was considered negligible).

In order to simulate reference images for all these test images, we need to ensure that reference images are similar to the test ones according to certain similarity measures (i.e., to have sufficient but not too high cross-correlation factor). At the same time, they also have to be different in several senses—with a different dynamic range, and containing some additional content not present in the image to be denoised (see the example in Figure 1). We cannot simply distort the original test image randomly, as this will be equivalent to adding a noise and $PSNR_{inp}$ decreasing. The use of a more complex simulation requires knowledge of the image information content formation, which is a priori unknown. Because of this, and based on thorough empirical study of multichannel images, we simulated the reference image as $I_{ij}^{ref} = 32\sqrt{I_{ij}^t + 0.5I_{ij}^{t180}}$, $i = 1, \dots, I_{Im}$, $j = 1, \dots, J_{Im}$, where I_{ij}^{t180} , $i = 1, \dots, I_{Im}$, $j = 1, \dots, J_{Im}$ denotes the same noise free test image rotated by 180° . Thus, as a reference image, we use a weighted sum of the original image with its copy rotated by 180° . This allows us to provide a correlation factor of the same level as for real-life multispectral data in channels 10–12.

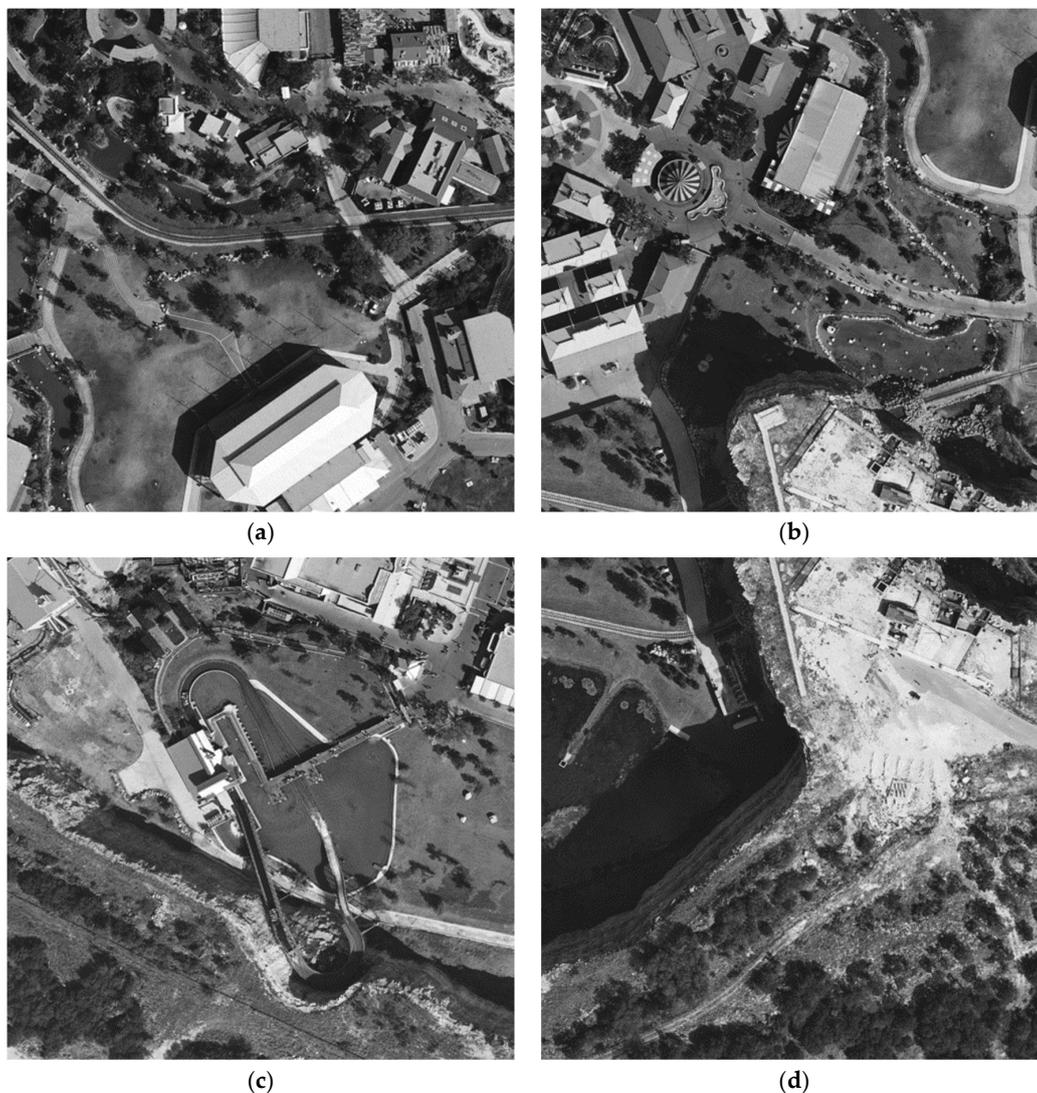


Figure 4. Remote sensing (RS) test images FR01 (a), FR02 (b), FR03 (c), and FR04 (d).

The obtained reference images for the test images FR01 and FR02 in Figure 4 are visualized in Figure 5 (note that the reference images are in the dynamic range, considerably different from those of the original range 0 . . . 255).

To characterize the efficiency of filtering, we used the following metrics. First, input PSNR is defined as

$$\text{PSNR}_{\text{inp}} = 10 \log_{10} \left(\text{DR}^2 / \sum_{i=1}^{I_{\text{Im}}} \sum_{j=1}^{J_{\text{Im}}} (I_{ij}^n - I_{ij}^t)^2 / (I_{\text{Im}} \times J_{\text{Im}}) \right) = 10 \log_{10} (\text{DR}^2 / \sigma^2), \quad (8)$$

where DR denotes the range of image representation and σ^2 is a noise variance (equivalent variance if noise is signal-dependent). Output PSNR is expressed as

$$\text{PSNR}_{\text{out}} = 10 \log_{10} \left(\text{DR}^2 / \sum_{i=1}^{I_{\text{Im}}} \sum_{j=1}^{J_{\text{Im}}} (I_{ij}^f - I_{ij}^t)^2 / (I_{\text{Im}} \times J_{\text{Im}}) \right) = 10 \log_{10} (\text{DR}^2 / \text{MSE}_{\text{out}}), \quad (9)$$

where MSE_{out} is the output mean square error (MSE). Effectiveness is then characterized by

$$\delta\text{PSNR} = \text{PSNR}_{\text{out}} - \text{PSNR}_{\text{inp}}. \quad (10)$$



Figure 5. Visualized reference images for the test images FR01 (a) and FR02 (b).

Alongside PSNR, we would like to analyze the visual quality of original (noisy) and filtered images. To do this, we propose using the metric PSNR-HVS-M (denoted later as PHVSM) [40]. Then, one has

$$\text{PHVSM}_{\text{inp}} = 10 \log_{10} (\text{DR}^2 / \text{MSE}_{\text{inp}}^{\text{HVS}}), \quad (11)$$

$$\text{PHVSM}_{\text{out}} = 10 \log_{10} (\text{DR}^2 / \text{MSE}_{\text{out}}^{\text{HVS}}), \quad (12)$$

$$\delta\text{PHVSM} = \text{PHVSM}_{\text{out}} - \text{PHVSM}_{\text{inp}}, \quad (13)$$

where $\text{MSE}_{\text{inp}}^{\text{HVS}}$ and $\text{MSE}_{\text{out}}^{\text{HVS}}$ are input and output MSEs, respectively, calculated while taking into account the aforementioned peculiarities of HVS.

Note that a filtering method can be considered good if it performs better than others for a wide set of test images and a wide range of noise variances (input PSNRs).

3. Results

3.1. Analysis of Simulation Data

The obtained data are presented in Table 2. To compare the performance of denoising techniques, we present input PSNR given by (8) and PSNR for outputs of four filters, namely, component-wise DCT filter (denoted as 2D), the proposed filtering with reference with linear correction (expressions (4) and (3), denoted as 3D_{C1}), and denoising with reference with nonlinear correction (expressions (5) and (3), denoted as 3D_{C2}, the second order polynomials with optimal parameters were employed here). Without loss of generality, the results were obtained for additive white Gaussian noise. Three values of noise variance σ^2 were analyzed: 10 (invisible noise), 25 (noise is visible in homogeneous image regions), and 100 (intensive noise).

Table 2. Simulation data.

Image	FR01		FR02		FR03		FR04		RS1		RS2	
Variants	PSNR	PHVSM	PSNR	PHVSM	PSNR	PHVSM	PSNR	PHVSM	PSNR	PHVSM	PSNR	PHVSM
	$\sigma^2 = 10$											
Input	38.14	45.66	38.15	45.65	38.12	45.99	38.13	44.89	38.13	42.33	38.13	41.96
2D	39.20	46.15	39.28	46.44	38.87	46.20	39.18	46.01	42.16	43.70	42.15	42.78
BM3D	39.69	46.64	39.72	46.76	39.26	46.45	39.59	46.40	42.68	44.41	42.69	43.68
3D _{C1}	42.06	48.93	42.06	49.34	42.22	49.73	42.23	48.69	45.51	48.13	45.59	47.99
3D _{C2}	44.13	51.79	44.24	51.85	43.98	52.12	44.27	51.37	45.95	49.10	45.87	48.53
	$\sigma^2 = 25$											
Input	34.15	40.28	34.16	40.27	34.16	40.43	34.14	39.72	34.13	37.52	34.15	37.25
2D	35.95	41.15	35.94	41.36	35.52	40.86	35.76	40.89	39.68	39.57	39.82	38.84
BM3D	36.60	41.65	36.59	42.00	36.06	41.44	36.27	41.47	40.17	40.31	40.26	39.60
3D _{C1}	39.19	44.31	39.14	44.69	39.19	44.89	39.31	44.29	42.79	43.93	42.85	43.66
3D _{C2}	40.53	46.83	40.61	46.82	40.29	46.79	40.60	46.54	43.01	44.44	43.00	43.84
	$\sigma^2 = 100$											
Input	28.15	32.55	28.14	32.49	28.12	32.51	28.14	32.17	28.13	30.71	28.12	30.54
2D	31.50	34.08	31.37	34.28	31.04	33.62	31.12	33.72	36.31	34.15	36.88	34.12
BM3D	32.35	34.78	32.21	35.08	31.68	34.25	31.71	34.29	36.78	34.81	37.28	34.73
3D _{C1}	34.99	38.10	34.78	38.20	34.67	38.26	34.74	38.03	38.98	38.10	39.05	37.54
3D _{C2}	35.55	39.47	35.52	39.61	35.15	39.23	35.33	39.20	39.01	38.17	39.11	37.50

The results for component-wise processing by BM3D filter are presented in Table 2 for comparison purposes. Note that BM3D is one of the best image filters that can be applied component-wise. One can see from the comparisons in Table 2 that BM3D slightly outperforms the 2D DCT-based filter, but the improvements due to employing denoising with a reference are far more significant.

An analysis of the data shows the following. The use of denoising with reference is always beneficial compared with 2D DCT-based filtering. The gain in PSNR is about 3 dB for AWGN variance $\sigma^2 = 10$ even if the reference image is transformed linearly. The use of nonlinear transformation of the reference image additionally provides 2 dB improvement. The benefits according to PHVSM are considerable too. While component-wise filtering improves this metric by only about 1 dB, filtering with linearly transformed reference provides about 3.5 dB improvement, and denoising with nonlinear transformation produces an additional improvement of about 2.5 dB. Thus, total improvement due to denoising with reference transformed nonlinearly reaches about 5 dB according to PSNR and about 5.5 dB according to PHVSM.

For noise variances $\sigma^2 = 25$ and $\sigma^2 = 100$, the situations and conclusions are similar. Although 2D DCT-based filtering improves quality of images according to both metrics, this improvement is not large for the test images FR01, FR02, FR03, and FR04, which contain fine details and textures. Effectiveness is better for the images RS1 and RS2. Meanwhile, denoising with references performs considerably better, although the benefits of nonlinear transformation of reference images are not essential, as for the case of $\sigma^2 = 10$.

This means that the method of denoising with reference performs well for different intensities of the noise (values of input PSNR) and different test images typical for remote sensing. The use of nonlinear transformation is preferable because performance is better. Note that determination

of parameters of transformations, either linear or nonlinear, does not take much time to compute, requiring one to solve a system of linear equations. This operation takes considerably less time than filtering itself, although DCT-based denoising is simple and fast as well.

The noisy test image FR04 (AWGN, $\sigma^2 = 100$) is presented in Figure 6. Noise is visible in homogeneous image regions. The output image for the 2D DCT-based filter is represented in Figure 7. Noise is suppressed, but edges and fine details are partly smeared. Improvement of visual quality is not obvious.



Figure 6. Noisy test image FR04 ($\sigma^2 = 100$, peak signal-to-noise ratio $(\text{PSNR})_{\text{inp}} = 28.11$ dB, $\text{PHVSM}_{\text{out}} = 32.17$ dB).



Figure 7. Output image for the 2D DCT-based filter ($\text{PSNR}_{\text{out}} = 31.11$ dB; $\text{PHVSM}_{\text{out}} = 33.70$ dB).

The results of image denoising using linearly and nonlinearly processed reference images are shown in Figures 8 and 9, respectively. The main difference compared with the image in Figure 7 is that edges and details are preserved better and, because of this, better visual quality is provided. For comparison purposes, we also give values of metrics for the original and denoised images.



Figure 8. Filtered noisy image by 3D-DCT ($\beta = 3.0$) with reference transformed by the first order polynomial ($\text{PSNR}_{\text{out}} = 34.73$ dB; $\text{PHVSM}_{\text{out}} = 38.05$ dB).



Figure 9. Filtered noisy image by 3D-DCT ($\beta = 3.0$) with reference transformed by second order polynomial ($\text{PSNR}_{\text{out}} = 35.34$ dB; $\text{PHVSM}_{\text{out}} = 39.24$ dB).

The denoising efficiency can be additionally improved if one uses two reference images and/or a BM3D filter instead of a DCT-based filter in the denoising with reference.

3.2. Application to Real Life Images

Let us see how good the filtering result is if denoising with reference is applied to a real-life image. The output for the 2D DCT-based filter has been already shown in Figure 2 and that image was partly smeared. The output of the proposed denoising technique with one nonlinearly transformed reference (second order polynomial was used) from channel #11 is shown in Figure 10a. The output for the case of using two nonlinearly transformed references numbers 11 and 12 (again, the second order polynomial was applied) is demonstrated in Figure 10b.

Both images are considerably “sharper” than the image in Figure 2 and more details are visible. Comparing the images in Figure 10 and the enlarged fragments in Figure 11, it is possible to state that the use of two reference images produces better visual quality of the processed image.

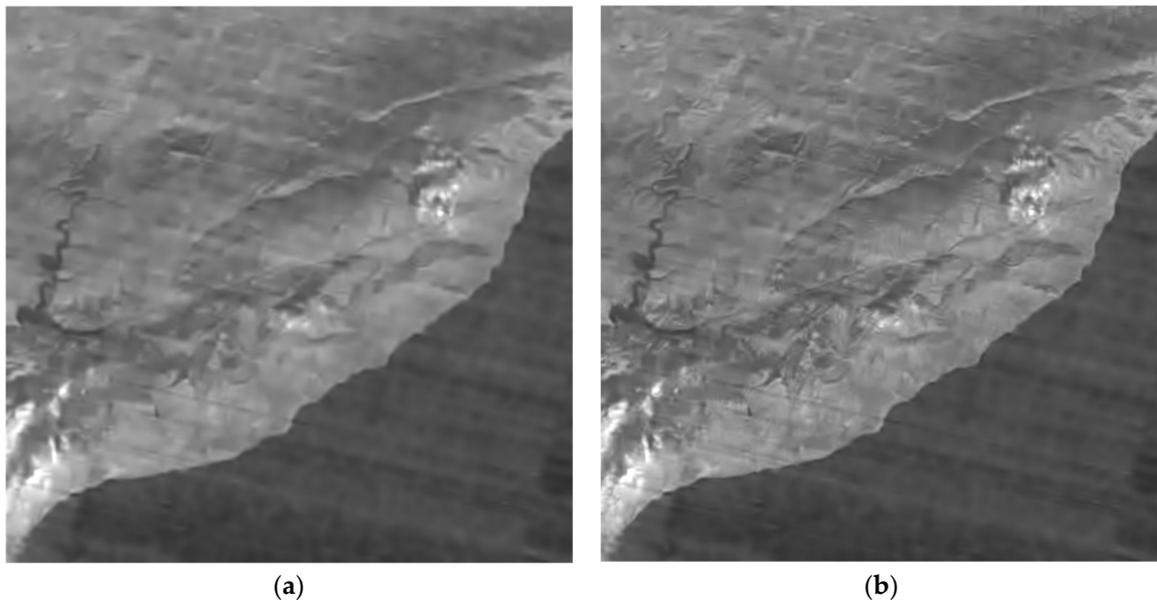


Figure 10. Output image for filtering with one (a) and two (b) nonlinearly transformed references (component image in channel #11 and component images in channels ##11 and 12).

The magnified difference image is shown in Figure 12. Comparing it to the image in Figure 1b, it is seen that noise (including strip-like artifacts) has been efficiently removed. The absence of visible regular structures in this image shows that almost no structural distortions were introduced into the output image by filtering.

In practice, one might be interested in how to decide what component image to choose among possible candidates. The strictly theoretical answer is that the component image that produces the smallest MSE (3) if the noisy image is not subject to VST, or the smallest MSE (6) or (7) if VST is applied, should be used. This means that all possible candidates have to be tried and the best one(s) has to be left for use in the proposed denoising method.



Figure 11. Enlarged fragment of output image for filtering with one (a) and two (b) nonlinearly transformed references (full images are presented in Figure 11).

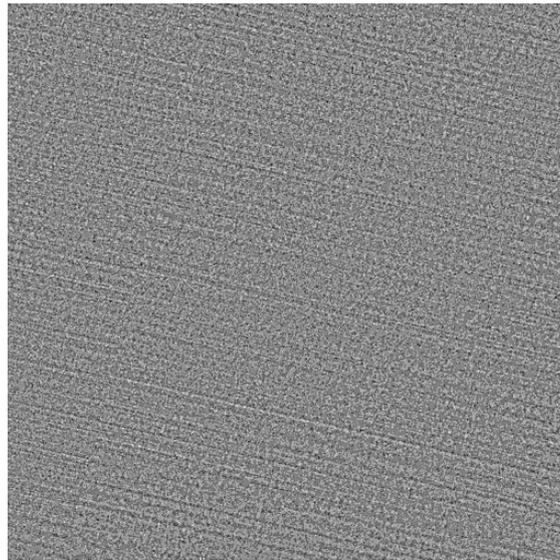


Figure 12. Magnified difference image for filtering with two nonlinearly transformed references (component images in channels ##11 and 12).

Meanwhile, in practice, this approach can be simplified. For example, if one knows that component images in a given channel are usually the most similar to the component images to be filtered, then it is possible to skip the choice of possible candidates and to set the fixed reference channel. For the considered case of multispectral Sentinel-2 data, the component images in the channel #10 are worth denoising. The component images in the channel #11 are worth using if one reference is applied. If two references are used, then one can employ the component images in channels ##11 and 12, as shown in the example above.

4. Conclusions

We considered the properties of component images acquired by Sentinel-2 multispectral sensor. It has been shown that there are component images in channel #10 for which denoising is expedient. We demonstrated that the use of the method of image denoising with reference can be a good solution in the sense of efficiency of noise suppression and simplicity of filtering. Both simulated and real-life data proving this are presented.

The method has several modifications where the use of nonlinear transformation of reference image(s) is preferable. Moreover, the use of two references instead of one provides additional benefits. The recommendations concerning selection of proper references for multispectral data are given.

Author Contributions: Conceptualization, V.L.; Data curation, M.U.; Formal analysis, B.V.; Investigation, S.A. and M.U.; Methodology, K.C.; Software, S.A.; Supervision, K.E.; Writing—original draft, V.L.; Writing—review & editing, K.E.

Funding: This research was partly funded by Ukrainian-French program “Dnipro”, grant number M/29-2018 (38015PA).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Schowengerdt, R.A. *Remote Sensing: Models and Methods for Image Processing*, 3rd ed.; Academic Press: San Diego, CA, USA, 2007.
2. Mielke, C.; Boshce, N.K.; Rogass, C.; Segl, K.; Gauert, C.; Kaufmann, H. Potential Applications of the Sentinel-2 Multispectral Sensor and the ENMAP hyperspectral Sensor in Mineral Exploration. *EARSEL Eproceedings* **2014**, *13*, 93–102.

3. Joshi, N.; Baumann, M.; Ehammer, A.; Fensholt, R.; Grogan, K.; Hostert, P.; Jepsen, M.R.; Kuemmerle, T.; Meyfroidt, P.; Mitchard, E.T.A.; et al. A Review of the Application of Optical and Radar Remote Sensing Data Fusion to Land Use Mapping and Monitoring. *Remote Sens.* **2016**, *8*, 70. [CrossRef]
4. Zhong, P.; Wang, R. Multiple-Spectral-Band CRFs for Denoising Junk Bands of Hyperspectral Imagery. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 2269–2275. [CrossRef]
5. First Applications from Sentinel-2A. Available online: http://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Sentinel-2/First_applications_from_Sentinel-2A (accessed on 10 January 2019).
6. Uss, M.; Vozel, B.; Lukin, V.; Chehdi, K. Local signal-dependent noise variance estimation from hyperspectral textural images. *IEEE J. Sel. Top. Signal Process.* **2011**, *5*, 469–486. [CrossRef]
7. Fan, Y.R.; Huang, T.Z.; Zhao, X.L.; Deng, L.J.; Fan, X. Multispectral Image Denoising Via Nonlocal Multitask Sparse Learning. *Remote Sens.* **2018**, *10*, 116. [CrossRef]
8. Lukin, V.; Abramov, S.; Krivenko, S.; Kurekin, A.; Pogrebnyak, O. Analysis of classification accuracy for pre-filtered multichannel remote sensing data. *Expert Syst. Appl.* **2013**, *40*, 6400–6411. [CrossRef]
9. Ponomaryov, V.; Rosales, A.; Gallegos, F.; Loboda, I. Adaptive vector directional filters to process multichannel images. *IEICE Trans. Commun.* **2007**, *90*, 429–430. [CrossRef]
10. Lukac, R.; Plataniotis, K.N. *Color Image Processing: Methods and Applications*; CRC Press: Boca Raton, FL, USA, 2006; 600p.
11. Astola, J.; Haavisto, P.; Neuvo, Y. Vector median filters. *Proc. IEEE* **1990**, *78*, 678–689. [CrossRef]
12. Manolakis, D.; Lockwood, R.; Cooley, T. On the Spectral Correlation Structure of Hyperspectral Imaging Data. In Proceedings of the IGARSS 2008—2008 IEEE International Geoscience and Remote Sensing Symposium, Boston, MA, USA, 7–11 July 2008; pp. II-581–II-584.
13. Mizutani, J.; Ogawa, S.; Shinoda, K.; Hasegawa, M.; Kato, S. Multispectral demosaicking algorithm based on inter-channel correlation. In Proceedings of the 2014 IEEE Visual Communications and Image Processing Conference, Valletta, Malta, 7–10 December 2014; pp. 474–477.
14. Deledalle, C.; Denis, L.; Tabti, S.; Tupin, F. MuLoG, or How to Apply Gaussian Denoisers to Multi-Channel SAR Speckle Reduction? *IEEE Trans. Image Process.* **2017**, *26*, 4389–4403. [CrossRef]
15. Ponomarenko, N.N.; Zelensky, A.A.; Koivisto, P.T. 3D DCT Based Filtering of Color and Multichannel Images. *Telecommun. Radio Eng.* **2008**, *67*, 1369–1392. [CrossRef]
16. Lukac, R.; Smolka, B.; Plataniotis, K.N.; Venetsanopoulos, A.N. Vector sigma filters for noise detection and removal in color images. *J. Vis. Commun. Image Represent.* **2006**, *17*, 1–26. [CrossRef]
17. Pizurica, A.; Philips, W. Estimating the probability of the presence of a signal of interest in multiresolution single- and multiband image denoising. *IEEE Trans. Image Process.* **2006**, *15*, 654–665. [CrossRef] [PubMed]
18. Kurekin, A.A.; Lukin, V.V.; Zelensky, A.A.; Koivisto, P.; Astola, J.; Saarinen, K.P. Comparison of component and vector filter performance with application to multichannel and color image processing. In Proceedings of the IEEE-EURASIP Workshop on NSIP, Antalya, Turkey, 20–23 June 1999; pp. 38–42.
19. Fevrale, D.V.; Ponomarenko, N.N.; Lukin, V.V.; Abramov, S.K.; Egiazarian, K.O.; Astola, J.T. Efficiency analysis of color image filtering. *EURASIP J. Adv. Signal. Process.* **2011**, *2011*, 41. [CrossRef]
20. Kozhemiakin, R.A.; Rubel, O.; Abramov, S.K.; Lukin, V.V.; Vozel, B.; Chehdi, K. Efficiency analysis for 3D filtering of multichannel images. In Proceedings of the SPIE Remote Sensing, Edinburgh, UK, 26–29 September 2016; p. 11.
21. Shah, C.A.; Watanachaturaporn, P.; Varshney, P.K.; Arora, M.K. Some recent results on hyperspectral image classification. In Proceedings of the IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data, Greenbelt, MD, USA, 27–28 October 2003; pp. 346–353.
22. Liu, P.; Huang, F.; Li, G.; Liu, Z. Remote-Sensing Image Denoising Using Partial Differential Equations and Auxiliary Images as Priors. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9*, 358–362. [CrossRef]
23. Zelinski, A.; Goyal, V. Denoising Hyperspectral Imagery and Recovering Junk Bands using Wavelets and Sparse Approximation. In Proceedings of the 2006 IEEE International Symposium on Geoscience and Remote Sensing, Denver, CO, USA, 31 July–4 August 2006; pp. 387–390.
24. Lukin, V.V.; Abramov, S.K.; Abramova, V.V.; Astola, J.T.; Egiazarian, K.O. DCT-based denoising in multichannel imaging with reference. *Telecommun. Radio Eng.* **2016**, *75*, 1167–1191. [CrossRef]
25. Yuan, Q.; Zhang, L.; Shen, H. Hyperspectral Image Denoising With a Spatial–Spectral View Fusion Strategy. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 2314–2325. [CrossRef]

26. Liu, X.; Bourennane, S.; Fossati, C. Denoising of Hyperspectral Images Using the PARAFAC Model and Statistical Performance Analysis. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 3717–3724. [[CrossRef](#)]
27. Wang, Y.; Niu, R.; Yu, X. Anisotropic Diffusion for Hyperspectral Imagery Enhancement. *IEEE Sens. J.* **2010**, *10*, 469–477. [[CrossRef](#)]
28. Chen, G.; Qian, S. Denoising of Hyperspectral Imagery Using Principal Component Analysis and Wavelet Shrinkage. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 973–980. [[CrossRef](#)]
29. Dabov, K.; Foi, A.; Katkovnik, V.; Egiazarian, K. Color Image Denoising via Sparse 3D Collaborative Filtering with Grouping Constraint in Luminance-Chrominance Space. In Proceedings of the 2007 IEEE International Conference on Image Processing, San Antonio, TX, USA, 16 September–19 October 2007; pp. I-313–I-316.
30. Lukin, V.V.; Abramov, S.K.; Abramova, V.V.; Astola, J.T.; Egiazarian, K.O. Denoising of Multichannel Images with References. *Telecommun. Radio Eng.* **2017**, *76*, 1719–1748. [[CrossRef](#)]
31. Abramov, S.K.; Abramova, V.V.; Lukin, V.V.; Egiazarian, K.O. Denoising of multichannel images with nonlinear transformation of reference image. *Telecommun. Radio Eng.* **2018**, *77*, 769–786. [[CrossRef](#)]
32. Dabov, K.; Foi, A.; Katkovnik, V.; Egiazarian, K. Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. *IEEE Trans. Image Process.* **2007**, *16*, 2080–2095. [[CrossRef](#)] [[PubMed](#)]
33. Uss, M.; Vozel, B.; Lukin, V.; Chehdi, K. Analysis of signal-dependent sensor noise on JPEG 2000-compressed Sentinel-2 multi-spectral images. In Proceedings of the SPIE Conference on Signal and Image Processing for Remote Sensing, Warsaw, Poland, September 2017; p. 12.
34. Rubel, O.; Lukin, V.; Abramov, S.; Vozel, B.; Pogrebnyak, O.; Egiazarian, K. Is Texture Denoising Efficiency Predictable? *Int. J. Pattern Recognit. Artif. Intell.* **2018**, *32*, 1860005. [[CrossRef](#)]
35. Rubel, O.; Kozhemiakin, R.A.; Abramov, S.K.; Lukin, V.V.; Vozel, B.; Chehdi, K. Performance prediction for 3D filtering of multichannel images. In Proceedings of the Image and Signal Processing for Remote Sensing XXI, Toulouse, France, 21 September 2015.
36. Meola, J.; Eismann, M.T.; Moses, R.L.; Ash, J.N. Modeling and estimation of signal-dependent noise in hyperspectral imagery. *Appl. Opt.* **2011**, *50*, 3829–3846. [[CrossRef](#)] [[PubMed](#)]
37. Ponomarenko, N.; Lukin, V.; Astola, J.; Egiazarian, K. Analysis of HVS-Metrics' Properties Using Color Image Database TID2013. In Proceedings of the ACIVS, Catania, Italy, 26–29 October 2015; pp. 613–624.
38. Pogrebnyak, O.B.; Lukin, V.V. Wiener discrete cosine transform-based image filtering. *J. Electron. Imaging* **2012**, *21*, 16. [[CrossRef](#)]
39. Solbo, S.; Eltoft, T. Homomorphic wavelet-based statistical despeckling of SAR images. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 711–721. [[CrossRef](#)]
40. Ponomarenko, N.; Silvestri, F.; Egiazarian, K.; Carli, M.; Astola, J.; Lukin, V. On between-coefficient contrast masking of DCT basis functions. In Proceedings of the Third International Workshop on Video Processing and Quality Metrics for Consumer Electronics, VPQM 2007, Scottsdale, AZ, USA, 25–26 January 2007.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).