Peer-reviewed version available at Remote Sens. 2019, 11, 611; doi:10.3390/rs11060611

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] National Aerospace University, Kharkov, Ukraine
[2] University of Rennes 1, Lannion, France
[3] Tampere University of Technology, Tampere, Finland
* Correspondence: lukin@ai.kharkov.com

Abstract: Multispectral remote sensing data may contain component images which are heavily corrupted by noise and 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 which 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. Examples of denoising of 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, airborne and spaceborne, etc. Modern remote sensing tends to improve spatial resolution of sensors and to make them multichannel, e.g. multi-polarization radar, hyperspectral, 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 to single-channel data. However, there exist the following problem in multichannel sensing - images in one or 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 it 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, etc. [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], it is worth exploiting inter-channel correlation of component images inherent for practically all types of multichannel images [11-14]. A question is how large is this 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 [9-11] and references therein). Originally, they have been oriented on removal of impulsive and mixed noises. Later, denoising methods based on orthogonal transforms have appeared with the

main application to color [15-17] images. Some of them have been modified to work with multichannel RS images [15, 18]. Necessity of such modifications appears since 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 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], etc. This makes filtering more complex and necessity 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, amount of jointly processed component images has to be either somehow optimized or chosen in a reasonable way.

A new group of methods of multichannel data filtering has appeared recently that can be treated as hybrid. They exploit inter-channel correlation in different ways. The main idea is that in multichannel RS data there can be the so-called "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 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 both in spatial and spectral views. The problem is that possible signal-dependent nature of the noise has not been taken into account. A method based on parallel factor analysis (PARAFAC) approach to denoising has been proposed in [26], but it also assumes an additive noise model. Anisotropic diffusion is applied for hyperspectral imagery enhancement in [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.

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, 23, 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 denoising of junk components in multispectral data, e.g., 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 resolution. There are three component images (##1, 9, and 10) that have resolution 60x60 m^2; six component images (## 5, 6, 7, 8A, 11, and 12) have resolution 20x20 m^2 whilst the remained four (##2, 3, 4, and 8) possess the best resolution of 10x10 m^2. This feature differs Sentinel-2 multispectral data from hyperspectral data partly discussed above that 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 has 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, 34] 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 saving time and resources for data pre-processing).

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 the following:

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

where I_{ij}^{t} denotes the true image value in an ij-th pixel, $n_{ij}(I_{ij}^{t})$ is the noise statistical properties of which are dependent on I_{ij}^{t}, 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 since spatial resolution and, respectively, number of pixels in each component image 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 that 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 signal-dependent (SD) component has been tested in [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 [33] (where specific effects appear due to a lossy compression) but we will further accept the general model (2).

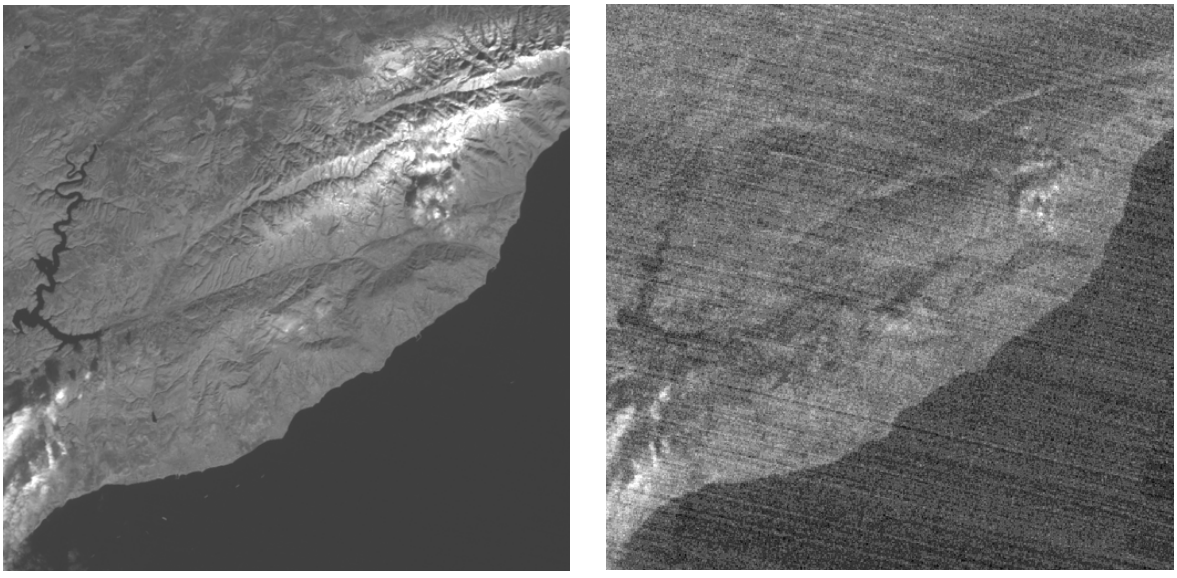
For this model, one can use the so-called equivalent noise variance that is equal to $\sigma_{eq}^{2} = \sum_{i=1}^{I_{Im}}\sum_{j=1}^{J_{Im}}\left(I_{ij}^{n} - I_{ij}^{t}\right)^{2} / \left(I_{Im} \times J_{Im}\right)$ if the true image is available. Alternatively, it can be estimated as $\sigma_{eq}^{2} = \sigma_{0}^{2} + k\sum_{i=1}^{I_{Im}}\sum_{j=1}^{J_{Im}} I_{ij}^{t} / \left(I_{Im} \times J_{Im}\right) \approx \hat{\sigma}_{0}^{2} + \hat{k}I_{mean}$ if the true image is not available, but quite accurate estimates $\hat{\sigma}_{0}^{2}$ and $\hat{k}$ have been 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.

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 one granule (set of multispectral data) are presented in Table 1. As one can see, practically in all component images, the equivalent variance is considerably larger than σ_{0}^{2}, although contribution of SD component is always smaller than that of SI component. The only exception is the component image in the channel #10, where σ_{0}^{2} is practically the same as the corresponding equivalent variance. This means that signal dependent nature of the noise has to be taken into account.

Note that equivalent variance of the noise is the smallest in component image #10. Then, one might think that this image is the least noisy. However, this conclusion is not correct, since we have not taken into account the range of image representation. Let us also analyze peak signal-to noise ratio. To avoid possible presence of hot pixels in data and bright points, consider below the so-called robust estimate of input PSNR determined as $PSNR_{inp}^{rob} = 10\log_{10}\left(D_{rob}^2/\sigma_{eq}^2\right)$, where $D_{rob} = I^{(p)} - I^{(r)}$, $p = 0.99I_{Im}J_{Im}$ and $r = 0.01I_{Im}J_{Im}$, $I^{(p)}$ and $I^{(r)}$ are the p-th and r-th order statistics of image values. The obtained values of $PSNR_{inp}^{rob}$ 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 1,a. Meanwhile, there is also an image in sub-band 10 for which $PSNR_{inp}^{rob}$ is only 11.6 dB and, thus, noise is visible (one example is given in Figure 1,b. As it is seen, noise is not white since specific diagonal structures are observed. One more observation is that these component images are similar to each other and cross-correlation factor $R_{\#10}$ for them is equal to 0.57.

Table 1. Noise parameters in component images of Sentinel-2.

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
$PSNR_{inp}^{rob}$	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



(**a**) (**b**)

Figure 1. Visualized fragments of Sentinel-2 images of size 512x512 pixels in the band #09 (a) and band #10 (b).

Cross-correlation factors for the component image in channel #10 and other component images are given in Table 1 (the last row). One can see that the correlation is low for component images (##1…4) that relate to visible range but it increases and exceeds 0.77 for the components number 11

Peer-reviewed version available at Remote Sens. 2019, 11, 611; doi:10.3390/rs11060611

and 12. If resolutions in the channel 10 and another channel image are different, the corresponding downsampling is applied before calculation of correlation factor.

We have processed the image in Figure 1,b 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.

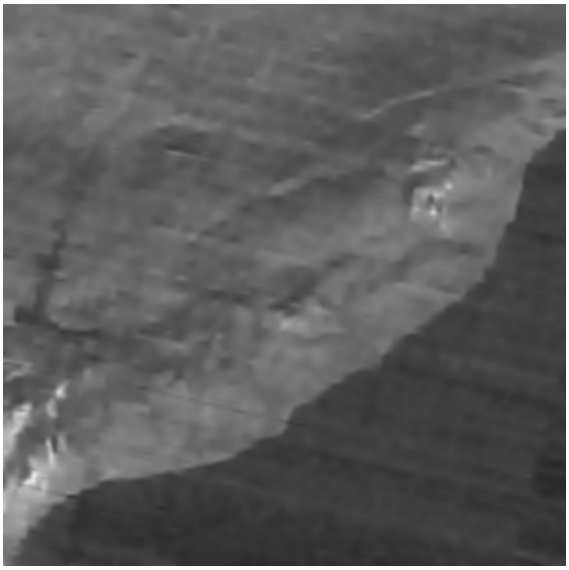


Figure 2. The output of DCT-based filter.

We have 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 $\mathrm{PSNR}_{\mathrm{inp}}^{\mathrm{rob}}$ and inter-channel correlation.

For image denoising with a reference, it is assumed that a reference image or a set of reference images $\mathrm{I}_{\mathrm{ijs}}^{\mathrm{ref}}, \mathrm{i}=1,...,\mathrm{I}_{\mathrm{Im}}, \mathrm{j}=1,...,\mathrm{J}_{\mathrm{Im}}, \mathrm{s}=1,...,\mathrm{S}$ 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 by 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.

Another assumption is that potential reference images are in some sense similar to $\mathrm{I}_{\mathrm{ij}}^{\mathrm{n}}, \mathrm{i}=1,...,\mathrm{I}_{\mathrm{Im}}, \mathrm{j}=1,...,\mathrm{J}_{\mathrm{Im}}$. It is known that similarity of images can be measured differently: mean square errors (MSE) between images, cross-correlation factor, etc. 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 discussion when and why each of them takes place.

The first practical case is that noise in $\mathrm{I}_{\mathrm{ij}}^{\mathrm{n}}, \mathrm{i}=1,...,\mathrm{I}_{\mathrm{Im}}, \mathrm{j}=1,...,\mathrm{J}_{\mathrm{Im}}$ is additive and then the main metric that describes similarity is

$$\mathrm{MSE}_{\mathrm{nrmod}} = \sum_{\mathrm{i}=1}^{\mathrm{I}_{\mathrm{Im}}} \sum_{\mathrm{j}=1}^{\mathrm{J}_{\mathrm{Im}}} \left(\mathrm{I}_{\mathrm{ij}}^{\mathrm{n}} - \mathrm{I}_{\mathrm{ij}}^{\mathrm{ref\,mod}} \right)^2 / \left(\mathrm{I}_{\mathrm{Im}} \times \mathrm{J}_{\mathrm{Im}} \right). \quad (3)$$

Here $\mathrm{I}_{\mathrm{ij}}^{\mathrm{ref\,mod}}, \mathrm{i}=1,...,\mathrm{I}_{\mathrm{Im}}, \mathrm{j}=1,...,\mathrm{J}_{\mathrm{Im}}$ defines the modified reference image which can be either linearly transformed as

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

188 or nonlinearly transformed as

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

189 where S_0, Δ_0 denote the parameters of linear least MSE regression (4) (case 1) and $\Psi(I_{ij}^{ref})$ defines
190 nonlinear transformation (case 2) that lead to minimizing MSE_{nrmod}.

191 Two other cases relate to the noise model described by (1) and (2). Then, if the noise is
192 signal-dependent, it is usually recommended to apply a proper homomorphic or variance stabilizing
193 transform (VST) to deal with an additive noise (although often non-Gaussian) in filtering [39, 40].
194 Advantage of this approach is that additive nature of the noise in an image to be denoised allows
195 applying a wider set of efficient filters [34]. As VST, the generalized Anscombe transform [39] or
196 logarithmic transform [40] can be used, depending on a type of signal-dependent noise one deals in
197 each particular case.

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

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

200 or

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

201 respectively.

202 Let us now recall how denoising with a reference is carried out for the simplest case of having
203 $I_{ij}^{n}, i = 1,...,I_{Im}, j = 1,...,J_{Im}$ and a properly chosen $I_{ij}^{ref\,mod}, i = 1,...,I_{Im}, j = 1,...,J_{Im}$. Then, the noisy and
204 the reference images are denoised jointly. A 2-point DCT is applied first in "vertical direction",
205 getting "sum" and "difference" images. The obtained images are filtered by the 2D DCT-based filter
206 or by BM3D with properly selected hard thresholds. After this, inverse 2-point DCT is applied and
207 the obtained first component is considered as the filtered image.

208 If VST is used, then the same operations are applied to $I_{ij}^{nVST}, i = 1,...,I_{Im}, j = 1,...,J_{Im}$ and
209 $I_{ij}^{ref\,mod}, i = 1,...,I_{Im}, j = 1,...,J_{Im}$ obtained by minimizing (6) or (7). The only difference is that the
210 denoised image has to be subject to inverse VST.

211 Also note that it is possible to have two or more modified reference images after 3- or more
212 point- DCT applied in vertical direction to decorrelate data. We have considered two reference
213 images instead of one in [31] and filtering has occurred to be more efficient in terms of standard
214 metrics, such as output PSNR and visual quality metrics, such as PSNR-HVS-M [41]. Besides, after 2-
215 or 3-point DCT, it is possible to apply component-wise different filters including standard DCT,
216 BM3D or others. Usually, if a given filter is more efficient in component-wise (single-channel)
217 denoising, its use is also beneficial in the considered denoising with a reference [30]. It is also worth
218 stressing that optimal (recommended) parameters of thresholds applied in DCT coefficient
219 thresholding have been determined in [24, 30-31]. These thresholds differ from those usually
220 recommended for the cases these filters are employed for noise removal in single channel images.
221 Thus, in our further studies, we will use just optimal thresholds.

222 *2.2. Performance Criteria*

223 We start analyzing performance of methods of image filtering with reference(s) for simulated
224 data [24, 30-31]. In our simulations, four test images typical for remote sensing, presented in Figure 3

Peer-reviewed version available at Remote Sens. 2019, 11, 611; doi:10.3390/rs11060611

and denoted as FR01, Fr02, FR03, FR04 and two high quality component images denoted RS1 and RS2 of AVIRIS hypercube of data have been used. Noise 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 test images to be similar to the test ones according to certain similarity measure (i.e., to have high cross-correlation factor). Same time, they also have to be different in several senses – with a different dynamic range, contain some additional content not present in the image to be denoised (see the example in Figure 1). Because of this, we have simulated the reference image as $I_{ij}^{ref} = 32\sqrt{I_{ij}^{t}} + 0.5I_{ij}^{t180}, i = 1,...,I_{Im}, j = 1,...,J_{Im}$ where $I_{ij}^{t180}, i = 1,...,I_{Im}, j = 1,...,J_{Im}$ denotes the same noise free test image rotated by 180º. This allows providing correlation factor of the same level as for real-life multispectral data in channels 10…12.

The obtained reference images for the test images FR01 and FR02 in Figure 3 are visualized in Figure 4 (note that the reference images are in the dynamic range considerably different from that of original range 0…255).

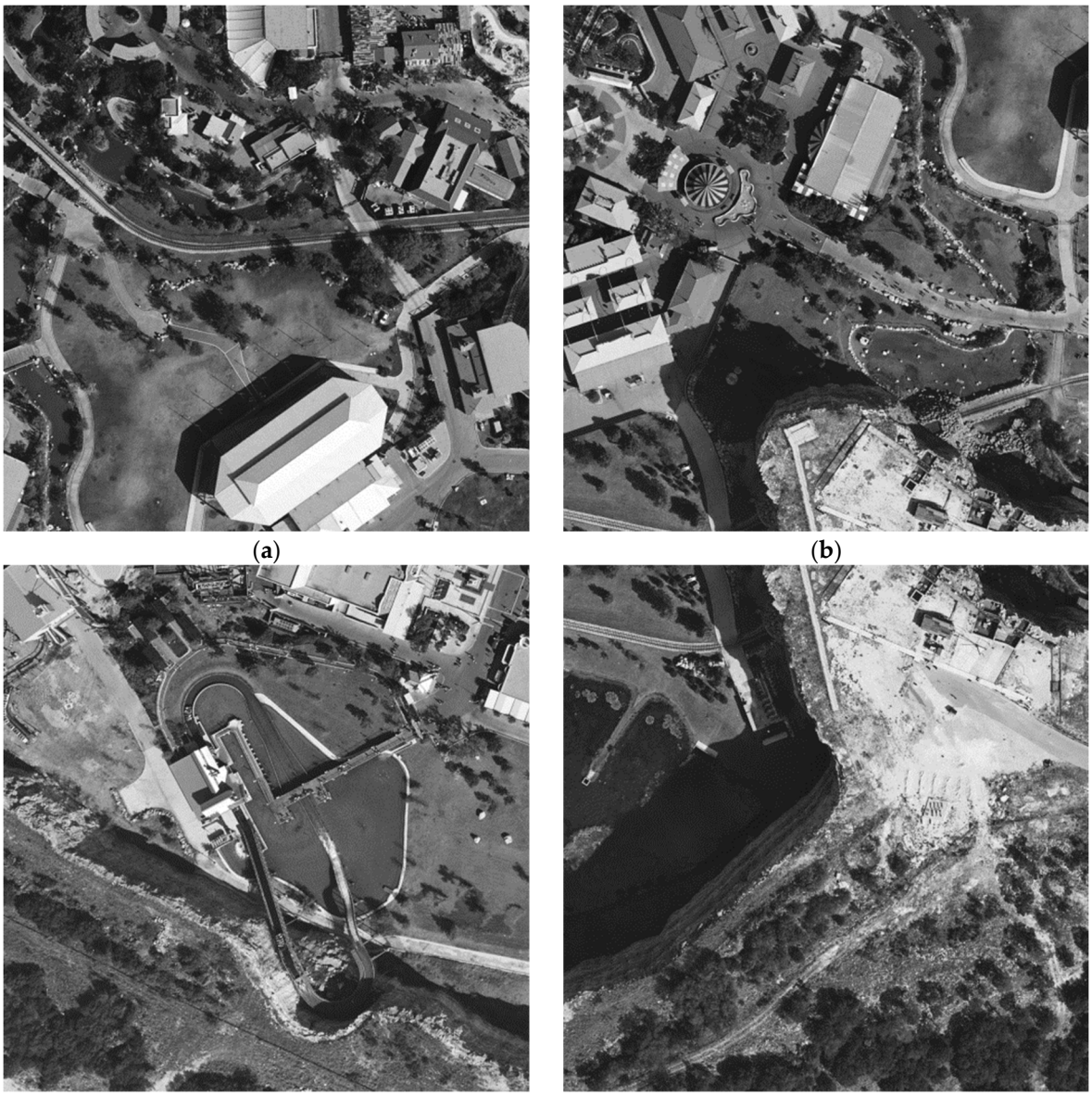


Figure 3. RS test images FR01 (a), FR02 (b), FR03 (c), and FR04 (d).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 January 2019 doi:10.20944/preprints201901.0142.v1

Peer-reviewed version available at Remote Sens. 2019, 11, 611; doi:10.3390/rs11060611

(a) (b)

Figure 4. Visualized reference images for the test images FR01 (a) and FR02 (b).

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

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

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

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

where $\mathrm{MSE}_{\mathrm{out}}$ is the output mean square error (MSE). Effectiveness is then characterized by

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

Alongside with PSNR, we would like to analyze visual quality of original (noisy) and filtered images. To do this, we propose to use the metric PSNR-HVS-M that takes into account two important properties of human vision system (HVS), namely, less sensitivity to distortions in high frequency components and masking effect of image texture and other heterogeneities [41]. Then, one has

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

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

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

where $\mathrm{MSE}_{\mathrm{inp}}^{\mathrm{HVS}}$ and $\mathrm{MSE}_{\mathrm{out}}^{\mathrm{HVS}}$ are input and output MSEs calculated with taking into account 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 have been employed here). Without loss of generality, the results have been obtained for additive white Gaussian noise. Three values of noise variance σ^2 have been analyzed: 10 (invisible noise), 25 (noise is visible in homogeneous image regions), and 100 (intensive noise).

The results for component-wise processing by BM3D filter are presented in Table 2 for comparison purposes. It is seen that this state-of-the-art filter slightly outperforms the 2D DCT-based filter but the results for the proposed denoising with reference are considerably better.

Table 2. Simulation data.

Image	FR01		FR02		FR03		FR04		RS1		RS2	
Variants	PSNR	PHVSM	PSNR	PHVSM	PSNR	PHVSM	PSNR	PHVSM	PSNR	PHVSM	PSNR	PHVSM
						σ^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
						σ^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
						σ^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

Analysis of data shows the following. The use of denoising with reference is always beneficial compared to 2D DCT-based filtering. The gain in PSNR is about 3 dB for AWGN variance σ^2=10 even if the reference image is transformed linearly. The use of nonlinear transformation of the reference image additionally provides 2 dB improvement. Benefits according to PHVSM are considerable too. Whilst component-wise filtering produces improvement of this metric by only about 1 dB, filtering with linearly transformed reference provides about 3.5 dB and denoising with nonlinear transformation produces additional improvement by 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 σ^2=25 and σ^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

Peer-reviewed version available at Remote Sens. 2019, 11, 611; doi:10.3390/rs11060611

considerably better although the benefits of nonlinear transformation of reference images are not essential as for the case σ^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 since performance is better. Note that determination of parameters of transformations, either linear or nonlinear, does not take much time to compute, requiring to solve a system of linear equations. This operation takes considerably smaller time than filtering itself, although DCT-based denoising is simple and fast as well.

The noisy test image FR04 (AWGN, σ^2=100) is presented in Figure 5. Noise is visible in homogeneous image regions. The output image for 2D DCT-based filter is represented in Figure 6. Noise is suppressed but edges and fine details are partly smeared. Improvement of visual quality is not obvious.

Figure 5. Noisy test image FR04 (σ^2=100, $PSNR_{inp}$ = 28.11 dB, $PHVSM_{out}$ = 32.17 dB).

Figure 6. Output image for 2D DCT-based filter ($PSNR_{out}$ = 31.11 dB; $PHVSM_{out}$ = 33.70 dB).

The results of image denoising using linearly and nonlinearly processed reference image are shown in Figures 7 and 8, respectively. The main difference compared to the image in Figure 6 is that

edges and details are preserved better and, due to this, better visual quality is provided. For comparison purpose, we also give values of metrics for the original and denoised images.

Figure 7. Filtered noisy image by 3D-DCT ($\beta = 3.0$) with reference transformed by the first order polynomial (PSNRout = 34.73 dB; PHVSMout = 38.05 dB).

Figure 8. Filtered noisy image by 3D-DCT ($\beta = 3.0$) with reference transformed by second order polynomial (PSNRout = 35.34 dB; PHVSMout = 39.24 dB).

The denoising efficiency can be additionally improved if one uses two reference images and/or BM3D filter instead of 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 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 has been used) from channel #11 is shown in Figure 9. The output for the case of using two nonlinearly transformed references numbers 11 and 12 (again the second order polynomial has been applied) is demonstrated in Figure 10.

Both images are considerably "sharper" than the image in Figure 2 and more details are visible. Comparing the images in Figures 9 and 10, it is possible to state that the use of two reference images produces better visual quality of the processed image.

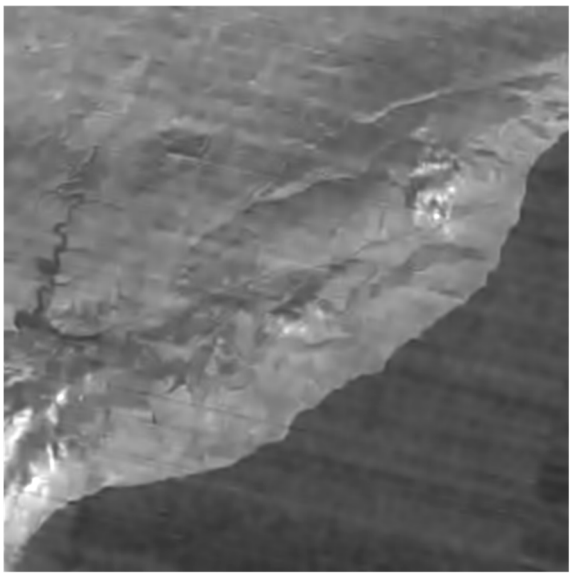


Figure 9. Output image for filtering with one nonlinearly transformed reference (component image in channel #11).

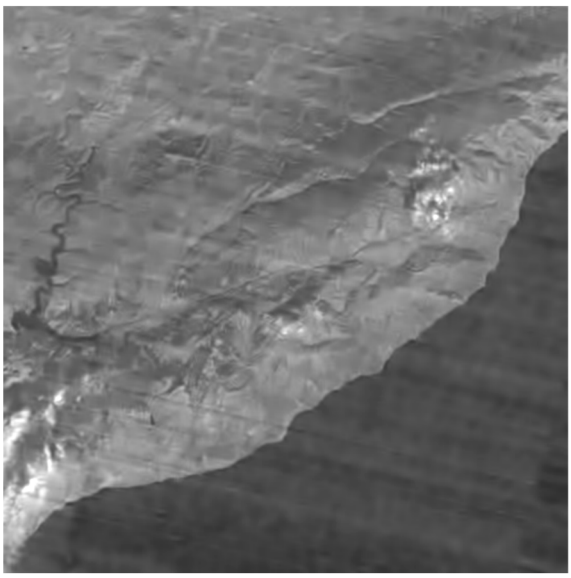


Figure 10. Output image for filtering with two nonlinearly transformed references (component images in channels ##11 and 12).

The magnified difference image is shown in Figure 11. Comparing it to the image in Figure 1,b, 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 have been introduced into the output image by filtering.

In practice, one might be interested 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) is 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.

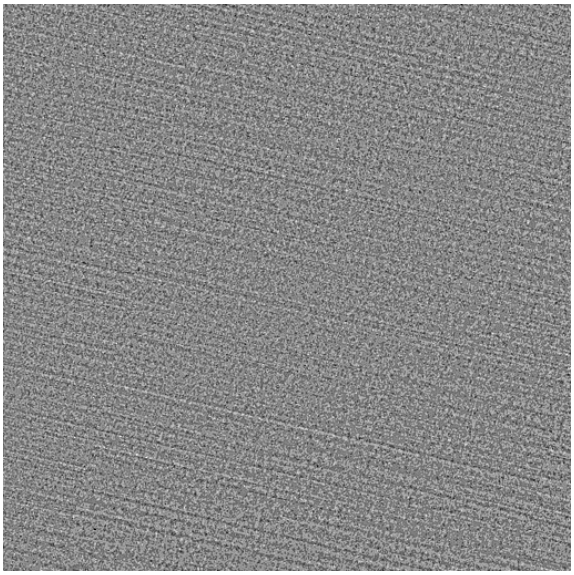


Figure 11. Magnified difference image for filtering with two nonlinearly transformed references (component images in channels ##11 and 12).

Meanwhile, in practice, the 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 have 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 have 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: Sergey Abramov was responsible for software and simulations, Mikhail Uss carried out analysis of component image properties, Vladimir Lukin put forward the idea of denoising the junk channel image and prepared the paper draft, Benoit Vozel was responsible for simulation result analysis, Kacem Chehdi proposed methodology of investigation, Karen Egiazarian carried out editing and supervision.

Funding: The work has been partly supported by the Project M/29-2018 (38015PA) of Ukrainian-French program “Dnipro”.

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; 515 p.
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, Proceedings of the EARSeL eProceedings, 13, 2/2014, pp. 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.; Reiche, J.; Ryan, C.M.; Waske, B. A Review of the Application of Optical and Radar Remote Sensing Data Fusion to Land Use Mapping and Monitoring. *Remote Sens.* **2016**, *8*, 70, 23 p.
4. Zhong, P.; Wang, R. Multiple-Spectral-Band CRFs for Denoising Junk Bands of Hyperspectral Imagery. *IEEE Transactions on Geoscience and Remote Sensing* **2013**, 51, 4, pp. 2269-2275.
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 Jan 2019).
6. Uss, M.; Vozel, B.; Lukin, V.; Chehdi, K. Local signal-dependent noise variance estimation from hyperspectral textural images, *IEEE Journal of Selected Topics in Signal Processing* **2011**, 5, 3, pp. 469-486.
7. Fan, Y.R.; Huang, T.Z.; Zhao, X.L.; Deng, L.J.; Fan, X. Multispectral Image Denoising Via Nonlocal Multitask Sparse Learning, *Remote Sensing*, **2018**, 10, 116, 13 p.
8. Lukin, V.; Abramov, S.; Krivenko, S.; Kurekin, A.; Pogrebnyak, O., Analysis of classification accuracy for pre-filtered multichannel remote sensing data. *Expert Systems with Applications* **2013,** 40 (16), pp. 6400-6411.
9. Ponomaryov, V.; Rosales, A.; Gallegos, F.; Loboda, I. Adaptive vector directional filters to process multichannel images. *IEICE Transactions on Communications* **2007**, 90 (2), pp. 429-430.
10. Lukac, R.; Plataniotis, K.N. *Color image processing; methods and applications*. Boca Raton: CRC Press, 2006, 600 p.
11. Astola, J.; Haavisto, P.; Neuvo, Y. Vector median filters. *Proceedings of the IEEE* **1990,** 78 (4), pp. 678-689.
12. Manolakis, D.; Lockwood, R.; Cooley, T. On The Spectral Correlation Structure of Hyperspectral Imaging Data. In Proceedings of IGARSS 2008 - 2008 IEEE International Geoscience and Remote Sensing Symposium, 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 2014 IEEE Visual Communications and Image Processing Conference, 7-10 Dec. 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 Transactions on Image Processing* **2017**, *26*, 4389-4403, doi:10.1109/TIP.2017.2713946.
15. Ponomarenko, N.N.; Zelensky, A.A.; Koivisto, P.T. 3D DCT Based Filtering of Color and Multichannel Images. **2008**, *67*, 1369-1392, doi:10.1615/TelecomRadEng.v67.i15.50.
16. Lukac, R.; Smolka, B.; Plataniotis, K.N.; Venetsanopoulos, A.N. Vector sigma filters for noise detection and removal in color images. *Journal of Visual Communication and Image Representation* **2006**, *17*, 1-26, doi:https://doi.org/10.1016/j.jvcir.2005.08.007.
17. Pizurica, A.; Philips, W. Estimating the probability of the presence of a signal of interest in multiresolution single- and multiband image denoising. *IEEE Transactions on Image Processing* **2006**, *15*, 654-665, doi:10.1109/TIP.2005.863698.
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, 1999, 1, pp. 38-42.
19. Fevralev, D.V.; Ponomarenko, N.N.; Lukin, V.V.; Abramov, S.K.; Egiazarian, K.O.; Astola, J.T. Efficiency analysis of color image filtering. *EURASIP Journal on Advances in Signal Processing* **2011**, *2011*, 41, doi:10.1186/1687-6180-2011-41.
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 SPIE Remote Sensing;10004, p. 11.
21. Shah, C.A.; Watanachaturaporn, P.; Varshney, P.K.; Arora, M.K. Some recent results on hyperspectral image classification. In Proceedings of IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data, 2003, 27-28 Oct. 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 Geoscience and Remote Sensing Letters* **2012**, *9*, 358-362, doi:10.1109/LGRS.2011.2168598.

23. Zelinski, A.; Goyal, V. Denoising Hyperspectral Imagery and Recovering Junk Bands using Wavelets and Sparse Approximation. In Proceedings of 2006 IEEE International Symposium on Geoscience and Remote Sensing, 31 July-4 Aug. 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. **2016**, *75*, 1167-1191, doi:10.1615/TelecomRadEng.v75.i13.30.
25. Yuan, Q.; Zhang, L.; Shen, H. Hyperspectral Image Denoising With a Spatial–Spectral View Fusion Strategy. *IEEE Transactions on Geoscience and Remote Sensing* **2014**, *52*, 2314-2325, doi:10.1109/TGRS.2013.2259245.
26. Liu, X.; Bourennane, S.; Fossati, C. Denoising of Hyperspectral Images Using the PARAFAC Model and Statistical Performance Analysis. *IEEE Transactions on Geoscience and Remote Sensing* **2012**, *50*, 3717-3724, doi:10.1109/TGRS.2012.2187063.
27. Wang, Y.; Niu, R.; Yu, X. Anisotropic Diffusion for Hyperspectral Imagery Enhancement. *IEEE Sensors Journal* **2010**, *10*, 469-477, doi:10.1109/JSEN.2009.2037800.
28. Chen, G.; Qian, S. Denoising of Hyperspectral Imagery Using Principal Component Analysis and Wavelet Shrinkage. *IEEE Transactions on Geoscience and Remote Sensing* **2011**, *49*, 973-980, doi:10.1109/TGRS.2010.2075937.
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 2007 IEEE International Conference on Image Processing, 16 Sept.-19 Oct. 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. **2017**, *76*, 1719-1748, doi:10.1615/TelecomRadEng.v76.i19.40.
31. Abramov, S.K.; Abramova, V.V.; Lukin, V.V.; Egiazarian, K.O. Denoising of multichannel images with nonlinear transformation of reference image. **2018**, *77*, 769-786, doi:10.1615/TelecomRadEng.v77.i9.30.
32. Dabov, K.; Foi, A.; Katkovnik, V.; Egiazarian, K. Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. *IEEE Transactions on Image Processing* **2007**, *16*, 2080-2095, doi:10.1109/TIP.2007.901238.
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 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? *International Journal of Pattern Recognition and Artificial Intelligence* **2018**, *32*, 1860005, doi:10.1142/s0218001418600054.
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 Image and Signal Processing for Remote Sensing XXI, Toulouse, France, September 21, 2015; pp. 96430D-96430D-96411.
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, doi:10.1364/AO.50.003829.
37. Ponomarenko, N.; Lukin, V.; Astola, J.; Egiazarian, K. Analysis of HVS-Metrics' Properties Using Color Image Database TID2013. In Proceedings of ACIVS, October 2015, Italy; pp. 613-624.
38. Pogrebnyak, O.B.; Lukin, V.V. Wiener discrete cosine transform-based image filtering. *Journal of Electronic Imaging* **2012**, *21*, 16.
39. Lukin, V.; Abramov, S.; Krivenko, S.; Kurekin, A.; Pogrebnyak, O. Analysis of classification accuracy for pre-filtered multichannel remote sensing data. *Expert Systems with Applications* **2013**, *40*, 6400-6411, doi:https://doi.org/10.1016/j.eswa.2013.05.061.
40. Solbo, S.; Eltoft, T. Homomorphic wavelet-based statistical despeckling of SAR images. *IEEE Transactions on Geoscience and Remote Sensing* **2004**, *42*, 711-721, doi:10.1109/TGRS.2003.821885.
41. Ponomarenko, N.; Silvestri, F.; Egiazarian, K.; Carli, M.; Astola, J.; Lukin, V. On between-coefficient contrast masking of DCT basis functions. CD-ROM Proceedings of VPQM, 2007; 4 p.