MULTISCALE DIRECTIONAL BILATERAL FILTER BASED FUSION OF SATELLITE IMAGES

Soner Kaynak[1], Deniz Kumlu[1,2] and Isin Erer[1]

[1]Faculty of Electrical and Electronic Engineering, Electronics and Communication Department, Istanbul Technical University, Maslak, Istanbul, 34469, Turkey,
{kaynaks, ierer}@itu.edu.tr
[2]Department of Electronics Engineering, Turkish Naval Academy, Tuzla, Istanbul, 34942, Turkey
dkumlu@dho.edu.tr

Abstract

The fusion of multispectral and panchromatic satellite images is widely used in remote sensing community to obtain high resolution images both in spatial and spectral domains. One approach is to obtain Panchromatic (PAN) image details through a multiresolution analysis (MRA) based decomposition such as wavelet transform and inject them to the low resolution Multispectral (MS) image. However, this process may lead to redundant detail injection, resulting to artifacts and overenhancement in the fusion results. In this paper, we propose a new fusion scheme based on the decomposition of the PAN image using a multidirectional bilateral filter. Obtained detail layers contain both resolution and direction information. By appropriate selection of these layers, it is possible to perform a proper detail injection. Visual and quantitative results for Quickbird images are presented to validate the proposed method.

1. INTRODUCTION

Remote sensing data provide inhomogeneous information in different methods that are spectral, spatial and temporal. However, this whole data set suggests different representations of the same physical environment. Data fusion is one of the best solutions to evaluate this remote sensing data [1]. Most remote sensing systems provide low-spatial-resolution multispectral (MS) images and low-spectral-quality panchromatic (PAN) images. The aim of the image fusion is to obtain a higher spatial resolution MS image that preserves the high spectral quality of the original MS image [2]. As a result of image fusion, a new image is obtained which is more suitable for other tasks of image processing such as human and machine detection, image fragmentation, feature extraction and object recognition [3].

A trous wavelet transform (ATWT) [4], additive wavelet luminance proportional (AWLP) [5], bilateral filter luminance proportional (BFLP) [6, 7, 8] are widely used multiresolution based fusion methods. In all the methods cited above, the detail layers of the high resolution PAN (HRP) image provided by the multiresolution decomposition process are added directly in ATWT or proportionally in AWLP and BFLP to the low resolution MS (LMS) image. The performance of these methods depends heavily on the amount of the details to be injected to LMS image: a high amount leads to artifacts, as well as spectral distortion while a less amount results in images with low spatial quality. Since it is widely known that *a trous* wavelet decomposition based methods ATWT and AWLP suffer from the redundant detail transfer, recently bilateral filter based methods have been proposed to enhance the decomposition [6].

The bilateral filter is an edge-preserving smoothing filter that works in both the spatial and range domains. It performs a two Gaussian kernel not only for spatial location of neighboring pixels but also for differences in the neighboring pixel intensity values [8]. Due to the use of Gaussian spatial and range kernels in the smoothing process, it is possible to control the transfer of image details to the detail layers. Thus, the use of bilateral filter instead of *a trous* wavelet transform in the decomposition of the HRP image gives us the opportunity to control detail injection process. Similar to wavelet based methods, in bilateral filter case the details can be injected directly: bilateral additive method (BFA) [6] or proportionally: Bilateral Filter Luminance proportional (BFLP) method as given in [6].

Motivated by the success of bilateral based approaches, we propose to explore the use of direction information for appropriate detail selection. For this purpose, we use the multiscale-multidirectional filter (MDFB) introduced in [9] to decompose the HRP image by the classical bilateral filter is extended to its multiscale form and shift-invariant scale subbands are obtained in different resolutions. Then, each scale subband is fed to directional filter banks (DFB) [9, 10] to give a higher number of detail planes which contain both scale and direction information. Then we present a new fusion scheme based on the appropriate selection of detail layers.

Thus, selection of injection details from these planes will provide fusion results with higher spatial resolution and fewer artifacts compared to other methods. Visual results and quantitative analysis for Quickbird images demonstrate the superiority of the use of MDBF method over the classical bilateral based approaches and other state-of-the art algorithms.

The paper is organized as follows. Section 2 presents a background information about MDFB and introduces the proposed MDFB based fusion method. Visual and quantitative results and concluding remarks are given in sections 3 and 4, respectively.

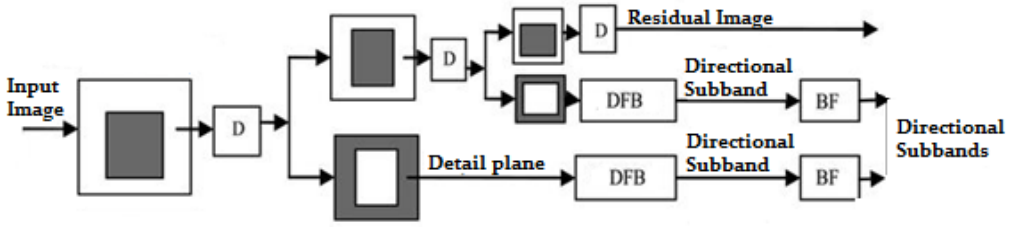

Fig. 1. Image decomposition block diagram of MDBF method

2. MDBF BASED FUSION METHOD

2.1. Review of MDFB

The bilateral filter output is defined as Eq. (1) by combination of spatial and range domain Gaussian filters as,

$$BF[I]_p = \frac{1}{W_p}\sum_{q\in S} G_{\sigma_s}(\|p-q\|)\, G_{\sigma_r}(\|I_p - I_q\|) I_q \qquad (1)$$

where I is the input image, p and q are the pixel locations, I_p and I_q are the intensity values of the Gaussian filter. G_{σ_s} is the Gaussian spatial filter with variance σ_s and G_{σ_r} is the Gaussian range filter with variance σ_r.
The normalization parameter w is given as,

$$W_p = \sum_{q\in S} G_{\sigma_s}(\|p-q\|)\, G_{\sigma_r}(\|I_p - I_q\|) \qquad (2)$$

where the Gaussian filter kernel is defined as,

$$G_\sigma(x) = \frac{1}{2\pi\sigma^2} exp\left(-\frac{x^2}{2\sigma^2}\right) \qquad (3)$$

G_{σ_s} is the Gaussian filter spatial component that reduces the effect of pixels with respect to their location information and G_{σ_r} is the Gaussian filter range component that reduces the influence of these values with respect to their intensity values where the pixel q is different from the pixel p and its neighboring pixels. If the base layer is filtered successively by a bilateral filter by halving/doubling the range/spatial parameters at each stage as used in *a trous* algorithm, a multiscale decomposition can be obtained as,

$$I = I_{rez} + \sum_{j=1}^{J} D_j[I] \qquad (4)$$

where j denotes decomposition level, I_{rez} residual image, namely the output of the last level and $D_j[I]$ the detail layer at level j.

For each decomposition level, detail layers are defined as,

$$D_j[I] = BF_{j-1}[I] - BF_j[I] \qquad (5)$$

Here $BF_0[I]$ is equal to input image I. Multi-scale Directional Bilateral Filter (MDBF) is then obtained by filtering each detail layer with a directional filter bank (DFB) as seen in Fig. 1. Since, MDBF has the ability to preserve edges while capturing the directional information. Inverse MDFB consists of simply inverse DFB of each directional layer and summation of the resulting images to the residual image of bilateral filtering step.

Algorithm 1 MDBF Based Image Fusion

Input A: LRM and HRP images
n: The LRM or HRP band to which the method is applied
***J*:** Number of decomposition levels (scale)
***L*:** Number of decomposition levels (2^l direction)
$\boldsymbol{\sigma_s}$ **and** $\boldsymbol{\sigma_r}$**:** Standard deviation of Gaussian filters
L1: number of selected direction dependent lower detail layers from the LRM.
L2: number of selected direction dependent lower detail layers from HRP.

Step 1: Decomposition of LRM and HRP images
Scale decomposition
for j = 1, …, J ***do***

$$BF\left[LRM_n^{j+1}(p)\right] = \frac{1}{w_p}\sum_{q\in S} G_{\sigma_s}(\|p-q\|)\, G_{\sigma_r}(\|I_p - I_q\|) LRM_n^j(q)$$

$$BF\left[HRP_n^{j+1}(p)\right] = \frac{1}{w_p}\sum_{q\in S} G_{\sigma_s}(\|p-q\|)\, G_{\sigma_r}(\|I_p - I_q\|) HRP_n^j(q)$$

for j = 2,…,J **do**

$$D^j[LRM] = BF^{j-1}[LRM] - BF^j[LRM]$$
$$D^j[HRP] = BF^{j-1}[HRP] - BF^j[HRP]$$

end for
end for
Directional decomposition
Apply DFB to $D^j[LRM]$ and $D^j[HRP]$ to obtain $D^{j,l}[LRM]$ *and* $D^{j,l}[LRM]$

Step 3: Fusion
for j = 2,…J and l= 1,…,L **do**

$$D_F^{j,l} = \sum_{j=2}^{J}\sum_{l=1}^{L2} D_{j,l}\,[\text{LRM}] + \sum_{j=2}^{J}\sum_{l=1}^{L3} D_{j,l}\,[\text{HRP}]$$

end for

Step 4: Reconstruction of HRM image.
Directional reconstruction
for j = 2,…,J **do**
Apply IDFB to $D_F^{j,l}$ to obtain D_F^j
end for
Scale reconstruction
for j = 2,…,J **do**

$$HRM_n = LRMresid\ddot{u}_n + \sum_{j=2}^{J} D_F^j$$

end for

2.2. MDBF Based Image Fusion Technique

MDBF was applied to obtain high-resolution images from Low Resolution MS (LRM) and High Resolution Pan (HRP) images as well as other pan sharpening methods. MDBF is performed separately for each band of LRM image. The images are primarily divided into directional subbands and approximation subbands via MDBF. Subsequently, the directional detail subbands and the approximation subband are fused in accordance with the given fusion rule. In order to reassemble a fused image, the inverse MDBF is applied on the fused directional detail subbands [9]. The proposed fusion method is summarized in Algorithm 1.

As shown in Fig. 1, first, the R, G, B bands of the low-resolution MS and the high-resolution PAN image are decomposed using the bilateral filter to obtain the detail planes. The residual layer of LRM is used as the base layer of the fused image. Then, detail planes of LRM and HRP images are applied to DFB to provide directional detail planes with both scale and direction information. In pansharpening methods, there is a trade-off between spatial improvement and spectral consistency the choice of detail planes to be used in the fusion process is of high importance. Overuse of details from HRP/LRM causes spectral/spatial loss of information in the fusion result. Since the last sub-planes have higher amplitude, they are taken as substitution from the PAN image, and the other planes are added as additive of PAN and MS. Then, according to the given fusion rule, the directional detail planes are merged as in Eq. (6) to obtain the fused directional detail planes as

$$D_F^{j,l} = \sum_{j=1}^{J} \sum_{l=1}^{L1} D_{j,l}\,[\mathrm{LRM}] \; + \sum_{j=1}^{J} \sum_{l=1}^{L2} D_{j,l}\,[\mathrm{HRP}] \qquad (6)$$

where *L1* and *L2* are the numbers of detail layers selected from LRM and HRP images, respectively.

Then, Inverse DFB is applied to the fused directional subbands $D_F^{j,l}$ to obtain fused detail layers D_F^{j}. The fused R, G, B bands are given by summing the residual image with the fused detail planes obtained by the inverse MDBF.

$$HRM_n = LRMresidu_n + \sum_{j=1}^{L1} D_F^{j}, \quad \mathrm{n}=1,2,3 \qquad (7)$$

3. EXPERIMENTAL RESULTS

3.1 Visual Results

The proposed method has been applied to various QuickBird images available at http://glcf.umd.edu [11]. The resolution of PAN images is 0.7 m while the resolution of MS images is 2.8 m. *1024x1024 PAN images* and *256x256* MS images are used. It is considered appropriate to use 2 detail planes and 16 directional detail planes in the light of obtained results.

After fusing the MS image, it is expected that the new image will have a spatial resolution of 0.7 m. However, since there is no MS in spatial resolution of 0.7 m, quantitative comparisons will not be accurate [7]. Thus, using the Wald protocol [12], PAN and MS images are reduced to 2.8 m and 11.2 m resolution with low pass filtering and decimation. The MDBF as well as other pansharpening algorithms are applied to these reduced resolution images and the fusion is compared to the original MS image at 2.8 m resolution. ATWT, AWLP and BFLP results are also included for comparisons purposes.

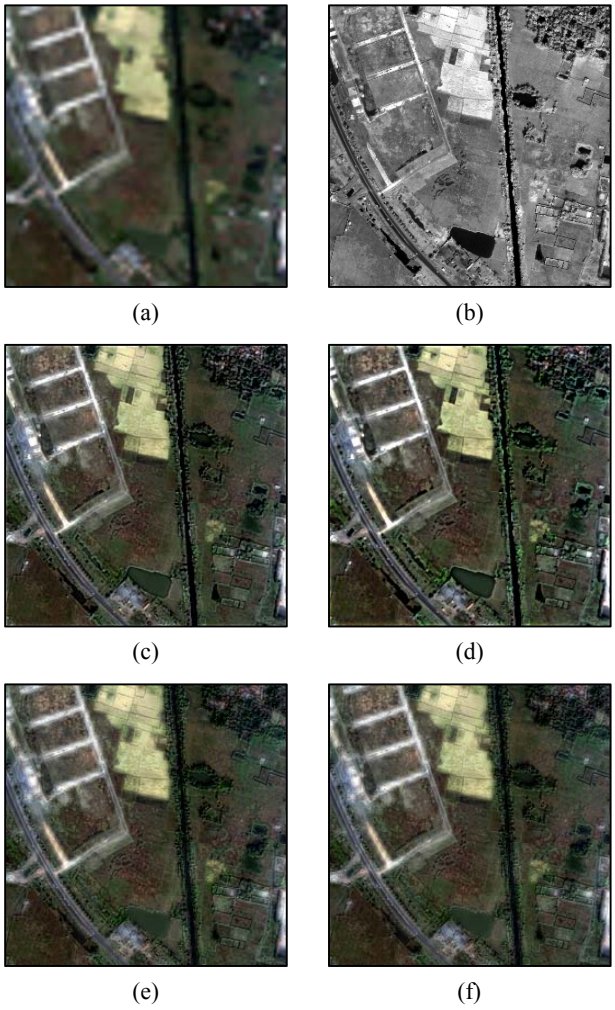

Fig. 2. MS, PAN images and fusion results (a) MS, (b) PAN, (c) ATWT, (d) AWLP, (e) BFLP, (f) MDBF

The decomposition results show that each directional detail plane has twice the width of the previous one due to the applied processes. For this reason, as one moves from direction 1 to direction 16, the amount of energy that it contains and thus the amount of detail will increase. In addition, due to the resolution difference in each direction, the directional detail planes of HRP image have a total amplitude of 10.783 times the directional detail planes of LRM image. If the MS component is used in the detail layer obtained from the sub planes with high amplitude, spatial information contained in the pan data causes interference and reduces spatial and spectral resolution. For this reason, the last two directional planes of LRM are discarded and remaining directional planes are added to the corresponding directional planes of HRP image.

Fig. 2 shows MS, PAN images and fusion results of ATWT, AWLP, BFLP and proposed method, respectively. One can observe that in ATWT and AWLP methods, redundant data is transferred to the image causing overenhancement in road and field borders. This effect is reduced in BFLP and MDBF which provide more natural images. The colors also seem more genuine showing a better preservation of the spectral information.

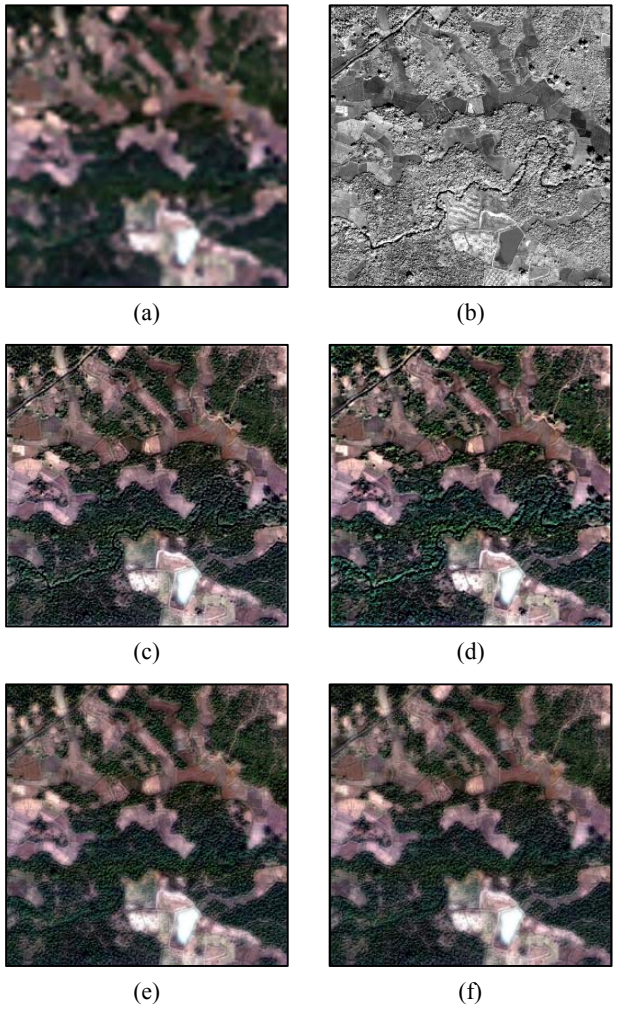

Fig. 3 MS, PAN images and fusion results (a) MS, (b) PAN, (c) ATWT, (d) AWLP, (e) BFLP, (f) MDBF

Fusion results for mountainous terrain are given in Figure 3. The results show that the ATWT and AWLP results are distorted in colors and that bilateral filter based methods give spectral results closer to the original MS. With the use of directional data, this problem is further reduced in MDBF. The vegetation parts are more natural in Fig. 3 (f).

3.2 Quantitative Results

Similarity between two images can be measured with the help of correlation coefficient. The correlation coefficient is between +1 and -1. A +1 of the coefficient means that the two views are equal. The correlation coefficient of -1 indicates that the two images are completely opposite to each other. The correlation coefficient is calculated by Eq. (8).

$$Corr(A,B) = \frac{\sum_{m,n}(Am,n-\mu A)(Bm,n-\mu B)}{\sqrt{\sum_{m,n}(Am,n-\mu A)^2(Bm,n-\mu B)^2}} \quad (8)$$

The above equation is the reference MS image with the fusion result of a single-band image. *Spatial_Corr* is the correlation between the fusion result and the reference PAN image. *Corr* and *Spatial_Corr* should be as close to 1 as possible [6].

RMSE is one of the very good spectral quality indicator. It corresponds to the square root of the average of the sum of the squares of the error between the fusion result and the original MS image. This measurement determines the average amount of distortion at each pixel. The RMSE calculation is made between the fusion bands of image bands related to each of the multiple spectral bands. *MS (m, n)* and *F (m, n)* are images of size *MxN (m, n)*. RMSE for single-band fusion to express the pixel is given in Eq. (9) [7].

$$RMSE = \sqrt{\frac{1}{MN}\sum_{i=1}^{M}\sum_{i=1}^{N}(A-B)^2}\mathrm{T} \quad (9)$$

Lower RMSE value means that the spectral properties are preserved in a better way. The relative dimensionless global error (ERGAS) value is given in Eq. (10) to calculate the global error between the MS image and the fusion image.

$$ERGAS = 100\frac{h}{l}\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{RMSE^2(MS,Fi)}{Mi^2}\right)} \quad (10)$$

Here, h and l are respectively the resolution values of the high and low spatial resolution images. While M_i is the average radiance of the reference MS image band. The ERGAS value should converge to "0" [12].

RM is defined in Eq. (11) to calculate the magnitude of the mean pixel value of the fusion image relative to the reference MS image.

$$RM = \%\frac{MS_{ort} - Füzyon_{ort}}{MS_{ort}} \quad (11)$$

The above relative slip value is for single-band fusion processes. The resulting image is close to the reference MS image at that time. Therefore, RM value is expected to be close to "0" [6].

The spectral angle mapper (SAM) value is the expression of the angle between two vectors. The SAM formulation is the same as in Eq. (12).

$$SAM = arccos\left(\frac{\langle MS,F\rangle}{\|MS\|_2\|F\|_2}\right) \quad (12)$$

where F and MS correspond to the coordinate values in a certain position of the fusion result and the original view, respectively. The above SAM value represents the angular difference in a certain coordinate of the fusion and reference view. SAM values are calculated and averaged for all pixels. If the value of the angle is small, it means that the inter-band relations of the resulting image with the reference MS image are similar. Therefore, it is expected that the value converges to "0" [6].

The spatial ERGAS (SERGAS) value is given in Eq. (13) to calculate the global error between the PAN image and the fusion image.

$$SERGAS = 100\frac{h}{l}\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{RMSE^2(HRP,Fi)}{Mi^2}\right)} \quad (13)$$

Here, h and l fall against the resolution values of the high and low spatial resolution images, respectively and M_i is the average radiance of each band. The SERGAS value should be as low as possible [12].

The quality index (Q) is calculated from a combination of correlation loss, brightness and contrast imperfections [12]. The formulation of the Q is given in Eq. (14).

$$Q = \frac{\sigma MSF}{\sigma MS\sigma MF}\frac{2\mu MS\mu F}{\mu^2 MS+\mu^2 F}\frac{\sigma MS\sigma MF}{\sigma^2 MS+\sigma^2 MF} \quad (14)$$

Q_Avg means the average of the quality index calculated for each band in a fusion image. While the Q value is calculated, a window of *NxN* is displayed on the image. The *N* value in this range can range from 16 to 128. The highest value that *Q_avg* can have is "1", which is the ideal result [13].

The quantitative results for scenes represented in Fig. 2 and 3 are given in Table 1 and 2, respectively. It can be observed that MDBF method achieves better results than all the other methods for most of the metrics. ERGAS and SERGAS results show that the proposed method is superior to BFLP in both spectral and spatial domains validating the use of direction information.

Table 1 Fusion results for scenario I

	ATWT	AWLP	BFLP	MDBF
Avg_Corr	0.9134	0.8788	0.9632	**0.9635**
Avg_Sp_Corr	0.9707	0.9767	0.8568	**0.8629**
Avg_RMSE	28.5006	36.2519	17.1257	**17.0673**
Avg_RM	1.7154	2.6204	0.0444	**0.0258**
SAM	7.6032	8.9566	1.9171	**3.0499**
ERGAS	7.1485	9.0796	4.2840	**4.2811**
Spatial_Ergas	11.7971	12.0156	12.0193	**11.9612**
Q_Avg	0.7203	0.6521	0.8222	**0.8202**
Q	0.9572	0.9184	0.9407	**0.9454**

Table 2 Fusion results for scenario II

	ATWT	AWLP	BFLP	MDBF
Avg_Corr	0.8765	0.8417	0.9402	**0.9422**
Avg_Sp_Corr	0.9847	0.9797	0.9081	**0.9188**
Avg_RMSE	36.1086	42.9295	23.1800	**22.8688**
Avg_RM	1.8026	2.7923	0.1549	**0.1635**
SAM	12.2865	13.5654	5.3091	**6.7388**
ERGAS	9.0325	10.7710	5.8167	**5.7205**
Spatial_Ergas	17.9505	17.8112	18.6593	**18.5801**
Q_Avg	0.7172	0.6714	0.8212	**0.8222**
Q	0.6375	0.5813	0.7376	**0.7442**

4. CONCLUSION

The greatest benefit of MDBF is the ability to use the direction information along the multiresolution decomposition. Due to this property, we can prevent redundant detail transfer to MS images. Thus, the proposed method does not suffer from overenhancement or artifacts which are common problems of multiresolution based pansharpening methods. Visual results and quantitative metrics demonstrate the superiority of the use of MDBF over the classical MRA based approaches where the direction information is not explored.

5. REFERENCES

[1] T. Taxt, A. H. Schistad, "Data fusion in remote sensing," *in Fifth Workshop on Data Analysis in Astronomy*, V. Di Gesu and L. Scarsi, Eds., Erice, Italy, Nov. 1996.

[2] C. Pohl, J. L. Genderen, "Multisensor image fusion in remote sensing: concepts methods and applications", *Int. Journal of Remote Sensing*, 1998, vol. 99, no. 5, pp. 823-854.

[3] Y. H. Tsai, Y. H. Lee, "Wavelet-based image fusion by adaptive decomposition", *Eighth International Conference on Intelligent Systems Design and Applications*, vol.2, 978-0-7695-3382-7/08, 2008 IEEE, pp283-287

[4] P. Dutilleux, An implementation of the "algorithme à trous" to compute the wavelet transform. *Proceedings du congrès ondelettes et méthodes temps-fréquence et espace des phases*, Marseille, 14-18 septembre 1987, Springer-Verlag Editors, pp. 298-304.

[5] J. Nuñez, X. Otazu, O. Fors, A. Prades, V. P. R. Arbiol, "Multiresolution-based image fusion with additive wavelet decomposition", *IEEE Trans. Geosci. Remote Sens.*, 1999, vol. 37, no. 3, pp. 1204-1211.

[6] N. H. Kaplan, I. Erer, "Bilateral filtering-based enhanced pansharpening of multispectral satellite images", *IEEE Geoscience and Remote Sensing Letters*, Volume: 11, Issue: 11, 06 May 2014.

[7] N. H. Kaplan, I. Erer, F. Elibol, Fusion of multispectral and panchromatic images by combining bilateral filter and IHS Transform, *20th European Signal Processing Conference* (EUSIPCO), 2012.

[8] N. H. Kaplan, I. Erer, "Bilateral filtering pyramid based pansharpening of multispectral sateliate images", *IEEE Geoscience and Remote Sensing Symposium (IGARSS 2012),* 2376-2379, 2012.

[9] H. Jianwen, L. Shutao, "The multiscale directional bilateral filter and its application to multisensor image fusion," *Information Fusion*, 2012 Volume 13, Issue 3, Pages 196–206.

[10] D. Kumlu, I. Erer, "Multiscale directional bilateral filter based clutter removal techniques in GPR image analysis*." IEEE International Geoscience and Remote Sensing Symposium*, 2017.

[11] *DigitalGlobe* (2005), QuickBird scene 000000185940_01 _P001, Level Standart 2A, 12/22/2009, DigitalGlobe, Longmont, Colorado.

[12] L. Wald, T. Ranchin, M. Mangolini., "Fusion of satellite images of different spectral resolutions: Assessing the quality of resulting images", *Photogramm. Eng. Remote Sens*., 1997, vol. 63, no. 6, pp. 691-699.

[13] A. Bovik, Z. Wang, "A universal image quality index" *IEEE Signal Proces*., 2002, Vol. 9 no. 3 pp. 81-84.