DECISION-BASED FUSION OF PANSHARPENED VHR SATELLITE IMAGES USING TWO-LEVEL ROLLING SELF-GUIDANCE FILTERING AND EDGE INFORMATION

F. Samadzadegan[1,*], A. Toosi[1], F. Dadrass Javan[1,2], A. Stein[2]

[1] School of Surveying and Geospatial Engineering, University College of Engineering, University of Tehran, Tehran, Iran – (samadz, ahmadtoosi71)@ut.ac.ir
[2] Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, The Netherlands – (f.dadrassjavan, a.stein)@utwente.nl

Commission VI, WG VI/3

KEY WORDS: Image Fusion, Pansharpening, Very-high-resolution Imagery, Guided Filter, Canny Edge Detector

ABSTRACT:

Pan-sharpening (PS) fuses low-resolution multispectral (LR MS) images with high-resolution panchromatic (HR PAN) bands to produce HR MS data. Current PS methods either better maintain the spectral information of MS images, or better transfer the PAN spatial details to the MS bands. In this study, we propose a decision-based fusion method that integrates two basic pan-sharpened very-high-resolution (VHR) satellite imageries taking advantage of both images simultaneously. It uses two-level rolling self-guidance filtering (RSGF) and Canny edge detection. The method is tested on Worldview (WV)-2 and WV-4 VHR satellite images on the San Fransisco and New York areas, using four PS algorithms. Results indicate that the proposed method increased the overall spectral-spatial quality of the base pan-sharpened images by 7.2% and 9.8% for the San Fransisco and New York areas, respectively. Our method therefore effectively addresses decision-level fusion of different base pan-sharpened images.

1. INTRODUCTION

Earth observation (EO) satellites such as WorldView, QuickBird and Gaofen, provide a low-spatial-resolution multispectral (MS) image of rich spectral content and a single-band high-spatial-resolution panchromatic (PAN) image without considerable spectral information. Fusing the MS and PAN images produces a high-resolution MS (called pan-sharpened) image of a relatively high spectral information content.

Several image fusion methods (also called pan-sharpening algorithms) have been proposed to integrate the above-mentioned images (Snehmani et al., 2017; Javan et al., 2021; Peijuan et al., 2021). Generally, every PS process can be performed on three levels: at the pixel level, the feature level, or the decision level (Belgiu and Stein, 2019). PS methods can be classified into five major categories of component substitutions (CS)-based methods, multi-resolution analysis (MRA)-based methods, variational optimization (VO)-based methods, hybrid methods and deep learning (DL)-based methods (Javan et al., 2021). Methods from each category have their strengths and weaknesses: some are good at preserving the spectral information of the MS image, while others focus more on transferring the spatial information of the PAN band to the MS image. In the first case, the pan-sharpened image largely has the spectral and color characteristics of the MS image, while in the latter case it may be more similar to the PAN image in terms of spatial features, for instance, sharpness of edges and spatial patterns.

Fusion of two or more pan-sharpened images being produced using different PS methods can be a good solution to enable us to take advantage of each base method, i.e. to enjoy the strength of each algorithm and on the other hand avoid their weaknesses. In this regard, Lou et al. (2013) purposed a decision-based method for the fusion of two pan-sharpened remotely sensing images. Their method was based on image segmentation and then locally selecting the best pan-sharpened image. Due to the segmentation

errors, their algorithms cannot always produce high-quality pan-sharpened images. Based on the object-based spectral quality assessment (Samadzadegan and Javan, 2011; Rodríguez-Esparragón et al., 2017; Toosi et al., 2020a) we proposed a decision-based fusion of multiple pan-sharpened images based on image multi-resolution segmentation and classification and showed that our new combined fusion strategy can increase the spectral quality of the initial pan-sharpened images by about 37%, without image spatial distortions and spectral artifacts (Toosi et al., 2020b).

Nowadays this kind of very-high-resolution (VHR) pan-sharpened images are widely used in applications where both spatial resolution and spectral information are required simultaneously (Mehravar et al., 2022). Urban mapping and land use/land cover (LULC) classification are two topics that require "high spatiality" and "high spectrality" information for these data (Javan et al., 2019; Gaur et al., 2015). Due to the wide range of applications more endeavors are needed to be conducted to take the advantage of two or more pan-sharpened images as much as possible. To do this, inspired by the guided filtering-based common pan-sharpening method developed by Meng et al. (2016), we propose a new method for decision-level fusion of two pan-sharpened images (produced by the means of two different basic PS algorithms) based on the edge-preserving rolling self-guidance filtering (RSGF) paradigm and edge detection.

The remainder of this paper is organized as follows: Section 2 is devoted to describing the proposed methodology, describing the concepts of RSGF, and the dataset. The results of the implementation phase and some discussions related to the obtained results are presented in Section 3. Finally, some concluding remarks are drawn in Section 4.

---

* Corresponding author

2. MATERIALS AND METHODS

2.1 Methodology

A diagram of the proposed method for decision-based fusion of two pan-sharpened images is presented in Figure 1. According to this flowchart, the initial PAN and MS images are fused using two different PS methods, called "Method 1" and "Method 2", while the output of the two PS methods are known as "PS 1" and "PS 2", respectively.

Next, the spectral and spatial qualities of the obtained pan-sharpened images are assessed compared to the initial MS and PAN images. The standard form of Wald's protocol (Wald et al., 1997) is commonly used to assess the spectral quality (Figure 2-a). Based upon the mentioned standard scheme the initial MS and PAN data are downscaled (with a scale factor of 1/4) and resampled and then fused and the fused product is compared to the downscaled images. In our study, we modify Wald's protocol (Figure 2-b), due to our specific decision-fusion paradigm. In this scheme, the MS image is sharpened using the PAN band and the obtained pan-sharpened image is rescaled and resampled to the 4×4 downer scale. Finally, it is compared with the initial MS and PAN data (Ghassemian, 2016).

After the spectral and spatial quality assessment, the pan-sharpened output with a relatively better spectral quality is denoted as PS_{Spc} while that with a relatively better spatial quality is denoted as PS_{Spt}. Then the histogram of the target PS_{Spt} image is matched to that of the reference PS_{Spc} image, and the output is named PS_{Spt-HM}. We chose this alternative in order to have more compatible outputs in the subsequent processing chain. We then minimize the objective function according to Eq. 1.

$$\Theta = argmin_M \sum_j d(M(P_j), Q_j) \qquad (1)$$

where $M(P_j)$ is the optimal color mapping function that aims to minimize the distance (d) between the histogram sets of PS_{Spc} and PS_{Spt}, named $P = \{p_i\}_{i=1}^k$ and $Q = \{q\}_{i=1}^k$, simultaneously. Thus, minimization of Θ leads to matching the histogram of PS_{Spt} to that of PS_{Spc}.

At the next step the obtained PS_{Spt-HM} is decomposed into its principal component as PC_1, PC_2, …, PC_n. In this way, a single-band image, i.e. PC_1, is obtained by reducing the data dimensionality that contains the majority of spatial information of the PS_{Spt-HM} image, with the lowest possible noise level. As an equation, the PCs are obtained as Eq. 2.

$$PC_n = \sum_{k=1}^m g_{nk} \times PS_{Spt-HM_k} \qquad (2)$$

where g_{nk} is the kth element of the ith eigenvector $g_n^t = (g_{11}, g_{12}, …, g_{1k}, …, g_{1m})$ (Liu and Mason, 2016).

The Canny edge detector (Canny, 1986) is used to extract edges from the PC1 resulting into the extracted edge map E. In addition, the edge-preserving rolling self-guidance filter (RSGF) (Zhang et al., 2014) is applied to PC1. The obtained result is a base layer (B) that contains both low frequency and strong edge information (Meng et al., 2016). The base layer may contain a low variety of slight edges. By subtracting the PC1 and B, smoothed by a Gaussian low-pass filter (GLF), the detail layer (D) is obtained (Eq. 3).

$$D = PC_1 - GLF * B \qquad (3)$$

where $GLF * B$ is referred to as the filtered base layer.

By applying multi-level thresholding (MLT) (Otsu, 1979) D is segmented into a desired detail mask (D_m) and an area containing no considerable details. Eq. 4 shows how this method looks for the optimum threshold to minimize the intra-class variance of image pixels.

$$\delta_w^2(t) = w_0 \times \delta_0^2(T) + w_1 \times \delta_1^2(T) \qquad (4)$$
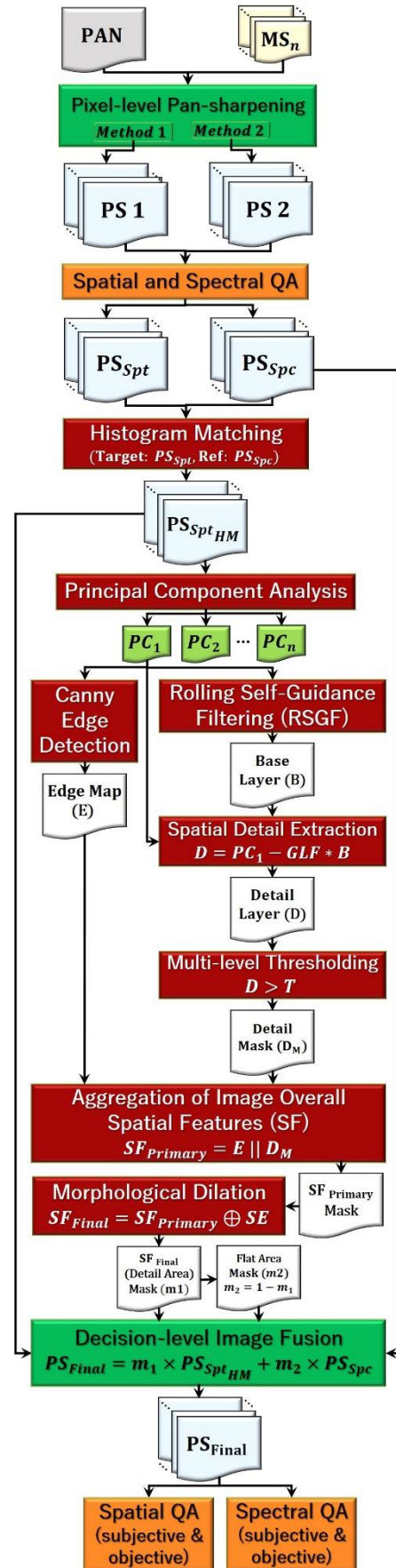$$D_m = D \mid D > T \qquad (5)$$

Figure 1. Flowchart of the proposed decision-fusion method

where $\delta_0^2(T)$ and $\delta_1^2(T)$ are variances of the two segments, and $w_0 = \sum_{i=0}^{T-1} p(i)$ and $w_1 = \sum_{i=0}^{L-1} p(i)$ are the probabilities of the two segments separated by an optimum threshold (T). The vertical bar | indicates conditional equality, i.e. conditional on those pixels of D where the $D>T$.
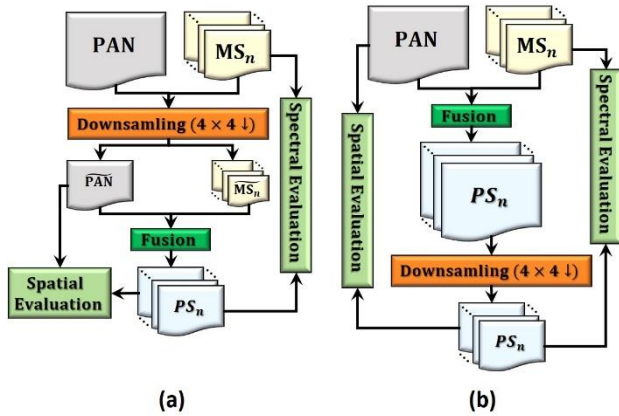

Figure 2. Wald's protocol: (a) the first (standard) algorithm, (b) the second algorithm

The overall spatial features of PC_1 are extracted by integrating the edge map E with the detailed map D_m using the OR logical operation as Eq. 6.

$$SF = E || D_m \qquad (6)$$

where || denotes the OR logical operator. Pixels that were recognized as edges in E or as details in Dm, will be present in the final SF image.

As in the VHR satellite image due to the high diversity of pixel values and high heterogeneousness of pixels around the edges and spatial details we dilate the obtained primary SF by morphological operations (Eq. 7) to obtain neighbor pixels around the edge and details pixels:

$$SF_{final} = SF_{primary} \oplus SE \qquad (7)$$

where $\oplus$ and SE denote the morphological dilation operation and the desired structural element, respectively, $SF_{primary}$ is the primary spatial features map and SF_{final} the dilated map. For brevity, SF_{final} is called the m_1 mask and its complement representing the flat area the m_2 mask ($m_2 = 1 - m_1$).

The m_1 and m_2 masks can be exerted to the PS_{Spt-HM} and PS_{Spc}, respectively according to the linear composition equation of Eq. 8, providing us the final pan-sharpened image, denoted as PS_{final}.

$$PS_{final} = m_1 \times PS_{Spt_{HM}} + m_2 \times PS_{Spc} \qquad (8)$$

Based upon the decision-based fusion paradigm of Eq. 8, PS_{final} takes the pixels related to spatial features from all spectral bands of PS_{Spt-HM}. This is because we consider the spatial characteristics of objects in these areas to be important and PS_{Spt-HM} better performs from the spatial quality point of view. Similarly, the reminder pixels which are related to the flat area with no noticeable spatial features are captured from the PS_{Spc} image and are assigned to PS_{final}.

During the final phase of the proposed method, the quality of the obtained PS_{final} is evaluated both subjectively and objectively. The subjective evaluation is done by the human (expert) visual system (HVS) while the latter is done using the quality assessment metric and according to the second form of Wald's protocol.

2.2 Rolling guidance filtering (RGF)

VHR satellite images contain a variety of spatial features including structures and object edges. Guidance filters as edge-preserving filters are utilized to perform edge-preserving smoothing on these kinds of images. This is usually conducted using the content of another image, called a guidance image, to influence the filtering process. Sometimes the target image can jointly be considered as the guidance image; in such a case the guidance filtering is called self-guidance filtering (Meng et al., 2016).

The edge-preserving guidance filtering involves an input image I, a guidance image G, and an output image I_{GF} (which is assumed to be a linear transformation of G in a local kernel w_k centered at pixel k (Eq. 9).

$$I_{GF_i} = a_k G_i + b_k, \forall i \in w_k \qquad (9)$$

where G_i and IGF_i stand for the ith pixel of the guidance image and output images in kernel wk. Also, w_k is a square kernel of size $(2R+1)\times(2R+1)$, where R is the radius of the kernel. Furthermore, a_s and b_k are constant linear coefficients in w_k and can be estimated by minimizing the squared difference between the input and output images (He et al., 2012; Gao et al., 2016). From a practical point of view, the guidance filtered images are usually estimated as Eq. 10.

$$I_{GF_i} = \bar{a}_i G_i + \bar{b}_i \qquad (10)$$

where $\bar{a}_i = \frac{1}{|w|}\sum_{k \in w_i} a_k$ and $\bar{b}_i = \frac{1}{|w|}\sum_{k \in w_i} b_k$.

Different types of guidance filters have been proposed so far which are used for PS tasks as well (Gao et al., 2016). One of the most effective methods was proposed by Zhang et al. (2014). They developed an iterative edge-preserving image guidance filtering method called rolling guidance filtering (RGF) that converges rapidly. This fast method is extensible to accommodate various types of image operations and achieves real-time performance and produces artifact-free results.

2.3 Data description

In this study, two VHR satellite imageries are used. Each dataset includes one PAN band and its corresponding MS images. The first dataset is a Worldview (WV)-2 image from the San Francisco-United States of America (USA) and the second one is a WV-4 from the New York (NY)-USA. The mentioned dataset are shown in Figure 3 in true color composite (TCC), i.e. RGB, and some pieces of information are provided in Table 1 related to them.

Data	Sat.*	Bands	GSD (MS/PAN)	Org. size** (PAN/MS)
San.	WV-2	8 (coastal, blue, green, yellow, red, red edge, NIR1, NIR2)	1.84/0.46 m	1000×1000 / 4000×4000 pixels
NY	WV-4	4 (red, green, blue, NIR)	1.24/ 0.31 m	16384×16164 / 4096×4041 pixels

* Sat.: Satellite, San.: San Francisco, NIR: near infra-red, GSD: Ground sampling distance, Org.: Original
** This is the initial size of the datasets. This study is conducted on the resized form of the original data, i.e. PAN: 1024×1024 and MS: 256×256 for both case studies
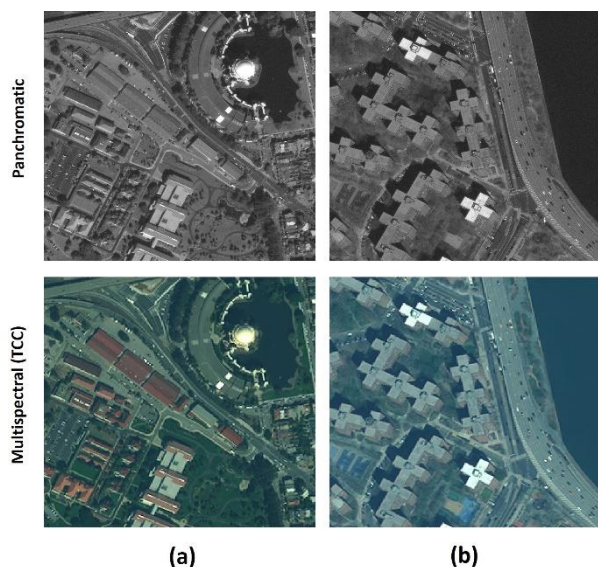
Table 1. Data descriptions

Figure 3. Presentation of datasets: (a) Resized PAN and MS image from San Francisco, (b) Resized PAN and MS image from NY

2.4 State-of-the-art PS methods

In this study, four PS methods are used: the Ehlers (Klonus and Ehlers, 2009), Gram-Schmidt (GS) (Laben and Brower, 2000), Smoothing Filter-based Intensity Modulation (SFIM) (Liu, 2000), and Brovey (Gillespie et al., 1987) methods. A brief description of these methods is provided in Table 2. It is to some degree irrelevant which PS methods is used in this study because it does not make much of a difference when implementing our method. We tried as much as possible to choose methods that behave differently in terms of performance based on the results of our former meta-analysis study (Javan et al., 2021).

Method	Description / Formula
Ehlers	In this method (which is also known as Fast Fourier-Transform IHS-Transform) image fusion is conducted in the Fourier space of the PAN and MS images.
Brovey	$PS_i = \left(\frac{MS_i}{\frac{1}{n}\sum_{i=1}^{n} MS_i}\right) \times PAN \qquad (11)$
GS	It is based on component substitution of the first GS component (GS1) of the MS image and the PAN band.
SFIM	$PS_i = \frac{PAN}{MS_{Mean}{}^*} \times MS_i \qquad (12)$

MS_{Mean}: a mean image that is produced by an average filter.

Table 2. Information on PS methods

3. RESULTS AND DISCUSSIONS

Ehlers and SFIM PS methods are applied to the San Francisco dataset and Brovy-GS pairs are used for PS of the NY images. It is noteworthy that as usually there is a ratio of 1:4 between the MS and PAN images so the MS data is first upsampled to the resolution of the initial PAN band ($\uparrow \widetilde{MS}_n$, 4×4) and then $\widetilde{MS}_n$ fused with the PAN information. Figure 4 presents the results of all PS methods in two case studies. Based on the second form of Wald's protocol the spectral and spatial quality assessment is conducted based on the multi-modal Spectral Angle Mapper (SAM) (Alparone et al., 2007) and the Spatial Correlation Coefficient (SCC) (Pushparaj and Hegde, 2017), respectively. These metrics as two of the most commonly used measures for

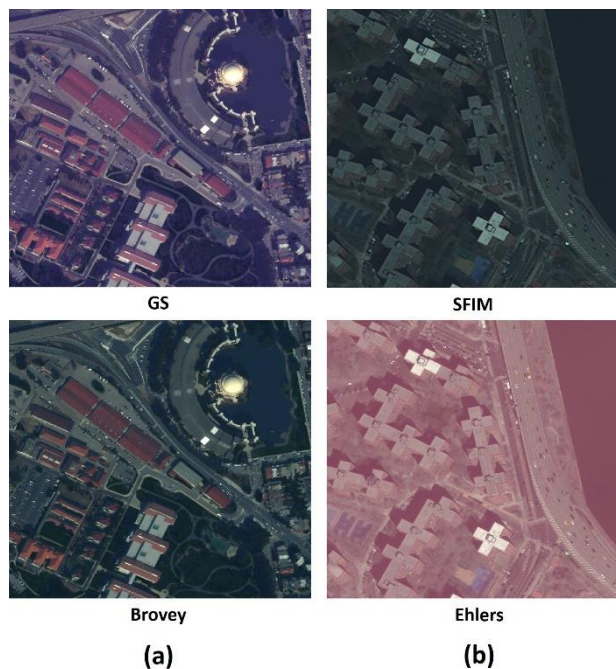spectral and spatial quality assessment are defined according to Eq. 13 and Eq. 14, respectively.

Figure 4. PS results: (a) San Francisco dataset, (b) NY dataset

$$SAM(I_r, I_f) = \arccos\left[\frac{\sum_{i=1}^{n}(I_r(i) - I_f(i))}{\sqrt{\sum_{i=1}^{n} I_r(i)}\sqrt{\sum_{j=1}^{n} I_f(i)}}\right] \qquad (13)$$

$$SCC(I_f, PAN) = \frac{\sum_{j=1}^{M}\sum_{i=1}^{N}(I_{f_i} - \bar{I}_f)(PAN_i - \overline{PAN})}{\sqrt{\sum_{j=1}^{M}\sum_{i=1}^{N}(I_{f_i} - \bar{I}_f)^2 \sum_{j=1}^{M}\sum_{i=1}^{N}(PAN_i - \overline{PAN})^2}} \qquad (14)$$

Where I_f and I_r referred to the fused and MS images, respectively. The bar symbol above some elements is related to the mean value of that element. Furthermore, M and N define the image dimensions. SAM values vary between 0 to 90 degrees where the *SAM=0*, as the ideal value, means that there is no spectral difference between the reference MS and pan-sharpened images. On the other hand, SCC measures the spatial similarity between the pan-sharpened image and the reference PAN band. The ideal SCC value is +1 where the spatial similarity is maximum and the value -1 refers to the minimum possible spatial similarity.

In the NY case, the Ehlers method has relatively better SCC while SFIM has a relatively better spectral performance so Ehlers and SFIM are known as PS_{Spt} and PS_{Spc}, respectively. In the case of the San Francisco dataset, GS produced a pan-sharpened image with relatively better SCC, and on the other hand, SAM for the Brovey method was closer to the ideal value, so here GS and Brovey are labeled as PS_{Spt} and PS_{Spc}, respectively. The results of the quality assessment are provided in Table 3, in which the superior methods are discriminated by bold font and the others are shown by the standard font.

Dataset	PS method	Quality metric	
		SAM	SCC
NY	Ehlers	1.3428	**0.9970**
	SFIM	**0.6485**	0.9584
San.	GS	3.6528	**0.9842**
	Brovey	**2.9791**	0.9667

Table 3. The results of spectral and spatial quality assessment

During the next step, the histogram of Ehlers and GS pan-sharpened images are matched to that of SFIM and Brovey, respectively. Then they are decomposed to their principal components using the PCA that their first principal components (PC_1) containing the maximum spatial information and the least possible noise are shown in Figure 5. The obtained PC1s are now ready for extraction of some important spatial information. First, the edges are extracted from both $PC1s$ using the Canny operator (Figure 6).
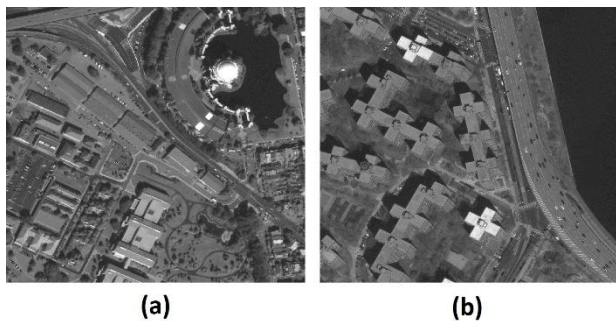

Figure 5. The first principal components (PC1): (a) San Francisco dataset, (b) NY dataset
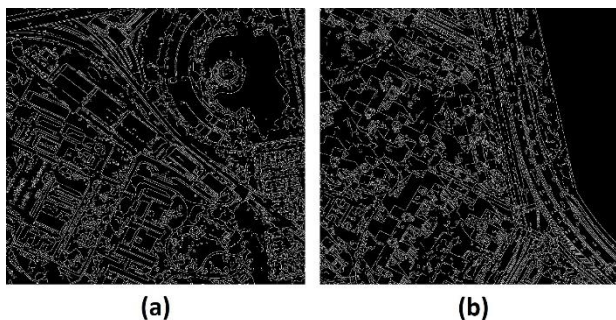
Figure 6. Edge maps (the white pixels denote the edges and the black background is related to non-edges pixels): (a) San Francisco dataset, (b) NY dataset

Rolling guided filter is applied to the PC1s, where in both cases the PCs themselves played the role of guidance images, i.e. the filtering was in the form of self-guidance. The parameters of RGF were set as default of the Zhang et al. (2014) paper as follows: $\sigma_s = 4$, $\sigma_r = 0.1$, $iteration = 4$, $GaussianPrecision = 0.05$. This led to the production of base layers (B) that contain both low-frequency information and strong edges (Figure 7).
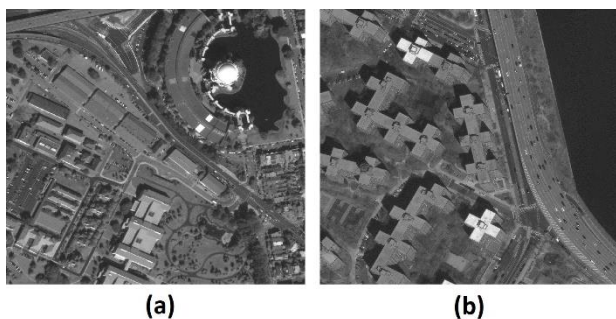

Figure 7. Base layers: (a) San Francisco dataset, (b) NY dataset

A Gaussian filter (with $\sigma = 5$) was applied to the obtained base layer (B) and the filtered output is subtracted from the $PC1s$. This operation leads to the extraction of detail layers (D). Figure 8 presents the obtained detail layers for both case studies. MLT provides us with the optimum thresholds for segmenting the significant features. For the San Francisco and NY cases, the thresholds were estimated as 22.04 and 3.95, respectively. Figure 9 presents the mentioned thresholds (T) on the histograms of detail layers and also the obtained masks which are produced by applying T to the detail layers.
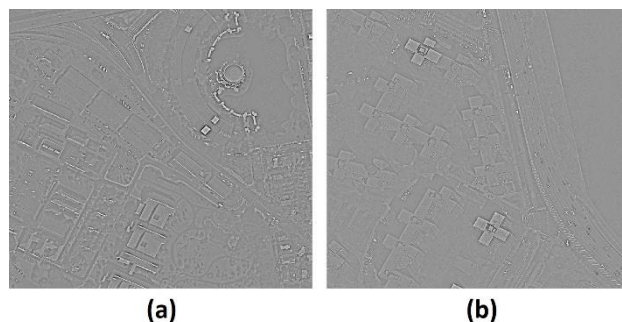

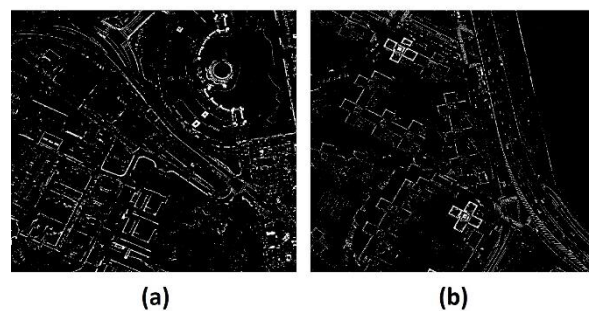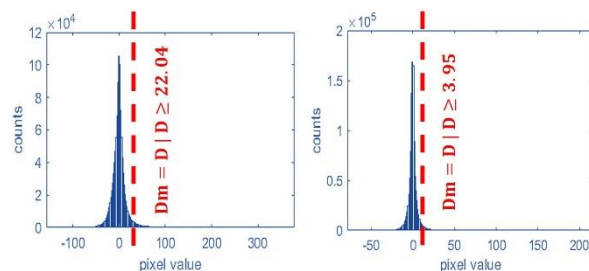**Figure 8.** Detail layers: (a) San Francisco dataset, (b) NY dataset

Figure 9. The results of MLT and the produced D_m: (a) San Francisco dataset, (b) NY dataset

The primary edge map (E) is integrated with the obtained detail mask (D_m) using the OR logical operation and then using a structural element with the size of 2, the results are dilated to contain the neighboring pixels around the edge, and details pixels as well. The final results which are called $SF_{final-mask}$ (or for brevity we call it m_1) and its complement (m_2) are shown in Figure 10. Note that m_1 and m_2 are masks that are related to the high-details and flat areas, respectively.

The linear combination of m_1, m_2, PS_{Spt-HM}, and PS_{Spt} (Eq. 8) produces the final PS image in a decision-based fusion scheme. The obtained decision-fused PS image for both case studies are shown in Figure 11. Here the concept of our decision-based fusion is better found. The "decision-based fusion" term indicates that the pixels of two base pan-sharpened images are transferred

to the final image with a decision that is made based on whether the pixels are part of the detailed class or part of the flat class.
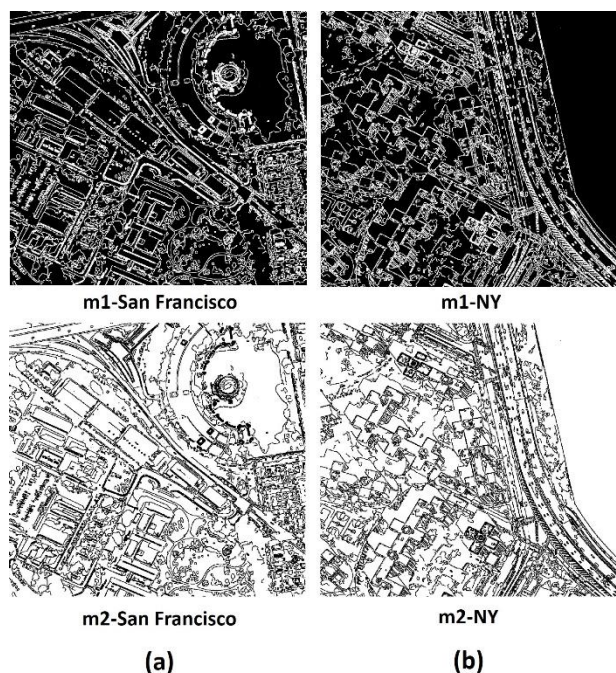

Figure 10. Masks of details area (m_1) and flat area (m_2): (a) San Francisco dataset, (b) NY dataset
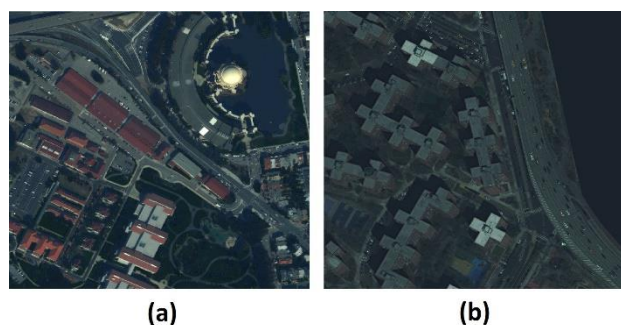

Figure 11. The final pan-sharpened images obtained by our decision-based fusion method: (a) San Francisco dataset, (b) NY dataset

As the subjective (visual) spectral and spatial quality assessment of the results, Figure 12 provides a comparison between the initial pan-sharpened images and the final decision-based fused images. By scrutinizing the results it can be seen that in the flat areas the final pan-sharpened outputs are similar to the PS_{Spc} images while in the details area they inherit the characteristics of the PS_{Spt-HM}.

Here the objective quality assessment is done using SAM and SCC metrics. For spatial quality assessment, according to the second scheme of the Wads' protocol, the obtained results are compared to the initial PAN images while for the spectral quality evaluation they are first downsampled to the resolution of initial MS bands, and then the evaluation is conducted between PS_{final} and MS. By calculating the above-mentioned metrics it is seen that for the San Francisco case study the evaluation results are as follows: SAM=3.0201 and SCC=0.9700. Furthermore, for the NY case, the results are as: SAM=0.7770 and SCC=0.9641. By
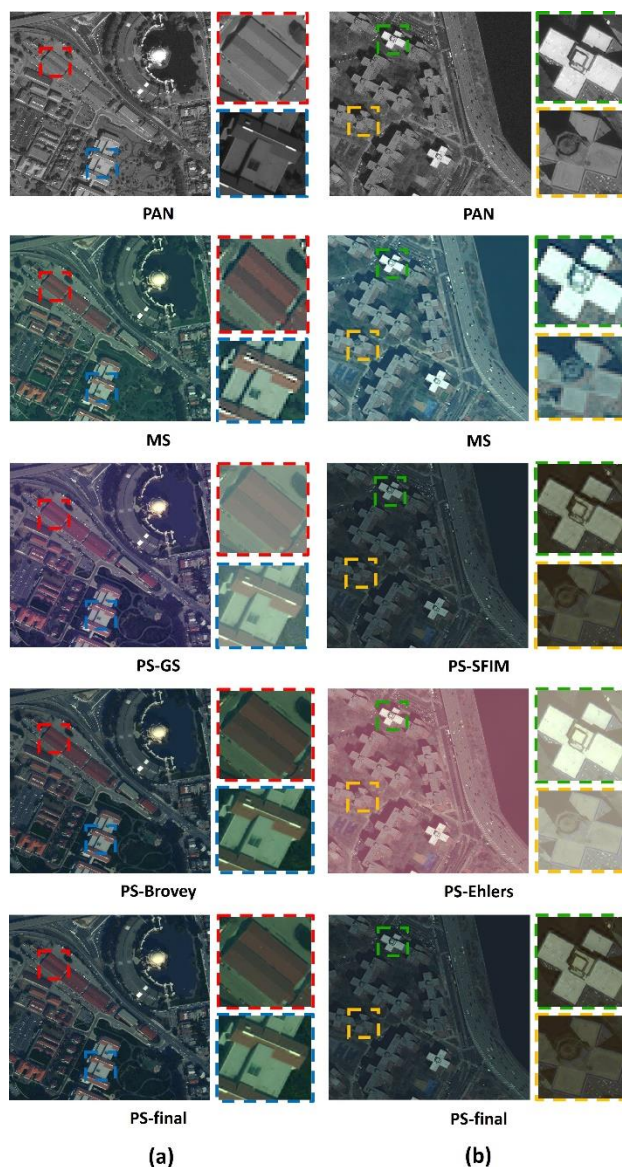

Figure 12. Subjective evaluation of final pan-sharpened images obtained by the decision-fusion method: (a) San Francisco dataset, (b) NY dataset

considering the results of Table 3 which declare SAM_{max}, SAM_{min}, SCC_{max}, and SCC_{min}, and also considering that the SAM ideal value is zero and the SCC ideal value is +1, we propose some indices as Eq. 15-17 which calculate the overall (pure) amount of spectral and spatial quality improvement of the final pan-sharpened imaged compared to the initial pan-sharpened images in percent (%) unit. The results indicate that for the San Fransico case the overall spectral-spatial quality of final pan-sharpened images improved by about 7.2 % compared to the initial Brovey and GS images. Also, for the NY case study, the overall spectral-spatial quality of the final decision-based fused image increased by about 9.8% in comparison to its base pan-sharpened images, i.e. Ehlers and SFIM.

$$QIP_{Spc} = \left(\left| \frac{SAM_{max} - \overline{SAM}}{SAM_{max}} \right| \times 100 \right) - \left(\left| \frac{SAM_{min} - \overline{SAM}}{SAM_{min}} \right| \times 100 \right) \quad (15)$$

$$QIP_{Spt} = \left(\left| \frac{\overline{SCC} - SCC_{min}}{SCC_{min}} \right| \times 100 \right) - \left(\left| \frac{SCC_{max} - \overline{SCC}}{SCC_{max}} \right| \times 100 \right) \quad (16)$$

$$OQIP = 0.5 \times (QIP_{Spc} + QIP_{Spt}) \quad (17)$$

Where QIP_{Spc}, QIP_{Spt}, and $OQIP$ stands for the improvement of spectral quality, the improvement of spatial quality, and the overall improvement of joint spectral-spatial quality (all in percent unit) of the final pan-sharpened image in comparison to the initial pan-sharpened images, respectively. the tilts symbol (~) denotes the metric value for the final pan-sharpened image while the *min* and *max* are referred to as the minimum and maximum value of SAM or SCC for the primary pan-sharpened images, respectively.

Since the performance of no two PS methods is the same, therefore our proposed method is applicable for fusing them at a decision level. The significant role of our proposed decision-fusion method is grasped when the two base PS methods have completely different performances in terms of preserving spectral information of the MS image and transmitting spatial information of the PAN band to the MS image. For better understanding, we consider the results of our meta-analysis on the performance of a variety of state-of-the-art PS methods (Javan et al., 2021). In the mentioned study a scatter plot has been produced in which the relative strength of different methods is illustrated in Figure 13. Our decision-level fusion method is so effective for the fusion of two pan-sharpened images that are located on the diagonal of this scatter plot (shown in yellow directions). As far as the methods located in these directions, our proposed method can produce more interesting results, as it can create new pan-sharpened data with both characteristics of the base images.
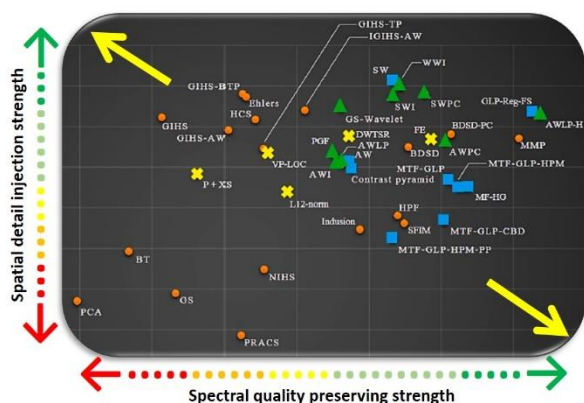

Figure 13. Comparison between a variety of PS methods (Javan et al., 2021)

4. CONCLUSIONS

In the current study, we proposed a new decision-based fusion method for the integration of two pan-sharpened VHR satellite imageries. The developed method is based on the extraction of significant spatial features using edge-preserving rolling self-guidance filtering and the edge information of images.

The performance of the proposed algorithm was tested using four state-of-the-art PS methods and two VHR satellite images. The results indicate that this decision-level fusion of pan-sharpened images can take the advantage of the base pan-sharpened images. It can effectively make a trade-off between the spectral and spatial quality of the base pan-sharpened images, i.e. it tries to reach the overall spectral-spatial quality of the final pan-sharpened image somewhere between the least and the most prior quality of the initial pan-sharpened images.

Future research can implement a decision-based fusion merging our method with the results of the object-level analysis of image spectral quality to propose a method that can decide

both based on the spatial fidelity and the local spectral quality of the image.

The result of our method can effectively be used in basic remote sensing research as it can create new data that could not be produced by any single PS method. At the end of this paper, we hope that the proposed method will open a new horizon in remote sensing image fusion.

REFERENCES

Alparone, L., Wald, L., Chanussot, J., Thomas, C., Gamba, P., & Bruce, L. M., 2007. Comparison of pansharpening algorithms: Outcome of the 2006 GRS-S data-fusion contest. *IEEE Transactions on Geoscience and Remote Sensing*, *45*(10), 3012-3021.

Belgiu, M., & Stein, A., 2019. Spatiotemporal image fusion in remote sensing. *Remote sensing*, *11*(7), 818.

Canny, J., 1986. A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, (6), 679-698.

Gao, Y., Zhang, Y., Li, X., Li, L., Gao, A., & Yu, P., 2016, December. A review of guided filter based pansharpening methods. In *4th International Conference on Information Systems and Computing Technology* (pp. 44-49). Atlantis Press.

Gaur, M. K., Moharana, P. C., Pandey, C. B., Chouhan, J. S., & Goyal, P., 2015. High resolution satellite data for land use/land cover mapping-a case study of Bilara Tehsil, Jodhpur district. *Annals of Arid Zone*, *54*(3&4), 125-132.

Ghassemian, H., 2016. A review of remote sensing image fusion methods. *Information Fusion*, *32*, 75-89.

Gillespie, A. R., Kahle, A. B., & Walker, R. E., 1987. Color enhancement of highly correlated images. II. Channel ratio and "chromaticity" transformation techniques. *Remote Sensing of Environment*, *22*(3), 343-365.

He, K., Sun, J., & Tang, X., 2012. Guided image filtering. *IEEE transactions on pattern analysis and machine intelligence*, *35*(6), 1397-1409.

Javan, F. D., Samadzadegan, F., Mehravar, S., & Toosi, A., 2019. A review on spatial quality assessment methods for evaluation of pan-sharpened satellite imagery. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *42*, 255-261.

Javan, F. D., Samadzadegan, F., Mehravar, S., Toosi, A., Khatami, R., & Stein, A., 2021. A review of image fusion techniques for pan-sharpening of high-resolution satellite imagery. *ISPRS journal of photogrammetry and remote sensing*, *171*, 101-117.

Klonus, S., & Ehlers, M., 2009, July. Performance of evaluation methods in image fusion. In *2009 12th International Conference on Information Fusion* (pp. 1409-1416). IEEE.

Laben, C. A., & Brower, B. V., 2000. *U.S. Patent No. 6,011,875*. Washington, DC: U.S. Patent and Trademark Office.

Liu, J. G., & Mason, P. J., 2016. *Image processing and GIS for remote sensing: techniques and applications*. John Wiley & Sons.

Liu, J. G., 2000. Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details. *International Journal of Remote Sensing*, *21*(18), 3461-3472.

Mehravar, S., Dadrass Javan, F., Samadzadegan, F., Toosi, A., Moghimi, A., Khatami, R., & Stein, A., 2022. Varying weighted spatial quality assessment for high resolution satellite image pan-sharpening. *International journal of image and data fusion*, *13*(1), 44-70.

Meng, X., Li, J., Shen, H., Zhang, L., & Zhang, H., 2016. Pansharpening with a guided filter based on three-layer decomposition. *Sensors*, *16*(7), 1068.

Otsu, N., 1979. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, *9*(1), 62-66.

Peijuan, W. A. N. G., ALGANCİ, U., & SERTEL, E., 2021. Comparative analysis on deep learning based pan-sharpening of very high-resolution satellite images. *International Journal of Environment and Geoinformatics*, *8*(2), 150-165.

Pushparaj, J., & Hegde, A. V., 2017. Evaluation of pan-sharpening methods for spatial and spectral quality. *Applied Geomatics*, *9*(1), 1-12.

Rodríguez-Esparragón, D., Marcello, J., Eugenio, F., García-Pedrero, A., & Gonzalo-Martín, C., 2017. Object-based quality evaluation procedure for fused remote sensing imagery. *Neurocomputing*, *255*, 40-51.

Samadzadegan, F., & Javan, F. D., 2011. New object level strategy for image fusion quality assessment of high resolution satellite imagery. *Photogrammetrie-Fernerkundung-Geoinformation*, *2011*(3), 123-133.

Snehmani, Gore, A., Ganju, A., Kumar, S., Srivastava, P. K., & RP, H. R., 2017. A comparative analysis of pansharpening techniques on QuickBird and WorldView-3 images. *Geocarto International*, *32*(11), 1268-1284.

Toosi, A., Dadrasjavan, F., & Samadzadegan, F., 2020a. Object Level Strategy for Spectral Quality Assessment of High Resolution Pan-sharpen Images. *Journal of Geomatics Science and Technology*, *10*(1), 97-110.

Toosi, A., Javan, F. D., Samadzadegan, F., & Mehravar, S., 2020b. Object-based spectral quality assessment of high-resolution pan-sharpened satellite imageries: new combined fusion strategy to increase the spectral quality. *Arabian Journal of Geosciences*, *13*(13), 1-17.

Wald, L., Ranchin, T., & Mangolini, M., 1997. Fusion of satellite images of different spatial resolutions: Assessing the quality of resulting images. *Photogrammetric engineering and remote sensing*, *63*(6), 691-699.

Zhang, Q., Shen, X., Xu, L., & Jia, J., 2014, September. Rolling guidance filter. In *European conference on computer vision* (pp. 815-830). Springer, Cham.