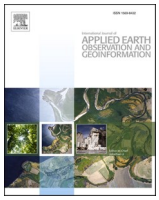
International Journal of Applied Earth Observations and Geoinformation 112 (2022) 102868

Contents lists available at ScienceDirect

International Journal of Applied Earth Observations and Geoinformation

journal homepage: www.elsevier.com/locate/jag



Quantifying the sensitivity of SAR and optical images three-level fusions in land cover classification to registration errors

Wenfu Wu [a], Zhenfeng Shao [a,b,c,*], Xiao Huang [d], Jiahua Teng [e], Songjing Guo [f], Deren Li [a,b]

[a] School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China
[b] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China
[c] Hubei Luojia Laboratory, Wuhan 430079, China
[d] Department of Geosciences, University of Arkansas, Fayetteville, AR 72701, USA
[e] Satellite Environment Protetion Key Laboratory of Satellite Remote Sensing, Satellite Application Center for Ecology and Environment, Ministry of Ecology and Environment, Beijing 100094, China
[f] Institute of Geophysics and Geomatics, China University of Geosciences (Wuhan), Wuhan 430074, China

ARTICLE INFO

Keywords:
SAR and optical (SAR-optical) images fusion
Registration errors
Sensitivity quantification
Land cover classification (LCC)

ABSTRACT

Synthetic aperture radar (SAR) and optical (SAR-optical) data are two important remote sensing data sources. Fusing them is expected to lead to complementary information that benefits land cover classification (LCC). The fusion of SAR-optical images occurs in three levels, i.e., pixel-level, feature-level, and decision-level. However, accurately registering SAR-optical images is still a challenge, and geometric registration errors will bring great uncertainty to LCC based on SAR-optical images fusion. Therefore, this study quantitatively evaluates the sensitivity of SAR-optical images three-level fusions to registration errors through simulation experiments. The results show that: (1) geometric registration errors affect LCC based on SAR-optical images fusion at three levels in a significant manner. Among them, feature-level fusion is the least sensitive to registration errors. (2) the fusion taken optical image as reference image presents better tolerance to registration errors than that based on SAR image. (3) the response of SAR-optical images fusion to registration errors in LCC of heterogeneous regions is greater than that in homogeneous regions. (4) during the LCC, the fusion of SAR-optical images with comparable spatial resolution has a higher tolerance of registration errors. (5) fusing SAR-optical data does not always guarantee the improvement of LCC compared to using optical data alone, depending on fusion levels, fusion methods and classifiers. We believe these findings can greatly benefit the research community in developing new SAR-optical images fusion methods to improve LCC in the future.

1. Introduction

Land cover classification (LCC) is of great significance to global climate warming, ecological environment and urban sprawl (Zhou et al., 2021; Shao et al., 2019; Shao et al., 2021). However, even today, classifying land cover accurately remains to be challenging, as land cover changes rapidly with complex trajectories under the influence of both human activities and natural processes (Joshi et al., 2016). Fortunately, with the development of remote sensing technology, richer data are becoming available for LCC. Fusing multi-source data can obtain complementary information of ground objects and improve LCC.

Synthetic aperture radar (SAR), as an active sensor, has cloud penetration capability and is able to obtain data all day and all weather due to its long wavelength. SAR captures the backscattering energy of ground objects, which is related to surface roughness, complex dielectric constant and moisture of ground objects (Moreira et al., 2013). Optical sensors receive reflections of sunlight from ground objects, usually covering the spectrum from ultraviolet to infrared, and belong to passive sensors. Although SAR sensors can obtain data around the clock and contain rich spatial information, they do not contain spectral information and suffer from speckle (Parrilli et al., 2011). Optical images contain rich spectral information, but the phenomenon of "same spectrum in different objects, same objects in different spectrum" always exists. Besides, optical images are easily disturbed by weather conditions such as clouds and fog. Therefore, fusing SAR and optical (SAR-optical) images can achieve information complementarity well. Numerous

* Corresponding author at: School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China.
E-mail addresses: wuwwf09140818@whu.edu.cn (W. Wu), shaozhenfeng@whu.edu.cn (Z. Shao), xh010@uark.edu (X. Huang), tengjh@secmep.cn (J. Teng), guosongjing@cug.edu.cn (S. Guo), drli@whu.edu.cn (D. Li).

https://doi.org/10.1016/j.jag.2022.102868
Received 4 March 2022; Received in revised form 3 June 2022; Accepted 10 June 2022
Available online 29 June 2022

efforts have been made to explore the potential of fusing SAR-optical data in LCC (Sukawattanavijit et al., 2017; Bai et al., 2021; Werner et al., 2014; Shao et al., 2016; Zhang et al., 2014; Lin et al., 2020).

In general, the SAR-optical images fusion occurs on three levels: decision-level, feature-level, and pixel-level. Decision-level fusion means that source images are first processed and classified respectively, and then the final LCC result is obtained by fusing the initial LCC results through decision rules. Some decision-level fusion methods, such as majority voting and the Dempster-Shafer (DS) theory (Zhang and Xu, 2018; Shao et al., 2016), have been proposed. Feature-level fusion refers to the process of extracting texture, shape and other features from source images first, then fusing them into new features through ensemble-learning or layer-stacking for LCC (Lin et al., 2020). Different from the above two fusion mechanisms, pixel-level fusion directly acts on the pixel value of source images to synthesize an image with richer information. According to a survey by Kulkarni and Rege (2020), pixel-level fusion methods include four categories. (1) Component substitution methods transform multispectral images to another spaces, and then a structural component of multispectral images is replaced by SAR images (Pal et al., 2007; Wang et al., 2005; Chen et al., 2010). (2) Multi-scale decomposition methods firstly decompose the spatial component of source images into multiple sub-bands using Wavelet transform, Shearlet transform and Contourlet transform, then fuse these sub-bands using fusion rules, and finally obtain the fused image through reconstruction (Garzelli, 2002; Wu et al., 2020; Nencini et al., 2007). (3) Hybrid methods are to combine the advantages of multiple methods to achieve better fusion images (Chen et al., 2010; Cheng et al., 2015). (4) Model-based methods can be divided into sparse representation based and variational based methods. Variational-based methods regard images fusion problem as an energy optimization task (Shao et al., 2020) while sparse representation-based methods represent source images sparsely to obtain sparse coefficients, and then these sparse coefficients are fused and reconstructed to obtain fused images (Huang et al., 2015). In general, hybrid methods are the most popular for SAR-optical images pixel-level fusion, due to their low computational complexity and satisfactory performance.

Although the processes and principles of SAR-optical images fusion differ at different levels, they share the common purpose of improving the accuracy of LCC. The optimal level of SAR-optical images fusion for LCC has been explored in some studies (Zhang et al., 2012; Zhang and Xu, 2018). However, these studies ignored the uncertainty caused by geometric registration errors. Although efforts have been made to register images from different sensors, registration remains to be challenging for heterogeneous sensors such as SAR and optical. We argue that quantifying the sensitivity of LCC based on SAR-optical images fusion at three levels to registration errors is of great significance, as it can provide valuable guidance for users to select appropriate fusion levels and methods given their own application requirements.

In this study, we quantify and compare the sensitivity of LCC based on SAR-optical images three-level fusions to geometric registration errors through simulation experiments. We select Intensity-Hue-Saturation (IHS) (Chen et al., 2010), discrete wavelet transforms (DWT) (Kulkarni and Rege, 2020), and IHS and gradient transfer fusion method (IHS-GTF) (Shao et al., 2020) as pixel-level fusion methods, and use their fusion results for LCC. As for feature-level fusion, we extract spectral features and texture features of SAR-optical images respectively to generate new features for LCC. For decision-level fusion, the initial LCC results of SAR-optical images are obtained respectively, and then the final LCC results are derived using DS theory. The choices of fusion methods evaluated in this study are made after careful consideration of their universality and performance.

2. Methods and datasets

2.1. Experimental datasets

In this study, we chose two scenes in Beijing and Wuhan, China, as our study areas (as shown in Fig. 1) (both with 970 × 970 pixels). The optical data used is from Sentinel-2A's Blue, Green and Red bands with a spatial resolution of 10 m, while the SAR data is from TerraSAR. The TerraSAR data used is HH polarized single-look complex (SLC) data with a spatial resolution of 3 m. The preprocessing steps including SLC to Ground Range Detected (GRD) conversion, speckle filtering, radiometric correction, and terrain correction are conducted using Sentinel Application Platform (SNAP) software. Detailed attributes of SAR-optical images used are shown in Table 1. The acquisition dates are slightly inconsistent between SAR-optical images, but this does not affect the validity of this study. Through careful visual interpretation of high-resolution Google Earth images, it is found that land cover has no obvious changes in both study areas.

2.2. Experiment design

In this study, the co-registered SAR-optical images with Root Mean Square Error (RMSE) less than one pixel will be used as reference images (without geometric registration errors) for subsequent geometric registration errors simulation experiments. To co-register SAR-optical images, we manually choose the intersection points of roads and corner points of buildings that can be well distinguished in the SAR-optical images as ground control points (GCPs) through visual interpretation, and repeatedly adjust the position of GCPs to make the RMSE between GCPs less than one pixel. Nearly 30 pairs of GCPs are finally selected. The RMSE of Wuhan and Beijing are 0.6166 and 0.6452 pixels, respectively. Then a third-order polynomial method is used to register SAR-optical images and the co-registered SAR-optical images have the same resolution of 3 m. Pixel shifting is one of the most serious consequences of geometric distortion. Referring to previous studies (Zhou et al., 2021; Tang et al., 2020), in this study, geometric registration errors are simulated by shifting from 1 to 15 pixels along the four directions: i.e., up, down, left, and right, respectively, and the size of the error is quantitatively measured by the shifting pixel distance. Then, the SAR-optical images with different registration errors are fused at three levels, and then the effects of registration errors on LCC based on SAR-optical images three-level fusion are discussed. Fig. 2 shows the overall experimental design of this study.

2.2.1. Pixel-level fusion experiment design

Three typical pixel-level fusion methods, including IHS transform method (Chen et al., 2010), DWT method (Kulkarni and Rege, 2020), and IHS-GTF method (Shao et al., 2020), are selected for comparative experiments.

IHS fusion method: It belongs to the category of component substitution. First, the multispectral image in RGB space is transformed into separated IHS space, and then the I component is replaced by SAR image, and the fused image is obtained by IHS inverse transformation.

DWT fusion method: It is a hybrid method that combines IHS transform and discrete wavelet transform. First, IHS transform on the multispectral image is performed, followed by multi-scale decomposition on the I component and SAR image, respectively, via discrete wavelet transform. Finally, the decomposed sub-bands are fused, and then the fused image is obtained by using inverse wavelet transform and inverse IHS transform.

In this study, the approximate sub-band and detail sub-bands of the fused image are the average value of the approximate sub-band of source images, and the detail sub-bands of source images corresponding to the maximum absolute value, respectively, whose mathematical expression is as follows:

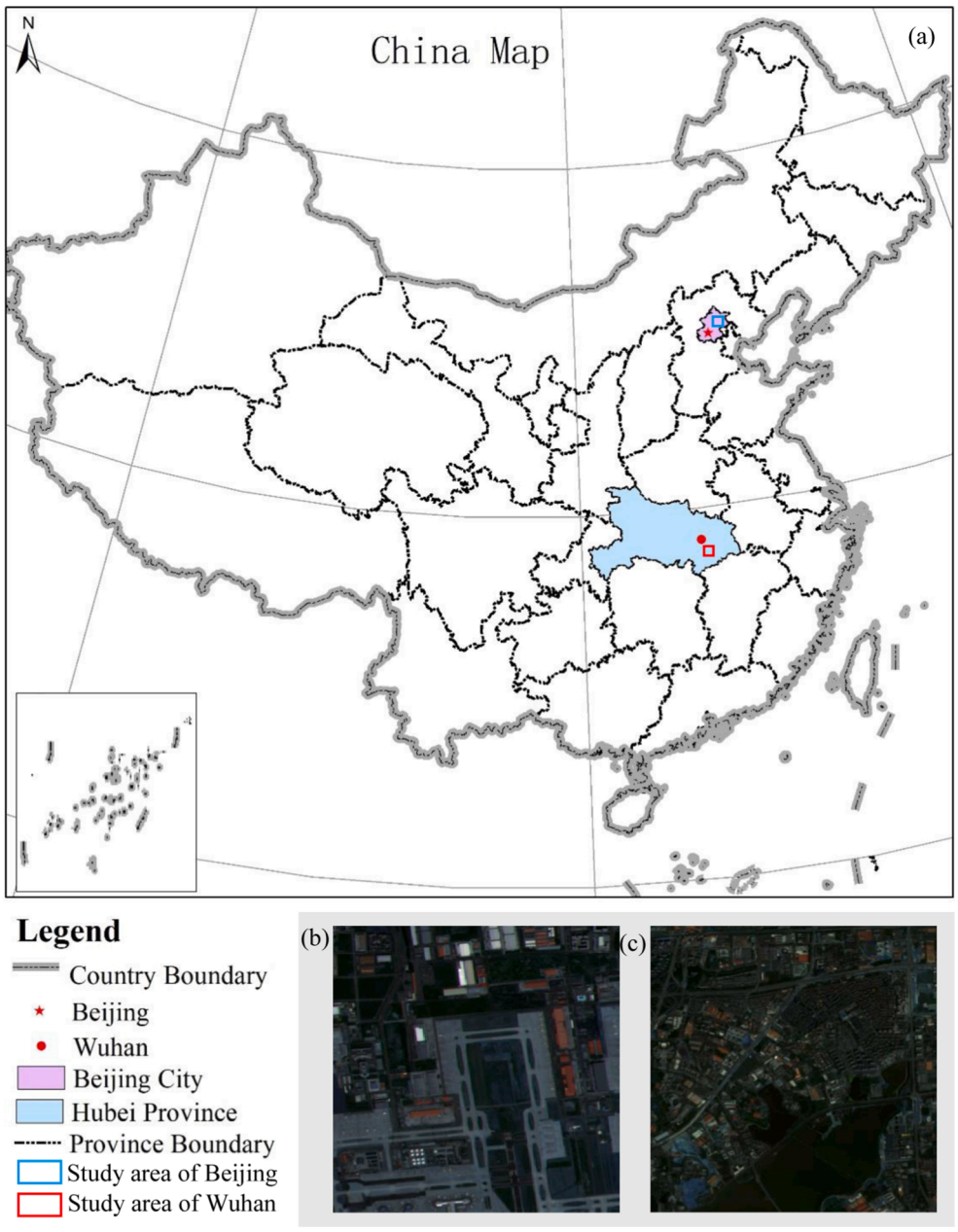


Fig. 1. (a) The geographical location of the study areas. (b-c) are Sentinel-2A pseudo-color images in Beijing and Wuhan study areas, respectively.

Table 1
Attributes of SAR-optical data used in this study.

Parameters	TerraSAR	Sentinel-2A
Product type	SLC	Level-1C
Polarization	HH	—
Mode	StripMap	—
Band	X band (9.6 GHz)	Band 2: 496.6 ± 98 nm Band 3: 560.0 ± 46 nm Band 4: 664.5 ± 39 nm
Spatial resolution	3 m	10 m
Acquisition date	February 19, 2014 (Beijing) August 11, 2015 (Wuhan)	August 27, 2015 (Beijing) August 18, 2016 (Wuhan)

$$A_{f_{x,y}} = \frac{A_{s_{x,y}} + A_{m_{x,y}}}{2}$$
$$D_{f^j_{x,y}} = D_{s^j_{x,y}}, if \left|D_{s^j_{x,y}}\right| \rangle \left|D_{m^j_{x,y}}\right|, else D_{f^j_{x,y}} = D_{m^j_{x,y}} \quad (1)$$

where $A_{f_{x,y}}, A_{s_{x,y}}$ and $A_{m_{x,y}}$ represent the approximate sub-band of the fused, SAR, and multispectral images, respectively. $D_{f^j_{x,y}}, D_{m^j_{x,y}}$ and $D_{s^j_{x,y}}$ are the detail sub-bands of the fused, SAR, and multispectral images at jth level of decomposition, respectively. (x, y) denotes a pixel positioned with x and y in the image. The level of decomposition, an important hyperparameter in the DWT fusion method, is empirically set to 4 in this study.

IHS-GTF fusion method: It is a variational-based method, which transforms SAR-optical images fusion into an optimization task and improves the problem of spectral distortion. It believes that spectral information of the fused image comes from optical image, and the spatial information should be provided by SAR-optical images together.

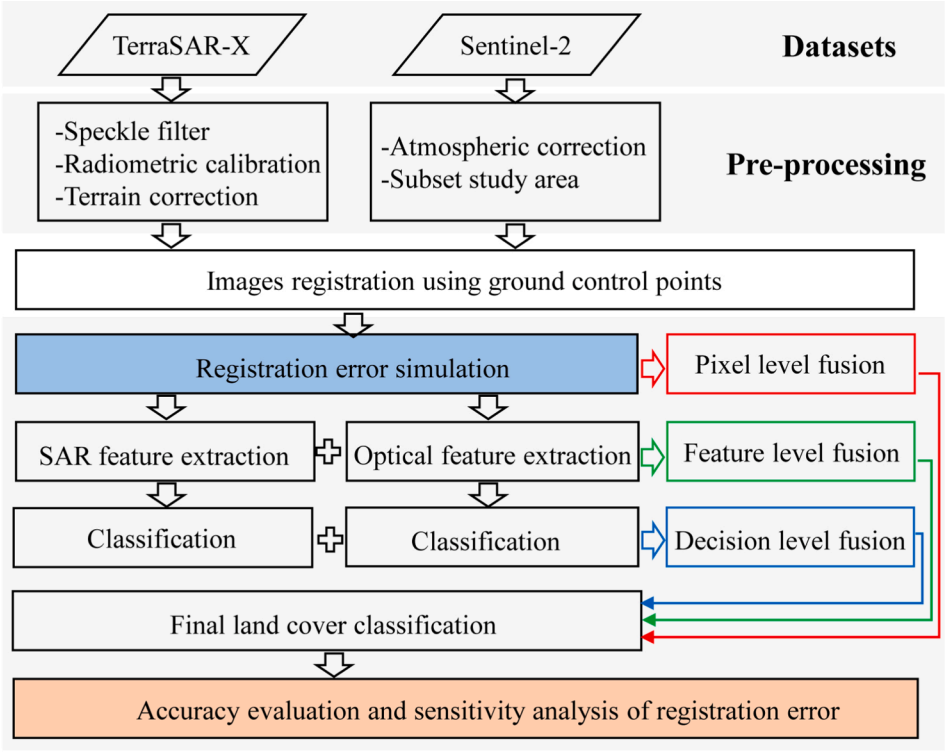


Fig. 2. The overall experimental design.

Therefore, the IHS-GTF method firstly fuses the SAR spatial information with the I component of the optical image through pixel saliency, and then generates the fused I component based on the total variational model. The fused image is finally obtained through inverse IHS transformation. Detailed procedures for the IHS-GTF algorithm are available in the literature (Shao et al., 2020).

To quantitatively evaluate pixel-level fusion results, three commonly used evaluation indices including structural similarity index measure (SSIM), spectral angle mapper (SAM) and peak signal to noise ratio (PSNR) are introduced in this study. SSIM and SAM measures the structural and spectral similarity between two images, respectively, with a value ranging from 0 to 1. PSNR evaluates images quality, and the larger the value, the smaller the distortion. Considering the goals of this study, we also conduct LCC on the fused results of IHS, DWT, and IHS-GTF.

2.2.2. Feature-level fusion experiment design

In feature-level fusion, features need to be extracted from SAR-optical images respectively. In this study, we only select the most typical and commonly used features. Studies have shown that texture features contain valuable information in SAR-optical images, which are beneficial to LCC (Zhang and Xu, 2018; Lin et al., 2020). Due to the study areas are typical urban scenes, four texture features including homogeneity, dissimilarity, entropy, and angular second moment suitable to urban landscapes are extracted from SAR-optical images respectively (Puissant et al., 2005). The size of the sliding window is experimentally set to 7 × 7. Besides, the band information of source images are also added as spectral features. Therefore, the final features number extracted from optical images is 15 (3 spectral features + 4 texture features × 3 bands), that from SAR image is 5 (1 spectral feature + 4 texture features × 1 band). Thus, the total feature number after fusion is 20.

2.2.3. Decision-level fusion experiment design

In decision-level fusion, we fuse the initial classification results of SAR-optical images based on DS theory to obtain the final classification results. During the initial classification, the features used are the spectral and texture features (the same as those used in feature-level fusion). DS theory, first proposed by Dempster in 1967 (Dempster, 2008) and further developed by Shafer (Shafer, 1976), is an uncertain inference method that consists of three main steps: identification framework construction, basic probability assignment, and DS composition. First, a recognition framework Θ is constructed for all possible results of the decision problem. Further, the basic probability distribution function named mass function assigns a probability to each hypothesis in the recognition framework, indicating the credibility of the hypothesis. In LCC, the recognition framework corresponds to land cover types, and the basic probability is obtained through the classifier. The mass function should meet the following requirements:

$$(1)m(\varnothing) = 0, (2)\sum_{A\subset\Theta} m(A) = 1 \tag{2}$$

where $m(\bullet)$ is the mass function and $\varnothing$ is the empty set.

For multi-source data, there exist several different basic probability distribution functions that need to be synthesized into a basic probability distribution by the DS composition rule. The DS composition rule between two-source data can be expressed as follows:

$$\begin{aligned} m(A) &= \frac{\sum_{A_i\cap B_i=A} m_1(A_i)m_2(B_i)}{1-K} \\ K &= \sum_{A_i\cap B_i=\varnothing} m_1(A_i)m_2(B_i) \end{aligned} \tag{3}$$

where $m_1(\bullet)$ and $m_2(\bullet)$ are mass functions for two data sources, and $m(\bullet)$ is the synthesized mass function.

To mitigate the conflict between evidence in the DS composition, we use the Murphy composition rule (Murphy, 2000). Murphy composition rule first averages the basic probability values of N data sources and then uses the DS composition rule to synthesize $N-1$ times. The final land cover type is determined as the type that corresponds to the maximum basic probability of synthesis.

2.3. Land cover classification and accuracy assessment

In this study, we select random forest (RF) algorithm as the classifier. RF is an ensemble learning algorithm widely used in remote sensing fields (Belgiu and Drăguţ, 2016). The two hyper-parameters in RF, i.e., the number of subdatasets and the maximum number of features, are set to 20 and 5, respectively and applied to all the cases in this study. 10-fold cross-validation is used in this study.

Referring to the commonly used classification scheme in urban areas, land cover is divided into five types in this study: bright impervious surface, dark impervious surface, bare soil, vegetation, and water. As the selected Beijing area does not contain water, the water category is removed in Beijing area. To train and test RF, we select a set of samples for each of the study areas through visual interpretation (as shown in Table 2). The overall accuracy (OA) and Kappa coefficient are used as indexes to evaluate LCC results.

3. Results

3.1. Pixel-level fusion results with geometric misregistration

3.1.1. Evaluating fusion results qualitatively and quantitatively

Fig. 3 shows the fusion results without geometric registration errors (reference) and with 15 pixels upward shifting. Compared with the reference results, the fusion results of IHS, DWT, and IHS-GTF with 15 pixels shifting have notable spectral distortion. Moreover, we can notice the shifting traces in the fusion results, making the fused images fuzzy. Fig. 4 and Fig. 5 quantitatively present the image fusion effects under different geometric registration errors. With the increase of registration errors, PSNR, SSIM, and SAM of the fusion results gradually become worse, and all fusion results with registration errors are worse than that of the reference results. We notice that when the registration error is less than 8 pixels, the fusion effects decrease rapidly with the increase of geometric registration errors. However, when the registration error is larger than 8 pixels, the fusion effect decreases slowly and tends to be stable with the increase of geometric registration error. Among the three pixel-level fusion methods, IHS has the most robustness to registration errors, but its fusion quality is the worst. Although the other two methods are sensitive to geometric registration errors, their fusion quality is better than IHS (especially the IHS-GTF).

3.1.2. Land cover classification based on pixel-level fusion images

Fig. 6 qualitatively shows the classification results of IHS, DWT, and IHS-GTF fusion images without geometric registration errors (reference) and with 15 pixels upward shifting. Compared with the reference results, the classification results with 15 pixels registration errors notably overestimate impervious surface in Wuhan area, with severe fragmentation issues. Besides, obvious misclassification phenomena can be found, where roads, bridges, and other impervious surfaces, are misclassified into vegetation and bare soils. Similarly, in Beijing area, the classification results with 15 pixels registration errors have obvious salt and pepper effects compared with the reference results. Fig. 7 and Fig. 8 quantitatively compare the LCC results of IHS, DWT, and IHS-GTF with different geometric registration errors from 1 to 15 pixels in four directions in Wuhan and Beijing areas, respectively. The OA and Kappa decrease with the increase of geometric registration errors. Among them, the results of IHS-GTF are the best, followed by DWT and IHS. And we notice that the OA and Kappa from IHS change slowly with the increase of registration errors, which again reflects the robustness of IHS to geometric registration errors. Compared with Wuhan area, the accuracies of the three fusion images change slower with the increase of registration errors in Beijing area. This may be related to the spatial patterns of the study area, which will be discussed in Section 4.1.

Table 2
Samples of each land cover type in Wuhan and Beijing areas.

Land cover types	Samples (pixels)	
	Wuhan	Beijing
Bright impervious surface	579	1301
Dark impervious surface	572	877
Bare soil	427	1010
Vegetation	415	1071
Water	564	—

3.2. Feature-level fusion results with geometric misregistration

Fig. 9 shows the LCC results without geometric registration errors and with 4, 8, 12, and 15 pixels upward shifting using SAR-optical images feature-level fusion in Wuhan and Beijing. When the shifting pixel is 4, we fail to observe notable differences in their classification results compared with the reference results by visual observation alone. However, with the increase of shifting pixels (i.e., to 8, 12, and 15), the differences between the classification results and the reference results start to manifest. Compared with the reference results, misregistration leads to poor integrity of ground objects, the phenomenon of “salt and pepper”, and misclassified ground objects. Fig. 10 quantitatively compares the classification results under different geometric registration errors from 1 to 15 pixels shifting in four directions in Wuhan and Beijing areas. We notice that the accuracies based on the feature-level fusion decrease with the increase of geometric registration errors. However, when the geometric registration error is within four pixels, their classification accuracies decrease slowly, which is consistent with the visual comparison of the classification results in Fig. 9.

3.3. Decision-level fusion results with geometric misregistration

Fig. 11 and Fig. 12 respectively present qualitative and quantitative LCC results of SAR-optical images decision-level fusion without geometric registration errors and with 4, 8, 12, and 15 pixels upward shifting in Wuhan and Beijing. A similar pattern to feature-level fusion can be observed. When the shifting pixel is 4, their classification results present unnoticeable visual differences compared to the reference results, which is confirmed by Fig. 12. With the increase of shifting pixels, the classification results start to deteriorate, with notable poor ground object integrity, ”salt-and-pepper” issues, and ground objects misclassification (Fig. 11). As shown in Fig. 12, similar to the feature-level fusion, the classification accuracies of decision-level fusion decrease as the geometric registration errors increase.

3.4. Sensitivity evaluation of SAR-optical images fusion to geometric registration errors

To quantitatively evaluate the sensitivity of SAR-optical images three-level fusion to geometric registration errors, we measure the sensitivity by differentiating classification accuracies against registration errors (Luo et al., 2021). The larger the absolute value of the derivative, the more sensitive the classification is to geometric registration errors. Table 3 shows the calculated sensitivity results, the lowest sensitivity highlighted in bold. We can find that feature-level fusion has the lowest sensitivity to geometric registration errors, followed by decision-level fusion and pixel-level fusion. Among the three pixel-level fusion methods used in this study, IHS has the lowest sensitivity to geometric registration errors, followed by IHS-GTF and DWT.

4. Discussion

4.1. Influence of spatial patterns

To investigate the effect of spatial patterns on the response of SAR-optical images fusion to geometric registration errors in LCC, Fig. 13 compares the pixel-level fusion results of Wuhan and Beijing under

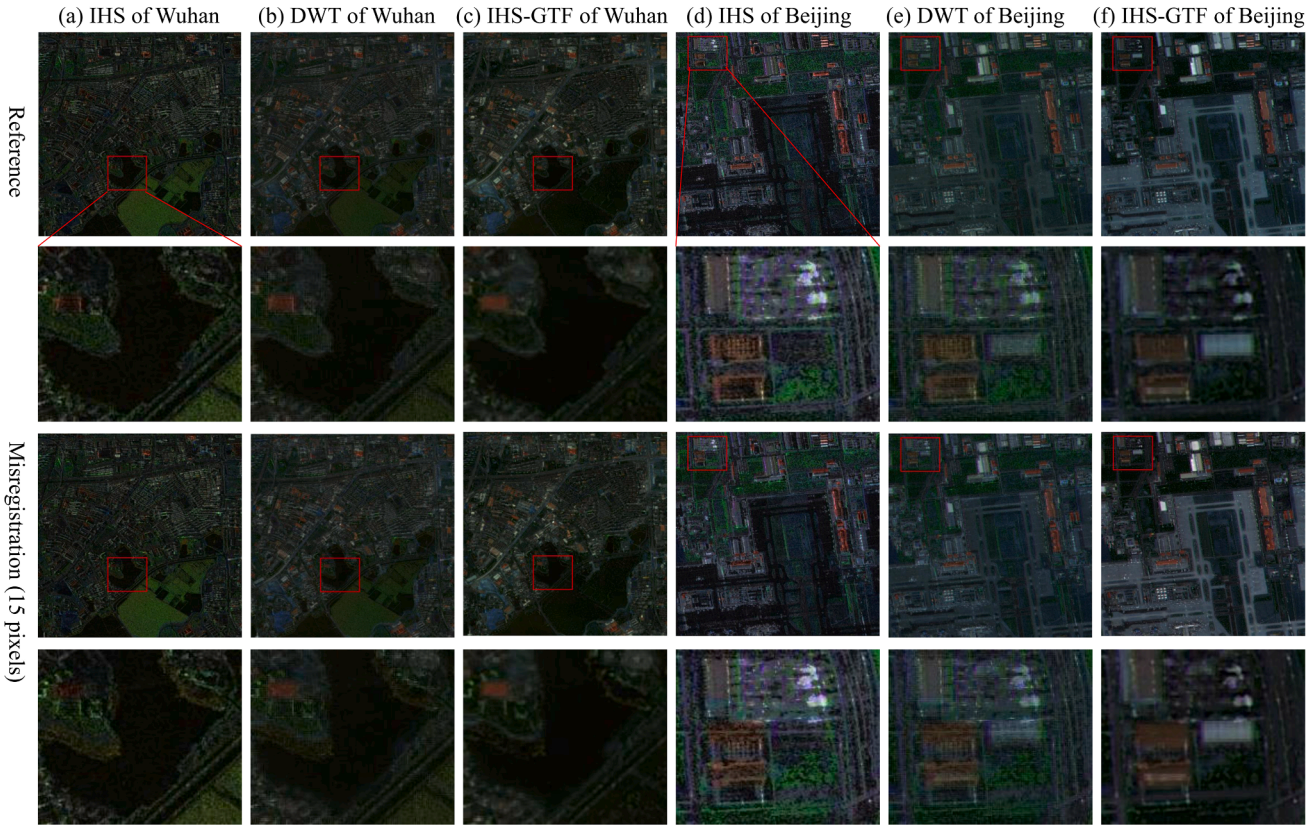


Fig. 3. Visual comparison of pixel-fusion results without geometric registration errors and with misregistration (15 pixels upward shifting): (a-c) are the IHS, DWT and IHS-GTF fusion results in Wuhan, respectively; (d-f) are the IHS, DWT and IHS-GTF fusion results in Beijing, respectively.

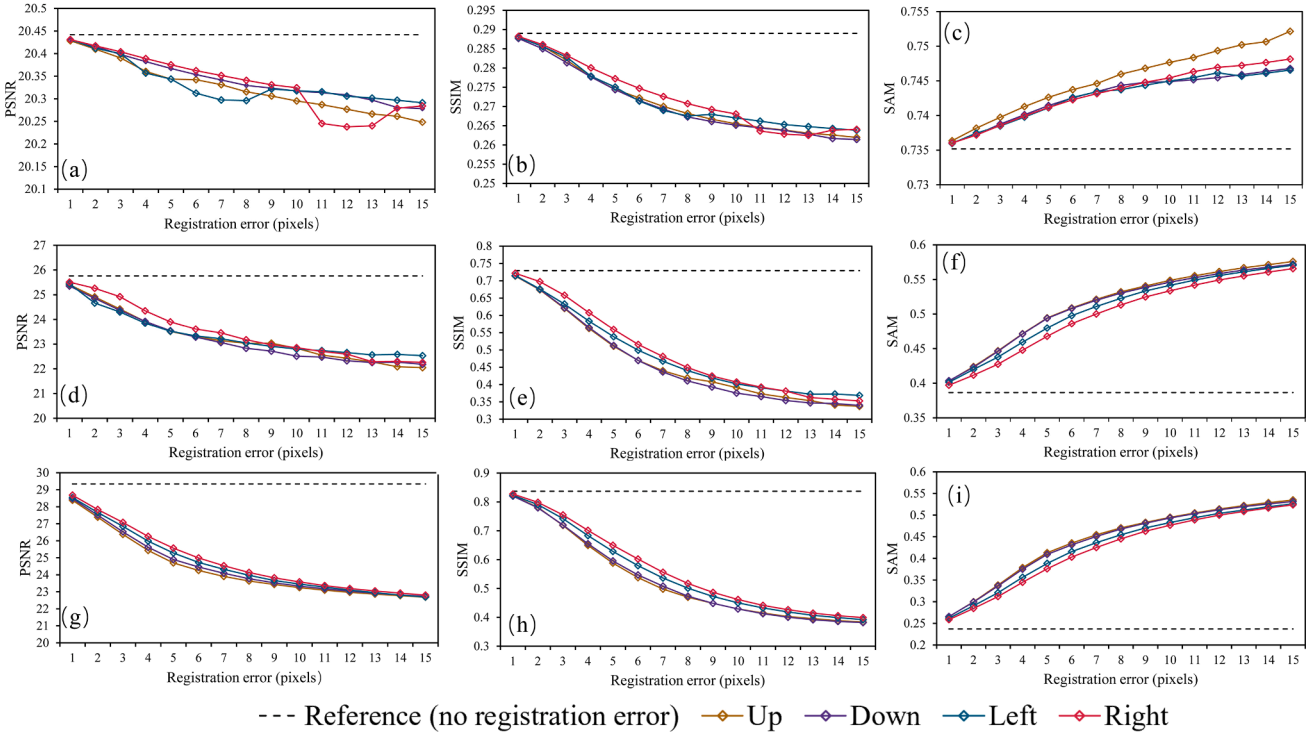


Fig. 4. Quantitative comparison of pixel-fusion results with geometric registration errors from 0 to 15 pixels in Wuhan area: (a-c) PSNR, SSIM and SAM of IHS; (d-f) PSNR, SSIM and SAM of DWT; (g-i) PSNR, SSIM and SAM of IHS-GTF.

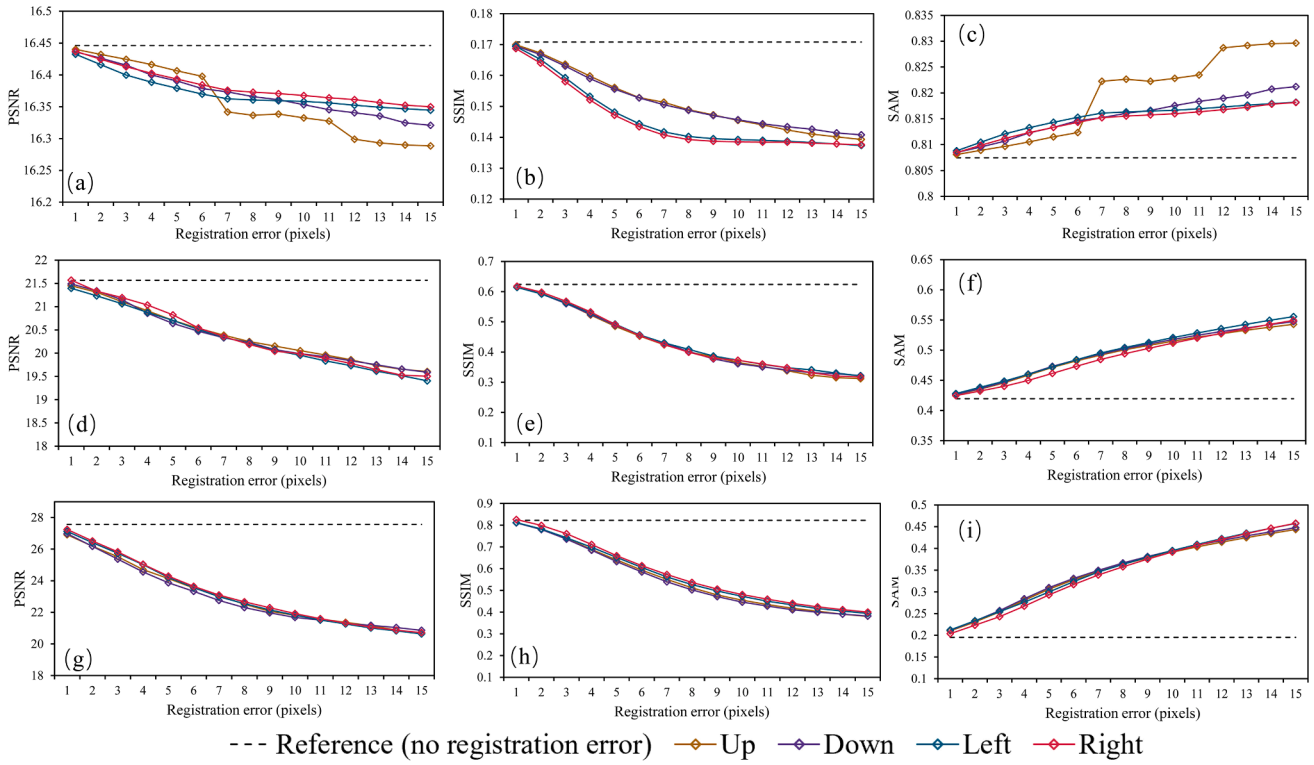


Fig. 5. Quantitative comparison of pixel-fusion results with geometric registration errors from 0 to 15 pixels in Beijing area: (a-c) PSNR, SSIM and SAM of IHS; (d-f) PSNR, SSIM and SAM of DWT; (g-i) PSNR, SSIM and SAM of IHS-GTF.

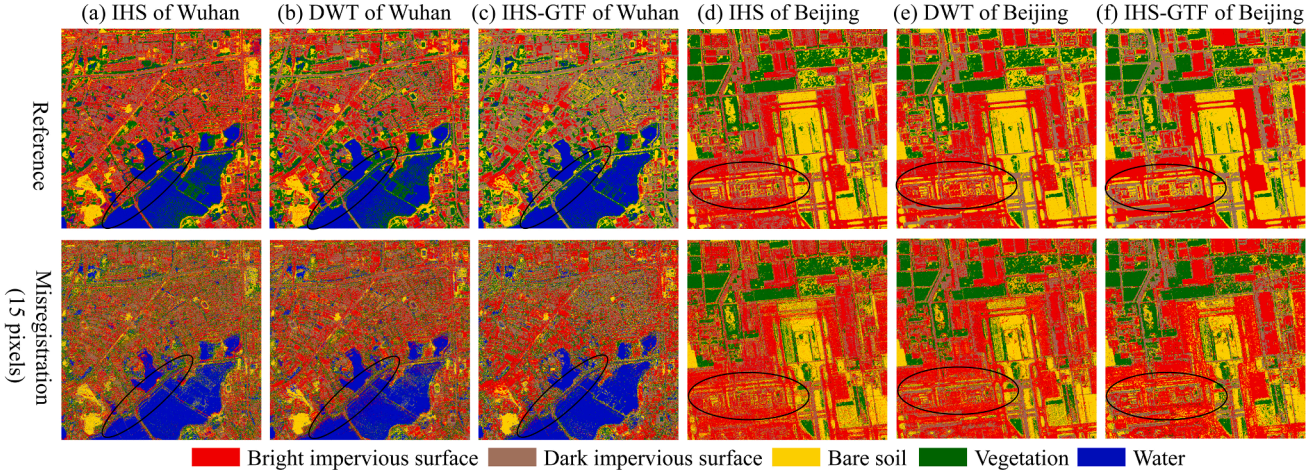


Fig. 6. Visual comparison of the LCC results of SAR-optical images pixel-level fusion without geometric registration errors and with misregistration (15 pixels upward shifting): (a-c) are the IHS, DWT and IHS-GTF LCC results in Wuhan, respectively; (d-f) are the IHS, DWT and IHS-GTF LCC results in Beijing, respectively.

different geometric registration errors (upward shifting) from 0 to 15 pixels. We observe that the IHS, DWT, and IHS-GTF fusion results in Wuhan are better than those in Beijing. To explain this phenomenon, we conduct histogram analysis on SAR-optical images, respectively, as shown in Fig. 14. Compared with Wuhan, greater differences between the histogram of the SAR image and that of the optical image can be found in Beijing, which makes source images distorted to some extent during fusion and affects the quality of fusion results. Besides, from Table 3, the sensitivity of LCC to geometric registration errors in Beijing is lower than that in Wuhan. We further analyze the influence of the spatial patterns on registration errors through via semi-variograms (the greater the semi-variogram, the greater the heterogeneity of spatial patterns). Fig. 15 shows the semi-variogram of the lag from 1 to 100 pixels in Wuhan and Beijing, which can be seen that the semi-variogram in Wuhan is greater than that in Beijing, indicating that Wuhan presents more heterogeneity. Thus, the LCC based on SAR-optical images fusion is more sensitive to geometric registration errors in heterogeneous scenes.

4.2. *Influence of the selection of reference image in pixel shifting*

In the previous experiments, the geometric registration errors are simulated by pixel shifting of the optical image based on SAR image (SAR-based for short). However, does the choice of reference image affect the response of SAR-optical images fusion to geometric registration errors? Therefore, this section takes optical image as the reference

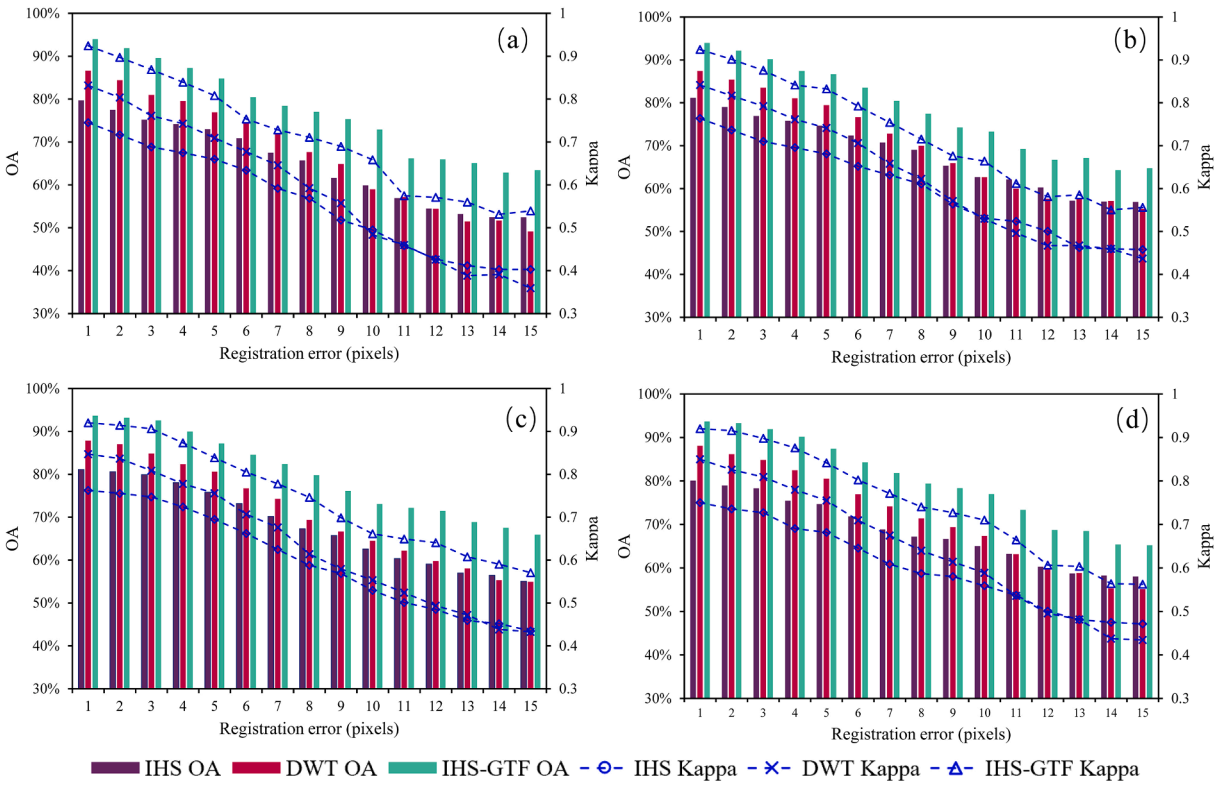


Fig. 7. Quantitative comparison of the IHS, DWT and IHS-GTF LCC results with different geometric registration errors from 1 to 15 pixels in Wuhan area: (a) Upward shifting; (b) Downward shifting; (c) Left shifting; (d) Right shifting.

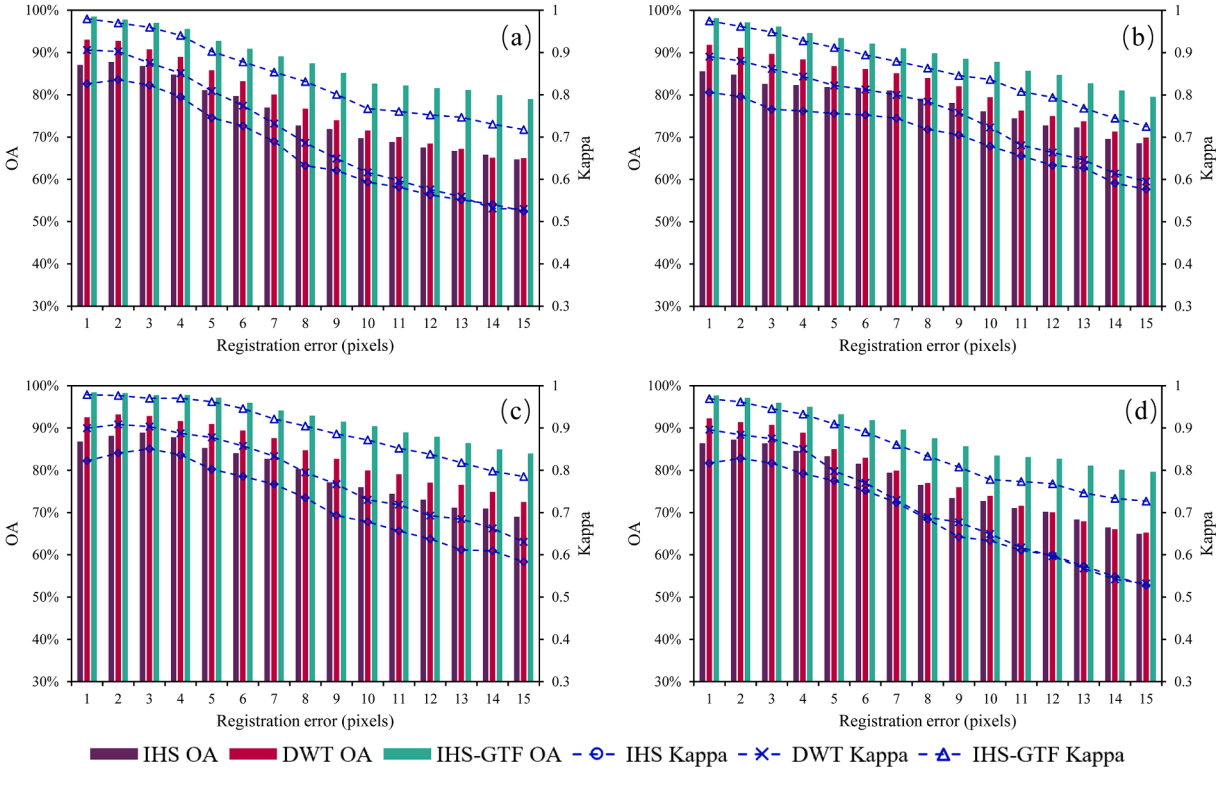


Fig. 8. Quantitative comparison of the IHS, DWT and IHS-GTF LCC results with different geometric registration errors from 1 to 15 pixels in Beijing area: (a) Upward shifting; (b) Downward shifting; (c) Left shifting; (d) Right shifting.

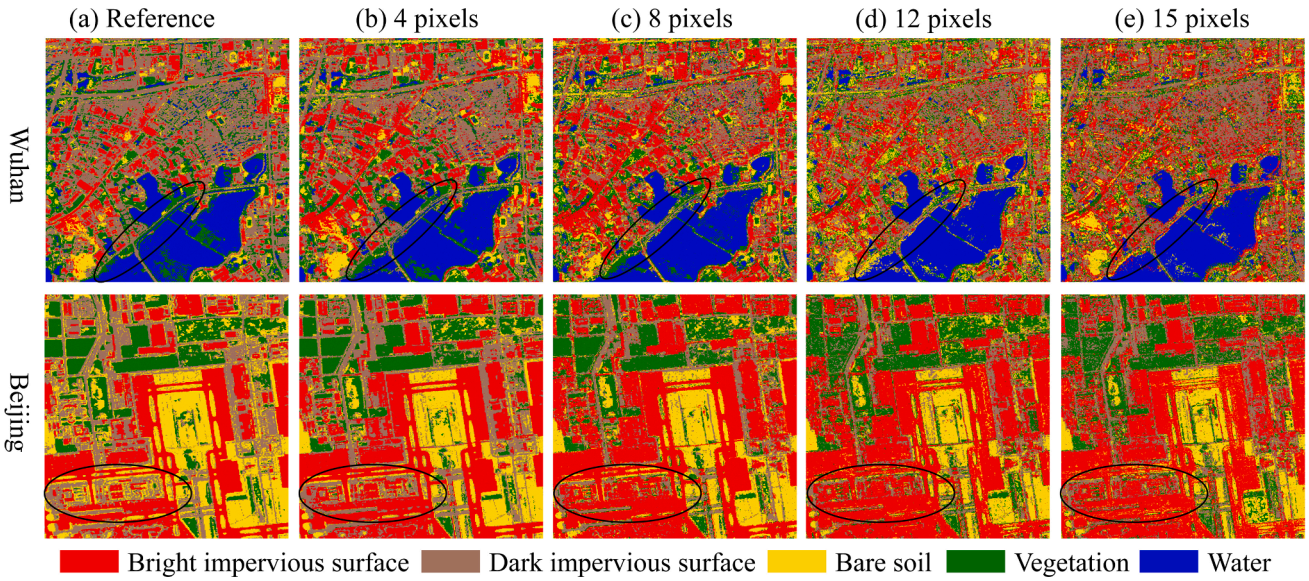


Fig. 9. Visual comparison of the LCC results of SAR-optical images feature-level fusion without geometric registration errors and with misregistration (4, 8, 12 and 15 pixels upward shifting) in Wuhan and Beijing: (a) are the reference results; (b-e) are the classification results with misregistration (4, 8, 12 and 15 pixels upward shifting), respectively.

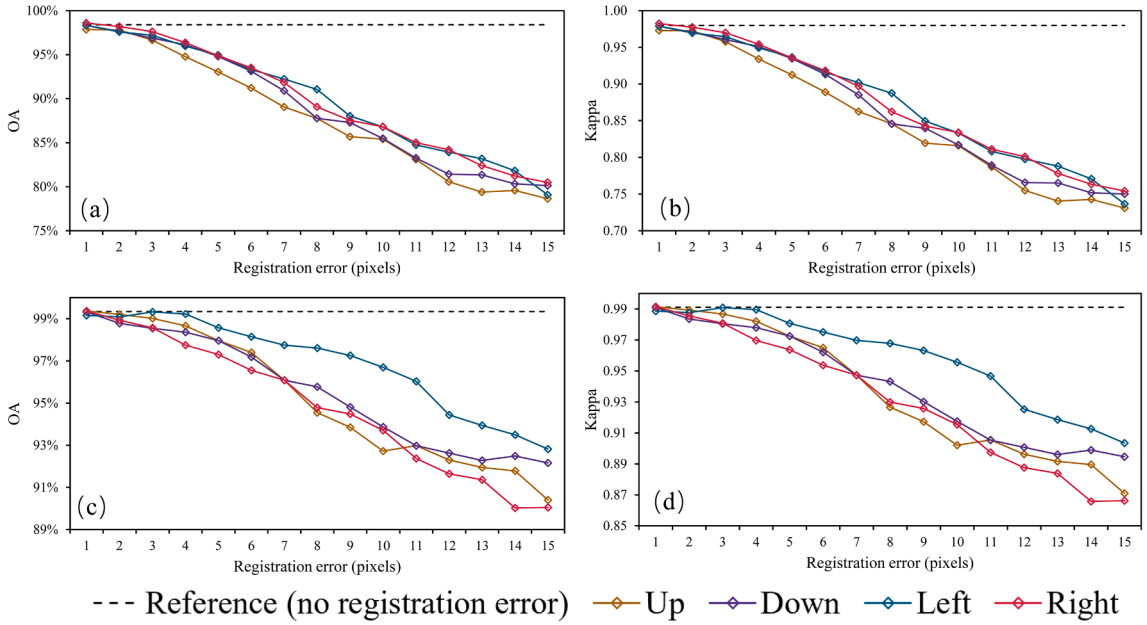


Fig. 10. Quantitative comparison of the LCC results of SAR-optical images feature-level fusion with different geometric registration errors from 0 to 15 pixels. (a-b) are OA and Kappa in Wuhan; (c-d) are OA and Kappa in Beijing.

image with shifting from 1 to 15 pixels in the four directions to simulate geometric registration errors (optical-based for short) and repeat the previous experiments. Fig. 16 shows the comparison of pixel-level fusion results of SAR-based and optical-based. Compared with the fusion results of SAR-based, PSNR, SSIM, and SAM of the fusion results from optical-based (except IHS method) present considerably concentrated distribution. Fig. 17 illustrates the comparison of LCC results of SAR-based and optical-based three-level fusions. The distribution of OA and Kappa of optical-based is more concentrated than that of SAR-based. According to the importance score of the features, in LCC based on SAR-optical images fusion, optical image because of its abundant spectral and spatial features plays a leading role, while SAR image plays an auxiliary role due to its imaging mechanism and speckle noise. Thus the optical-based fusion is less affected by geometric registration errors than SAR-based fusion in LCC.

4.3. *Influence of the spatial resolution*

To explore the influence of the spatial resolution on the response of SAR-optical images fusion to geometric registration errors in LCC, this section selects Sentinel-1 and Sentinel-2 images with similar imaging dates and spatial resolution (both 10 m) for repeated experiments. The polarization of Sentinel-1 is selected the same as TerraSAR used in this study (HH polarization), and the bands used in Sentinel-2 are RGB bands as well. Fig. 18 quantitatively shows the LCC results of Sentinel-1 and Sentinel-2 images three-level fusions under different registration errors. We find that the response of Sentinel-1 and Sentinel-2 three-level fusion to registration errors in LCC is basically the same as that of TerraSAR and

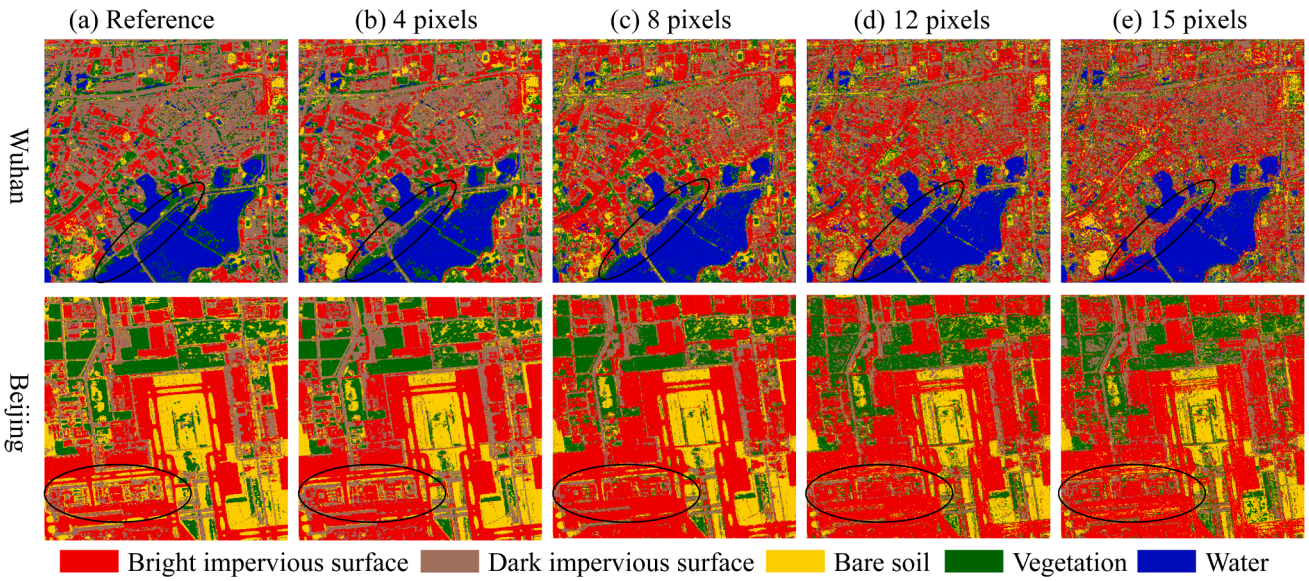


Fig. 11. Visual comparison of the LCC results of SAR-optical images decision-level fusion without geometric registration errors and with misregistration (4, 8, 12 and 15 pixels upward shifting) in Wuhan and Beijing: (a) are the reference results; (b-e) are the classification results with misregistration (4, 8, 12 and 15 pixels upward shifting), respectively.

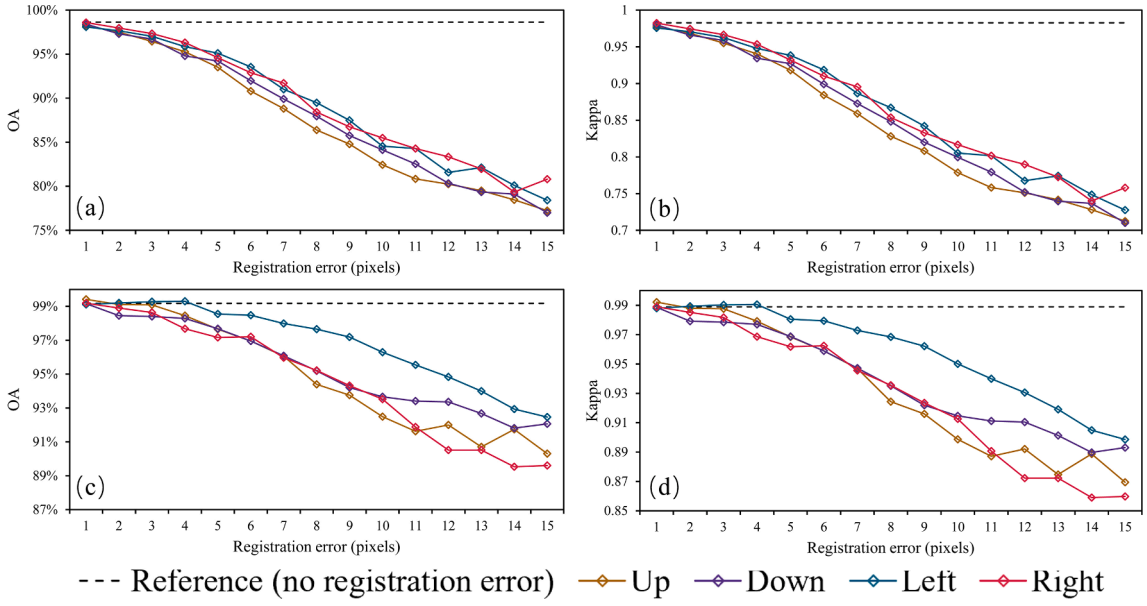


Fig. 12. Quantitative comparison of the LCC results of SAR-optical images decision-level fusion with different geometric registration errors from 0 to 15 pixels: (a-b) are OA and Kappa in Wuhan; (c-d) are OA and Kappa in Beijing.

Table 3
Sensitivity of SAR-optical images fusion to geometric registration errors.

	Wuhan				Beijing			
OA	Up	Down	Left	Right	Up	Down	Left	Right
IHS	0.0215	0.0184	0.0203	0.0173	0.0178	0.0118	0.0142	0.0164
DWT	0.0282	0.0249	0.0253	0.0246	0.0219	0.0157	0.015	0.0205
IHS-GTF	0.0236	0.0227	0.0213	0.0218	0.0149	0.0127	0.0105	0.014
Feature-level fusion	**0.015**	**0.0145**	**0.0135**	**0.013**	**0.0066**	0.0057	**0.0045**	**0.0069**
Decision-level fusion	0.0162	0.0158	0.0147	0.0143	0.007	**0.0055**	0.0047	0.0074
Kappa	Up	Down	Left	Right	Up	Down	Left	Right
IHS	0.027	0.0232	0.0256	0.0217	0.0239	0.0159	0.0191	0.0221
DWT	0.0356	0.0313	0.0319	0.031	0.0295	0.0211	0.0202	0.0277
IHS-GTF	0.0297	0.0286	0.0269	0.0274	0.02	0.0171	0.0141	0.0188
Feature-level fusion	**0.0189**	**0.0182**	**0.017**	**0.0172**	**0.0089**	0.0077	**0.0061**	**0.0093**
Decision-level fusion	0.0204	0.02	0.0186	0.018	0.0094	**0.0074**	0.0063	0.01

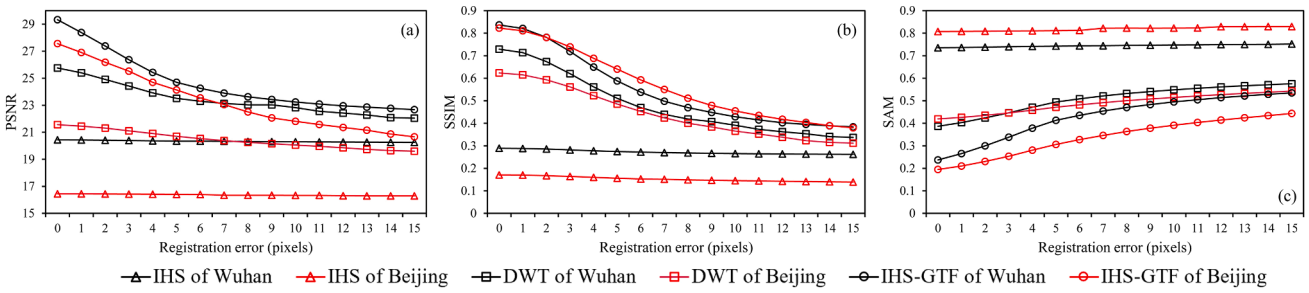


Fig. 13. Comparison of IHS, DWT, and IHS-GTF pixel-fusion results of Wuhan and Beijing under different geometric registration errors (upward shifting) from 0 to 15 pixels. (a) PSNR; (b) SSIM; (c) SAM.

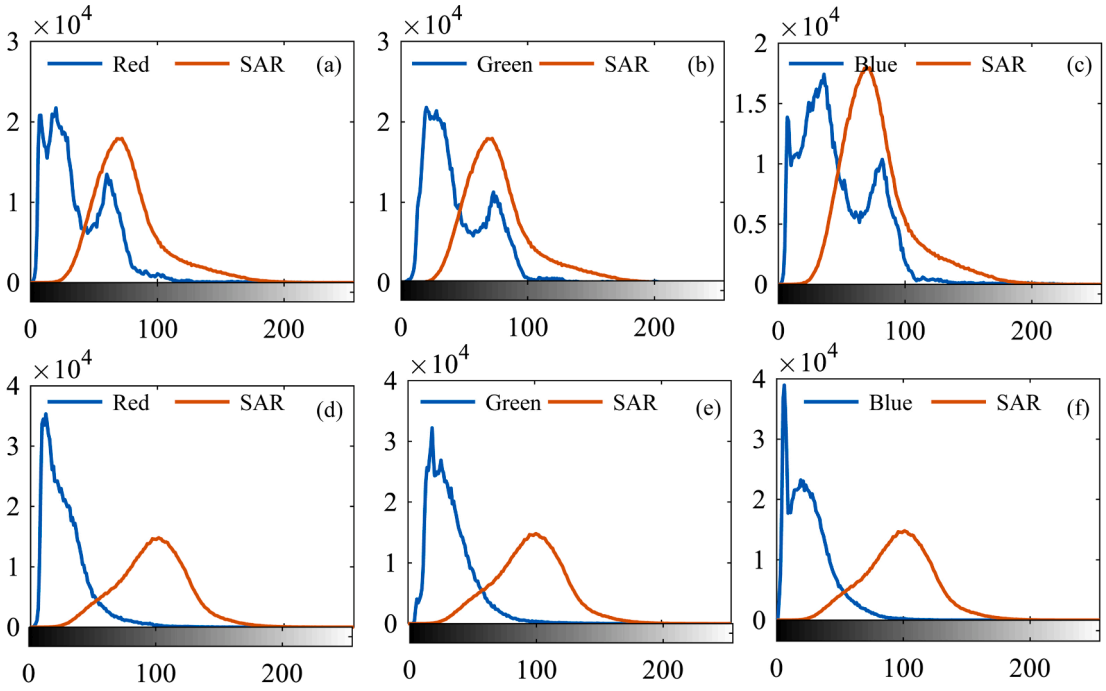


Fig. 14. Histogram analysis of optical image and of SAR image in the study areas. (a-c) Beijing; (d-f) Wuhan.

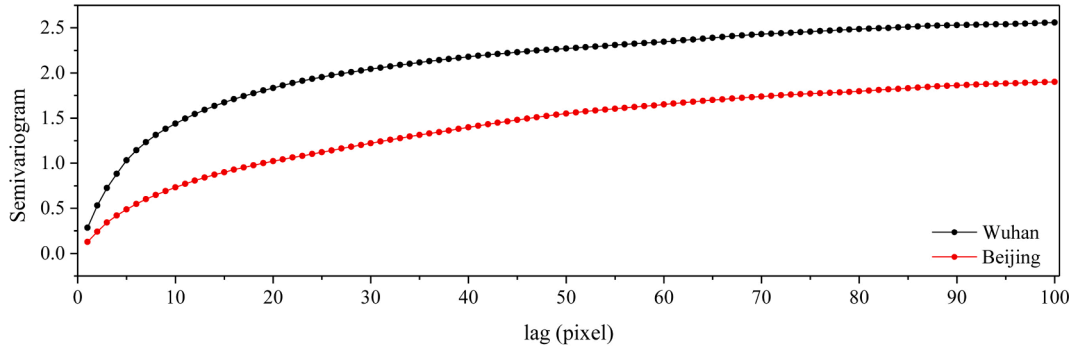


Fig. 15. Semi-variograms of Wuhan and Beijing areas.

Sentinel-2 images fusion, that is, the LCC accuracy decreases with the increase of geometric registration errors and pixel-level fusion is the most sensitive to geometric registration errors. However, their response patterns are different, especially the feature-level and decision-level fusions. Compared with TerraSAR and Sentinel-2 images fusion, Sentinel-1 and Sentinel-2 images fusion is less sensitive to registration errors, and the LCC accuracy decreases rapidly at first, then slows down and finally becomes stable with the increase of registration errors. At the same time, combined with sensitivity calculation, the LCC based on the fusion of SAR-optical images with comparable spatial resolution has a higher tolerance of registration errors.

4.4. Fusion level selection and guidance for future research

One of the goals of this study is to guide the application of SAR-optical images fusion in LCC. Thus, it is necessary to quantify the

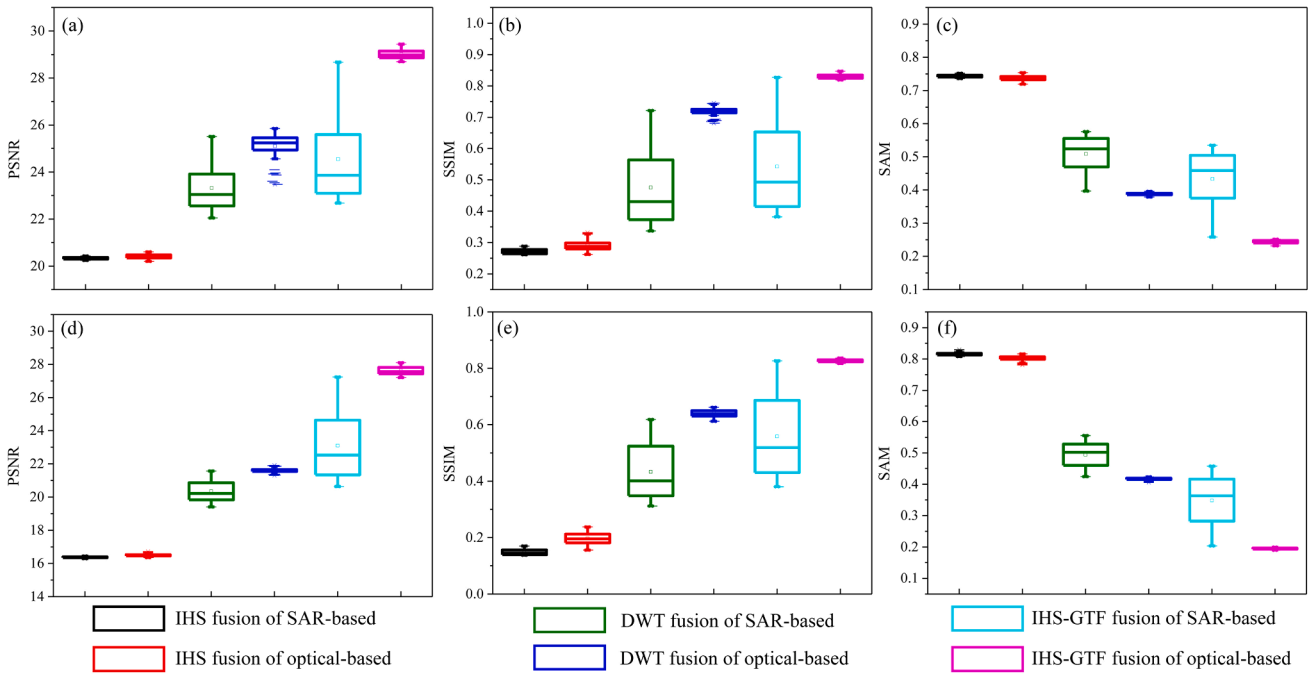


Fig. 16. Comparison of pixel-level fusion results of SAR-based and optical-based. (a-c) are PSNR, SSIM and SAM in Wuhan area, respectively; (d-f) are PSNR, SSIM and SAM in Beijing area, respectively.

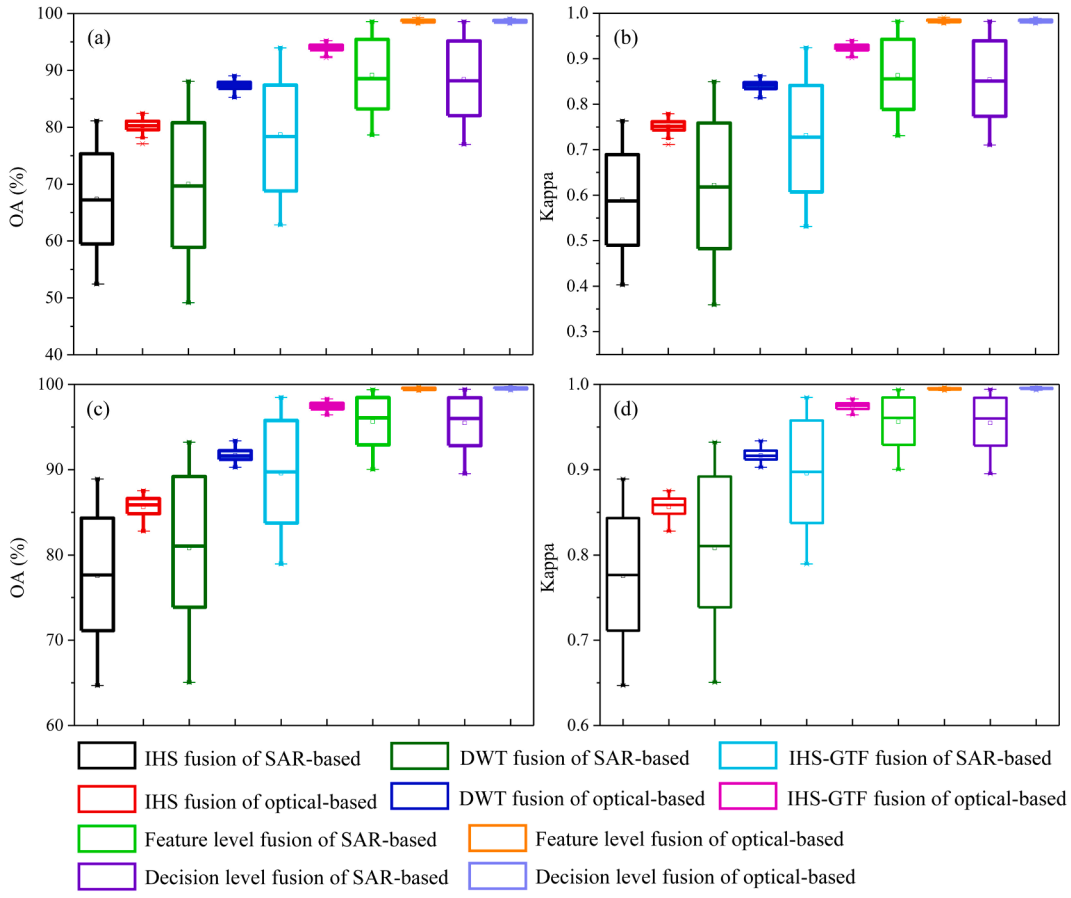


Fig. 17. Comparison of LCC results of SAR-based and optical-based. (a-b) are OA and Kappa of Wuhan area, respectively; (c-d) are OA and Kappa of Beijing area, respectively.

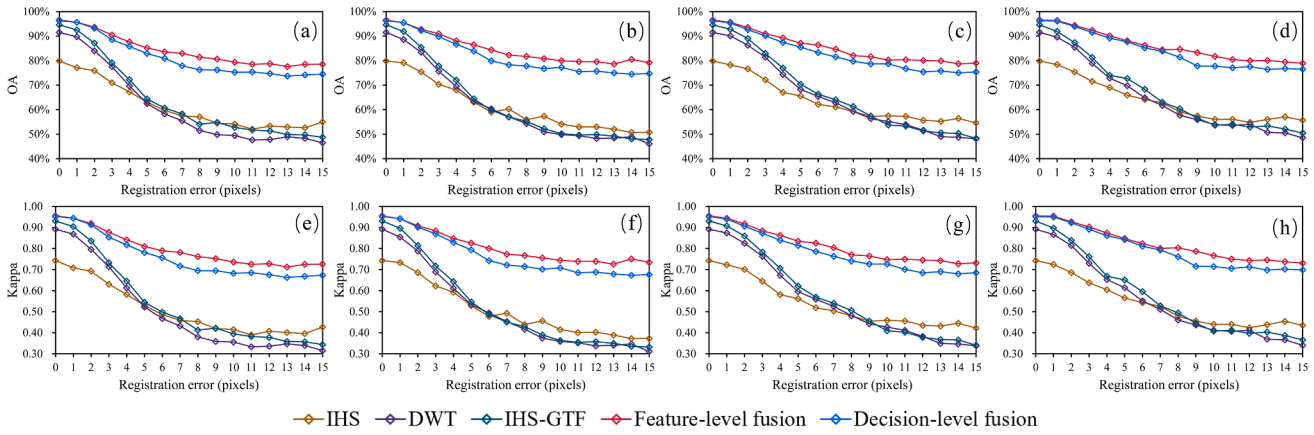


Fig. 18. OA and Kappa of the LCC results based on Sentinel-1 and Sentinel-2 images three-level fusions: (a) and (e) are upward shifting; (b) and (f) are downward shifting; (c) and (g) are left shifting; (d) and (h) are right shifting.

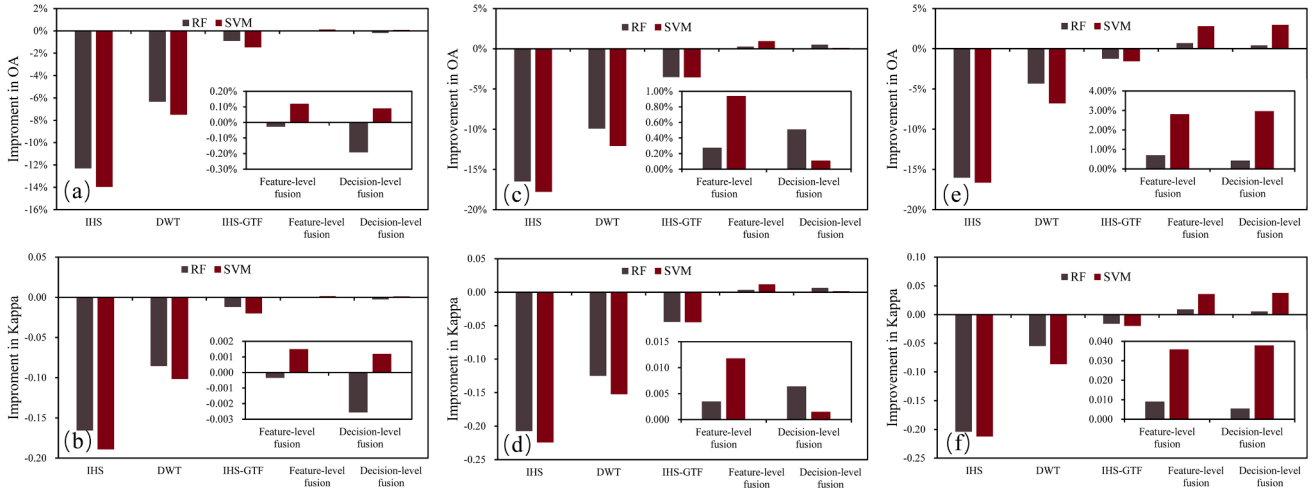


Fig. 19. Improvement of LCC based on SAR-optical images fusion at different levels (without registration errors); (a-b) are TerraSAR and Sentinel-2 images fusion in Beijing; (c-d) are TerraSAR and Sentinel-2 images fusion in Wuhan; (e-f) are Sentinel-1 and Sentinel-2 images fusion.

improvement of LCC at different fusion levels without registration errors. Support vector machine (SVM) is added as another classifier for LCC in this section, where the kernel function is RBF kernel function, its penalty coefficient C and kernel parameter g are set to −4 and 0.02, respectively. To better reflect the improvement of LCC by SAR-optical images fusion, we compare LCC results at different fusion levels with that using optical data alone, as shown in Fig. 19. We find that both feature-level fusion and decision-level fusion can improve the accuracy of LCC, but the degree of improvement is different (in most cases, feature-level fusion improves the most), which is consistent with the previous study by Zhang et al. (2018). But, the SAR-optical images pixel-level fusion does not significantly improve LCC. Besides, the improvement degree of SVM and RF is also different. In conclusion, the improvement of LCC based on SAR-optical images fusion depends on a variety of factors, such as fusion levels, fusion methods, and classifier. In combination with the tolerance of registration errors, the feature-level fusion is the most ideal level for SAR-optical images fusion in LCC.

It should be noted that this study did not fully consider all the influential factors. In this study, geometric registration errors are simply simulated by pixel shifting. However in practice, there are more complex geometric errors between sensors, such as image scaling, rotation and tilting, which will bring great uncertainty to images fusion. Besides, when the spatial resolution of SAR-optical images is inconsistent, scale conversion through resampling will also bring uncertainty. These errors should be carefully considered in future studies of SAR-optical images fusion according to the enlightening experimental results of this study.

5. Conclusion

SAR and optical data are two widely used remote sensing data sources. Fusing them can obtain complementary information that is helpful in LCC. In LCC based on SAR-optical images fusion, geometric registration errors between SAR-optical images needs to be considered. Therefore, this study explores the sensitivity of SAR-optical images three-level fusion in LCC to registration errors through simulation experiments.

Our results show that geometric registration errors affect LCC based on SAR-optical images three-level fusion in a significant manner. Among them, feature-level fusion is the least sensitive to registration errors. Second, the selection of reference images in fusion affects the response to registration errors as well. The fusion based on the optical image has better tolerance of geometric registration errors. Third, the response of SAR-optical images fusion to registration errors in LCC of heterogeneous regions is greater than that in homogeneous regions. Forth, during the LCC, the fusion of SAR-optical images with comparable spatial resolution has a higher tolerance of registration errors. Finally, fusing SAR-optical data does not always guarantee improvement of LCC compared to using optical data alone. The improvement degree depends on fusion levels, fusion methods and classifiers. We believe these findings can facilitate determining reasonable SAR-optical images fusion processes

and fusion levels, providing valuable guidance for the research community to develop new fusion methods in the future.

Funding

This work was supported in part by the National Natural Science Foundation of China under Grants 42090012, Guangxi science and technology program (GuiKe2021AB30019), the Special Fund of Hubei Luojia Laboratory (220100009), in part by 03 special research and 5G project of Jiangxi Province in China (20212ABC03A09); Zhuhai industry university research cooperation project of China (ZH22017001210098PWC); Key R & D project of Sichuan Science and technology plan (2022YFN0031); Zhizhuo Research Fund on Spatial–Temporal Artificial Intelligence (Grant No. ZZJJ202202).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank the editor and anonymous reviewers for their constructive comments and suggestions that improved the manuscript.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.jag.2022.102868.

References

Bai, Y., Sun, G., Li, Y., Ma, P., Li, G., Zhang, Y., 2021. Comprehensively analyzing optical and polarimetric sar features for land-useland-cover classification and urban vegetation extraction in highly-dense urban area. Int. J. Appl. Earth Obs. Geoinf. 103, 102496.

Belgiu, M., Drăguţ, L., 2016. Random forest in remote sensing: A review of applications and future directions. ISPRS J. Photogram. Remote Sens. 114, 24–31.

Chen, S., Zhang, R., Su, H., Tian, J., Xia, J., 2010. Sar and multispectral image fusion using generalized ihs transform based on à trous wavelet and emd decompositions. IEEE Sens. J. 10, 737–745.

Cheng, J., Liu, H., Liu, T., Wang, F., Li, H., 2015. Remote sensing image fusion via wavelet transform and sparse representation. ISPRS J. Photogram. Remote Sens. 104, 158–173.

Dempster, A.P., 2008. Upper and lower probabilities induced by a multivalued mapping, in: Classic works of the Dempster-Shafer theory of belief functions. Springer, pp. 57–72.

Garzelli, A., 2002. Wavelet-based fusion of optical and sar image data over urban area. International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences 34, 59–62.

Huang, B., Li, Y., Han, X., Cui, Y., Li, W., Li, R., 2015. Cloud removal from optical satellite imagery with sar imagery using sparse representation. IEEE Geosci. Remote Sens. Lett. 12, 1046–1050.

Joshi, N., Baumann, M., Ehammer, A., Fensholt, R., Grogan, K., Hostert, P., Jepsen, M.R., Kuemmerle, T., Meyfroidt, P., Mitchard, E.T., et al., 2016. A review of the application of optical and radar remote sensing data fusion to land use mapping and monitoring. Remote Sensing 8, 70.

Kulkarni, S.C., Rege, P.P., 2020. Pixel level fusion techniques for sar and optical images: A review. Information Fusion 59, 13–29.

Lin, Y., Zhang, H., Lin, H., Gamba, P.E., Liu, X., 2020. Incorporating synthetic aperture radar and optical images to investigate the annual dynamics of anthropogenic impervious surface at large scale. Remote Sens. Environ. 242, 111757.

Luo, B., Minnett, P.J., Nalli, N.R., 2021. Infrared satellite-derived sea surface skin temperature sensitivity to aerosol vertical distribution? field data analysis and model simulations. Remote Sens. Environ. 252, 112151.

Moreira, A., Prats-Iraola, P., Younis, M., Krieger, G., Hajnsek, I., Papathanassiou, K.P., 2013. A tutorial on synthetic aperture radar. IEEE Geoscience and remote sensing magazine 1, 6–43.

Murphy, C.K., 2000. Combining belief functions when evidence conflicts. Decision support systems 29, 1–9.

Nencini, F., Garzelli, A., Baronti, S., Alparone, L., 2007. Remote sensing image fusion using the curvelet transform. Information Fusion 8, 143–156.

Pal, S., Majumdar, T., Bhattacharya, A.K., 2007. Ers-2 sar and irs-1c liss iii data fusion: A pca approach to improve remote sensing based geological interpretation. ISPRS J. Photogram. Remote Sens. 61, 281–297.

Parrilli, S., Poderico, M., Angelino, C.V., Verdoliva, L., 2011. A nonlocal sar image denoising algorithm based on llmmse wavelet shrinkage. IEEE Trans. Geosci. Remote Sens. 50, 606–616.

Puissant, A., Hirsch, J., Weber, C., 2005. The utility of texture analysis to improve per-pixel classification for high to very high spatial resolution imagery. Int. J. Remote Sens. 26, 733–745.

Shafer, G., 1976. A mathematical theory of evidence. Princeton University Press.

Shao, Z., Fu, H., Fu, P., Yin, L., 2016. Mapping urban impervious surface by fusing optical and sar data at the decision level. Remote Sensing 8, 945.

Shao, Z., Fu, H., Li, D., Altan, O., Cheng, T., 2019. Remote sensing monitoring of multi-scale watersheds impermeability for urban hydrological evaluation. Remote Sens. Environ. 232, 111338.

Shao, Z., Wu, W., Guo, S., 2020. Ihs-gtf: A fusion method for optical and synthetic aperture radar data. Remote Sensing 12, 2796.

Shao, Z., Wu, W., Li, D., 2021. Spatio-temporal-spectral observation model for urban remote sensing. Geo-Spatial Information Science 24, 372–386.

Sukawattanavijit, C., Chen, J., Zhang, H., 2017. Ga-svm algorithm for improving land-cover classification using sar and optical remote sensing data. IEEE Geosci. Remote Sens. Lett. 14, 284–288.

Tang, Y., Wang, Q., Zhang, K., Atkinson, P.M., 2020. Quantifying the effect of registration error on spatio-temporal fusion. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 13, 487–503.

Wang, Z., Ziou, D., Armenakis, C., Li, D., Li, Q., 2005. A comparative analysis of image fusion methods. IEEE transactions on geoscience and remote sensing 43, 1391–1402.

Werner, A., Storie, C.D., Storie, J., 2014. Evaluating sar-optical image fusions for urban lulc classification in vancouver canada. Canadian Journal of Remote Sensing 40, 278–290.

Wu, W., Guo, S., Cheng, Q., 2020. Fusing optical and synthetic aperture radar images based on shearlet transform to improve urban impervious surface extraction. J. Appl. Remote Sens. 14, 024506.

Zhang, H., Xu, R., 2018. Exploring the optimal integration levels between sar and optical data for better urban land cover mapping in the pearl river delta. International journal of applied earth observation and geoinformation 64, 87–95.

Zhang, H., Zhang, Y., Lin, H., 2012. A comparison study of impervious surfaces estimation using optical and sar remote sensing images. Int. J. Appl. Earth Obs. Geoinf. 18, 148–156.

Zhang, Y., Zhang, H., Lin, H., 2014. Improving the impervious surface estimation with combined use of optical and sar remote sensing images. Remote Sens. Environ. 141, 155–167.

Zhou, D., Xiao, J., Frolking, S., Liu, S., Zhang, L., Cui, Y., Zhou, G., 2021. Croplands intensify regional and global warming according to satellite observations. Remote Sens. Environ. 264, 112585.

Zhou, J., Chen, J., Chen, X., Zhu, X., Qiu, Y., Song, H., Rao, Y., Zhang, C., Cao, X., Cui, X., 2021. Sensitivity of six typical spatiotemporal fusion methods to different influential factors: A comparative study for a normalized difference vegetation index time series reconstruction. Remote Sens. Environ. 252, 112130.