Shadow Detection and Reconstruction of High-Resolution Remote Sensing Images in Mountainous and Hilly Environments

Zhenqing Wang, Yi Zhou, Futao Wang, Shixin Wang, Gang Qin and Jinfeng Zhu

***Abstract*—The undulating terrain in mountainous and hilly regions results in a greater variety and complexity of shadows. Efficient methods for shadow detection and reconstruction in high-resolution remote sensing images are particularly important in such hilly areas. The accurate detection of shadow masks is a prerequisite for shadow reconstruction. By utilizing the features of high hue and low intensity in shadow areas, an initial spectral ratio is constructed based on the CIELCh color space model. Simple linear iterative clustering is employed to perform superpixel segmentation on the image, and the segmented results are spatially constrained to reconstruct the initial spectral ratio. Afterwards, an automatic multi-level global thresholding approach is applied to obtain the shadow mask and eliminate the influence of interfering objects. For shadow reconstruction, the segmented superpixels are treated as the smallest processing units. Similar neighboring objects have similar ambient light intensities. Based on this, we propose a shadow reconstruction method, which compensates shadow superpixels using adjacent non-shadow superpixels and determines compensation weights based on their similarity. Furthermore, the shadow boundaries are dilated to obtain penumbra, and mean filtering is performed to compensate for the illumination in the penumbra. Finally, the proposed method is qualitatively and quantitatively compared with existing shadow detection and reconstruction methods. Experimental results demonstrate that this method can accurately detect shadows in high-resolution remote sensing images in mountainous and hilly environments, and effectively reconstruct the spectral information of shadow areas. This has significant implications for subsequent feature extraction and further analysis in mountainous and hilly regions.**

***Index Terms*—Hilly environment; High resolution remote sensing; Shadow detection; Shadow reconstruction**

I. INTRODUCTION

High-resolution remote sensing images efficiently capture information about areas of interest and have played a significant role in various fields such as urban development, environmental management, and disaster risk reduction. However, high-resolution remote sensing images are severely affected by shadows. Shadow coverage results in significant loss of information in the corresponding areas, greatly impacting subsequent feature extraction and analysis [1]. The detection and reconstruction of shadow regions not only visually contribute to shadow restoration but also provide assistance in tasks such as road extraction, building detection, and impervious surface studies [2].

From the perspective of shadow imaging mechanisms, shadows can be primarily classified into two parts: cast shadows and self-shadows. As shown in Figure 1, cast shadows occur when an object obstructs direct sunlight, while self-shadows are generated when an object itself blocks sunlight. Cast shadows can further be divided into umbra and penumbra. Umbra results from complete blockage of direct sunlight, whereas penumbra forms due to partial obstruction of direct sunlight. In high-resolution remote sensing images, the shadows present mainly consist of umbra and penumbra. In plain areas, shadows primarily consist of building shadows. However, in mountainous and hilly areas, due to its special topography, in addition to building shadows, slope shadows and cloud shadows also occupy a large proportion, as shown in Figure 2. The existence of multiple types of shadows makes shadow information restoration methods particularly important in mountainous and hilly environments.

The process of shadow information restoration involves two main components: shadow detection and shadow reconstruction. Shadow reconstruction heavily relies on the results of shadow detection. Currently, shadow detection methods can generally be categorized into three types: physics-based methods, feature-based methods, and machine learning-based methods [3]. Physics-based methods utilize mathematical models constructed using satellite sensor positions, solar

This paragraph of the first footnote will contain the date on which you submitted your paper for review, which is populated by IEEE. This work was supported in part by the Finance Science and Technology Project of Hainan Province under Grant ZDYF2021SHFZ103, and in part by the National Key Research and Development Program of China under Grant No.2021YFB3901201, Grant 2022YFC3006400 and Grant 2022YFC3006402. *(Zhenqing Wang and Yi Zhou contributed equally to this work.)(Corresponding authors: Futao Wang)*

Zhenqing Wang and Futao Wang are with the Key Laboratory of Earth Observation of Hainan Province, Hainan Aerospace Information Research Institute, Sanya 572029, China, and also with the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China, and also with the University of Chinese Academy of Science, Beijing 100049, China (e-mail: wangzhenqing19@mails.ucas.ac.cn; wangft@aircas.ac.cn).

Yi Zhou, Shixin Wang, Gang Qin and Jinfeng Zhu are with the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China, and also with the University of Chinese Academy of Science, Beijing 100049, China (e-mail: zhouyi@aircas.ac.cn; WangSX@irsa.ac.cn; qingang20@mails.ucas.ac.cn; zhujf201735@aircas.ac.cn).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully e
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3338976

azimuth angles, and digital elevation models (DEM) to obtain shadow information. For instance, Luo et al. first obtained rough shadows through DEM and then optimized shadow features using support vector machines [1]. However, physics-based methods require camera poses, illumination directions, and high-resolution DEMs, which are not easily obtainable [4]. Feature-based methods typically involve image feature extraction and segmentation. Shadow regions exhibit higher hue and lower intensity compared to non-shadow regions. These algorithms primarily transform remote sensing images from the RGB model to the hue, saturation, and intensity (HSI) space or equivalent spaces. Based on this, Tsai et al. [5] compared and analyzed the ratio of hue to intensity in different color spaces, and used the Otsu algorithm to obtain the shadow mask. Silva et al. [6] transformed the image to the CIELCh model, proposed an improved shadow index, and applied multi-level thresholding to obtain the shadow mask. Zhou et al. [7] combined the color space with the near-infrared band, proposed an enhanced shadow index to highlight shadows, and employed a thresholding segmentation method to obtain the shadow mask. In recent years, machine learning, particularly deep learning networks, has rapidly developed, providing new inspiration for shadow extraction. Many researchers have designed various effective deep learning methods for shadow detection and achieved better results than traditional methods [8-12]. Unfortunately, current machine learning algorithms primarily focus on general images. Due to the limitations of satellite revisit periods and weather conditions, it is challenging to acquire shadow images and corresponding shadow-free images from the same location to create shadow datasets for remote sensing image shadow detection.

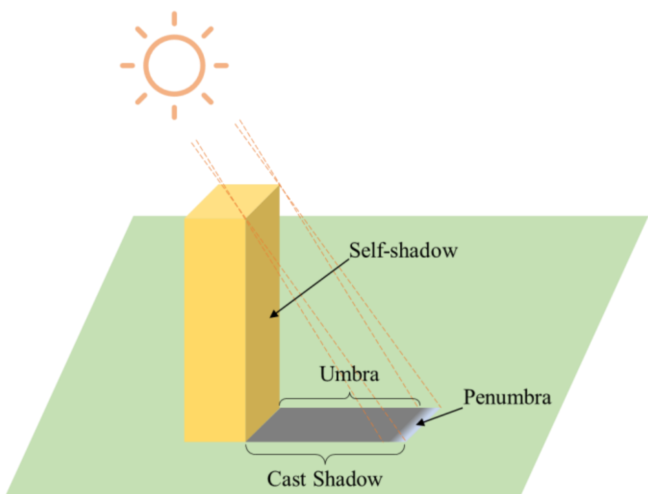


Fig. 1. Imaging mechanism of shadows.

The spectral information of non-shadow regions serves as the primary source for shadow information reconstruction, and utilizing this information is crucial for shadow removal. Common methods for shadow compensation include histogram matching, linear correlation correction, intensity-based methods, and machine learning-based methods. Histogram matching methods align the shadow histogram to the non-shadow histogram to achieve information reconstruction [13-16]. Linear correlation correction methods employ linear correlation functions to reconstruct the shadow information by adjusting the intensity of shadow pixels based on the statistical characteristics of the corresponding non-shadow regions [17-19]. Intensity-based methods reconstruct shadows based on the illumination model [6,20]. Shadow regions are formed by ambient light, while non-shadow regions are composed of ambient light and direct sunlight. The information of shadow regions can be recovered by the ratio of direct sunlight to ambient light. In recent years, various deep learning models, such as Generative Adversarial Networks (GAN) [8,10,12,21], have been widely used for generating shadow-free images. GAN-based methods utilize a generator G and a discriminator D to obtain shadow-free images, where G is trained to generate shadow-free images and D evaluates the similarity between the generated shadow-free images and the reference images. However, deep learning requires a large amount of training data, and the training cost to obtain a generalized model is significant. Besides considering which model to use for shadow information compensation, another key issue in shadow removal is where to select the information from. Some researchers model the entire image's non-shadow and shadow regions, but this may lead to excessive compensation of the shadow region's information [5,6]. To address this issue, in some studies, modeling is performed on shadow pixels and their neighborhood non-shadow pixels [7,20]. In summary, for mountainous and hilly remote sensing images with complex backgrounds, effective and accurate methods for shadow reconstruction are still lacking.

Fig. 2. Components of shadows in mountainous and hilly Areas.

Despite extensive research on shadow detection and reconstruction, several limitations persist: (1) The accuracy of shadow detection results still has room for improvement due to the influence of dark objects and water bodies, among other objects. (2) Non-shadow samples used for shadow region information restoration are crucial for achieving effective reconstruction. However, selecting the most suitable non-shadow samples is a challenging task. (3) The increasing volume of remote sensing data calls for a fully automated method for shadow detection and reconstruction. Semi-automatic methods are no longer suitable for handling large-scale shadowed remote sensing images.

In order to efficiently address complex shadow scenarios in mountainous and hilly regions, a fully automated method for shadow detection and reconstruction has been developed. This research makes the following key contributions: (1) A novel

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully e
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3338976

shadow index is proposed by combining the CIELCh color model and superpixel clustering techniques. (2) A method is introduced for the reconstruction of shadow superpixels based on the similarity-weighted compensation of adjacent non-shadow superpixels, taking into account the similarity of environmental lighting conditions between neighboring similar objects. (3) Through validation on remote sensing images containing various types of shadows, the proposed method has achieved state-of-the-art results in shadow detection and reconstruction.

II. DATA

The data used in this study is sourced from the GaoFen-2 satellite. GaoFen-2 is the first domestically developed optical remote sensing satellite in China with a spatial resolution better than 1 meter for civilian use. It is equipped with two sensors: a high-resolution 1-meter panchromatic sensor and a 4-meter multispectral sensor, providing a spatial resolution of up to 0.8 meters at the nadir. GaoFen-2 images offer the advantages of high spatial resolution and wide coverage, enabling detailed observation of ground information over large areas. GaoFen-2 has been utilized in various important applications, including disaster assessment and ground information extraction. After radiometric calibration, geocoding, and image fusion, the multispectral imagery with a resolution of 0.8 meters is obtained.

The study area comprises two hilly counties in China, namely Hengdong County in Hunan Province and Daning County in Shanxi Province. Hengdong County is characterized by a predominantly hilly terrain, with additional plains and mountainous areas. Daning County, on the other hand, features a landscape of gullies and undulating mountain ranges. Partial remote sensing images of the two regions are shown in Figure 3.

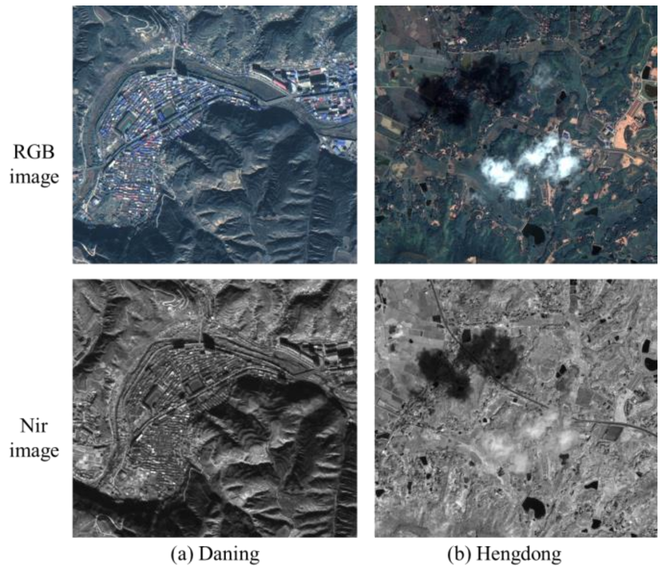


Figure 3. Part of the remote sensing image of the study area. The first row is RGB images, and the second row is Nir images.

To quantitatively evaluate the performance of shadow detection, three sub-areas were selected to create ground truth values for shadow detection. These sub-areas include typical examples of slope shadows, building shadows, and cloud shadows. Multiple types were chosen to ensure the validity of the model performance evaluation. The creation of ground truth values involved the visual interpretation of multi-temporal images within the same regions. Domain experts were invited to review the results, and revisions were made based on their feedback. The final ground truth values are presented in Figure 4.

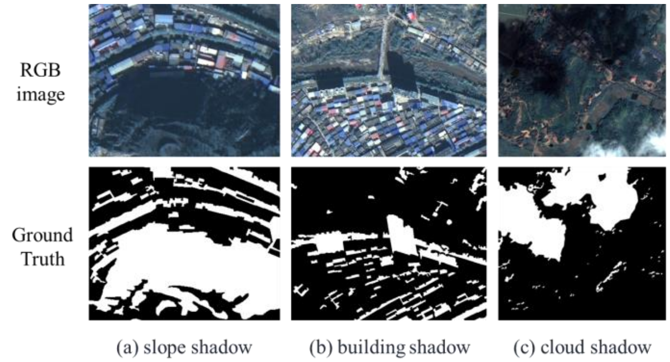


Figure 4. Ground truth of shadowed areas. In a ground truth row, white pixels represent shadows and black pixels represent non-shadows.

III. METHODOLOGY

A. Shadow detection

Building upon the spectral ratio framework proposed by Silva et al. [6], we perform superpixel reconstruction of the spectral ratio. This approach enables a finer delineation of boundaries in the shadow detection results.

Initially, the RGB color space of the remote sensing image is mapped to the CIELCh color space. The CIELCh color space represents the polar coordinate representation of the CIELAB color space, which was designed by the Commission Internationale de l´Eclairage (CIE). The CIELCh color space is device-independent and is derived from the CIE tristimulus values CIEXYZ, describing colors in terms of Lightness, chroma, and hue.

Convert RGB to CIEXYZ color space:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4124564 & 0.3575761 & 0.1804375 \\ 0.2126729 & 0.7151522 & 0.0721750 \\ 0.0193339 & 0.1191920 & 0.9503041 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{1}$$

The CIELAB color space can be obtained from the CIEXYZ:

$$L = \begin{cases} 116(\frac{Y}{Y_n})^{\frac{1}{3}} - 16 & \text{if } \frac{Y}{Y_n} > 0.008856 \\ 903.3(\frac{Y}{Y_n}) & else \end{cases} \tag{2}$$

$$a = 500(f(\frac{X}{X_n}) - f(\frac{Y}{Y_n})) \tag{3}$$

$$b = 200(f(\frac{Y}{Y_n}) - f(\frac{Z}{Z_n})) \tag{4}$$

where

$$f(x) = \begin{cases} x^{\frac{1}{3}} & \text{if } x > 0.008856 \\ 7.787x + \frac{16}{116} & else \end{cases} \tag{5}$$

In the above equation, X_n, Y_n, Z_n are the reference values of the standard D65 light source, that is, $X_nY_nZ_n$ = {95.047, 100.00, 108.883}.

The CIELAB color space can be transformed into the CIELCh color space by a simple geometric transformation:

$$C = \sqrt{a^2 + b^2} \tag{6}$$

$$h = \text{atan2}(b, a)$$

$$h = \begin{cases} h + 360? & \text{if } h < 0^{\circ} \\ h - 360? & \text{if } h \geq 360^{\circ} \end{cases} \tag{7}$$

Due to the absence of direct sunlight, shadow areas exhibit features of high hue and low intensity. The ratio between the hue and intensity of pixels is calculated to accentuate the differences in values between shadow and non-shadow regions.

$$SR = \frac{h+1}{L+1} \tag{8}$$

Here, *SR* represents the spectral ratio image, while h and L are normalized to the range of [0,1] prior to calculation. In the spectral ratio image, pixels within shadow regions will have higher values compared to pixels in non-shadow regions.

To address the issues of fine noise and the preservation of shadow edge integrity when directly thresholding the spectral ratio of shadows, superpixel reconstruction is performed on the spectral ratio image. Firstly, the simple linear iterative clustering (SLIC) [22] algorithm is employed for superpixel segmentation of the remote sensing image. SLIC significantly reduces the number of distance calculations in optimization by limiting the search space to regions proportional to the size of superpixels, while providing control over the compactness of the superpixels. Fewer superpixels can result in rough boundaries, while a larger number increases computational complexity. To address this, an initial generation of a larger number of superpixels is followed by merging similar adjacent superpixels, as illustrated in Figure 5. The superpixels obtained from SLIC are converted into a region adjacency graph, and regions with similar colors are gradually merged to obtain the final superpixels.

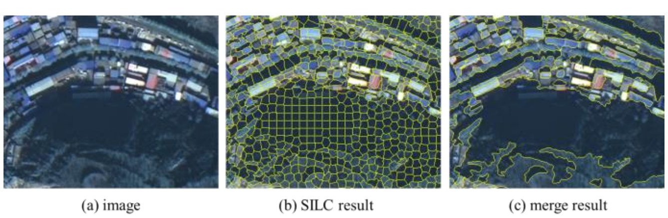


Figure 5. Superpixel segmentation and similar superpixel merging.

Subsequently, the spectral ratio image is reconstructed by taking the mean value of all pixels within each superpixel as the value of that superpixel. Assuming the remote sensing image is segmented into n superpixels, with each superpixel containing m_i pixels, the reconstructed pixel ratio can be expressed as follows:

$$SR_{sp} = \bigcup_{i}^{n} (\sum_{j}^{m_i} (SR_j) / m_i) \tag{9}$$

where SR_j represents the *SR* value of pixel *j*.

To achieve fully automatic shadow extraction, the threshold needs to be determined automatically. Dark objects, water bodies, and shadows often lie on the same side of the spectral ratio histogram. If a binary thresholding method is used, there is a risk of erroneously extracting dark objects and water bodies. By employing an automatic multi-level global thresholding approach, multiple thresholds are obtained, and the maximum threshold value is selected as the shadow threshold. Thresholding the superpixel-reconstructed spectral ratio image into a binary mask allows for better segmentation of shadows from dark objects and water bodies.

B. Shadow reconstruction

Shadows in remote sensing images consist of umbra and penumbra. The umbra is influenced solely by the intensity of ambient light, while the penumbra is composed of the intensity of both ambient light and dynamically attenuated direct sunlight. Shadow reconstruction is divided into two parts: umbra compensation and penumbra post-processing.

For any pixel *i* in the image, the value I_i can be expressed as the product of the intensity of the light (L_i) and the reflectance of the pixel (R_i):

$$I_i = L_i R_i \tag{10}$$

Light intensity comprises direct intensity and environmental intensity: direct intensity originates from solar radiation, while environmental intensity is primarily caused by sky scattering [23]. The pixel values in non-shadow and shadow regions of the image can be calculated as follows:

$$\begin{aligned} I_{unshaded,i} &= (L_d + L_e)R_i \\ I_{shaded,i} &= (k_i L_d + L_e)R_i, k_i \in [0,1) \end{aligned} \tag{11}$$

Where $I_{unshaded,i}$ represents the value of a non-shadow pixel, $I_{shaded,i}$ represents the value of a shadow pixel, L_d is the direct intensity, Le is the environmental intensity. R_i is the reflectance of the pixel, and ki represents the attenuation factor of direct intensity. If a pixel is solely illuminated by environmental light, i.e., it is in the umbra region, k_i is equal to 0. If a pixel is illuminated by both environmental light and direct light, i.e., it is in the penumbra region, k_i takes a value between 0 and 1.

The pixel values in shadow regions can be re-illuminated based on the ratio between direct intensity and environmental intensity [20]. For any pixel *i*, the ratio *r* between environmental intensity and direct intensity can be expressed as:

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully e
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3338976

$$r=\frac{L_{unshaded}-L_{shaded}}{L_{shaded}}=\frac{(L_d+L_e)R_i\text{-}(k_iL_d+L_e)R_i}{(k_iL_d+L_e)R_i}=\frac{L_d\text{-}k_iL_d}{L_e+k_iL_d}\approx\frac{L_d}{L_e} \tag{12}$$

The final shadow reconstructed pixel value can be expressed as:

$$I_{ShadowReconstruction,i}=(\tfrac{r+1}{kr+1})I_i \tag{13}$$

The reconstructed pixel value $I_{ShadowReconstruction,i}$ can be simplified as the shadow value multiplied by $(r + 1)$ when $k = 0$ (i.e., the original pixel I_i is in the umbra).

According to the First Law of Geography, proximity implies a closer association between objects. Each superpixel is most likely to share similar attributes with its neighboring superpixels. In the absence of shadow influence, adjacent superpixels receive similar solar radiation. Therefore, the most accurate light intensity reconstruction for a shadow superpixel can be achieved by using neighboring non-shadow superpixels. When objects are occluded by shadows, the overall intensity decreases, but the trend of the histogram remains consistent. Therefore, the closer the similarity in histogram trends among adjacent superpixels, the greater their contribution. Consequently, umbra compensation can be performed through the following steps:

(1) To begin with, all shadow superpixels are selected as a preparatory step for subsequent processing.

(2) Get the set $SPnear$ = {$SP_{near,1}$, $SP_{near,2}$, $SP_{near,3}$, …, $SP_{near,S}$} of adjacent non-shaded superpixels of a specific shadow superpixel SP_{shaded}, S is the number of adjacent non-shaded superpixels of SP_{shaded}.

(3) Calculate the histogram H_{shaded} and H_{near} of SP_{shaded} and superpixels in SP_{near}, and use the Bhattachary coefficient (indicated by ρ in the formula) to calculate the similarity of the histograms, and then calculate the weight W_{near} of each superpixel:

$$H_{shaded} = hist\left(SP_{shaded}\right) \tag{14}$$

$$H_{near} = \begin{Bmatrix} hist\left(SP_{near,1}\right),\ hist\left(SP_{near,2}\right),\ hist\left(SP_{near,3}\right), \\ \ldots,\ hist\left(SP_{near,S}\right) \end{Bmatrix} \tag{15}$$

$$B_{near} = \begin{Bmatrix} \rho\left(H_{shaded},\ SP_{near,1}\right), \\ \rho\left(H_{shaded},\ SP_{near,2}\right), \\ \rho\left(H_{shaded},\ SP_{near,3}\right), \\ \ldots,\ \rho\left(H_{shaded},\ SP_{near,S}\right) \end{Bmatrix} \tag{16}$$

$$W_{near} = \frac{\left(1-\ B_{near}\right)}{\sum_i^S\left(1-B_{near}\right)}*S \tag{17}$$

(4) Calculate the intensity of SP_{shaded} and the intensity of the adjacent non-shadowed area to obtain the ratio:

$$L_{shaded} = \sum_i^P SP_{shaded,i} \,/\, P \tag{18}$$

$$L_{unshaded} = \frac{\sum_i^S\sum_j^Q(SP_{near,i,j}\times W_{near,i})}{S\times Q} \tag{19}$$

Among them, P is the number of pixels of SP_{shaded}, and Q is the number of pixels of superpixels in SP_{near}.

(5) Calculate the ratio r between the environmental intensity and direct intensity for SP_{shaded}, leading to the result of $SP_{ShadowReconstruction}$ after umbra compensation of SP_{shaded}:

$$SP_{ShadowReconstruction} = (\frac{L_{unshaded}-L_{shaded}}{L_{shaded}}+1)SP_{shaded} \tag{20}$$

Repeat steps (2)-(5) for the remaining shadow superpixels until all shadow superpixels have undergone illumination compensation.

The penumbra lies in the transition zone between umbra and non-shadow regions. The shadows extracted using automatic multilevel thresholding only include partial areas of the penumbra. Darker portions of the penumbra are extracted as shadows and may exhibit overcompensation after umbra compensation. Brighter portions of the penumbra are not extracted and thus do not undergo illumination compensation, resulting in undercompensation. Therefore, post-processing is required for the penumbra region.

The boundary of the extracted shadow is expanded, and the obtained buffer is considered to be a penumbra:

$$Penumbra = ShadowBoundary \oplus kernel \tag{21}$$

Dilation "grows" the value in *ShadowBoundary* according to the structure kernel. After that, average filtering is performed on the penumbra area to alleviate the anomaly of illumination compensation in the penumbra area. Get the final result image:

$$I_{result} = \begin{cases} K_b * I_{ShadowFree} & if\ \ in\ penumbra \\ I_{ShadowFree} & else \end{cases} \tag{22}$$

where K_b is the kernel of mean filtering.

IV. Results and Discussion

In this section, the proposed shadow detection and reconstruction method is tested on the three subregions described in Section II. To verify the effectiveness and robustness of the proposed method, we conduct a comparative analysis with existing state-of-the-art shadow detection and reconstruction methods. These methods are proposed by Silva et al. [6], Zhou et al. [7], and Liu et al. [24], respectively. The method of Liu et al. [24] is a shadow detection method based on deep learning. Since the corresponding shadow-free image of the shadow area is difficult to obtain, it is difficult to carry out shadow reconstruction based on deep learning. Therefore, deep learning-based methods are not included in the shadow reconstruction comparison methods.

A. Performance and comparative analysis of shadow detection

Visualizing the results of different shadow detection methods allows for a direct comparison of their performance. Figure 1 displays the detection results of three methods for slope shadows, building shadows, and cloud shadows. By comparing them with the ground truth, it is evident that our method

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully e
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3338976

outperforms the others significantly. For slope shadows, all methods are capable of identifying large contiguous shadow areas. However, the other three methods are prone to false positives in small-scale slopes and transition areas between shadows, as shown in the red-boxed region in Figure 6. For building shadows, Zhou et al. [7]'s method misses many small-scale building shadows. All four methods successfully detect large-scale building shadows, but our method achieves higher precision, as demonstrated in the green-boxed region in Figure 6. For cloud shadows, Silva et al. [6]'s method incorrectly identifies gaps between most cloud shadows as cloud shadows, as indicated by the blue-boxed region in Figure 6.

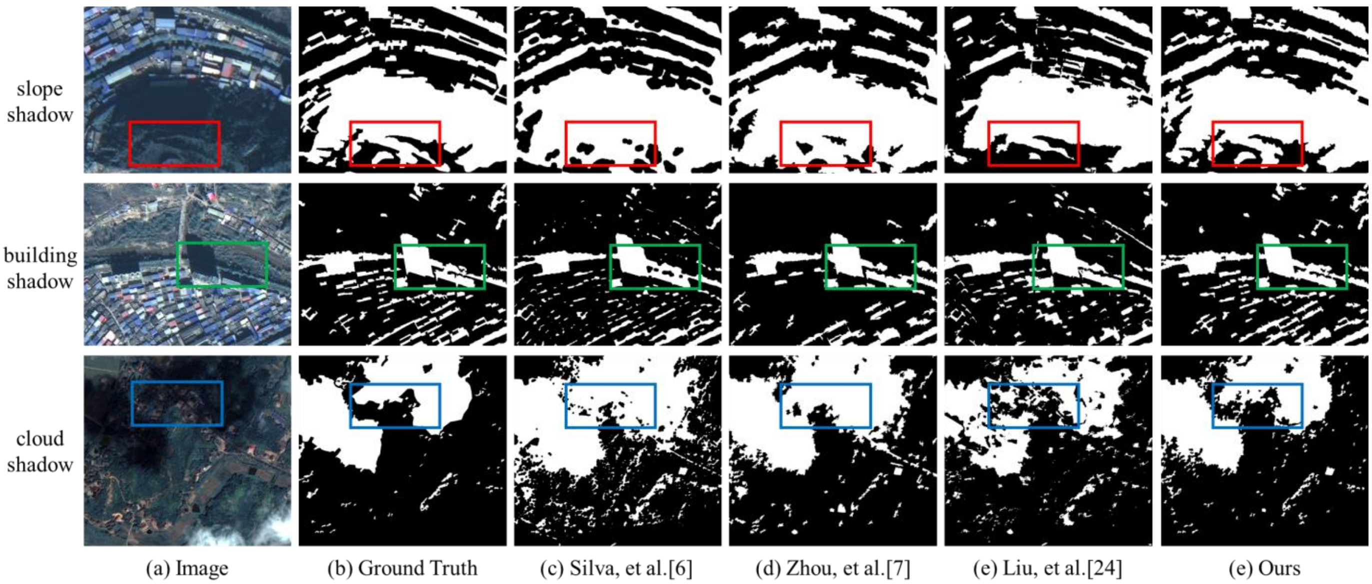


Fig. 6. Detection results of shadow regions. white pixels represent shadows and black pixels represent non-shadows. (a) RGB images, (b) ground truth for shadows, (c) shadow results detected using the method proposed by Silva et al. [6], (d) shadow results detected using the method proposed by Zhou et al. [7], (e) shadow results detected using our proposed method.

In order to quantitatively analyze the performance of different methods, the quality of the results is evaluated using precision, recall, and F_1 score. Precision refers to the proportion of true positive samples among all positive samples in the shadow detection results. Recall refers to the proportion of successfully detected shadow positive samples among all positive samples in the ground truth. The F_1 score is the harmonic mean of precision and recall, calculated as follows:

$$F_1 = \frac{2*(Precision*Recall)}{Precision+Recall} \tag{23}$$

The detection accuracy of mountain shadow is calculated and presented in Table I. Our method achieves the F_1 score of 94.77%, which is 4.39%, 6.30% and 9.32% higher than the other three methods, respectively. The improvement in precision is more significant than the improvement in recall, mainly due to fewer false positives at the boundary between shadows and slopes. The detection accuracy of building shadow is shown in Table II. Our method achieves an F_1 score of 89.66%. The performance in building shadow detection is not as good as that in mountain shadow detection, mainly because it is more challenging to identify scattered small-area building shadows compared to concentrated large-area mountain shadows. This is also the reason why the other three methods only achieve F_1 scores of 78.02%, 71.94% and 69.18%, respectively. The detection accuracy of cloud shadow is presented in Table III. Our method still achieves the highest F_1 of 93.05%. Although the recall rate is 3.07% and 3.32% lower than the methods of Silva et al. [6] and Zhou et al. [7], there is a significant improvement in precision, which is 26.21% and 16.82%, respectively. Methods based on deep learning strongly rely on training data, while the training data of Liu et al. [25]'s method is aerial images. This is most likely the reason why the method performs poorly on satellite data. Our shadow detection method demonstrates the highest performance in all three shadow scenarios, providing a fundamental basis for subsequent shadow reconstruction.

TABLE I
DETECTION PERFORMANCE OF SLOPE SHADOWS

	Precision (%)	Recall (%)	F_1(%)
Silva et al. [6]	85.72	95.56	90.38
Zhou et al. [7]	84.01	93.44	88.47
Liu et al. [24]	89.14	82.06	85.45
Ours	93.49	96.08	94.77

This article has been accepted for publication in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. This is the author's version which has not been fully e
content may change prior to final publication. Citation information: DOI 10.1109/JSTARS.2023.3338976

TABLE II
DETECTION PERFORMANCE OF BUILDING SHADOWS

	Precision (%)	Recall (%)	F_1 (%)
Silva et al. [6]	83.18	73.46	78.02
Zhou et al. [7]	79.04	66.02	71.94
Liu et al. [24]	66.31	72.31	69.18
Ours	88.42	90.94	89.66

TABLE III
DETECTION PERFORMANCE OF CLOUD SHADOWS

	Precision (%)	Recall (%)	F_1 (%)
Silva et al. [6]	65.34	97.66	78.30
Zhou et al. [7]	74.73	97.91	84.77
Liu et al. [24]	75.27	79.59	77.37
Ours	91.55	94.59	93.05

B. Performance and comparative analysis of shadow reconstruction

Qualitative analysis allows for subjective evaluation of the shadow reconstruction results. Figure 7 presents the shadow reconstruction results in the RGB band. All three methods demonstrate different levels of shadow reconstruction. In the red box region of Figure 7, our method clearly and distinctly reconstructs the buildings, with no high contrast between the shadow edges and adjacent non-shadow edges. In the green box region of Figure 7, our method achieves better recovery of vegetation information obscured by buildings compared to the other two methods. In the blue box region of Figure 7, both our method and the method proposed by Silva et al. [6] effectively restore the buildings under cloud shadows. However, the method proposed by Zhou et al. [7] adjusts the spectral information of the buildings to be similar to vegetation, making it difficult to distinguish the buildings. The method proposed by Silva et al. [6] tends to overcompensate in thin cloud shadows, resulting in overly bright regions and a visually abrupt effect. Figure 8 presents the shadow reconstruction results in the NIR band, further highlighting the superior visual quality of our method.

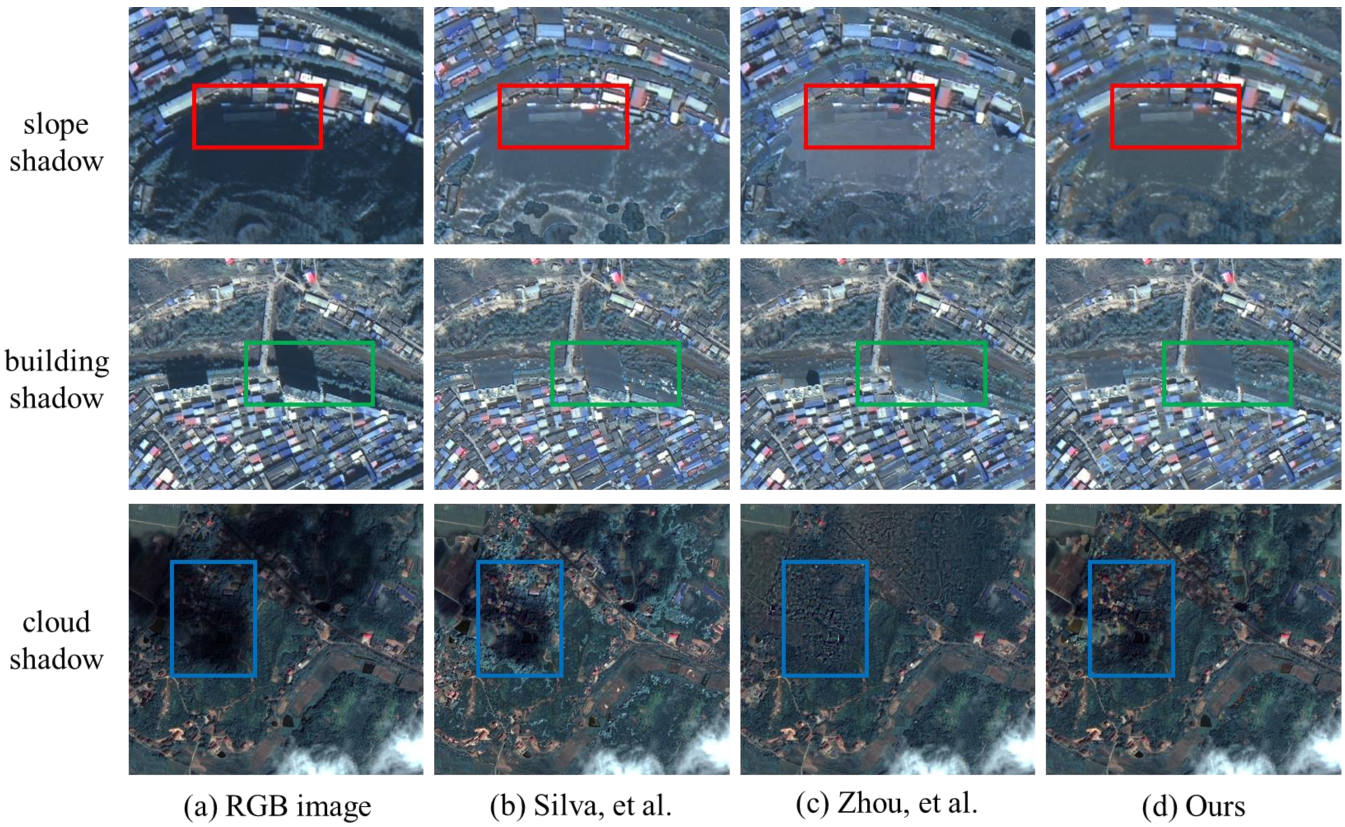


Figure 7. Reconstruction results of the RGB bands for shadow. (a) RGB images, (b) the reconstruction results using the method proposed by Silva et al. [6], (c) the reconstruction results using the method proposed by Zhou et al. [7], (d) the reconstruction results using our proposed method.

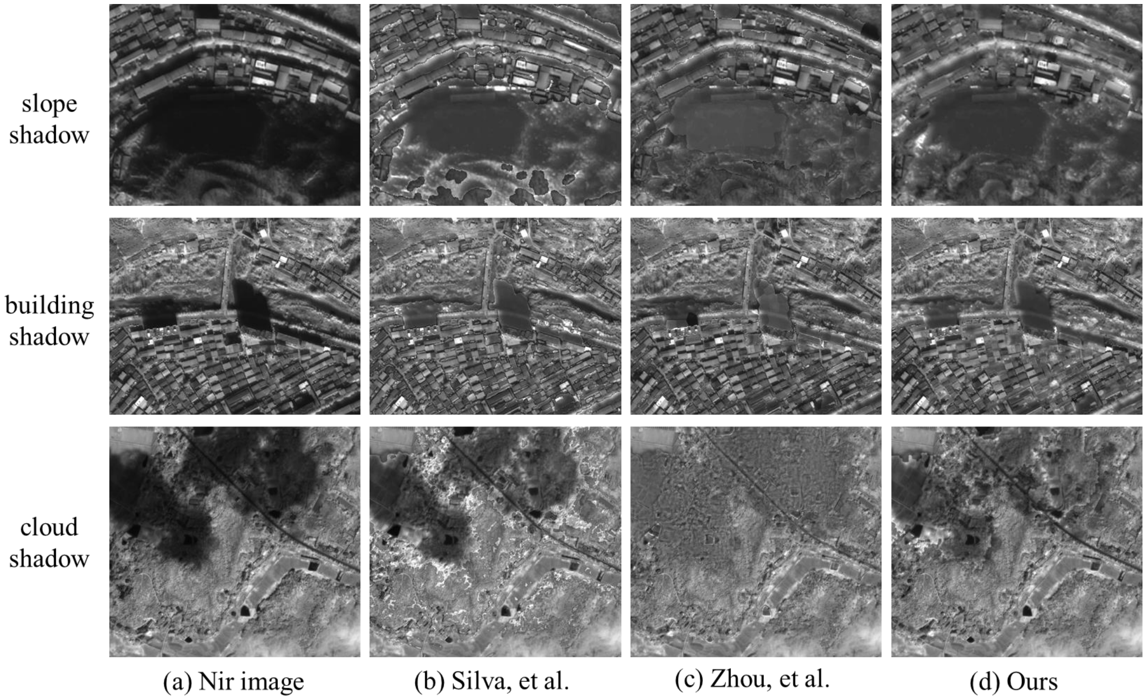


Figure 8. Reconstruction results of the NIR band for shadow. (a) Nir images, (b) the reconstruction results using the method proposed by Silva et al. [6], (c) the reconstruction results using the method proposed by Zhou et al. [7], (d) the reconstruction results using our proposed method.

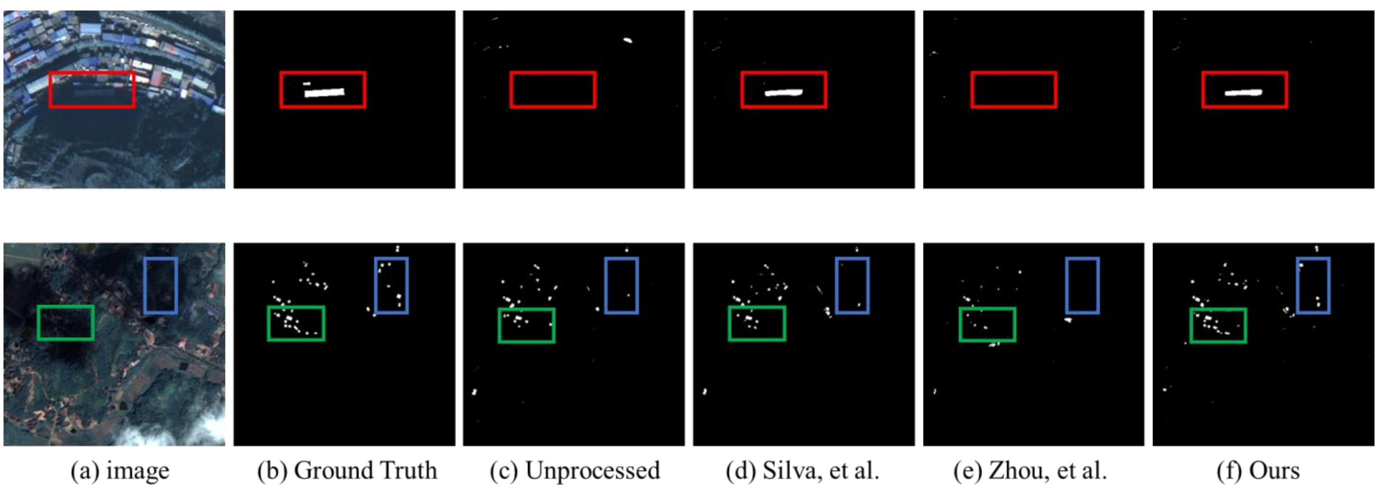


Figure 9. Qualitative results of building extraction after different shadow reconstruction methods. (a) RGB images, (b) Ground truths of buildings in shadow areas, (c) Buildings extracted from the image without shadow reconstruction, (d) Buildings extracted after shadow reconstruction using the method proposed by Silva et al. [6], (e) Buildings extracted after shadow reconstruction using the method proposed by Zhou et al. [7], (f) Buildings extracted after shadow reconstruction using our proposed method.

Due to the unavailability of shadow-free images of the same area at the same time, direct quantitative evaluation of shadow reconstruction is not possible. Therefore, we conducted building extraction experiments on the shadow-reconstructed images using the building extraction model proposed by Wang et al. [25] (The shaded area of the study area contains many buildings.). Building extraction results were used as an indirect quantitative assessment of the shadow reconstruction performance. Table IV presents the quantitative results of building extraction after different shadow reconstruction methods. Compared to the original unprocessed remote sensing image, Silva et al. [6]'s method achieved a 3.59% (F_1) improvement in building extraction. Zhou et al. [7]'s method had a negative impact on building extraction, possibly due to the stretching of building colors to resemble surrounding objects. Although color adjustments were made to shadows from a visual perspective, the adjustment magnitude was too large, leading to significant spectral information deviation. Our method achieved significant improvements in building extraction within shadow regions. The accuracy, recall, and F_1

were improved by 14.47%, 30.95%, and 24.83% respectively. Visualization of the building extraction results after different shadow reconstruction methods is shown in Figure 9. It is evident from the highlighted rectangular boxes in the figure that our method produced favorable results.

In addition, we use filtering operations when processing the penumbra, resulting in a smoother connection between the shadow area and the non-shadow area. Although this visually alleviates the abruptness of the shadow edges, it also loses some details and texture information. If dynamic optical compensation is performed on shadow edges, detailed information may be preserved, but the time cost will be high. How to better handle the penumbra area requires further exploration and attempts.

TABLE IV
QUANTITATIVE RESULTS OF BUILDING EXTRACTION AFTER DIFFERENT SHADOW RECONSTRUCTION METHODS

	Precision (%)	Recall (%)	F_1 (%)
Unprocessed	58.59	38.05	46.14
Silva et al. [6]	60.72	42.11	49.73
Zhou et al. [7]	29.69	11.94	17.03
Ours	73.06	69.00	70.97

C. Algorithm efficiency and limitations

For shadow detection and reconstruction algorithms, in addition to accuracy, efficiency is also an important indicator to measure the quality of the algorithm. Since the image size processed by the deep learning model at a time is 512×512, we uniformly use remote sensing images of 512×512 size as the calculation object. The running time of different algorithms is calculated, as shown in Table V.

TABLE V
DETECTION AND RECONSTRUCTION TIME OF DIFFERENT ALGORITHMS

	Detection time (s)	Reconstruction time (s)
Silva et al. [6]	0.3	0.1
Zhou et al. [7]	0.8	3.5
Liu et al. [24] (using CPU)	6.7	-
Liu et al. [24] (using GPU)	0.2	-
Ours	4.1	1.2

It can be found that the shadow detection method we proposed is the slowest in efficiency compared to other non-deep learning methods. The operation of merging adjacent similar superpixels is time-consuming and takes up most of the time in shadow detection. But it also has two advantages: (1) It reduces the complexity of shadow reconstruction, making the reconstruction efficiency higher than the method of Zhou et al. [7]; (2) It keeps the reconstruction information of adjacent similar shadow superpixels consistent. The deep learning method used by Liu et al. [24] takes the longest time for shadow detection when using CPU operations, but is most efficient when accelerated by GPU. In the future, we can consider writing a GPU running program based on Compute Unified Device Architecture (CUDA) to accelerate the operation of our method.

D. Sensitivity analysis

In superpixel segmentation, the number of superpixels is a key parameter. For a specific image, the number of pixels in a single superpixel determines the number of superpixels. The number of pixels of a single superpixel in the experiment is set to 200. We additionally conducted comparative experiments where the number of pixels in a single superpixel was 100, 150, 250, and 300 respectively. The experimental results are shown in Table VI.

TABLE VI
RESULTS CORRESPONDING TO DIFFERENT NUMBERS OF PIXELS

Number of pixels	Precision (%)	Recall (%)	F_1 (%)
100	91.54	94.62	93.05
150	91.52	94.58	93.02
200	91.47	94.40	92.91
250	90.25	93.11	91.66
300	89.04	91.57	90.29

It can be seen that the smaller the number of pixels in a single superpixel, the more refined the shadow segmentation result is. When the number of pixels is less than 200, the performance improvement is slow, but the running speed will slow down a lot. So, 200 is a reasonable value after considering performance and efficiency.

When binarizing the spectral ratios, we used an automated multi-level global thresholding method. The number of categories set in the experiment is 4. We conducted additional comparative experiments in categories 2, 3, 5, and 6. The experimental results are shown in Table VII.

TABLE VII
RESULTS CORRESPONDING TO DIFFERENT NUMBERS OF CLASSES

Number of classes	Precision (%)	Recall (%)	F_1 (%)
2	92.73	90.25	91.47
3	91.92	92.01	91.96
4	91.47	94.40	92.91
5	91.48	94.38	92.91
6	91.52	93.24	92.37

It can be seen that a smaller number of categories corresponds to poorer shadow segmentation results. This is because a small number of categories will cause dark objects, water bodies, and other objects similar to shadows to be segmented into shadows. Excessive number of categories will also have side effects, because some relatively bright shadows may be segmented into non-shadows. Therefore, setting the analog number to 4 is a moderate value.

V. Conclusion

Aiming at the common occurrence of hill and hillside environments, including slope shadows, building shadows, and cloud shadows, we propose a fully automatic method for shadow detection and reconstruction. For the shadow detection task, a shadow index is constructed based on the spectral characteristics of shadows. The image's color space is transformed into CIELCh, and the initial spectral ratio is computed using hue and luminance. To achieve clear shadow boundaries, the image is subjected to superpixel segmentation, and the resulting segments are merged based on their similarity. The merged superpixels are used to reconstruct the initial spectral ratio, and an automatic multi-level global threshold is applied to extract the shadow mask from the spectral ratio. Considering that nearby objects have higher correlations, non-shadow superpixels adjacent to shadow superpixels are used as references for shadow reconstruction. The contribution weights for information recovery are weighted based on the correlations between superpixels, ensuring that the reconstructed shadows restore their original spectral information as accurately as possible. For special cases of partial shadows, a mean filtering method is employed to reduce the abrupt transition between reconstructed shadows and non-shadows. In order to validate the generalization and effectiveness of the proposed method, detection and reconstruction experiments are conducted on the three types of shadows. Through comparisons with existing state-of-the-art methods, the proposed method demonstrates its applicability in shadow detection and reconstruction of high-resolution remote sensing images in hill and hillside environments, contributing to the further development and application of remote sensing in such areas. For the convenience of researchers, the code in the experiment will be available online at https://github.com/WangZhenqing-RS/Shadow-Detection-and-Reconstruction-RSI.

References

[1] Luo, H.; Wang, L.; Shao, Z.; Li, D. Development of a multi-scale object-based shadow detection method for high spatial resolution image. Remote Sens. Lett. 2015, 6, 59–68.

[2] Sabri, M.A.; Aqel, S.; Aarab, A. A multiscale based approach for automatic shadow detection and removal in natural images. Multimed. Tools Appl. 2019, 78, 11263–11275.

[3] Zhang, H.; Sun, K.; Li, W. Object-Oriented Shadow Detection and Removal From Urban High-Resolution Remote Sensing Images. IEEE T rans. Geosci. Remote Sens. Lett. 2014, 52, 6972–6982.

[4] Liu, X.; Hou, Z.; Shi, Z.; Bo, Y .; Cheng, J. A shadow identification method using vegetation indices derived from hyperspectral data. Int. J. Remote Sens. 2017, 38, 5357–5373.

[5] Tsai, V .J.D. A comparative study on shadow compensation of color aerial images in invariant color models. IEEE Trans. Geosci. Remote Sens. Lett. 2006, 44, 1661–1671.

[6] Silva, G.F.; Carneiro, G.B.; Doth, R.; Amaral, L.A.; De Azevedo, D.F.G. Near real-time shadow detection and removal in aerial motion imagery application. ISPRS J. Photogramm. Remote Sens. Lett. 2018, 140, 104–121.

[7] Zhou, T.; Fu, H.; Sun, C.; Wang, S. Shadow Detection and Compensation from Remote Sensing Images under Complex Urban Conditions. Remote Sens. 2021, 13, 699.

[8] Qu, L.; Tian, J.; He, S.; Tang, Y .; Lau, R.W.H. DeshadowNet: A Multi-context Embedding Deep Network for Shadow Removal. In Proceedings of the Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 2308–2316.

[9] Morales, G.; Huamán, S.G.; Telles, J. Shadow Removal in High-Resolution Satellite Images Using Conditional Generative Adversarial Networks. In Proceedings of the 2018 Annual International Symposium on Information Management and Big Data, Lima, Peru, 3–5 September 2018; pp. 328–340.

[10] Wang, J.; Li, X.; Yang, J. Stacked Conditional Generative Adversarial Networks for Jointly Learning Shadow Detection and Shadow Removal. In Proceedings of the Computer Vision and Pattern Recognition, Salt Lake City , UT, USA, 18–23 June 2018; pp. 1788–1797.

[11] Dong, G.; Huang, W.; Smith, W.A.; Ren, P . A shadow constrained conditional generative adversarial net for SRTM data restoration. Remote Sens. Environ. 2020, 237, 111602.

[12] Ding, B.; Long, C.; Zhang, L.; Xiao, C. ARGAN: Attentive Recurrent Generative Adversarial Network for Shadow Detection and Removal. Proc. IEEE Int. Conf. Comput. Vis. 2019.

[13] Dare P M. Shadow Analysis in High-Resolution Satellite Imagery of Urban Areas[J]. Photogrammetric Engineering Remote Sensing of Environment,2005,71 (2): 169-177.

[14] Highnam R, Brady M. Model-based image enhancement of far infrared images[J]. IEEE Transactions on Pattern Analysis Machine Intelligence,1997,19 (4): 410-415.

[15] Sarabandi P, Yamazaki F, Matsuoka M, et al. Shadow detection and radiometric restoration in satellite high resolution images[C]. international geoscience and remote sensing symposium,2004: 3744-3747.

[16] Vicente T F Y, Hoai M, Samaras D. Leave-One-Out Kernel Optimization for Shadow Detection and Removal[J]. IEEE Transactions on Pattern Analysis Machine Intelligence,2018, 40 (3): 682-695.

[17] Lorenzi L, Melgani F, Mercier G. A complete processing chain for shadow detection and reconstruction in VHR images[J]. IEEE transactions on geoscience remote Sensing,2012, 50 (9): 3440-3452.

[18] Liu W, Yamazaki F. Object-based shadow extraction and correction of high-resolution optical satellite images[J]. IEEE Journal of Selected Topics in Applied Earth Observations Remote Sensing of Environment,2012, 5 (4):1296-1302.

[19] Xiao C, Xiao D, Zhang L, et al. Efficient shadow removal using subregion matching illumination transfer[C]. Computer Graphics Forum,2013: 421-430.

[20] Guo, R., Dai, Q., Hoiem, D., 2013. Paired regions for shadow detection and removal. IEEE Transactions on Pattern Analysis and Machine Intelligence 35 (12), 2956–2967.

[21] Yarlagadda S K, Guera D, Montserrat D M, et al. Shadow Removal Detection and Localization for Forensics Analysis[C]. international conference on acoustics speech and signal processing,2019: 2677-2681.

[22] Achanta, R.; Shaji, A.; Smith, K.; Lucchi, A.; Fua, P.; SüSstrunk, S. SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Trans. Pattern Anal. Mach. Intell. 2012, 34, 2274–2282.

[23] Zhang L, Zhang Q, Xiao C. Shadow Remover: Image Shadow Removal Based on Illumination Recovering Optimization[J]. IEEE Transactions on Image Processing, 2015, 24 (11): 4623-4636.

[24] Liu D, Zhang J, Wu Y, Zhang Y. A shadow detection algorithm based on multiscale spatial attention mechanism for aerial remote sensing images. IEEE Geoscience and Remote Sensing Letters, 2021, 1-5.

[25] Wang, Z.; Zhou, Y.; Wang, F.; Wang, S.; Qin, G.; Zou, W.; Zhu, J. A Multi-Scale Edge Constraint Network for the Fine Extraction of Buildings from Remote Sensing Images. Remote Sens. 2023, 15, 927.

Zhenqing Wang was born in Shandong, China, in 1997. He received the B.S. degree in geographic information science from China University of Petroleum., Qingdao, China. He is currently working toward the Ph.D. degree at Airspace Information Research Institute, Chinese Academy of Sciences, Beijing, China and University of Chinese Academy of Science, Beijing, China.

His research interests include remote sensing semantic segmentation and object detection based on deep learning.

Yi Zhou received the B.S. degree in natural resources from Nanjing University, Nanjing, China, in 1986.

She is currently a Researcher with the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China. Her research interests include remote sensing of environment and disaster, remote sensing of urban systems and urbanization, and quantitative monitoring and evaluation of environmental remote sensing.

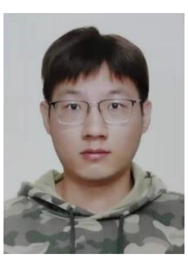
Futao Wang received the B.S. degree in resource and environmental engineering from Shandong University of Technology, Zibo, China, in 2005, the M.S. degree in mapping and geographic information from Guilin University of Technology, Guilin, China, in 2008, and the Ph.D. degree in cartography and geographical information system from the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China, in 2011.

He is currently an Associate Research Fellow with the Aerospace Information Research Institute, Chinese Academy of Sciences. His research interests include remote disaster remote sensing monitoring and assessment.

Dr. Wang is currently a member of the Youth Innovation Promotion Association of Chinese Academy of Sciences, and a Reviewer for International Journal of Disaster Risk Science and Journal of Remote Sensing.

Shixin Wang received the B.S. degree in natural resources from Nanjing University, Nanjing, China, in 1987, and the M.S. degree in cartography and geographical information system from Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China, in 1990.

He is currently a Researcher with Aerospace Information Research Institute, Chinese Academy of Sciences. His research interests include remote sensing of environment and disaster.



Gang Qin was born in Henan, China in 1997. In 2020, he obtained an undergraduate academic degree in software engineering from Zhengzhou University. Currently, he is pursuing a Ph.D. at the Aerospace Information Research Institute, Chinese Academy of Sciences, and the University of Chinese Academy of Sciences in Beijing, China.

His research interests are in remote sensing image classification and knowledge graph.

Jinfeng Zhu received the B.S. degree in geographic information system from Northwest University, Xi’an, China, in 2008, the M.S. degree in Earth system science from Lanzhou University, Lanzhou, China, in 2011, and the Ph.D. degree in mapping and geographical information system from the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China, in 2014.

He is currently an Associate Research Fellow with the Aerospace Information Research Institute, Chinese Academy of Sciences. His research interests include remote sensing of disaster and environment.