Object-oriented detection of building shadow in TripleSat-2 remote sensing imagery

Yuxuan Liu,[a] Yuchun Wei,[a,b,c,]* Shikang Tao,[a] Qiuping Dai,[b]
Wenyao Wang,[a] and Mengqi Wu[a]
[a]Nanjing Normal University, School of Geography, Nanjing, China
[b]Ministry of Education, Nanjing Normal University, Key Laboratory of
Virtual Geographic Environment, Nanjing, China
[c]Nanjing Normal University, Jiangsu Center for Collaborative Innovation in
Geographical Information Resource Development and Application, Nanjing, China

Abstract. The projection of objects on the earth's surface caused by the sunlight produces shadows. They are inevitable in high-spatial-resolution satellite remote sensing images and reduce the accuracy of change detection, land cover classification, target recognition, and many more applications. Dark-colored land covers in these satellite images, such as water bodies, road, and soil, appear with similar spectral properties as those of shadows and often result in difficulty in shadow detection, especially in complex urban settings. We propose an object-oriented building shadow extraction method and tested it using six selected study areas from TripleSat-2 satellite imagery with 3.2-m spatial resolution. The method's main steps include (1) selecting six image features that can highlight the shadow information and then segment the image based on edge; (2) extracting shadow region based on multiple object features; and (3) masking nonbuilding shadow regions by the shadow and dark object separation index, image features including spectral, textural, and geometric features, and contextual information. The average precision, recall, and F1-score of the shadow detection were 85.6%, 88.6%, and 87.0%, respectively, and the ranges were 73.0% to 91.0%, 76.6% to 94.1%, and 74.7% to 91.2%, respectively. Compared with multiscale segmentation, edge-based segmentation is more efficient and helpful to completely and accurately extract shadows. © *The Authors. Published by SPIE under a Creative Commons Attribution 4.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI.* [DOI: 10.1117/1.JRS.14.036508]

1 Introduction

High-spatial-resolution satellites (HSRS) images provide rich detailed information about land covers and have been used widely in different remote sensing applications.[1] However, the shadows in this type of image are obvious and cannot be ignored,[2] such as building shadows, vegetation shadows, river bank shadows, road shadows, and bridge shadows, consequently, shadow detection has become an indispensable image processing content in remote sensing information extraction.[1,3–5]

Shadow appears in two different effects in HSRS images. On the one hand, shadows can be regarded as a kind of useful information in three-dimensional (3-D) reconstruction, building position recognition, and height estimation.[6,7] On the other hand, shadows present adverse influence in information extraction and applications,[8–10] such as reducing the accuracy of change detection and land cover classification. More obvious negative effects appear in cities, where many shadows are distributed and are often confused with dark land covers.

Scholars are devoted to improving the detection accuracy of shadow in remote sensing images[11–15] and have proposed many methods that can be divided into four types:[4] model-based,

---

*Address all correspondence to Yuchun Wei, E-mail: weiyuchun@njnu.edu.cn

physics-based, machine learning, and property-based. In a multispectral HSRS image with only four bands and a spatial resolution between 1 and 5 m, the obvious shadow limits the application of model-based,[16,17] physics-based,[4,18] and machine learning methods[19–24] because these methods need more parameters, auxiliary data, training samples, etc.

Property-based methods were widely used and have the advantage of not requiring *in situ* or ancillary data, taking into account the shadow properties generally directly deduced from the image data. Some examples of the method are simple histogram thresholding, invariant color model, and object-oriented algorithm. The used features include the image spectrum,[13,25,26] texture,[11,27–30] edge,[31–35] spatial context,[14,36,37] etc. Typical spectral features include the C3 component;[38] the difference between intensity and saturation of the hue, saturation, and intensity (HSI) model;[39] the normalized index;[40] the hue and brightness in color spaces such as HSI, hue, saturation, and value (HSV), hue, chroma, and value (HCV), luminance, hue, and saturation (YIQ), and luma, blue-difference chroma, and red-difference chroma (YCbCr);[40,41] the CIELCh color space.[42] Texture features include the gray level co-occurrence matrix,[27,29] local binary pattern,[29] edge,[28] and Gabor filter.[30] Edge features include the ratio edge[32] and weighted edge gradient ratio.[35] Spatial context features use the morphological building index and morphological shadow index to first determine the region of the building and shadow and then extract shadow by the distance between the building and shadow.[14,36,37] Shadow information can also be directly segmented by the threshold methods, such as the mean of the two peaks in the histogram[8] and the Otsu method maximizing the difference between the classes[43] or by manual selection.[44]

Most of the property-based methods are pixel-based and often work well. Studies suggest that the image features combined with the direction information and the positional relationship between the shadow and the object can improve shadow detection accuracy.[31,34] However, these methods are scene dependent and have a low degree of automation.

Taking advantage of high spatial resolution to segment the image pixels into "objects" and extract more information can often obtain higher precision.[45,46] This object-oriented method received a lot of attention in the shadow detection of HSRS images, and multiscale segmentation is most often used.[15,47–49] Related tools were provided by the typical remote sensing software such as ENVI,[50] eCognition,[51] Definiens,[52] and ERDAS Imagine.[53] However, object-oriented image processing often requires human–computer interaction and many tests to determine the relevant parameters and is difficult to convert small shadow patches into independent objects.[15]

In order to improve the accuracy and the automation degree of shadow detection in multispectral HSRS imagery, taking the most widely distributed building shadows in the city as the subject, this paper constructs a fast building shadow detection method by coupling the advantages of object-oriented and property-based methods. The new method uses the orthorectification remote sensing image as the input and can mask dark land cover that easily confused with building shadow. Technical innovations of the method include: (1) using multiple image features edge-based image segmentation instead of multiscale segmentation to obtain a more complete object, (2) using remote sensing indices to separate shadow from nonbuilding shadow regions, and (3) integrating the spectrum features with texture, geometric morphological, and context information to further remove the nonbuilding shadow areas.

The content is organized as follows. Section 2 describes the data set. Section 3 describes the proposed method in detail. Section 4 discusses the shadow detection result and accuracy. Section 5 gives the conclusions.

2 Dataset

Satellite remote sensing images of urban region are often covered by various land covers and the number, length, and brightness of shadows in such images can vary widely. Water bodies, roads, and soils appear dark and their spectra are similar to that of the shadows in an image, increasing the difficulty of shadow detection.

The TripleSat-2 multispectral image on November 27, 2016, in Nanjing, China, was used in this paper. The TripleSat-2 orbital altitude is 651 km, its spatial resolution is 3.2 m, and
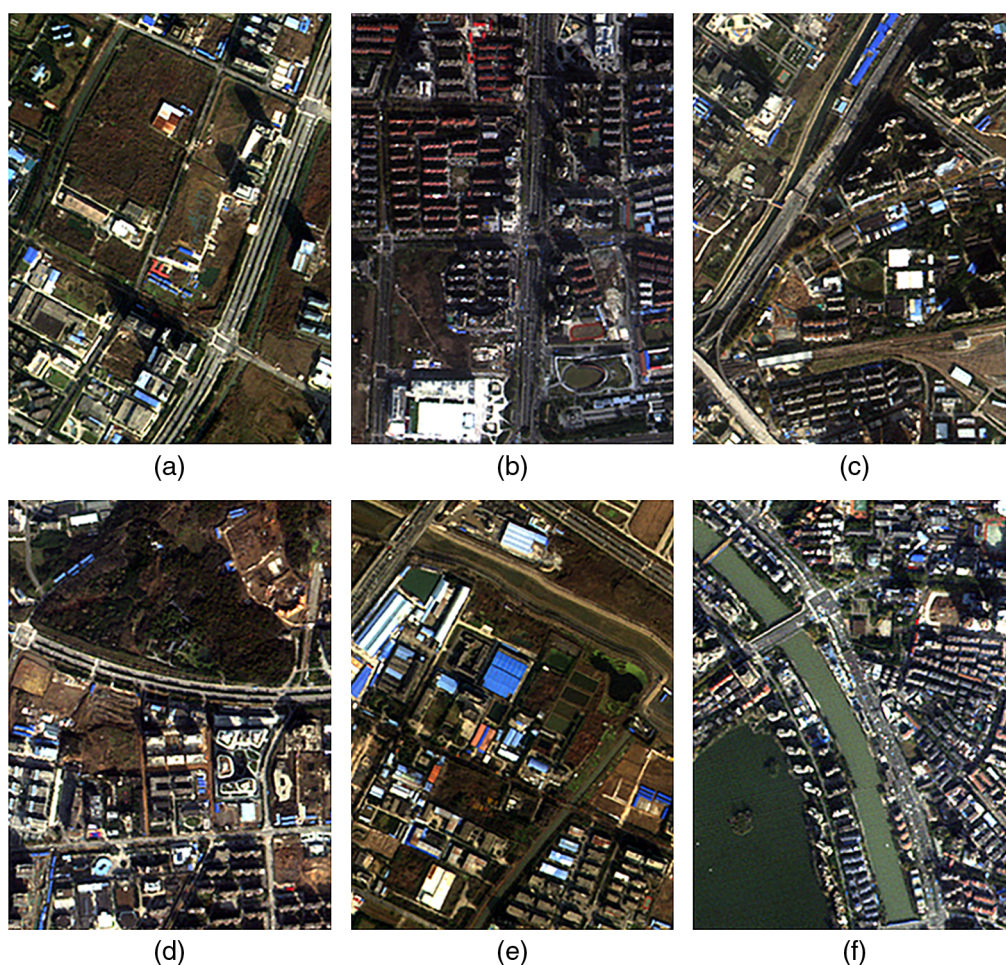
Fig. 1 True-color composite images with 2% linear stretch, six typical cases from (a)–(f).

it includes four bands: blue (B), green (G), red (R), and near-infrared (NIR). Six typical cases in one image were selected and are shown in Fig. 1. The case image size is 400 pixels × 300 pixels. The distribution characteristics of the shadows and dark land covers are shown in Table 1 and are used in Sec. 3.3.

Among the dark land covers, water body is the first to be easily confused with building shadows, as seen in Figs. 1(a), 1(c), 1(e), and 1(f). Next is asphalt road and dark vegetation, as seen in Figs. 1(a), 1(b), 1(d), and 1(e). The shadows in (a), (c), and (d) are longer and easier to detect, whereas those in (b), (e), and (f) are short and are difficult to detect.

Table 1 Dark land covers in case images.

Case	Dark water body	Dark roads, soils, and buildings
(a)	Yes	Buildings and soils
(b)	No	Soils and roads
(c)	Yes	Buildings
(d)	No	Buildings and soils
(e)	Yes	Soils
(f)	Yes	No

3 Proposed Method

As the dark area formed by the sun's rays project onto an opaque object, shadow is divided into a self-shadow (the part of the object that is not illuminated by direct light) and a cast shadow (the shadow projected by the object based on the direction of the light source),[54] and the latter can be further divided into the umbra and penumbra. The direct light is completely obscured in the umbra, but only a small portion of the direct light is blocked in the penumbra. The penumbra is usually located between the umbra and the nonshadow and may be ambiguous,[4] resulting in a blurry boundary between the shaded and nonshaded areas.

In satellite remote sensing image, due to the limitations of the quality and spatial resolution, the presence of a cast shadow is obvious, while that of self-shadow is often weak. In a cast shadow, the penumbra accounts for a small portion only[8] and is difficult to distinguish from the umbra. Therefore, the cast shadow was extracted in this paper.

The object-oriented shadow method consists of four parts: (1) extracting image features that can enhance the shadow information and then segmenting the image based on edge detection to obtain the object; (2) extracting the suspected shadow regions based on the statistical properties of the object; (3) refining the shadow combined the spectrum, texture, and geometric features with spatial context information to remove other nonbuilding shadows; and (4) postprocessing.

The orthorectification image with digital number (DN) is used as the input. The four bands were denoted as B, G, R, and NIR in order of increasing central wavelength.

3.1 Object Segmentation

3.1.1 Feature selection and extraction

The selected feature should highlight the shadow information and be positive to the image segmentation to obtain objects with the closed boundary. Previous studies[38,39,55] have shown that (1) the brightness (intensity) of the shadow area is low due to the occlusion of sunlight; (2) the atmospheric Rayleigh scattering effect is sensitive to the short-wavelength blue and violet bands, but insensitive to the NIR band; and (3) the C3 component of the shadow area has a high value and a high saturation. Therefore, six image features were selected from C1C2C3, the HSV color space, and the remote sensing index: C3, $RATIO_{H\_V}$, $RATIO_{S\_V}$, NIR, the normalized difference vegetation index (NDVI), and the visible atmospherically resistant index (VARI).

(1) C3

The C3 component of the C1C2C3 color space highlights the difference between shadow and greenish objects (water bodies, green areas, plastic runways, etc.)[54]

$$C3 = \arctan\left[\frac{B}{\max(R, G)}\right].$$ (1)

(2) $RATIO_{H\_V}$ and $RATIO_{S\_V}$

Shadow appears more clearly in the HSV color space than in the RGB color space. RGB is converted to HSV according the following equations:[56]

$$H = \begin{cases} \theta & (g \geq b) \\ 2\pi - \theta & (g < b) \end{cases},$$ (2)

$$S = 1 - \frac{3 * \min(r, g, b)}{r + g + b},$$ (3)

$$V = \frac{r + g + b}{3},$$ (4)

where $\theta = \arccos \frac{(r-g)+(r+g)}{2*\sqrt{(r-g)^2+(r-b)*(g-b)}}$, and r, g, and b are the normalized R, G, and B band values, respectively. The values of H, S, and V are all in the range from 0 to 1.

Based on the relationship between H and V in the HSV color space, Tsai constructed the RATIO$_{H\_V}$ index[41]

$$\text{RATIO}_{H\_V} = \frac{H+1}{V+1}. \tag{5}$$

Based on the relationship between S and V, Ma et al.[57] constructed the normalized saturation-value difference index (NSVDI)

$$\text{NSVDI} = \frac{S-V}{S+V}. \tag{6}$$

Referring to the selected S and V in NSVDI, the ratio of S to V can be formed

$$\text{RATIO}_{S\_V} = \frac{S}{V}. \tag{7}$$

A comparison in Fig. 2 showed that RATIO$_{S\_V}$ can highlight the shadow better than NSVDI. In Fig. 2, (a) is a local part of Fig. 1(b), (b)–(e) are images of RATIO$_{S\_V}$, NSVDI, and their edges. As seen in Fig. 2, RATIO$_{S\_V}$ better highlights the edges of the shadow and dark road.

Figure 3 shows the relationship between RATIO$_{S\_V}$ and NSVDI in Figs. 2(b) and 2(c). At higher values, the RATIO$_{S\_V}$ index rises rapidly and presents a better discrimination.
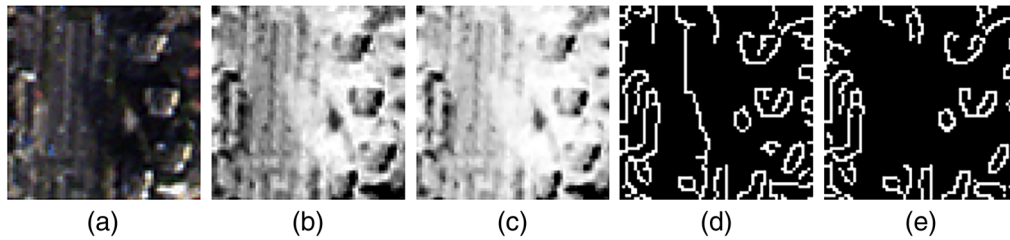

(a)   (b)   (c)   (d)   (e)

Fig. 2 Comparison of RATIO$_{S\_V}$ and NSVDI: (a) local image with true-color composite, (b) RATIO$_{S\_V}$, (c) NSVDI, (d) edge of (b), and (e) edge of (c).
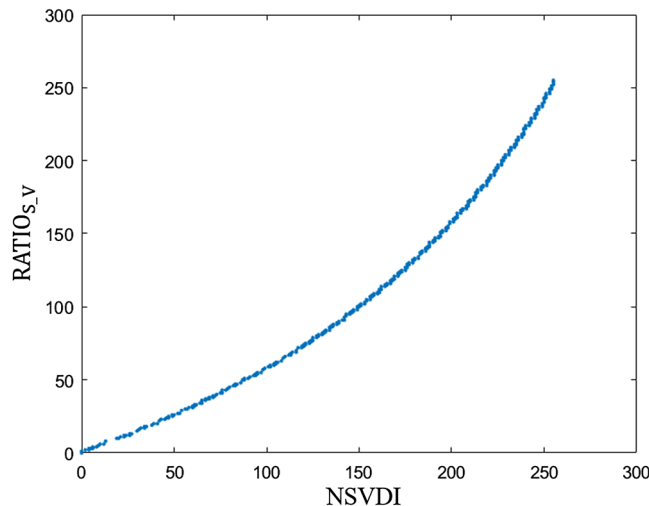

Fig. 3 Relationship between NSVDI and RATIO$_{S\_V}$.

(3) NIR, NDVI, and VARI

In the NIR band, it is easier to distinguish shadow from nonshadow than in the visible band.[4] NDVI[58] can highlight the DN difference between a shadow area and vegetation.[59] VARI is less sensitive to atmospheric effects[60] and is used as a supplement to NDVI

$$\text{NDVI} = \frac{\text{NIR} - \text{R}}{\text{NIR} + \text{R}}, \tag{8}$$

$$\text{VARI} = \frac{\text{G} - \text{R}}{\text{G} + \text{R}}. \tag{9}$$

3.1.2 *Image segmentation*

The image features C3, $\text{RATIO}_{\text{H\_V}}$, $\text{RATIO}_{\text{S\_V}}$, NIR, NDVI, and VARI were first linearly stretched by 2% to enhance the information, and then their value ranges were changed to the range [0, 255].

The image segmentation procession is shown in Fig. 4. First, the edges of each image feature are detected and then the union set of these edges is taken as the image edge to ensure that the shadow region is not missed. Pixels are connected to get the segmented image if their edges touch.

The Canny operator[61] in MATLAB® 2018b[62] is used to detect the edges of each image feature. The edge's union is postprocessed by the skeleton and bridge operator[62] with a window size of 3×3 to make the edge width one pixel. Finally, a pixel search is made around the outermost image boundary and the current pixel is set as an edge if it touches the edge.
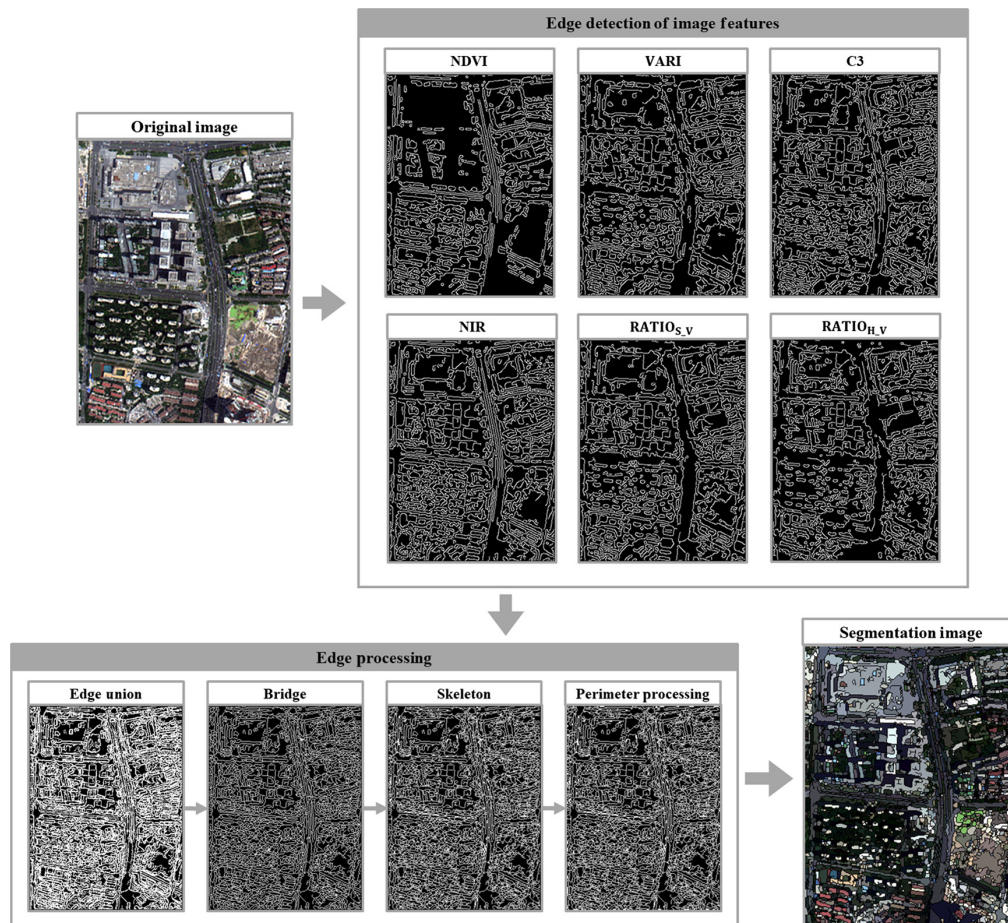
The nonedge areas are the image object.

Fig. 4 Example of image segmentation for shadow detection.

3.2 Suspected Shadow Area Detection

The shadows are initially detected by the image features C3, NSVDI, and NDVI. The obvious shadows should have higher C3, lower NSVDI, and NDVI.

First, the average of each band value is calculated as the attribute of the image object, then the C3, NSVDI, and NDVI of the object are calculated. The Otsu method[43] is used to determine the threshold to divide the feature into a high-value area and a low-value area. The intersection of the low-value area in NDVI and NSVDI with the high-value area of C3 was used as an initial shadow area S0.

3.3 Nonbuilding Shadow Area

A nonbuilding shadow area is shadow caused by objects other than buildings, consisting of dark land cover such as water bodies, buildings, soil, and roads. It is divided into two types: (1) type 1, the spectrum differs from that of building shadow, including dark buildings, dark soil, and dark roads. (2) Type 2, the spectrum is similar to that of the shadow, including dark water bodies and very dark land cover.

Type 1 area was extracted by the spectral features first; then S0, after being masked with type 1, was used to get the type 2 area by the multirule.

3.3.1 Type 1 area

Rayleigh scattering is inversely proportional to the wavelength. The smallest difference in radiant energy between the shaded and nonshaded area appears in the blue band and the largest in the NIR band.[63] In addition, the shaded area has a higher saturation (S) and a lower value (V) in the HSV color space.[39] Based on these conclusions, the shadow and dark object separation index (SDSI) is proposed to expand the difference between the shadow and the nonshadow:

$$\mathrm{SDSI} = \alpha \frac{\mathrm{B}}{\mathrm{NIR}} + (1 - \alpha) \frac{\mathrm{S}}{\mathrm{V}}, \tag{10}$$

where B and NIR are the DNs of band B and band NIR in the original image, S and V are calculated by Eqs. (3) and (4), and the range of values of B/NIR and S/V are all normalized to 0 to 1. α is an environmental parameter indicating the proportion of the shaded area in an image, $\alpha \in [0, 1]$.

Confusion between building shadows and dark roads can be reduced by using a different α. Taking the original image in Fig. 4 as an example, the shadow with the typical α is shown in Fig. 5. Figure 5(a) ($\alpha = 0$) has the best dark removal effect. With an increase of α, the effect worsens, and the building shadows on some vegetation are removed when $\alpha \geq 0.8$, as in Figs. 5(f) and 5(g). Section 4.2 further discusses the effect of α on the accuracy of shadow detection.

3.3.2 Type 2 area

After masking the type 1 area from S0, the type 2 area is extracted by integrating spectral, textural, geometrical, and spatial context information.

In a high-spatial-resolution image, the shadows spectra often have the highest similarity to the spectrum of dark water bodies. According to the object boundary shape, a water body is divided into slender and nonslender.

(1) Slender water body

A river is an example of a slender water body, as shown in Fig. 1(c). Its shape can be described by the ratio d of the long-axis length l_1 and the short-axis length l_2 of the object in image
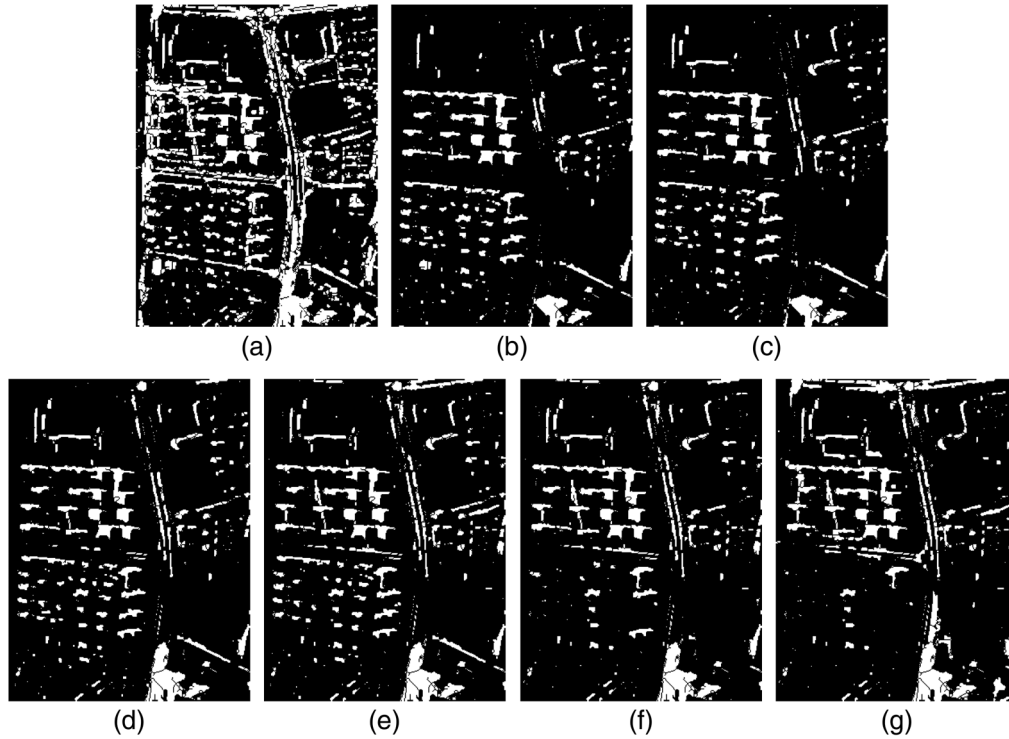
Fig. 5 Shadow distribution based on the typical α and SDSI. (a) Suspected shadow area S0 of image in Fig. 4, (b) $\alpha = 0$, (c) $\alpha = 0.2$, (d) $\alpha = 0.4$, (e) $\alpha = 0.6$, (f) $\alpha = 0.8$, and (g) $\alpha = 1$.

$$d = \frac{l_1}{l_2}. \tag{11}$$

For a slender water body, $d > T_d$ and $l_1 > T_l$. The threshold T_d is set to 10 and T_l is set to 50 here.

(2) Nonslender water body

A nonslender water body is shown in Figs. 1(a), 1(e), and 1(f), where (c) is the river, (e) is the pool, and (f) is the lake.

Spectral and textural features are used to extract this kind of water body. Compared with the shadow spectrum, the water body spectrum often appears as a larger value in the G band and a smaller value in the NIR band. Based on this, some remote sensing indices were designed to distinguish water bodies, such as (G-NIR)/(G + NIR) (normalized difference water index, NDWI),[64] G/NIR,[64] and G-NIR.

It is also possible to extract a water body by segmenting the G band and NIR band. At first, the image was segmented by the Otsu method[43] into the high- and low-value areas, then the intersection of the high-value area of G and the low-value area of NIR can be taken as the water body.

In general, the pixel value in the area of a water body tends to be uniform. This can be described by the sum of the standard deviations (SSD) of the pixel values of each band, where the SSD is low in water bodies and high in shadow areas.

Four methods are used to extract the shadow when a nonslender water body exists. (1) The NDWI low-value area intersects with the SSD high-value area. (2) The G/NIR low-value area intersects with the SSD high-value area. (3) The G-NIR low-value area intersects with the SSD high-value area. (4) The G low-value area is combined with the NIR high-value area, and then intersects with the SSD high-value area. The results are shown in Fig. 6.

For shorter shadows in cases (e) and (f), NDWI and G/NIR are more suitable to distinguish shadows with water bodies. For longer shadows in case (a), method 4 is more suitable. More test cases containing water bodies validate this conclusion.
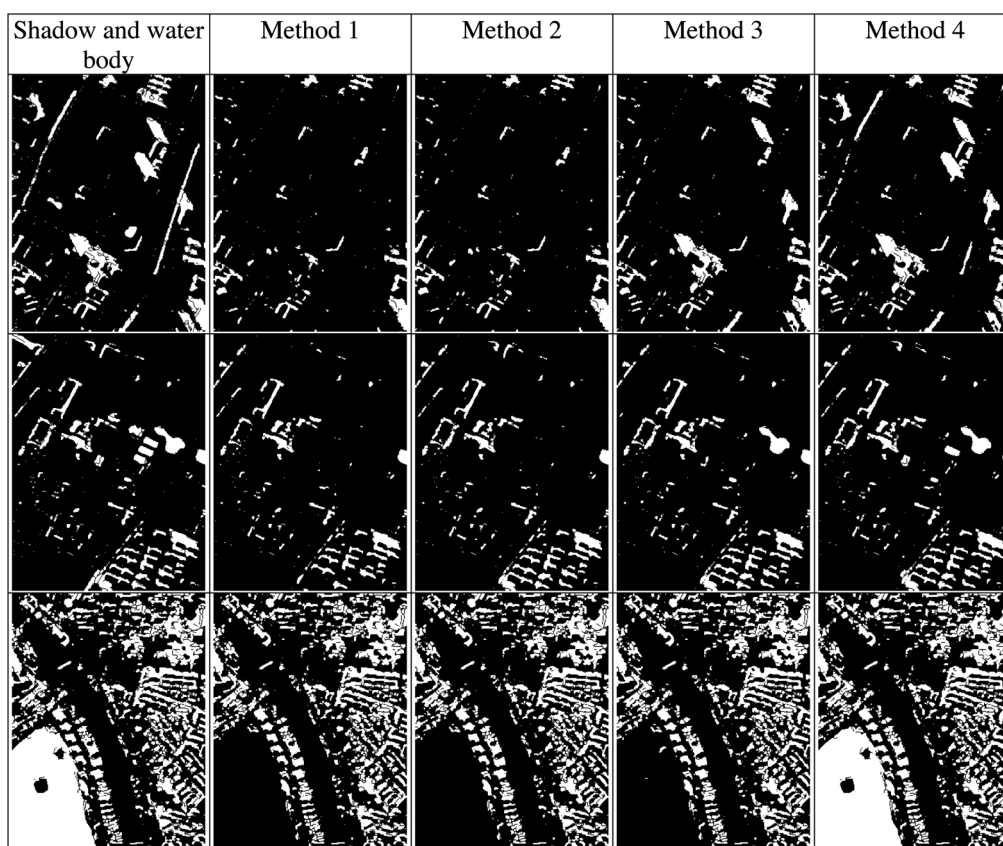
Shadow and water body	Method 1	Method 2	Method 3	Method 4

Fig. 6 Water body extraction of case images (a), (e), and (f) from top to bottom by the four methods from left to right. In case (a), ponds distributed in the upper left and middle, and rivers in the upper left and right; in case (e), pools distributed in the middle; and in case (f), lake distributed in the lower left.

After masking the water body, the shadow accuracy in Fig. 6 is decreased in the order of method 1, method 2, method 3, and method 4. Method 4 may not extract the water body and the other methods may remove some shadows, but can extract the water body more completely.

The appropriate method depended on the image features of a specific image. For example, method 4 is the best for case image (a), whereas method 1 or 2 is the best for case images (e) and (f). As described in Sec. 3.1.1, the ratio G/NIR is more distinguishable than NDWI in the high-value areas. More multispectral test images of TripleSat-2 showed that method 2 or method 4 often achieves better accuracy.

In a city, buildings and water bodies are often surrounded by green plants, trees often grow on both sides of a road, and a building shadow is adjacent to the building. Therefore, the following rules were constructed (S1 denotes the image after removal of the water body):

(1) *Extract vegetation coverage (VC) area.* NDVI > 0 is regarded as VC. In addition, the blue building and the vegetation appear similar in the NDVI image in some cases, and the obvious blue building area can be eliminated by $B < T_v$, where threshold T_v is a value between the max DN of the vegetation area and the min DN of the blue building area. T_v is not used in our six case images.

(2) To eliminate edges, the VC is dilated by $5 * 5$ to obtain VC_{dilate}, and S1 is dilated by $3 * 3$ to obtain I_{dilate}.

(3) VC_{dilate} and I_{dilate} are combined to obtain VI, which represents the collection of vegetation, shadows, and water bodies.

(4) Traversing through region S1, most pixels in VI (usually 95% or more) are water and are assigned as 0, otherwise, the pixel is a shadow and assigned as 1.

The image S1 processed by the above steps is expressed as image S2.

3.4 *Postprocessing*

To obtain a more unbroken shadow distribution, S2 is processed as follows:

(1) Remove any small patches with an area smaller than T_a, where $T_a = 9$ (3*3, basic window size) in case images.
(2) Perform a morphology closing operation.
(3) Fill in the holes in each region with $T_h < 30$, where T_h is the maximum pixel numbers of the hole.
(4) Pixel search at the outermost boundary of the image and set the current pixel as an edge by the four-neighbor rule.

3.5 *Threshold*

The difference between the bright object and the shadow was segmented by the Otsu method,[43] including the segmentation of NDVI, NSVDI, and C3 in Sec. 3.2, and SDSI, G and NIR, and SSD in Sec. 3.3.

The specific thresholds include T_d and T_l in Sec. 3.3, and T_a and T_h in Sec. 3.4. These thresholds need to be specified for a specific image.

4 Results and Discussion

4.1 *Shadow Detection Accuracy*

The shadow detection workflow is divided into four steps: step 1 image segmentation; step 2 extraction of the shadow area; step 3 nonbuilding shadow area removal; and step 4 postprocessing. Step 3 can be further divided into step 3.1, removal of type 1 area, and step 3.2, removal of type 2 area. The necessary steps include step 1, step 2, and step 4; step 3 is optional.

In the images in Fig. 1, cases (a), (c), and (e) were processed by step 1, step 2, step 3, and step 4; cases (b) and (d) by step 1, step 2, step 3.1, and step 4; and case (f) by step 1, step 2, step 3.2, and step 4.

The parameter α of SDSI in step 3 is 0.5, which means that the two components in SDSI are equally important. The images containing water bodies are cases (a), (e), and (f). According to Sec. 3.3.2, the water body removal method in case (a) is method 4, and that of cases (e) and (f) use method 2.

Taking manually identified shadows as the ground truth, precision (P_s), recall (U_s), and F1-score (F_1) were used to evaluate the detection accuracy as follows:

$$P_s = \frac{TP}{TP + FN}, \tag{12}$$

$$U_s = \frac{TP}{TP + FP}, \tag{13}$$

$$F_1 = 2\frac{P_s U_s}{P_s + U_s}, \tag{14}$$

where TP denotes the number of pixels correctly detected as shadows, FN denotes the number of pixels in which shadows were detected as nonshadows, and FP denotes the number of pixels in which nonshadows were detected as shadows.

The results of the shadow detection in six case images are shown in Fig. 7 followed by the four steps. The shadow ground truth was obtained by visual interpretation and is shown in Fig. 8.

In Fig. 7, step 1 shows that the segmented image is consistent with the source image, and step 2 shows that SDSI can better eliminate the interference of other dark land covers. In step 3, all water bodies in the images, except in (a), are better removed. In step 4, the pond in the upper left
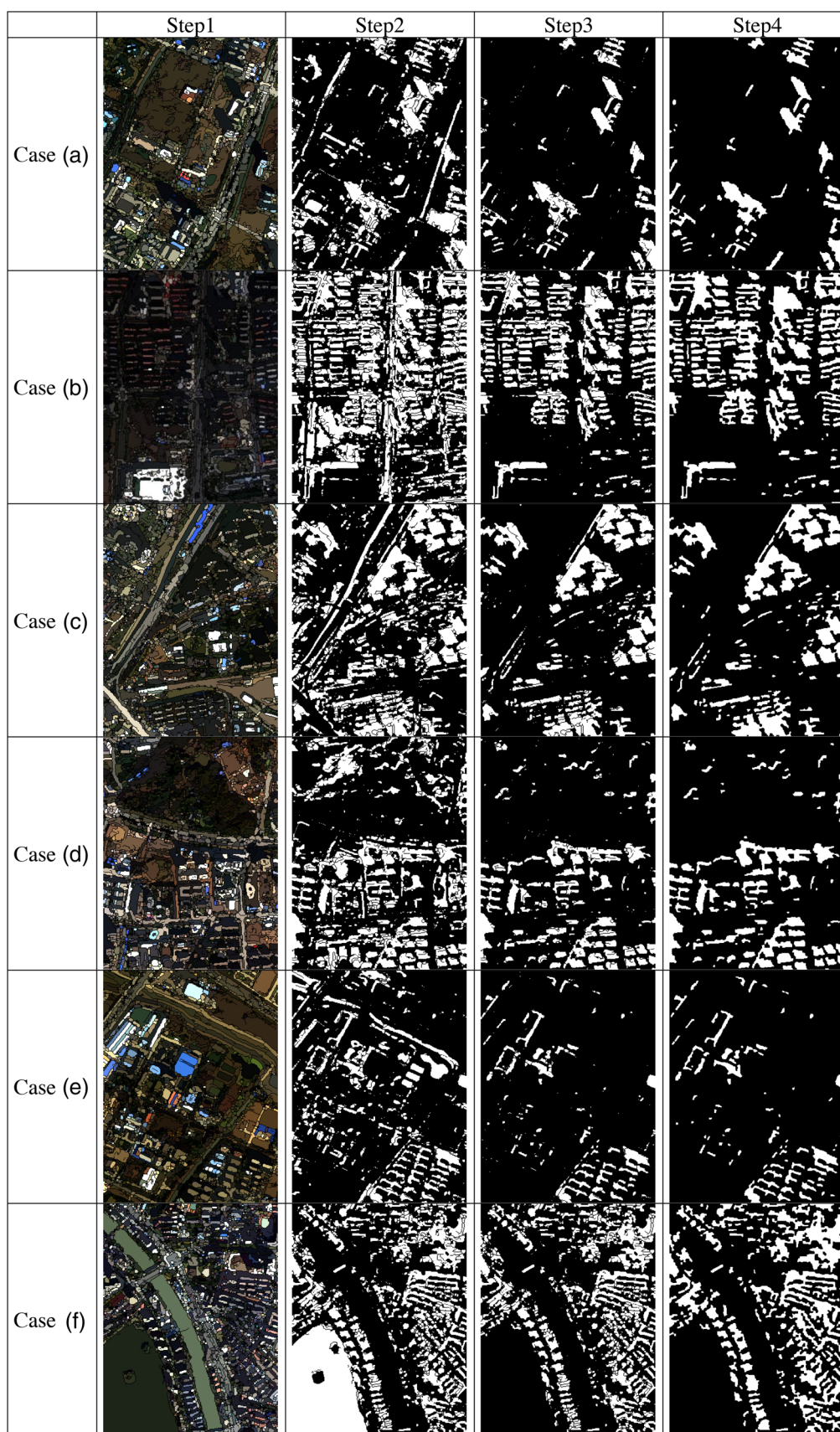
	Step1	Step2	Step3	Step4
Case (a)				
Case (b)				
Case (c)				
Case (d)				
Case (e)				
Case (f)				

Fig. 7 Shadow detection procession for the four steps from left to right, for cases (a)–(f) from top to bottom.
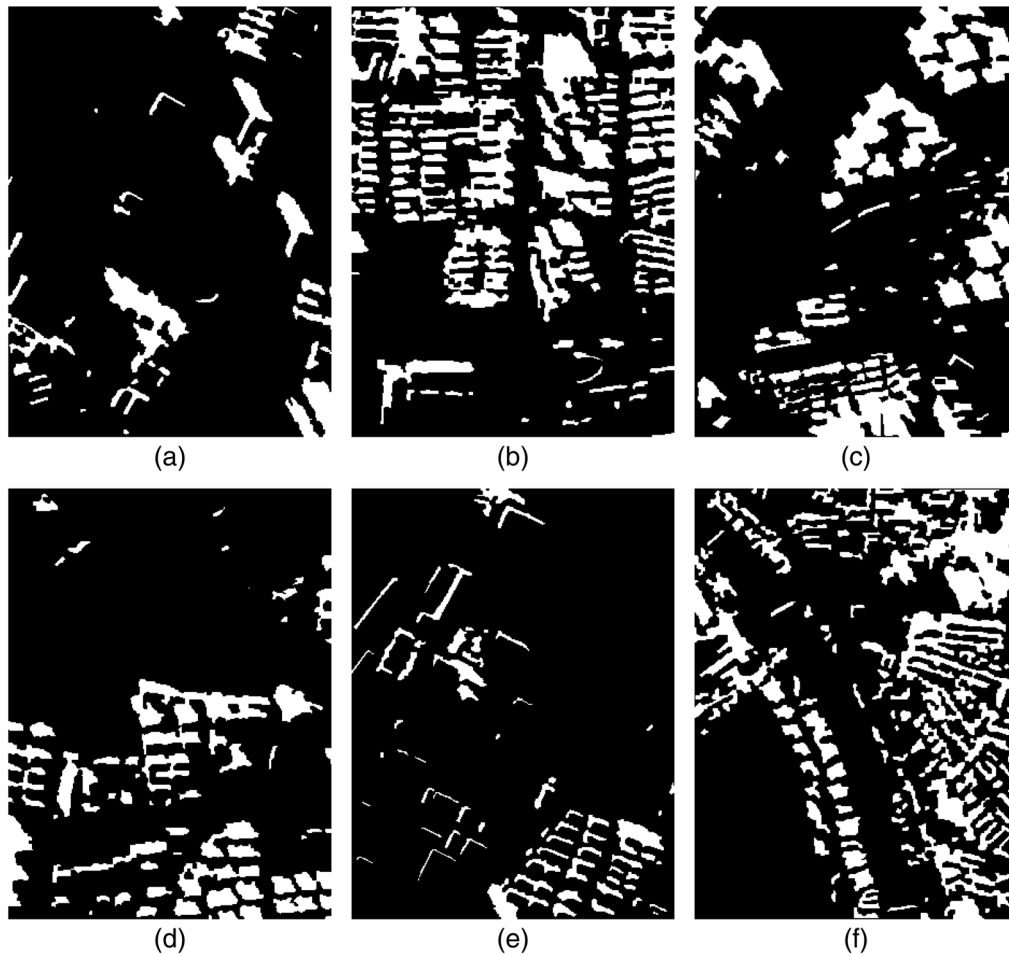
Fig. 8 The actual shadows in cases (a)–(f) based on visual interpretation.

corner of case (a) is removed by the rule of contextual information, and the river in cases (a) and (e) is removed by the rule of geometric shape.

The shadow detection accuracy is shown in Table 2. The average precision, recall, and $F1$-score were 85.6%, 88.6%, and 87%, respectively, and the ranges were 73.0% to 91.0%, 76.6% to 91.4%, and 74.7% to 91.2%, respectively.

According to the $F1$-score and visual comparison, cases (b) and (f) have the best accuracy without obvious errors or omissions. The omissions in cases (a) and (e) mainly appear in the area of light shadows and small shadows. Overextraction of the vegetation shadow appears in cases (a), (c), and (d). Case (c) also extracts the road shadow.

Table 2 Shadow detection accuracy of case image (%).

Case	Precision	Recall	$F1$-score
(a)	88.5	89.0	88.8
(b)	90.9	91.5	91.2
(c)	91.0	88.4	89.7
(d)	84.4	94.1	89.0
(e)	73.0	76.6	74.7
(f)	85.5	91.7	88.5

The error of the precision is mainly related to the vegetation shadow. The error of the recall is mainly due to the fact that the penumbra is often blurred on the image and is difficult to identify and extract. Owing to the skeleton and other operations in image segmentation, small shadows of <3 pixels in case images cannot be extracted.

The lowest accuracy appears in (e) where the shadow is short and small and there is a large area of dark soil and a dark water body as interference.

4.2 *Parameter Assignment*

The shadow detection of Sec. 4.1, step 3 and step 4, are affected by parameters. Step 3 needs to determine the parameter α of SDSI. Step 4 needs to assign the number of pixels of the largest fragment and the largest hole according to the actual image.

The following is a discussion of SDSI and its parameter α.

In remote sensing images, the sensitivity of dark land cover in different color spaces is different. We integrated an RGB color space and HSV color space to deal with this complex situation, and two components, B/NIR and S/V, were combined in SDSI. More tests showed that S/V can better highlight the difference between shadows and dark roads in summer images, whereas B/NIR performs better in winter images.

Parameter α indicates the proportion of the shaded area representing the degree of influence of the segmentation of B/NIR and S/V on the result. The larger α is, the greater the influence of B/NIR on the result is, and the smaller the influence of S/V on the result is. The parameter α is changed from 0 to 1 with an increment of 0.1, and its relationship with shadow detection accuracy is shown in Fig. 9.

SDSI can also be used as an index for directly separating shadows from other dark objects. Without the steps in Sec. 3.2, the shadow detection accuracy corresponding to different α is shown in Fig. 10.

Figure 9 shows that the sensitivity of SDSI varies in different images, and the precision of shadow extraction varies greatly with different α. Figure 10 shows that SDSI can also achieve
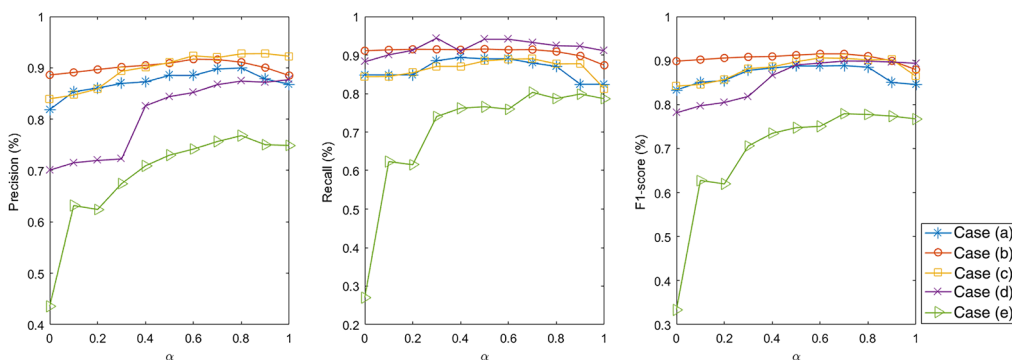

Fig. 9 Sensitivity of α in SDSI to shadows for cases (a)–(e) after suspected shadow detection.
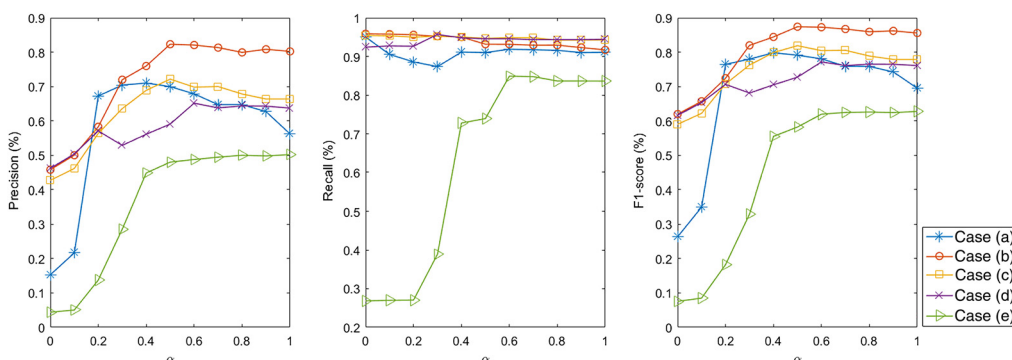

Fig. 10 Sensitivity of α in SDSI to shadow for cases (a)–(e) ignoring suspected shadow.

better results in images [e.g., cases (b) and (e)] without objects that are easily confused with shadows (e.g., water body), indicating that SDSI can also be used as an index of shadow detection.

4.3 Comparison with Results of Multiscale Segmentation

The multiscale segmentation algorithm is one of the most widely used and is included in typical remote sensing software tools. Unlike the top-down image segmentation in this paper, multiscale segmentation is bottom-up.

We use eCognition[51] to handle the multiscale segmentation. The simple yet robust estimation of scale parameter (ESP) tool enables fast and objective parametrization when performing image segmentation,[65] but for our experimental image with a spatial resolution of 3.2 m, it produced an insufficient segmentation. Thus, we still choose the segmentation parameters manually. After several tests, four multiscale segmentation parameters including scale, band weight, smoothness, and compactness are determined manually according to the visual display of a color image, and they are 40, 1:1:1:4, 0.1, and 0.5, respectively. The shadow distribution based on this segmentation is shown in Fig. 11 and the precision is listed in Table 3.
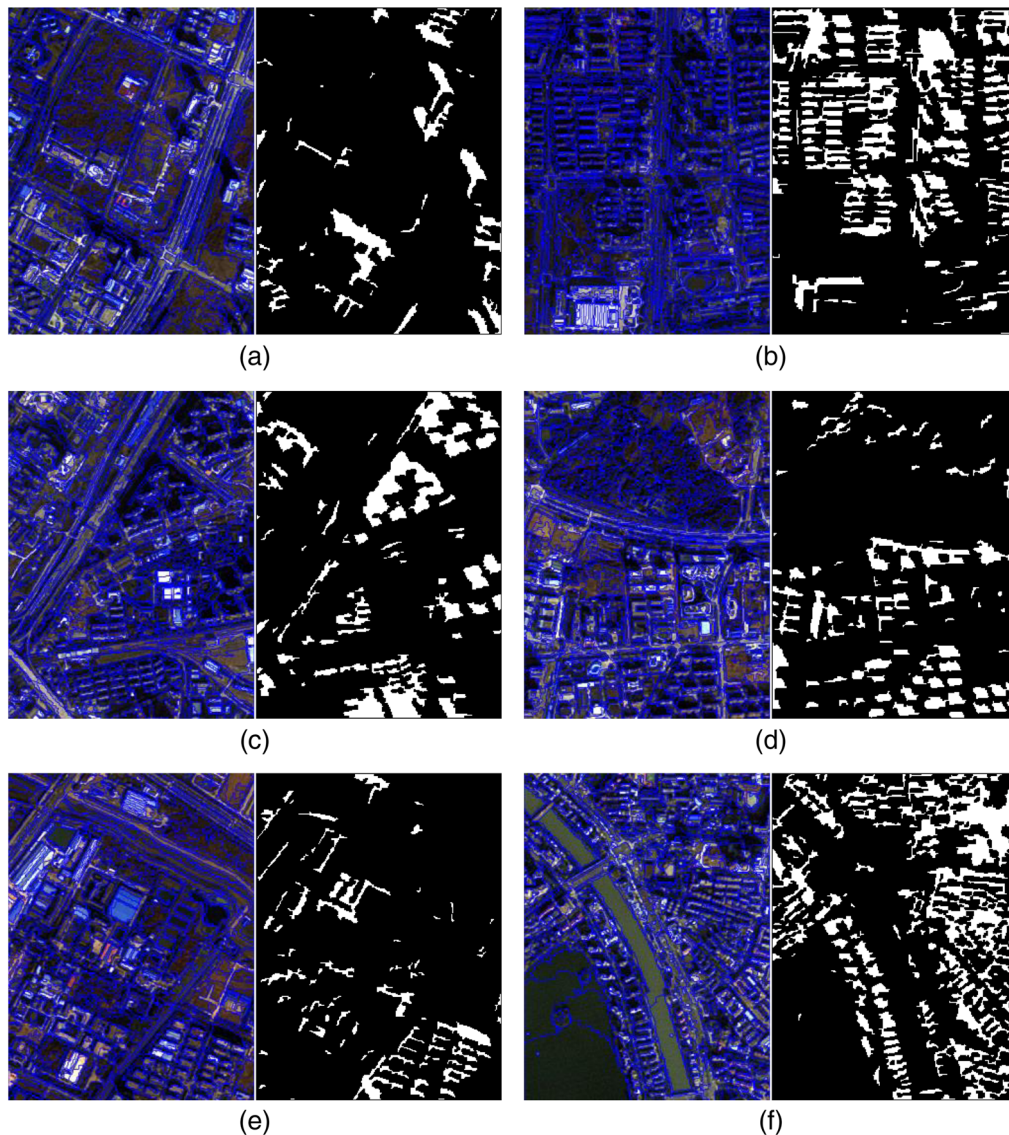

Fig. 11 Multiscale segmentation of image and shadow detection for cases (a)–(f).

Table 3 Shadow detection accuracy after multiscale segmentation (%).

Case	Precision	Recall	F1-score
(a)	81.3	68.5	74.4
(b)	87.6	71.3	78.6
(c)	86.6	62.7	72.7
(d)	82.1	69.1	75.0
(e)	55.5	61.1	58.2
(f)	76.6	69.0	72.6

The errors and omissions of shadows in Fig. 11 are consistent with that in Sec. 4.1. However, compared with Figs. 7, 8, and 11, multiscale segmentation cannot extract different scales of patches at the same time due to the different scales of various ground features, which easily leads to the omission of shadows and light shadows.

The average precision, recall, and F1-score of a shadow based on multiscale segmentation are 78.3%, 67%, and 71.9%, respectively, and their ranges are 55.5% to 87.6%, 61.1% to 71.3%, and 58.2% to 78.6%, respectively. Compared with the results in Table 2, the accuracy of the method is lower and the range is larger.

In high-spatial-resolution image applications, multiscale segmentation often achieves better results. However, for the shadow detection with 3.2-m spatial resolution image in this paper, the performance of multiscale segmentation is poor. The disadvantage of multiscale segmentation is that it requires many manual adjustments, and these adjustments need to take into account the scales of different regions in the image, which requires more experience in practical applications. In our method, the closed and sharp edge is used to determine the boundary of the object, which ensures the integrity of the ground object after segmentation and results in better performance.

5 Conclusions

A method and workflow of building shadow detection in HSRS remote sensing image based on different image features and edge-based image segmentation were proposed in this paper. Our method is highly automated and has a fast computation speed and the performance of making small-size shadows into independent objects is better than that of multiscale segmentation.

The six case images from the TripleSat-2 multispectral sensor showed that average F1-score of shadow detection accuracy was 87%, indicating that the proposed method has a good reference value for shadow detection in a high-spatial-resolution remote sensing image. The extracted shadow can be used as a mask for invalid areas or for uncertainty analysis and can provide effective help for change detection, remote sensing classification, etc.

The adjustable parameter involved in this method is α of SDSI, which needs to be adjusted according to the image. The empirical value is 0.5.

The availability of the method and workflow for other sensor images needs more validation, including how to better combine the advantages of edge segmentation with those of region segmentation to detect shadows with higher accuracy. This needs further research.

Acknowledgments

References

1. S. Musleh, M. Sarfraz, and L. Niepel, "A comparative study on shadow detection methods based on features," in *Int. Conf. Comput. Sci. and Eng.*, pp. 1–6 (2018).
2. Y. Chen et al., "Shadow information recovery in urban areas from very high resolution satellite imagery," *Int. J. Remote Sens.* **28**(15), 3249–3254 (2007).
3. H. Luo et al., "Development of a multi-scale object-based shadow detection method for high spatial resolution image," *Remote Sens. Lett.* **6**(1), 59–68 (2015).
4. K. R. M. Adeline et al., "Shadow detection in very high spatial resolution aerial images: a comparative study," *ISPRS J. Photogramm. Remote Sens.* **80**, 21–38 (2013).
5. A. Shahtahmassebi et al., "Review of shadow detection and de-shadowing methods in remote sensing," *Chin. Geogr. Sci.* **23**(4), 403–420 (2013).
6. T. Kim, T. A. Javzandulam, and T. Y. Lee, "Semiautomatic reconstruction of building height and footprints from single satellite images," in *IEEE Int. Geosci. and Remote Sens. Symp.* (2007).
7. J. I. Shun-Ping and X. X. Yuan, "A method for shadow detection and change detection of man-made objects," *J. Remote Sens.* **11**, 323–329 (2007).
8. P. M. Dare, "Shadow analysis in high-resolution satellite imagery of urban areas," *Photogramm. Eng. Remote Sens.* **71**(2), 169–177 (2005).
9. Y. Li, P. Gong, and T. Sasagawa, "Integrated shadow removal based on photogrammetry and image analysis," *Int. J. Remote Sens.* **26**(18), 3911–3929 (2005).
10. W. Zhou et al., "Object-based land cover classification of shaded areas in high spatial resolution imagery of urban areas: a comparison study," *Remote Sens. Environ.* **113**(8), 1769–1777 (2009).
11. Y. Han et al., "Shadow detection based on texture features in gray sequence images," *Opt. Precis. Eng.* **21**(11), 2931–2942 (2013).
12. G. D. Finlayson et al., "On the removal of shadows from images," *IEEE Trans. Pattern Anal. Mach. Intell.* **28**(1), 59–68 (2006).
13. X. Zhang, F. Chen, and H. He, "Shadow detection in high resolution remote sensing images using multiple features," *Acta Autom. Sin.* **42**(2), 290–298 (2016).
14. X. Huang and L. Zhang, "Morphological building/shadow index for building extraction from high-resolution imagery over urban areas," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **5**(1), 161–172 (2012).
15. H. Zhang, K. Sun, and W. Li, "Object-oriented shadow detection and removal from urban high-resolution remote sensing images," *IEEE Trans. Geosci. Remote Sens.* **52**(11), 6972–6982 (2014).
16. R. B. Irvin and D. M. McKeown, "Methods for exploiting the relationship between buildings and their shadows in aerial imagery," *IEEE Trans. Syst. Man Cybern.* **19**(6), 1564–1575 (1989).
17. G. Tolt, M. Shimoni, and J. Ahlberg, "A shadow detection method for remote sensing images using VHR hyperspectral and LIDAR data," in *IEEE Int. Geosci. and Remote Sens. Symp.*, Vancouver, BC, Canada, pp. 4423–4426 (2011).
18. M. Cameron and L. Kumar, "Diffuse skylight as a surrogate for shadow detection in high-resolution imagery acquired under clear sky conditions," *Remote Sens.* **10**(8), 1185 (2018).
19. A. M. Melesse and J. D. Jordan, "A comparison of fuzzy vs. augmented-ISODATA classification algorithms for cloud-shadow discrimination from Landsat images," *Photogramm. Eng. Remote Sens.* **68**(9), 905–911 (2002).
20. Y. Du, J. Li, and Y. Wang, "Shadow detection in high-resolution remote sensing image based on improved K-means," in *8th Int. Conf. Internet Multimedia Comput. and Service*, pp. 281–286 (2016).
21. T. P. Wu and C. K. Tang, "A Bayesian approach for shadow extraction from a single image," in *Tenth IEEE Int. Conf. Comput. Vision*, Vol. 1, pp. 480–487 (2005).
22. M. D. Levine and J. Bhattacharyya, "Removing shadows," *Pattern Recognit. Lett.* **26**(3), 251–265 (2005).
23. J. F. Lalonde, A. A. Efros, and S. G. Narasimhan, "Detecting ground shadows in outdoor consumer photographs," *Lect. Notes Comput. Sci.* **6312**, 322–335 (2010).

24. J. Liu, F. Tao, and D. Li, "Shadow detection in remotely sensed images based on self-adaptive feature selection," *IEEE Trans. Geosci. Remote Sens.* **49**(12), 5092–5103 (2011).
25. Y. Mostafa and A. Abdelhafiz, "Accurate shadow detection from high-resolution satellite images," *IEEE Geosci. Remote Sens. Lett.* **14**(4), 494–498 (2017).
26. W. Jiao et al., "Shadow detection based on FLICM clustering algorithm and color space for remote sensing images," *Geogr. Geo-Inf. Sci.* **34**(5), 37–41 (2018).
27. H. Xia and P. Guo, "A shadow detection of remote sensing images based on statistical texture features," *J. Remote Sens.* **15**(4), 778–791 (2011).
28. K. Jiang et al., "Adaptive shadow detection using global texture and sampling deduction," *IET Comput. Vision* **7**(2), 115–122 (2013).
29. F. P. S. Luus, F. van den Bergh, and B. T. J. Maharaj, "Adaptive threshold-based shadow masking for across-date settlement classification of panchromatic QuickBird images," *IEEE Geosci. Remote Sens. Lett.* **11**(6), 1153–1157 (2014).
30. L. Chen and L. Zhang, "An automatic image shadow detection method based on texture operator," *Comput. Appl. Software* **32**(2), 207–210 (2015).
31. H. Guo, Q. Xu, and B. Zhang, "Shadow extraction of building based on multiple constraints," *Geomatics Inf. Sci. Wuhan Univ.* **30**(12), 1059–1062 (2005).
32. W. Zhang, X. Z. Fang, and X. Yang, "Moving cast shadows detection based on ratio edge," in *18th Int. Conf. Pattern Recognit.*, pp. 73–76 (2006).
33. K. J. Lin and X. D. Wan, "Shadow detection algorithm based on edge information and illumination direction," *Comput. Eng.* **35**(20), 192–193, 196 (2009).
34. W. Lui et al., "A shadow detection algorithm based on statistical and edge properties," *Acta Photonica Sin.* **38**(7), 1872–1876 (2009).
35. B. Pan et al., "Shadow detection in remote sensing images based on weighted edge gradient ratio," in *IEEE Geosci. and Remote Sens. Symp.*, pp. 505–508 (2014).
36. Q. Fu et al., "Building extraction and its height estimation over urban areas based on morphological building index," *Remote Sens. Technol. Appl.* **30**(1), 148–154 (2015).
37. L. I. Jiménez, J. Plaza, and A. Plaza, "Efficient implementation of morphological index for building/shadow extraction from remotely sensed images," *J. Supercomput.* **73**(1), 482–494 (2017).
38. E. Salvador, A. Cavallaro, and T. Ebrahimi, "Cast shadow segmentation using invariant color features," *Comput. Vision Image Understanding* **95**(2), 238–259 (2004).
39. A. M. Polidorio et al., "Automatic shadow segmentation in aerial color images," in *16th Brazilian Symp. Comput. Graphics and Image Process.*, pp. 270–277 (2003).
40. K. L. Chung, Y. R. Lin, and Y. H. Huang, "Efficient shadow detection of color aerial images based on successive thresholding scheme," *IEEE Trans. Geosci. Remote Sens.* **47**(2), 671–682 (2009).
41. V. J. D. Tsai, "A comparative study on shadow compensation of color aerial images in invariant color models," *IEEE Trans. Geosci. Remote Sens.* **44**(6), 1661–1671 (2006).
42. G. F. Silva et al., "Near real-time shadow detection and removal in aerial motion imagery application," *ISPRS J. Photogramm. Remote Sens.* **140**, 104–121 (2018).
43. N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst. Man Cybern.* **9**(1), 62–66 (1979).
44. F. Yamazaki, W. Liu, and M. Takasaki, "Characteristics of shadow and removal of its effects for remote sensing imagery," in *IEEE Int. Geosci. and Remote Sens. Symp.*, pp. IV-426–IV-429 (2009).
45. F. Yuan and M. E. Bauer, "Mapping impervious surface area using high resolution imagery: a comparison of object-based and per pixel classification," in *Am. Society For Photogramm. and Remote Sens. Annu. Conf. Proc.*, Reno, Nevada (2006).
46. Y. Chen et al., "Hierarchical object oriented classification using very high resolution imagery and LIDAR data over urban areas," *Adv. Space Res.* **43**(7), 1101–1110 (2009).
47. Z. Pu, L. Yang, and J. Bai, "Shadow detection and removal based on object-oriented method in high spatial resolution remote sense image," *Remote Sens. Technol. Appl.* **23**(6), 735–738 (2008).
48. T. Shuai, H. Zhang, and L. Zhang, "The object-based method of shadow detection in high-resolution remote sensing imagery," *Acta Photonica Sin.* **44**(12), 1228002 (2015).

49. N. Mo et al., "Deshadowing of urban airborne imagery based on object-oriented automatic shadow detection and regional matching compensation," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **11**(2), 585–605 (2018).
50. Esri China Information Technology Co., Ltd., "ENVI," 2019, http://www.enviidl.com.
51. Beijing Tuyuan Technology Co. Ltd., "eCognition," 2019, http://www.ecognition.com.
52. Definiens AG, "Definiens," 2019, https://www.definiens.com.
53. Intergraph Inc., "ERDAS Imagine," 2019, http://www.erdas.com.
54. V. Arévalo, J. González, and G. Ambrosio, "Shadow detection in colour high-resolution satellite images," *Int. J. Remote Sens.* **29**(7), 1945–1963 (2008).
55. J. Huang, W. Xie, and L. Tang, "Detection of and compensation for shadows in colored urban aerial images," in *Fifth World Cong. Intell. Control and Autom. (IEEE Cat. No.04EX788)*, Hangzhou, China, pp. 3098–3100 (2004).
56. A. R. Smith, "Color gamut transform pairs," in *ACM SIGGRAPH Comput. Graphics*, Los Angeles, pp. 12–19 (1978).
57. H. Ma, Q. Qin, and X. Shen, "Shadow segmentation and compensation in high resolution satellite images," in *IEEE Int. Geosci. and Remote Sens. Symp.*, pp. II-1036–II-1039 (2008).
58. J. W. Rouse et al., "Monitoring vegetation systems in the Great Plains with ERTS," Nasa Special Publication 351, 309 (1973).
59. D. Cai et al., "Study on shadow detection method on high resolution remote sensing image based on HIS space transformation and NDVI index," in *18th Int. Conf. Geoinformatics*, pp. 1–4 (2010).
60. A. A. Gitelson et al., "Novel algorithms for remote estimation of vegetation fraction," *Remote Sens. Environ.* **80**(1), 76–87 (2002).
61. J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-8**, 679–698 (1986).
62. The MathWorks Inc., "MATLAB (R2018a)," 2019, https://www.mathworks.com.
63. G. Duan et al., "Shadow extraction based on characteristic components and object-oriented method for high-resolution images," *J. Remote Sens.* **18**(4), 760–770 (2014).
64. S. K. McFeeters, "The use of the normalized difference water index (NDWI) in the delineation of open water features," *Int. J. Remote Sens.* **17**(7), 1425–1432 (1996).
65. L. Drăguţ, D. Tiede, and S. R. Levick, "ESP: a tool to estimate scale parameter for multi-resolution image segmentation of remotely sensed data," *Int. J. Geogr. Inf. Sci.* **24**(6), 859–871 (2010).

Yuxuan Liu is currently pursuing her BS degree at Nanjing Normal University, Nanjing, China. Her research interests include spatiotemporal geographic information science and remote sensing digital image processing.

Yuchun Wei received his PhD in physical geography from Nanjing Normal University, Nanjing, China, in 2001. He is currently a professor at Nanjing Normal University. His professional interests include remote sensing information extraction and analysis, and the research points are change detection of urban areas based on the high-spatial-resolution remote sensing imagery.

Biographies of the other authors are not available.