An Automatic Shadow Detection Method for VHR Remote Sensing Orthoimagery

Qiongjie Wang [1], Li Yan [1,*], Qiangqiang Yuan [1] and Zhenling Ma [2]

[1]   School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China; joanwang@whu.edu.cn (Q.W.); yqiang86@gmail.com (Q.Y.)

[2]   Research Center for Ocean Mapping and Applications, Shanghai Engineering Research Center of Hadal Science and Technology, College of Marine Sciences, Shanghai Ocean University, Shanghai 201306, China; zlma@shou.edu.cn

*   Correspondence: lyan@sgg.whu.edu.cn; Tel.: +86-27-6877-8166

Abstract: The application potential of very high resolution (VHR) remote sensing imagery has been boosted by recent developments in the data acquisition and processing ability of aerial photogrammetry. However, shadows in images contribute to problems such as incomplete spectral information, lower intensity brightness, and fuzzy boundaries, which seriously affect the efficiency of the image interpretation. In this paper, to address these issues, a simple and automatic method of shadow detection is presented. The proposed method combines the advantages of the property-based and geometric-based methods to automatically detect the shadowed areas in VHR imagery. A geometric model of the scene and the solar position are used to delineate the shadowed and non-shadowed areas in the VHR image. A matting method is then applied to the image to refine the shadow mask. Different types of shadowed aerial orthoimages were used to verify the effectiveness of the proposed shadow detection method, and the results were compared with the results obtained by two state-of-the-art methods. The overall accuracy of the proposed method on the three tests was around 90%, confirming the effectiveness and robustness of the new method for detecting fine shadows, without any human input. The proposed method also performs better in detecting shadows in areas with water than the other two methods.

Keywords: shadow detection; matting; VHR; DSM; geometrical method; skeleton operation

1. Introduction

Shadows exist in most very high resolution (VHR) orthoimagery (with a resolution of up to 1 m) and lead to more complex and detailed land-cover features, especially in urban areas [1]. Shadow areas usually have incomplete spectral information, lower intensity, and fuzzy boundaries, which seriously affect the subsequent interpretation [2–6]. However, dealing with shadows always generates costs in labor and/or time [7]. Hence, accurate and automatic shadow detection is an active area of research [8,9].

Most of the research into shadow detection is based on natural imagery and videos [10], and relatively few studies have addressed shadow detection in VHR orthoimagery with compound land-cover features and a broader variety of features co-existing in a scene [11,12]. The current methods for shadow detection can be divided into three types [11–13]: (1) property-based methods [9,13–20]; (2) geometrical methods [14,17–20]; and (3) machine learning methods [15,16,21,22].

The property-based methods are the most popular approach as the imagery is the only required input. Property-based methods identify shadows using the statistical characteristics of brightness, color, texture, gradient, and edges. In the property-based methods [23–26], an initial shadow mask is

generated by a thresholding method, among which Otsu's thresholding method [27] is the most widely used approach, augmented with various morphological operators. A representative property-based approach deploying invariant color models is Tsai's spectral method (TSM) [26], which transforms the red, green, and blue (R-G-B) color images into different invariant color spaces. Otsu's thresholding method detects the shadowed areas based on the ratio of the hue and intensity components in the image. Based on Tsai's work, Chung et al. [25] presented a successive thresholding scheme, instead of a global thresholding process. However, according to Adeline's survey paper [11], Tsai's method performs better than Chung's algorithm. Liu et al. [9] used the properties derived from shadow samples to generate a feature space and calculate decision parameters; a series of transformations is then used to separate the shadowed and non-shadowed regions. Using the physical properties of the blackbody radiator model, Makarau et al. [13] put forward a method based on direct and scattered light having different temperatures. Ngo et al. [28] applied Otsu's thresholding method to three binary channels, i.e., C3, ExG, and L. Based on Dempster–Shafer theory, the influence indexes of vegetation and shadow are combined pixel by pixel to improve the detection quality. Markov random fields (MRF) is used to increase the constraint of adjacent pixel information, and the minimum energy method is used to obtain the optimal segmentation result, i.e., shadow areas, vegetation areas, and other areas. Li et al. [29] employed a matting technique to extract accurate soft shadow masks, with manually identified shadowed and non-shadowed areas as the input. After the manual marking process, the shadow detection performance of Li's matting method (LMM) is very good. By the use of interactive brushing assisted by a Gaussian mixture model, Xiao et al. [30] proposed a method to simultaneously detect disconnected shadow areas, with little user interaction.

Geometric methods establish a precise model using a scenario, object, and light source that are not affected by image quality and the material reflectance. These methods require prior knowledge when constructing the model, such as a digital surface model (DSM) and the shoot time of the image. Tolt et al. [17] proposed a straightforward line-of-sight analysis with a DSM to detect large shadowed and non-shadowed areas, and trained a support vector machine (SVM) supervised classifier with the detected shadow areas to refine the results. Giles [19] and Steven et al. [20] proposed similar methods that trace a straight line from every pixel to the solar position at the time of image acquisition, to determine if a topographic obstruction exists between a point on the land surface and the sun. Once the line is obstructed, the pixel is marked as shadow. These two methods require a digital elevation model (DEM), the shoot time, and the coordinates of the scene. A similar idea of judging the coverage of occlusion is described in the work of Aliakbarpour et al. [31]. Prasath et al. [32] also applied a similar idea to georeferenced SAR images, and they proposed an automatic method for radar shadow detection from SAR images utilizing geometric projections along with a low-resolution DEM. For a mountainous area, instead of a DEM, Li et al. [18] applied a DSM, which contains the height of objects, such as building and bridges, to determine the shadows by ray tracing. Khan et al. [21] employed multiple convolutional deep neural networks (ConvNets) to obtain the most relevant features of the shadows, and fed the learned features to a conditional random field (CRF) model to extract the final shadow mask. Guo et al. [15] proposed a shadow processing method for natural scenes by segmenting the image with the mean shift algorithm, and estimated the confidence of each region being shadowed by the use of a trained classifier. A soft matting technique is then applied to produce smooth shadow boundaries. Lorenzi et al. [16] selected samples of shadowed and non-shadowed counterparts from different regions as initial values, which are fed into the classifier to obtain a binary shadow mask.

Unfortunately, many of the shadow detection methods based on natural scenes fail when processing complex remote sensing images. Compared to natural images, remote sensing images have a larger range, as more shadows exist in a scene and they contain more complex features. Hence, significant problems are encountered when automatically detecting shadows in VHR remote sensing images [33]. High-accuracy detection results are essential in the final shadow compensation process. However, in the current methods, there is a tradeoff between accuracy and automaticity [26]. Penumbras do not appear in low-resolution images, and thus penumbra detection is seldom

applied [29]. However, in VHR images, this should be considered as the width of the penumbra can be more than five pixels, which can seriously affect the compensation accuracy in the final output. Moreover, the property-based methods, in general, are not independent of the data in the image. Therefore, these methods are only applicable to images obtained under certain conditions, and they may fail on images obtained at different times and under varying atmospheric conditions. For shadowed and non-shadowed areas having similar attribute information, such as low-brightness non-shadows (e.g., dark asphalt on roads and black roofs) and high-brightness shadowed regions (e.g., windows), or bluish objects (e.g., water), misjudgment often occurs [26,34]. A robust automatic matching method for shadowed areas and the corresponding non-shadowed areas in VHR imagery begs further research [7].

The geometrical methods are independent of the quality of the remote sensing images and the material reflectance. However, obtaining a priori information has been a major hindrance to the development of the geometric methods. Fortunately, automatic 3D photogrammetry software, such as Acute3D, Pix4D, etc., now provides a richer and more reliable foundation for the application of geometrical methods. In this paper, a simple and effective automatic shadow detection method for complex urban environments is presented, which combines the advantages of the property-based and geometrical methods. The sun position and a DSM are required to rebuild the geometric relationship between the surface and the sun at the time the raw images were captured, thereby obtaining a coarse binary shadow mask. The next step is to process the mask and determine the initial labels for the shadowed and non-shadowed areas using morphological operators, and refine the shadow areas with the matting algorithm. Since all the procedures are processed directly from the input data, the proposed method can be performed on VHR orthoimagery with few preset parameters. In the experimental part, using manually delineated ground-truth data, the proposed method is compared with two state-of-the-art methods. The shadow removal results in three types and scales of shadow situations confirm that the performance of the proposed method is better than the other two methods, from both visual and quantitative aspects.

The rest of the paper is organized as follows. Section 2 details the proposed shadow detection method. Section 3 provides the experimental results and analysis for different resolutions of VHR images. Section 4 analyzes the possible error sources of the proposed method. Section 5 describes the considerations in the process of parameter selection. Section 6 concludes the paper.

2. The Proposed Shadow Detection Method

Most of the high-accuracy shadow detection methods require the samples of shadows and their corresponding non-shadowed areas to be manually collected before the process of detection, and most of the automatic methods fail on certain shadow-like features, such as water and dark asphalt roads [26,29,33]. The proposed method is processed in three steps and benefits from the automation of the spectral-based methods and the high accuracy of the geometric methods. The procedure of the proposed geometry-assisted shadow detection method is shown in Figure 1.

The geometrical model of the proposed method (Figure 1a) is composed of a DSM/3D model (Figure 1b), the solar position (Figure 1c), and the VHR images (Figure 1d). With the input of the DSM and the solar position, the geometrical methods are largely automated, and can produce highly accurate results when identifying shadows (as shown in Figure 1e). This is the first step, which is described in detail in Section 2.1. Due to the technical limits of DSM production at the current time, DSMs usually have a slightly lower horizontal accuracy at building boundaries than VHR imagery. In addition, DSMs may include some processing artifacts caused by moving cars, trees, and water, which can cause shadow detection failure [18]. To overcome this drawback, a matting method is used in the third step.
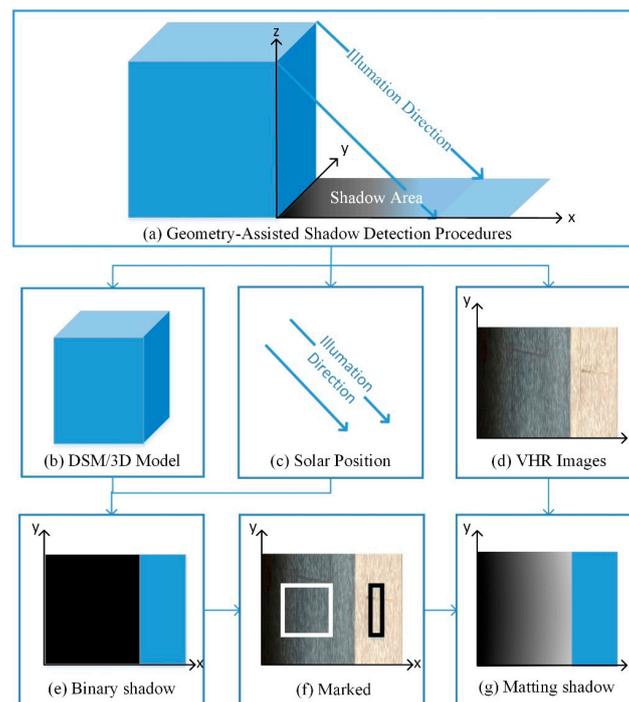
Figure 1. The procedure of the proposed geometry-assisted shadow detection method.

Interactive digital matting is an accurate method that can be applied to extract shadowed areas in handheld camera pictures or remote sensing imagery of a small area. Compared to natural imagery, VHR remote sensing imagery has a lower feature resolution and a much larger range, which results in the shadowed areas being smaller and more fragmented. Marking all the shadowed and non-shadowed areas in VHR images manually, as in the matting method of Li et al. [29] (LMM), would incur significant time costs. Meanwhile, directly applying a matting method on these images may fail because this does not distinguish the shadowed and non-shadowed areas in sufficient detail. In the proposed method, to balance the time and labor, every shadowed and non-shadowed area in the image is delineated with a single mark, which is assisted by morphological erosion and skeleton processing. This is the second step of the proposed method, which is described in detail in Section 2.2. After the morphological processing, the proposed method delineates the shadowed/non-shadowed areas automatically with a shadow mask generated by the geometric method. As shown in Figure 1f, the white and black lines represent the marks of the shadowed and non-shadowed areas, respectively. Figure 1g shows that the marks are then fed into the matting machine, and a final soft shadow mask with high accuracy and low time cost is obtained. This is the final step of the proposed method, which is discussed in Section 2.3.

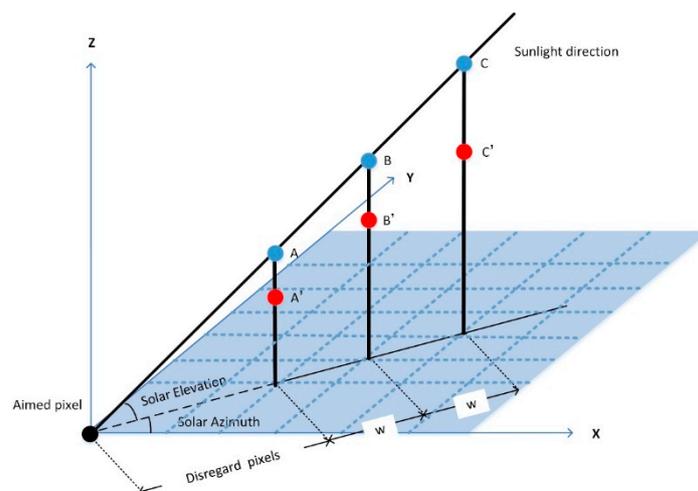### 2.1. Geometrical Shadow Detection

The first step is to produce a binary shadow mask (Figure 1e) from the DSM/3D model (Figure 1b) and the solar position (Figure 1c). Based on the assumption that the sun is a parallel light source, the sunlight is traced by the solar elevation and azimuth, which are calculated by the shoot time and the geographic coordinates of the imagery.

As shown in Figure 2, we draw a line from the target pixel to the sun, and calculate the interpolated elevation values (blue points) for the pixels the line passes through. The width w is calculated (as shown in Equation (1)) based on the grid width of the DSM that is recorded in the header file of the data and the solar azimuth angle. A three-dimensional view of Figure 2 is shown in Figure 3.

$$w = \frac{\text{Grid width of the DSM}}{\cos(\text{Solar azimuth})} \tag{1}$$
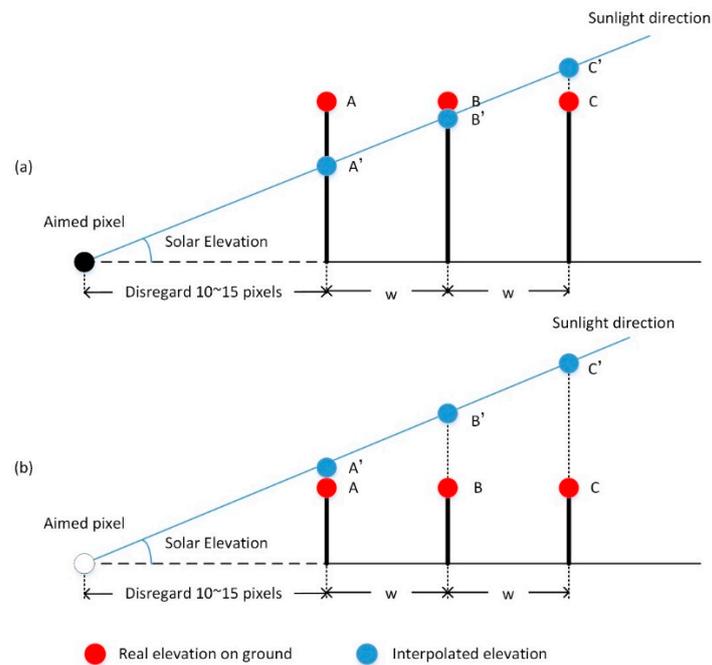
Figure 2. The process of geometric shadow detection: (**a**) judging the target pixel as shadow (black point); and (**b**) judging the target pixel as non-shadow (white point).


Figure 3. The process of geometric shadow detection in a three-dimensional view.

We then compare the height of the real elevation values (red points) and the interpolated elevation values (blue points) in the order of distance from the current pixel (in the order of A, B, C, as shown in Figures 2 and 3) and we disregard a certain number of pixels (equivalent to a 1-m ground distance) closest to the current pixel to reduce noise such as DSM elevation errors (the error sources are discussed in Section 4). Detailed notes on parameter selection are provided in Section 5.1. Once the real elevation is above the interpolated elevation (in the case of B and B' in Figure 2a), we stop the cycle and assign the target pixel as a shadowed pixel (the black point in Figure 2a); otherwise, we assign the target pixel as a non-shadowed pixel (the white point in Figure 2b). Through calculating the pixels one by one, a binary shadow mask is obtained.

2.2. Mask Processing with Morphological Operators

Before obtains a finer shadow mask, representative marks of every shadowed/non-shadowed area are required. An automatic way to obtain the required marks is introduced in this step, which corresponds to the process shown in Figure 1e,f. As the boundary of the original binary shadow mask is not as reliable as the center, the mask is further processed to obtain purer marks, using an erosion operation with a 10-by-10 "circle" structuring element to eliminate noisy and inaccurate boundaries (the parameter selection is detailed in Section 5.2).

As shown in Figure 4, areas A and B are exemplary non-convex boundary shapes, which are representations of the complex shapes in the shadow results, such as lakes, or trees on a street. The erosion operation is used to "erase" the inaccurate boundaries of the shadowed areas (as shown in Figure 4a,b). A similar procedure is used to yield candidate non-shadowed areas (as shown in Figure 4d,e). Every candidate black/white pixel in Figure 4b,e indicates a region of shadowed/non-shadowed area. This step is designed to select representative pixels of each region for faster matting processing. By applying a skeleton operation, a linear shadowed/non-shadowed area delineation is derived. The skeleton operation can remove pixels on the boundaries of areas, but does not allow areas to break apart. Overlaying the shadowed/non-shadowed skeleton (the white/black lines in Figure 4c,f) on the raw image, we obtained the input image to be used in the next step. This operation preserves the key shape characteristics for the matting and greatly decreases the computational cost of the matting.
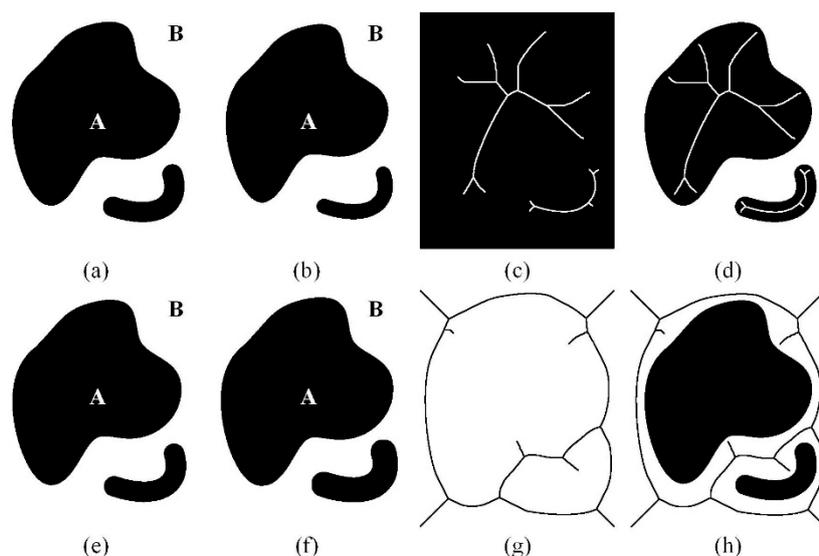
Figure 4. The process of marking the shadowed/non-shadowed areas with erosion and skeleton operations: (**a**) original shadowed area A; (**b**) shadowed area A after the erosion operation; (**c**) white marks of shadowed area A after the skeleton operation; (**d**) superimposed marks of the shaded area on the original shadowed result; (**e**) original non-shadowed area B; (**f**) non-shadowed area B after the erosion operation; (**g**) black marks of non-shadowed area B after the skeleton operation; and (**h**) superimposed marks of the lit area on the original non-shadowed result.

2.3. Shadow Matting

Interactive digital matting is an accurate way to extract areas of interest in imagery. In a matting method, an image *I* is considered to be composed of a foreground image *F* and a background image *B*. Note that, in the proposed case, the foreground image *F* is equivalent to the shadowed areas, and the background image *B* represents the non-shadowed areas. Based on the proposed shadow delineation method, accurate shadow regions are extracted as the foreground objects from the image, where α stands for the probability of each pixel belonging to absolute shadow, ranging from 0 to 1. We briefly

describe the process below. Detailed notes on parameter selection and the proof of the theorem can be found in Levin's article [35,36].

The linear combination of the foreground and background is undertaken as follows:

$$I_i = \alpha_i F_i + (1 - \alpha_i) B_i. \tag{2}$$

where I_i, α_i, F_i, B_i denote the values of I, α, F, B at pixel i. Assuming F and B are constant over a small window around each pixel, then Equation (2) can be rewritten as:

$$\alpha_i \approx a I_i + b, \ \forall i \in w. \tag{3}$$

where $a = \frac{1}{F-B}$, $b = -\frac{B}{F-B}$, and w is a small image window, which we set as 3×3 in our experiments. This relation suggests finding α, a, and b by minimizing the following cost function:

$$J(\alpha, a, b) = \sum_{j \in I} \left(\sum_{i \in w_j} \left(\alpha_i - a_j I_i - b_j \right)^2 + \epsilon a_j^2 \right), \tag{4}$$

where w_j is a small image window around pixel j, and ϵ is the weight of the regularization term on a. According to Theorem 1 in [35]:

$$J(\alpha) = \min_{a,b} J(\alpha, a, b) = \alpha^T L \alpha. \tag{5}$$

where L is an $N \times N$ matrix, whose (i, j)th element is:

$$\sum_{k|(i,j) \in w_k} \left(\delta_{ij} - \frac{1}{|w_k|} \left(1 + (I_i - \mu_k) \left(\Sigma_k + \frac{\epsilon}{|w_k|} I_3 \right)^{-1} (I_j - \mu_k) \right) \right) \tag{6}$$

where δ_{ij} is the Kronecker delta, $|w_k|$ is the number of pixels in this window, μ_k is a 3×1 mean vector of the colors in the window w_k, Σ_k is a 3×3 covariance matrix, and I_3 is a 3×3 identity matrix.

To extract the final mask α, matching the shadowed/non-shadowed constraints that we obtained in Section 2.2, we minimize the following energy function, which is firstly proposed by Levin [35], and a final shadow mask is obtained.

$$\alpha = \text{argmin} \alpha^T L \alpha + \lambda \left(\alpha^T - b^T \right) D(\alpha - b), \tag{7}$$

where D is a diagonal matrix whose diagonal elements are 1 for the marked pixels (as described in Section 2.2) and 0 for the others, and b is the vector containing the specified α values for the marked pixels, and 0 for the others. Parameter λ is usually set as a large number, and we followed the value Levin used in [35], i.e., 100. The influence of λ is discussed in detail in Section 5.3.

The former item $\alpha^T L \alpha$ is to ensure that α is under the assumption that the background (non-shadowed areas) and foreground (shadowed areas) are locally smooth in a small window around each pixel. The usage of the latter item $\left(\alpha^T - b^T \right) D(\alpha - b)$ is to add the influence of the marked pixels to the final results, which constrains the differences between the marked α (0 or 1) and the final α (ranging from 0 to 1) to be as small as possible. The initial value of $(\alpha - b)$ for the constrained pixels is 0, and the values of the others are not. The final α values of the constrained pixels may be very close or equal to the corresponding b, and for the other pixels, they vary from 0 to 1. As the cost is quadratic in α, the global minimum can be solved by differentiating and setting the derivatives to 0, which is equal to solving the sparse linear system as follows:

$$(L + \lambda D)\alpha = \lambda b \tag{8}$$

With the three steps illustrated above, the final soft shadow mask α is obtained. The values in α represent the probabilities of belonging to absolute shadow. The experimental results of the proposed method and two state-of-the-art methods (TSM [26] and LMM [27]) are discussed in the next section.

3. Experiments and Comparisons

The proposed method was implemented on a personal computer with an Intel Core i5-2500K 3.3 GHz CPU, using MATLAB in the Windows 7 environment. Two state-of-the-art methods, as proposed in [26,27], were also implemented for comparison and evaluation. As Table 1 shows, TSM is used as a representative of the automatic spectral-based shadow detection approach [11,26], and LMM is applied as it is both accurate and representative of the matting methods using manually delineated shadowed and non-shadowed areas [27]. The results of LMM and the proposed method are judged as soft shadow in the matting process, which means that the value of each pixel in the shadow result ranges from 0 to 1, instead of 0 or 1 in the binary shadow result. For a comparison and evaluation of the methods, an effective thresholding method is needed to convert the soft shadow into a binary shadow. In our tests, Otsu's thresholding method was applied to automatically obtain the suitable threshold value and the binary shadow mask.

Table 1. The comparative methods used in this paper.

Method	Automatic	Property	Process	Result
TSM	Yes	Spectral	Ratio + thresholding	Binary shadow
LMM	No	Spectral	Matting	Soft shadow
Proposed	Yes	Geometric + spectral	Ray tracing + matting	Soft shadow

The data used in the experiments are shown in Table 2. In test 1, the data were composed of a 0.25 m DSM generated from Trimble Inpho 5.3 software and a 0.25-m orthophoto computed by Trimble Inpho OrthoVista from aerial imagery of an urban area of Vaihingen, Germany [37]. The data in test 2 were a 0.09 m DSM acquired from aerial airborne laser scanning (ALS) and a 0.09 m orthophoto computed by Trimble INPHO OrthoVista from aerial imagery of an urban area of Vaihingen. Data of tests 1 and 2 were provided by the German Society of Photogrammetry, Remote Sensing and Geoinformation (DGPF) project [38]. The solar elevation/azimuth was $41°/107°$. The data used in test 3 were from a 0.06 m orthophoto and a 0.06 m DSM processed by Pix4D of an urban area of Zurich, Switzerland. The raw aerial images were provided by the ISPRS/EuroSDR project. The solar elevation/azimuth was $55°/195°$, respectively.

Table 2. The test datasets.

Property	Test 1	Test 2	Test 3
Resolution	0.25 m	0.09 m	0.06 m
Channels	NIR, red and green	NIR, red and green	red, green, and blue
Camera type	Intergraph/ZI DMC	Intergraph/ZI DMC	Leica RCD30
Solar elevation/azimuth	$41°/107°$	$41°/107°$	$55°/195°$
Main features	Water, trees, buildings	Buildings, trees	Roads, buildings, trees
Location	Vaihingen, Germany	Vaihingen, Germany	Zurich, Switzerland

The data in the test 1 images contained water features, which are often incorrectly detected by the current automatic shadow detection methods [26]. The images of test 2 contained the common urban features, such as shadows from concentrated high-rise buildings and trees. The test 3 images contained the common urban features such as roads, buildings, and trees.

The three bands of the optical imagery were NIR, red, and green in order in tests 1 and 2, and red, green, and blue in test 3. As the bands of the imagery in tests 1 and 2 were NIR-R-G, the plants in

both Sections 3.1 and 3.2 are shown in red. The experimental results of the three methods working on the R-G-B channels are shown in Section 3.3. Three sets of images were used to test the accuracy and performance of the three approaches on different sources and different scales of data.

3.1. Water Areas

Figure 5a shows the original imagery of test 1. In this test, the main challenge for the shadow detection methods is the river. Water has a lower intensity and bluish color, which is similar to the characteristics of shadows. As shown in the yellow box in Figure 5b (result of TSM), water causes false detection for the automatic spectral-based shadow detection methods. Figure 5c shows the shadow detection result of LMM. With manually delineated shadowed and non-shadowed areas, the shadow boundaries in the result of LMM are similar to those in the original imagery. Figure 5d shows the result of the proposed method, where it can be seen that most of the shadows are correctly detected. However, there are still some problems. For example, some of the wooded areas are falsely discerned as shadow, and some small building shadows are miss-detected. Figure 5e shows the manually selected shadow mask, which is considered as the true shadow on the ground. Using the shadow compensation method described in [26], Figure 5f–i shows the results obtained with the shadow masks displayed in Figure 5b–e. The right images in Figure 5 are zoomed images of the region within the green boxes of the left images.

To empirically compare the effects of the different methods, four well-known metrics—producer's accuracy, user's accuracy, overall accuracy and F-score—are used to evaluate the shadow detection results [11,39]. As mentioned at the beginning of this section, an automatic thresholding method (Otsu's thresholding method [27]) is applied on the soft shadow mask to obtain a binary value. The metrics are defined as follows.

The producer's accuracy is used to measure how accurately the method detects the true shadowed and non-shadowed pixels.

$$\tau_s = \frac{TP}{TP + FN} \tag{9}$$

$$\tau_n = \frac{TN}{TN + FP} \tag{10}$$

where τ_s represents the producer's accuracy of the shadowed pixels and τ_n represents the non-shadowed pixels. The true positive (TP) is the number of shadowed pixels correctly discerned when compared to the manually selected mask, and true negative (TN) is the number of non-shadowed pixels that are correctly detected and identified by the true values from the manual mask. False negative (FN) is the number of true shadow pixels judged as non-shadowed, and false positive (FP) is the number of non-shadowed pixels that are falsely detected as true shadow pixels.

The user's accuracy is used to indicate the probability of the correctness of the detected shadowed and non-shadowed pixels.

$$\sigma_s = \frac{TP}{TP + FP} \tag{11}$$

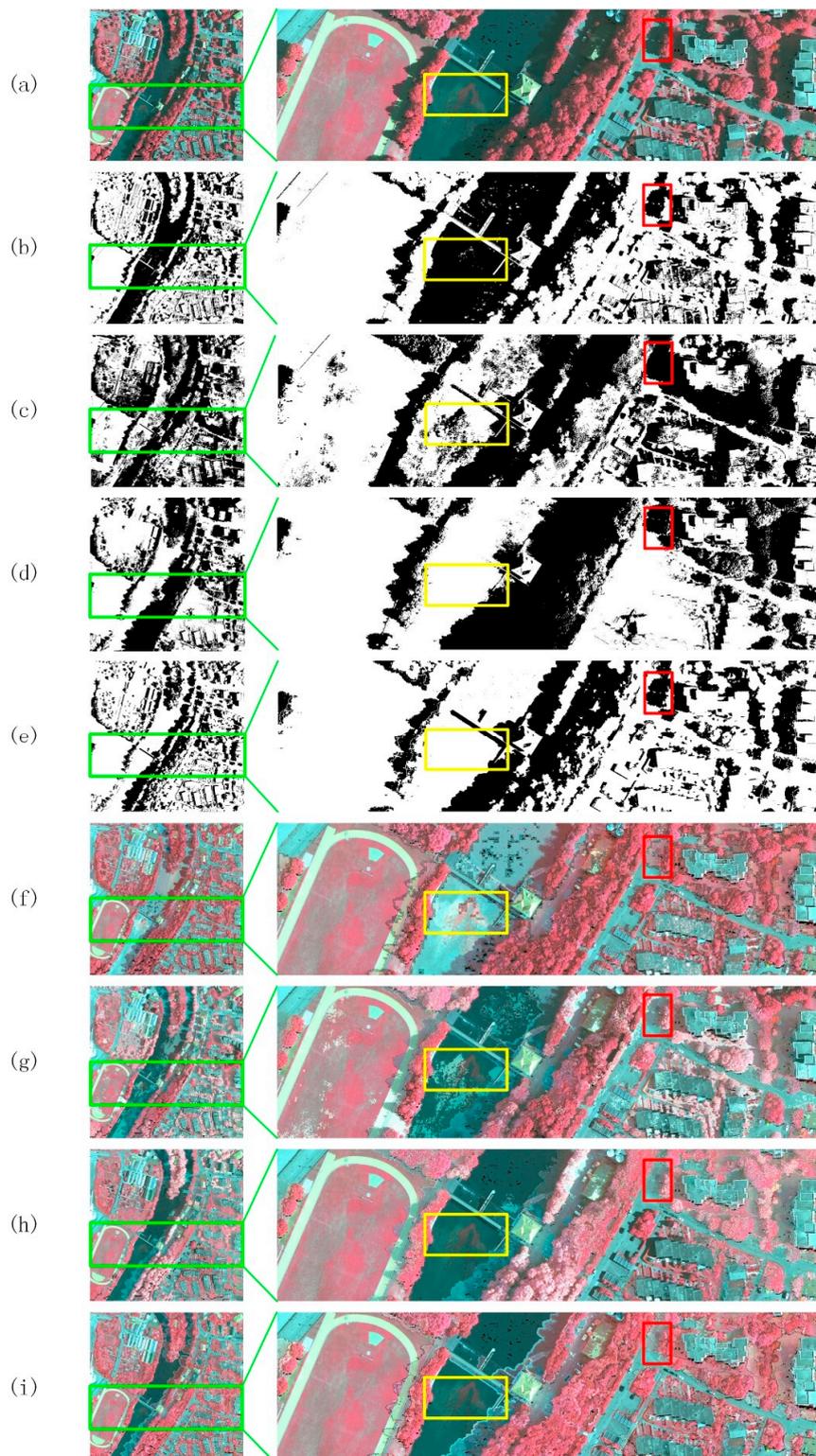$$\sigma_n = \frac{TN}{TN + FN} \tag{12}$$

where σ_s represents the user's accuracy of the shadowed pixels and σ_n represents the non-shadowed pixels.

The overall accuracy ρ refers to the percentage of correctly detected shadowed and non-shadowed pixels.

$$\rho = \frac{TP + TN}{TP + TN + FN + FP} \tag{13}$$

The F-score is a balanced way to judge the user's accuracy and the producer's accuracy.

$$F = 2\frac{\tau_s \cdot \sigma_s}{\tau_s + \sigma_s} \tag{14}$$

Figure 5. Results of the shadow detection and removal methods in test 1, and zoomed images of the green boxes: (**a**) original imagery; (**b**) result of TSM; (**c**) result of LMM; (**d**) result of the proposed method; (**e**) manually selected shadow mask; and (**f**–**i**) shadow compensation results based on the shadow masks in (**b**–**e**).

The accuracy measurements are shown in Table 3, where the bold numbers represent the highest values and the underlined numbers represent the lowest. From the table, it can be seen that TSM obtains an acceptable producer's accuracy performance for τ_n (82.0%) and a user's accuracy for σ_s of 65.8%. This method, however, delivers the worst performance in producer's accuracy performance for τ_s (83.3%) and user's accuracy for σ_n (92.2%). LMM obtains an acceptable producer's accuracy for τ_s (86.2%) and user's accuracy for σ_n (92.6%), and the worst producer's accuracy for τ_n (71.7%) and user's accuracy for σ_s (55.9%). The proposed method obtains the best performance for all the six metrics. The producer's accuracy for τ_s is 89.7%, the producer's accuracy for τ_n is 82.2%, the user's accuracy for σ_s is 67.7%, and the user's accuracy for σ_n is 95.0%. For the overall accuracy, the proposed method obtains a value of 84.4%, TSM ranks second at 82.4%, and LMM obtains an acceptable value of 75.9%. For the F-score, the proposed method ranks first with 77.1%, TSM ranks second with 73.5%, and LMM obtains 67.8%.

Table 3. Shadow detection accuracy measurements of test 1.

Method	Producer's Accuracy (%)		User's Accuracy (%)		Overall (%)	F-Score (%)
	τ_s	τ_n	σ_s	σ_n	ρ	F
TSM	<u>83.3</u>	82.0	65.8	<u>92.2</u>	82.4	73.5
LMM	86.2	<u>71.7</u>	<u>55.9</u>	92.6	<u>75.9</u>	<u>67.8</u>
Proposed	**89.7**	**82.2**	**67.7**	**95.0**	**84.4**	**77.1**

From the zoomed images in Figure 5, it can be seen that the river is falsely detected as shadow by TSM and LMM in Figure 5b,c, whereas the proposed method correctly judges it as non-shadow (Figure 5d). Thus, the proposed method has an outstanding advantage in detecting shadows in areas with water when compared to the other two methods. To verify the effect of the shadow masks in the shadow removal, the shadow removal method proposed in [26] and the five indicators proposed in [40] are used to analyze the shadow removal results of the three methods.

The shadow removal method in [26] was carried out by matching the cumulative distribution function (CDF) of the shadowed pixels and their surrounding non-shadowed pixels. From the red boxes in the upper right part of the zoomed images in Figure 5f–h, it can be seen that the shadow area near the trees is correctly discerned as shadow by all three methods, but the compensation results are quite different. The result of TSM is the worst of the three, and it falsely discerns cement as trees. Both the results of LMM and the proposed method display an improved spectral quality for the objects in the red boxes. Overall, the proposed method performs the best of the three methods in the visual interpretation.

Table 4 shows the statistics of the five indicators of mean, standard deviation, entropy, gradient, and hue deviation index (HDI) [40]. The first four indicators are commonly used in the evaluation process. Detailed descriptions of them can be found in Shen's paper [40]. The HDI is used to measure the variance of hue between the original image and the processed image, which is calculated as follows:

$$\text{HDI} = \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} \min[|H_{ori}(x,y) - H_{de}(x,y)|, 1 - |H_{ori}(x,y) - H_{de}(x,y)|]}{M \times N} \times 100\% \qquad (15)$$

where H_{ori} represents the hue value of the pixel located at (x, y) from the original image, and H_{de} represents the hue value of the same pixel from the processed image. M and N are the size of the image.

**Table 4.** Quantitative analysis of the methods in test 1.

Image	Mean	Standard Deviation	Entropy	Gradient	HDI
Original	127, 125, 125	62, 52, 53	7.6, 7.3, 7.4	9.6, 8.4, 8.9	0
TSM	158, 145, 147	52, 53, 51	7.5, 7.3, 7.4	17.1, 12.7, 13.5	0.06
LMM	152, 156, 156	56, 51, 49	7.4, 7.2, 7.2	9.9, 12.6, 12.0	0.05
Proposed	150, 152, 153	59, 47, 46	7.6, 7.4, 7.4	11.0, 11.0, 11.0	0.04

In Table 4, the three numbers in each cell represent the values of the three channels (IR-R-G for tests 1 and 2, R-G-B for test 3). The mean intensities of the image processed by the three methods are all increased, which is due to the removal of the dark shadow areas. At the same time, the three methods have very close results in standard deviations, meaning that the contrast of the entire image is lower than the original imagery. Information entropy denotes the information richness of the image; the higher the better. All the methods yield good results, with almost the same information entropy values as the original image. The gradient refers to the detail contrast in the image. All three methods enhance the detail contrast, which may be caused by the features under the shadow. An HDI of less than 1% means that the shadow removal process can effectively preserve the hue.

3.2. City Center Areas

In test 2, an area in the center of Vaihingen was selected, which contained the common features found in urban areas, such as crowded buildings, trees, and cars. Figure 6a shows the original image. The shadow detection results of TSM, LMM, and the proposed method are respectively shown in Figure 6b–d. Figure 6b displays false detection of dark objects, where cars and other artifacts are incorrectly identified as shadowed areas. As the buildings in the scene have complex roofs, manually delineating the shadowed and non-shadowed areas is a very time-consuming process. The delineation for the shadowed/non-shadowed areas in Figure 6c,d is shown in Figure 7a,b. In Figure 7a, the shadowed and non-shadowed areas of every building are manually delineated. The results shown in Figure 6c may not be accurate enough, and thus a finer delineation of the shadowed and non-shadowed areas is needed. According to [29], a finer delineation can lead to a more accurate shadow mask. Many buildings exist in an image of a city center, but a single mark for each shadowed and non-shadowed area for each building is sufficient in the proposed method. Figure 6d shows the automatically generated result of the proposed method, which contains less noise and clearly represents the shadowed/non-shadowed areas, obtained without any human input. Figure 6e shows the manually delineated shadow mask. Figure 6f–i shows the compensated results obtained using the shadow compensation method described in Section 3.1, with the shadow masks displayed in Figure 6b–e. The right images in Figure 6 are zoomed images of the region within the green boxes of the left images.

The accuracy evaluation is shown in Table 5, where TSM ranks second in the user's accuracy and the producer's accuracy. LMM obtains the best performance for two out of the six metrics, with a producer's accuracy for τ_s of 96.7% and a user's accuracy for σ_n of 97.1%. At the same time, LMM delivers the worst producer's accuracy for τ_n of 66.9% and user's accuracy for σ_s of 63.3%. The proposed method obtains the best score in producer's accuracy for τ_n of 93.0%, user's accuracy for σ_s of 87.9%, and an acceptable performance in terms of producer's accuracy for τ_s with a score of 87.3%. The user's accuracy for σ_n ranks third, 4.6% lower than the best result. In terms of overall accuracy, TSM obtains a value of 91.2%, the proposed method ranks second at 90.5%, and LMM obtains an acceptable score of 77.9%. For the F-score, TSM ranks first with 88.6%, the proposed method ranks second with 87.5%, and LMM obtains 76.5%.
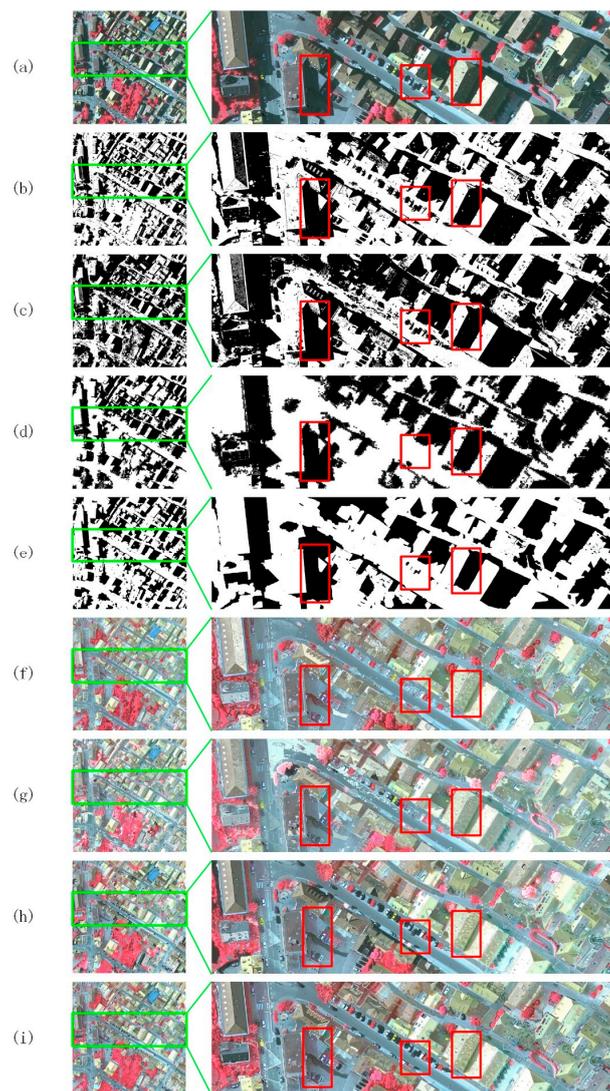
Figure 6. Results of the shadow detection and removal methods in test 2, and zoomed images of the green boxes: (**a**) original imagery; (**b**) result of TSM; (**c**) result of LMM; (**d**) result of the proposed method; (**e**) manually delineated shadow mask; and (**f**–**i**) shadow compensation results based on the shadow masks in (**b**–**e**).
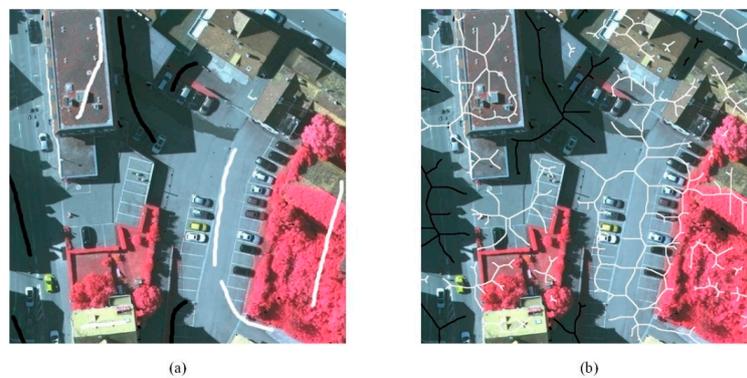
Figure 7. Shadow delineation: (**a**) manually delineated; and (**b**) automatically generated by the proposed method. The black/white lines mark the shadowed/non-shadowed areas.

Table 5. Shadow detection accuracy measurements for test 2.

Method	Producer's Accuracy (%)		User's Accuracy (%)		Overall (%)	F-Score (%)
	τ_s	τ_n	σ_s	σ_n	ρ	F
TSM	91.3	91.2	86.0	94.7	**91.2**	**88.6**
LMM	**96.7**	<u>66.9</u>	<u>63.3</u>	**97.1**	<u>77.9</u>	<u>76.5</u>
Proposed	87.3	**93.0**	**87.9**	92.5	90.5	87.5

According to the statistics shown in Table 6, the mean intensity and the standard deviation values for the three methods show the same tendencies as the results in Table 4 for test 1. For the information entropy, the proposed method delivers the best result of the three methods. All three methods enhance the detail contrast, with low HDI values. Compared to Table 4, the overall trend between the results of test 2 and test 1 is almost unchanged.

Table 6. Quantitative analysis of the methods in test 2.

Image	Mean	Standard Deviation	Entropy	Gradient	HDI
Original	126, 126, 125	62, 59, 58	7.6, 7.4, 7.5	5.8, 5.7, 5.7	0
TSM	166, 163, 164	45, 52, 50	7.1, 7.2, 7.2	9.4 9.6, 9.6	0.05
LMM	154, 150, 151	47, 50, 46	7.1, 7.1, 7.1	7.4, 8.7, 8.2	0.04
Proposed	165, 159, 159	51, 53, 50	7.6, 7.5, 7.5	8.4, 8.5, 8.5	0.05

For the visual interpretation, the red boxes in Figure 6f–h show the compensation results of the shadow masks of the three methods. As the shadow mask of TSM is hard shadow (binary shadow mask), the ground is overcompensated in the left red box in Figure 6b, and the black cars are falsely compensated, as shown in the middle red box in Figure 6b. With limited manual delineation, the compensation result of LMM misses some shadowed areas, such as the objects in the two red boxes shown in Figure 6e. From the three red boxes shown in Figure 6h, the proposed method accurately detects the shadows of the buildings, without being influenced by the cars and artifacts, obtaining a more stable and balanced shadow removal performance than the other two methods.

3.3. RGB Orthophotos

The test region in test 3 was an urban area of Zurich, which also contained the common urban features such as buildings, roads, parks, and cars. Figure 8a shows the original image. The shadow detection results of TSM, LMM, and the proposed method are respectively shown in Figure 8b,d. Figure 8b displays false detection, with trees, cars, and other artifacts incorrectly identified as shadowed areas. The results shown in Figure 8c contain a lot of noise. Figure 8d shows the automatically generated result of the proposed method, which shows a clear representation for the shadowed/non-shadowed areas, and fewer undetected shadows for the small areas, obtained without any human input. Figure 8e shows the manually delineated shadow mask. Figure 8f–i shows the compensated results obtained using the shadow compensation method described in Sections 3.1 and 3.2, with the shadow masks displayed in Figure 8b–e. The right images in Figure 8 are zoomed images of the region within the green boxes of the left images.
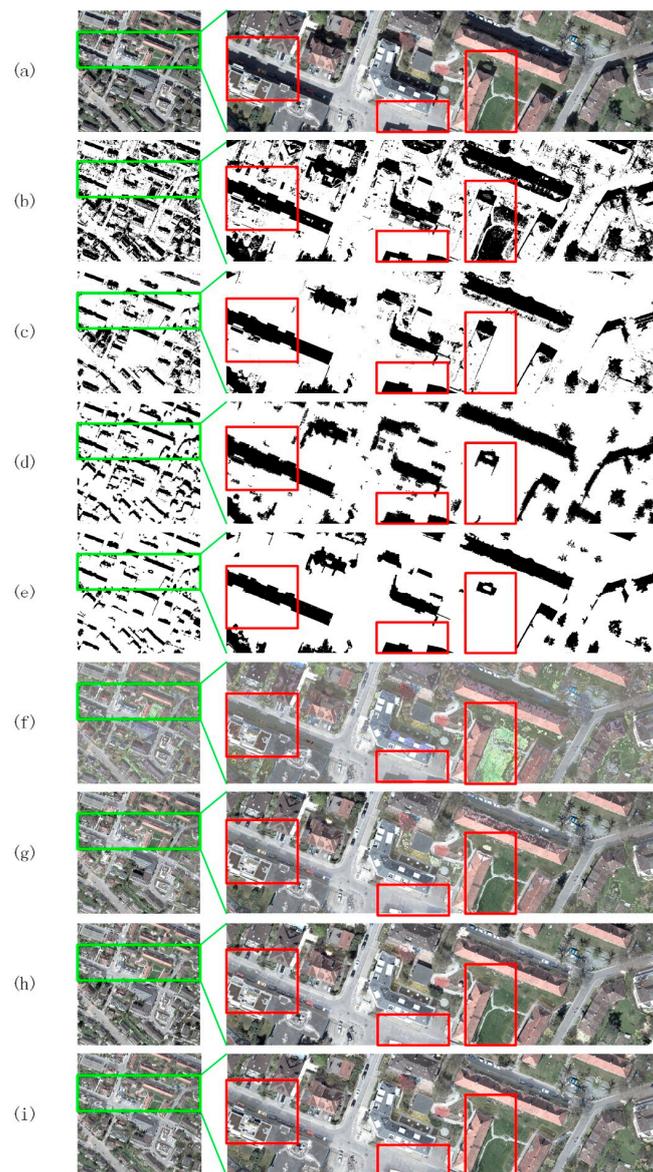
Figure 8. Results of the shadow detection and removal methods in test 3, and zoomed images of the green boxes: (**a**) original imagery; (**b**) result of TSM; (**c**) result of LMM; (**d**) result of the proposed method; (**e**) manually delineated shadow mask; and (**f**–**i**) shadow compensation results based on the shadow masks in (**b**–**e**).

The accuracy evaluation is shown in Table 7, where TSM obtains the best performance for the producer's accuracy for τ_s of 96.3% and user's accuracy for σ_n of 99.0%. At the same time, TSM delivers the worst producer's accuracy for τ_n with 76.8%, and a user's accuracy for σ_s of 46.6%. LMM ranks second in the user's accuracy and the producer's accuracy. The proposed method obtains the best score in the producer's accuracy for τ_n with 94.7%, user's accuracy for σ_s with 74.4%, and an acceptable performance in terms of user's accuracy for σ_n with a score of 95.2%. The producer's accuracy for τ_s ranks third. In terms of overall accuracy, the proposed method ranks first at 91.6%, LMM obtains an acceptable score of 91.2%, and TSM obtains a value of 80.0%. For the F-score, the proposed method ranks first at 80.0%, LMM obtains a value of 76.6%, and TSM obtains a value of 61.9%.

Table 7. Shadow detection accuracy measurements of test 3.

Method	Producer's Accuracy (%)		User's Accuracy (%)		Overall (%)	F-Score (%)
	τ_s	τ_n	σ_s	σ_n	ρ	F
TSM	**96.3**	<u>76.8</u>	<u>46.6</u>	**99.0**	<u>80.0</u>	<u>61.9</u>
LMM	85.3	92.4	69.4	96.9	91.2	76.6
Proposed	<u>76.3</u>	**94.7**	**74.4**	<u>95.2</u>	**91.6**	**80.0**

4. Discussion on the Error Sources of the Proposed Geometric Method

As shown in Figure 1, the shadow areas detected by the proposed geometric method depend on three factors: the solar position, the DSM, and the VHR remotely sensed images. These three factors are also the three main error sources for the shadow detection results, the error between the true solar position and the geographic coordinates calculated by the middle point of the whole image, the error between the true height of the scene and the DSM, and the error in the registration of the images and the DSM [11]. Three errors were simulated with the data of test 1, and their contributions to the final error of the results are discussed below.

4.1. DSM Height Accuracy

The accuracies of the geometric-assisted shadow detection methods are influenced by the accuracy of the geometric models, especially at the boundary of the detected shadow mask. As the height accuracies of the DSM are usually not as good as its horizontal accuracies, the noise was only added to the height values of the DSM to analyze the extent of the errors generated by the inaccuracy of the geometric models. Taking the data in test 1 as an example, the maximum height difference was 70 m. Considering that the height accuracy of the DSM was at a sub-meter level, Gaussian white noise was added with a signal-to-noise ratio (SNR) of 18, 21, 24, 28, 31, and 38, corresponding to height accuracies of 1 m, 0.5 m, 0.25 m, 0.1 m, 0.05 m, and 0.01 m, respectively.

$$\text{SNR} = 10 \log_{10}\left(\frac{Ps}{Pn}\right) \tag{16}$$

where Ps, Pn represent the power of the signal and the noise, respectively. The larger is the SNR, the less the noise of the DSM. For SNR > 38, the error produced by the inaccuracy of the DSM is less than 1% for the shadow detection result. Considering the original shadow mask to be the true values of the shadowed and non-shadowed areas, the errors generated by the height of the DSM are shown in Table 8.

Table 8. The error generated by the height accuracy of the DSM.

Type	18 (%)	21 (%)	24 (%)	28 (%)	31 (%)	38 (%)
False shadow	0.24	0.24	0.18	0.12	0.08	0.04
False non-shadow	0.32	0.33	0.23	0.14	0.10	0.04
Total	0.56	0.57	0.41	0.26	0.18	0.08

From Table 8, we can conclude that the height inaccuracy of the DSM can cause at most 0.57% overall error in the final shadow detection result. A DSM with a height accuracy of 1 m can contribute 0.56% to the total error. When the height accuracy of the DSM reaches 0.1 m or higher, the influence can be controlled within 0.3% (as 0.26%, 0.18%, and 0.08%, shown in Table 8).

4.2. Solar Position

In test 1, the solar elevation was 41° and the solar azimuth was 107°, which were calculated from the shoot time and the geographic coordinates of the middle pixel in the image. The data were obtained

from the DGPF project, which took place between 09:50 and 10:20 on 24 July 2008. Considering the flight areas and the total time of the project, the solar azimuth angle can vary over 5° and the solar elevation angle can vary over 3° [11]. An error of ±5° was added to the original solar elevation angle (E) and the solar azimuth angle (A), respectively. As shown in Figure 9, the shadow detection result of test 1 is shown in the middle of the figure as the true value of the shadows. The green areas refer to the newly increased shadow, and the red areas refer to the vanished shadow compared with the true values. In the following discussion, we keep the same meaning for the two colors. From left to right in Figure 9, the solar elevation angle increases. The larger the solar elevation angle, the smaller the overall shadow areas. From top to bottom of Figure 9, the solar azimuth angle increases. The solar azimuth angle decides the orientation of the shadows, and the shadows in each column show a gradual movement from the bottom left to the top right. The percentages of the errors generated by the different solar positions are listed in Table 9. From the table, it can be concluded that the error in solar elevation angle has more effect on the shadow detection result, which can reach 5.5% at ±5°. An error in the solar azimuth angle of ±5° will lead to at most 4.5% falsely detected shadowed areas. The error of the solar position can bring at most 7.5% inaccuracy to the final result.
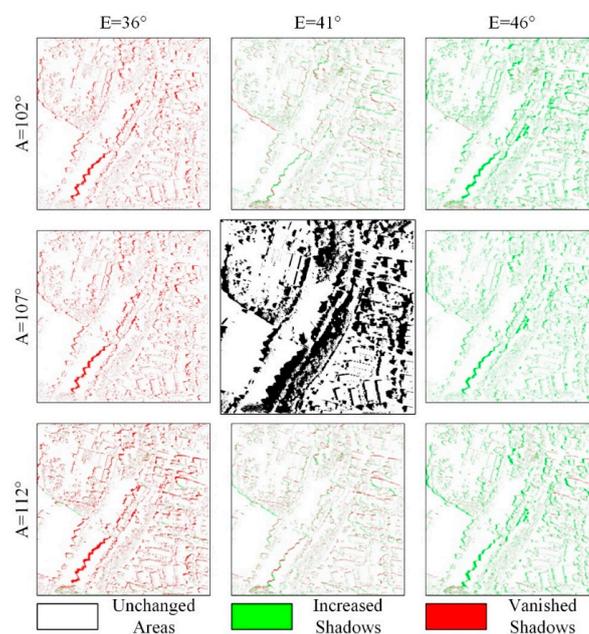

Figure 9. The influence of error in the solar position.

Table 9. The error generated by the solar position.

Solar Azimuth (°)	Solar Elevation (°)					
	36		41		46	
	Green (%)	Red (%)	Green (%)	Red (%)	Green (%)	Red (%)
102	1.05	5.71	2.51	1.76	6.29	0.54
107	5.42	0	0	0	5.01	0
112	0.70	6.87	1.83	2.56	5.24	0.92

The error described in this section can be decreased by the use of accurate records of the capture time and the coordinates of the images. With the addition of inertial measurement unit (IMU)/GPS data, the accuracy of the solar position can show a qualitative improvement.

4.3. Registration of the Image and the DSM

As the registration of the image can reach a sub-pixel level, ±1 pixel error was added to the coordinates of the image. The registration between the image and the DSM was then processed. Compared with the original shadow detection result, the influence of a registration error of ±1 pixel is shown in Table 10, where it can be seen that the maximum error (Row − 1, Column − 1) resulting from this process is 4%.

Table 10. The error generated by the registration.

Pixel Error	Row + 1		Row + 0		Row − 1	
	Green (%)	Red (%)	Green (%)	Red (%)	Green (%)	Red (%)
Column + 1	1.51	1.70	1.26	1.34	1.96	1.61
Column + 0	1.24	1.32	0	0	1.41	1.22
Column − 1	1.84	1.78	1.44	1.45	1.96	1.76

The difference between the result of (Row − 1, Column − 1) and the original (Row + 0, Column + 0) is shown in black in Figure 10, where it can be seen that the falsely detected pixels are usually located at the boundaries of the shadows.
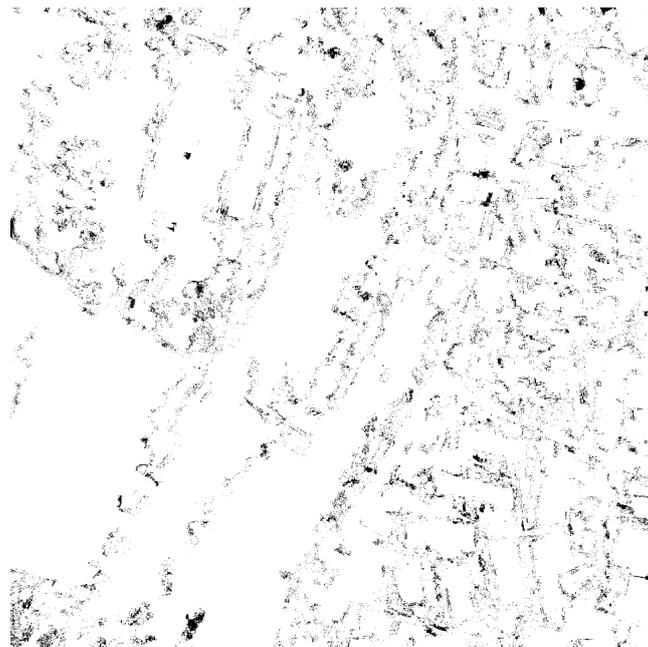
Figure 10. The difference between the registration results of (Row − 1, Column − 1) and (Row + 0, Column + 0).

4.4. Other Possible Error Sources

There are some other possible sources of error, beside the three factors discussed above. As these sources are likely to have only minor effects on the results and are hard to quantify, they are simply listed below, for the completeness of the analysis. For the purpose of comparison, the manually marked shadow mask is considered as the correct values of the shadows, which is an approach that is widely accepted in the studies of shadow detection. However, the human interpretation process may result in reduced accuracy, especially at the boundaries of the shadow patches and the very small areas. The blanks in the DSMs and the orthoimagery may also bring some uncertainty to the final result.

5. Notes on the Parameter Selection

In this section, some notes on the parameter selection of the proposed shadow detection method are given. Our intention here is to help peer researchers who want to reproduce the method with their own dataset to produce a better result.

5.1. The Disregarded Pixels in the Geometric Shadow Detection

In the first step of the proposed shadow detection method described in Section 2.1, pixels that are equivalent to a 1 m ground distance are disregarded in the geometric shadow detection process. We take the data from test 2 as an example to discuss the use of this parameter. The size of the image is 2779 × 2779, and the resolution is 0.09 m. According to the resolution of the image, we disregarded 11 pixels in the geometric shadow detection process.

Figure 11a shows the shadow detection result before using this parameter, where it can be seen that some areas are falsely detected as shadow, especially on the roofs. The reason for this is that the boundaries of the DSM are not accurate enough. The slight ups and downs on the building boundaries lead to the stripes, which are the falsely detected shadowed lines, on the output shadow mask. After adding the parameter (set as 10, 15, and 20), the results are shown in Figure 11b–d, respectively. The large red boxes in Figure 11 are the zoomed images. In Figure 11, it can be seen that the stripe noise is largely controlled after adding this parameter. According to our test experience, 10 or 15 pixels can achieve good results, with less noise and less miss-detection of the shadows at the boundaries. This result is consistent with the number of disregarded pixels (11 pixels) that we use in the experiment, which is equivalent to a 1 m ground distance.
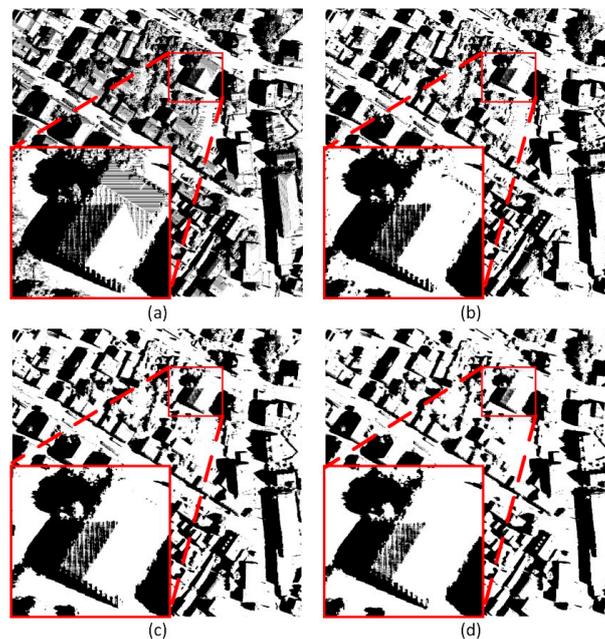

Figure 11. The results with different numbers of disregarded pixels in the geometric shadow detection process (the image resolution is 0.09 m): (**a**) original result; (**b**) disregarding 10 pixels; (**c**) disregarding 15 pixels; and (**d**) disregarding 20 pixels.

For the shadow detection process of VHR imagery, the number of the disregarded pixels is quite small compared with the number of the pixels that are used to judge whether the target pixel is a shadowed pixel or not (usually 500–1000, depending on the resolution of the imagery). It is worth sacrificing these pixels to balance the error produced by the noise. In the real world, especially in urban areas, shadow is usually caused by the objects within 20–50 m. Most shadows are caused by

buildings. However, the boundaries of the buildings are not accurate enough in DSMs. The closer the object next to the target pixel, the more influence it will have on the detection result. This can be easily explained by the principle of light tracking.

As Figure 12 shows, the blue pixels are the height recorded in the DSM, and the red pixels A′, B′, C′, and D′ represent the true positions of pixels A, B, C, and D, respectively. This shows how the height error of the pixels affects the judgment of the target pixel. Suppose the solar elevation angle is around 30°. If the height error occurs in pixel A or B, the light to the target pixel will be obstructed and the target pixel will be judged as a shadowed pixel. When the solar elevation angle is decreased, the height error of pixel C and pixel D may lead to false judgment of the target pixel. We can conclude that the closer the error pixel is to the target pixel, the higher the possibility of it leading to a wrong judgment. If the minimum solar elevation angle existing in the VHR imagery is 20°, the maximum height error of the pixel will be 0.5 m (with ground control points). The safe range from the error pixel to the target pixel is 1.37 m. As the resolution of the imagery is 0.09 m, this is equivalent to 15 pixels. However, this is under an extreme condition. In practice, the safe range of 1 m is enough to decrease the noise.
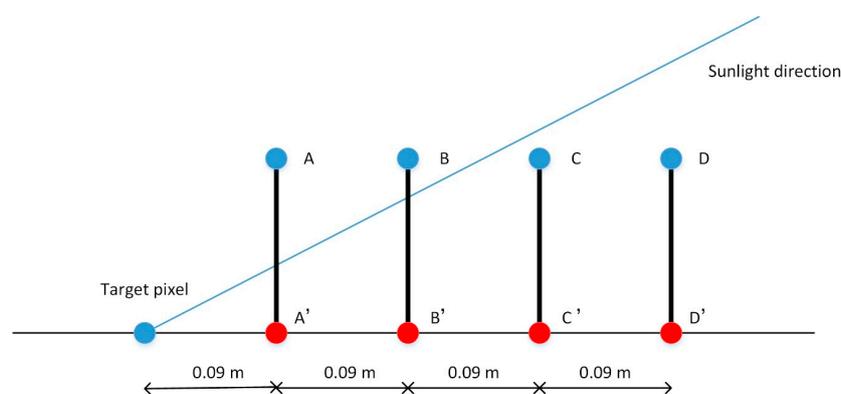
Figure 12. The principle of the geometric shadow detection process.

In the experiments, the resolution of the data was 0.06 m, 0.09 m, and 0.25 m, respectively. For the safe range of 1 m, the corresponding real range on the surface is shown in Table 11. In the corresponding experiments, the numbers of pixels shown in Table 11 were disregarded, which means that the influence of the surrounding objects within 1 m of the target pixel was ignored.

Table 11. The disregarded pixels corresponding to 1 m in different resolutions.

Resolution	0.06 m	0.09 m	0.25 m
Disregarded pixels	16	11	4

5.2. The Size of the Erosion Operator

In the second step of the proposed method described in Section 2.2, the size of the erosion operator was set as 10×10. From the analysis in Section 5.1, it can be deduced that this process means ignoring the influence of 1 m outside the boundaries of the target shadows/non-shadows. The process of size selection for the erosion operator is detailed here. As analyzed in Section 4, the boundary of the original hard-shadow mask is not as reliable as the center of the shadow mask. As shown in Figure 9, the main error source—the error in the solar position—can lead to at most ±10 pixels error in the shadow detection results. In this way, we set 10×10 as the original size. We take test 2 as an example, where 5×5, 10×10, and 15×15 were set as the size of the erosion operator to verify which setting performs the best.

The final skeleton marks on the original images and the corresponding shadow detection results are shown in Figure 13. The size of the cut image is 601 × 911, and the resolution is 0.09 m. From the figures, it can be seen that 10 × 10 performs the best in keeping the correctness of the marks, as well as ignoring acceptable small areas of true shadows in the different resolutions. The erosion operation is aimed to reducing the influence of the false shadows at the hard-shadow boundary on the subsequent matting process. This is a tradeoff between the true shadows in the small areas and the false shadows at the boundaries of large shadowed areas. Due to the principle of the matting method, which considers the marks as the absolute true values, the correctness of the shadow marks is preferred to the loss of the small shadow areas.
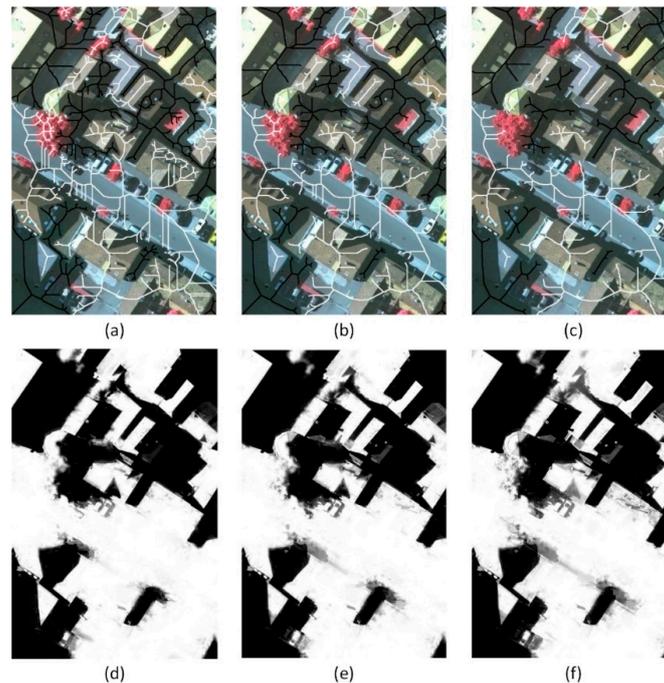
Figure 13. The final skeleton marks on the original images and the shadow detection results: (**a**) skeleton operation results of 5 × 5; (**b**) skeleton operation results of 10 × 10; (**c**) skeleton operation results of 15 × 15; (**d**) shadow mask of 5 × 5; (**e**) shadow mask of 10 × 10; and (**f**) shadow mask of 15 × 15.

5.3. The Influence of λ on the Shadow Detection Results

In the final step of the proposed method described in Section 2.3, the parameter λ was set as 100. According to our tests, the influence of λ on the final α is small. Some typical α with different values of λ are listed as Table 12 shows. The first column is the serial number of the randomly selected points of the image. The next five columns are the calculated α values of different values of λ. The values in seventh column, which is named as b, are the specified α values for the marked pixels and 0 for the others. The last column is the diagonal elements of D, for which the value is 1 for the marked pixels and 0 for the others. According to the table, we can conclude that the larger the value of λ, the smoother the final α value. For α of the same pixel, the biggest difference caused by λ varying from 1 to 10,000 is less than 5%. More than 99% of the differences are less than 1%. The α figure shows no difference in visual interpretation.

Table 12. The influence of λ on the final α.

No.	λ = 1	λ = 10	λ = 10²	λ = 10³	λ = 10⁴	b	D
1	**7.7%**	**8.4%**	**8.5%**	**8.5%**	**8.5%**	**0**	**0**
2	8.7%	9.3%	9.4%	9.5%	9.5%	0	0
3	6.8%	7.4%	7.5%	7.6%	7.6%	0	0
4	8.0%	8.6%	8.8%	8.8%	8.8%	0	0
5	11.6%	12.2%	12.3%	12.3%	12.3%	0	0
6	10.6%	11.2%	11.4%	11.4%	11.4%	0	0
7	9.6%	10.3%	10.4%	10.4%	10.4%	0	0
8	10.6%	11.2%	11.4%	11.4%	11.4%	0	0
9	**90.5%**	**91.2%**	**91.4%**	**91.4%**	**91.4%**	**0**	**0**
10	92.8%	93.4%	93.6%	93.6%	93.6%	0	0
11	92.2%	92.9%	93.1%	93.1%	93.1%	0	0
12	93.9%	94.8%	95.0%	95.1%	95.1%	0	0
13	96.3%	96.7%	96.8%	96.9%	96.9%	0	0
14	100.0%	100.0%	100.0%	100.0%	100.0%	1	1
15	100.0%	100.0%	100.0%	100.0%	100.0%	1	1
16	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**100.0%**	**1**	**1**
17	**0.0%**	**0.0%**	**0.0%**	**0.0%**	**0.0%**	**0**	**1**
18	0.0%	0.0%	0.0%	0.0%	0.0%	0	1

From above table, α values of No. 1, No. 9, No. 16, and No. 17 are shown in Figure 14, where point No. 1 and point No. 9 are the unconstrained pixels, point No. 16/point No. 17 is the marked shadowed/non-shadowed pixel, respectively. As can be seen from the Figure 14, with the increase of λ (from 1 to 10,000), the value of α (ranging from 0 to 1) tends to be stable. When λ is equal to 100, we can get a robust value of α. Thus, we suggest our peer researchers use λ = 100 directly in their work.
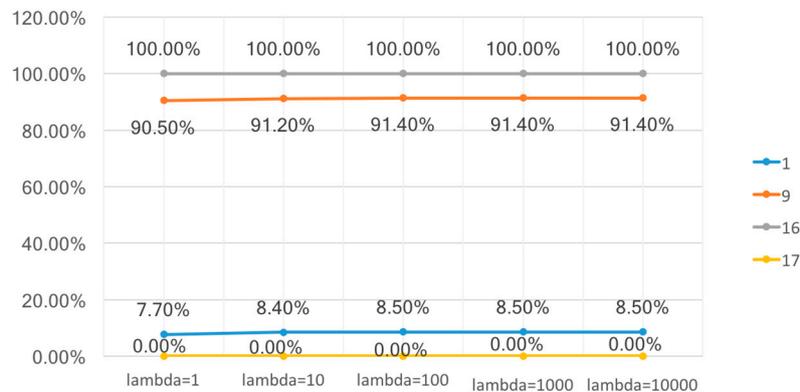
Figure 14. The influence of λ.

6. Conclusions

VHR orthoimagery and the corresponding DSMs can now be easily obtained from unmanned aerial systems (UASs) or aerial images acquired by the traditional airborne platforms. In this paper, a novel geometrical method is proposed to automatically detect shadows in urban areas with DSM data, assisted by VHR orthophotos with IR-R-G and R-G-B bands, respectively. The proposed geometric method is able to effectively locate shadows, with fewer false alarms. The binary shadow result is then processed by morphological operators to delineate the candidate shadowed areas for the next step. Finally, a matting method is applied to identify the detailed boundaries of shadowed areas, using the shadow delineation generated in the previous process. The proposed method prevents interference from low-intensity features such as water, moving cars, and processing artifacts, which are falsely discerned as shadow by many of the spectral methods, and it detects accurate boundaries of shadows at the same time. Compared with two of the state-of-the-art methods—LMM [29] and TSM [26]—the

proposed method shows a robust and outstanding performance, with an overall accuracy of 90% in complex urban areas. However, the proposed method still has some problems that need to be overcome. Some small shadowed areas may be ignored in the skeleton operation process, which can lead to miss-detection in the final result. According to the discussion in Section 4, the error sources of the proposed method can be summarized into three types: the error of the height accuracy of the DSM (<1%), the error of the solar position (7.5%), and the error of the registration (4%). In our future work, we will include geometric information from the DSM to further research a quantitative compensation method for the shadows of remote sensing imagery.

Author Contributions: Q.W. proposed the method and performed the experiments; Q.W. and L.Y. conceived and designed the experiments; and L.Y., Q.Y., and Z.M. wrote the manuscript. All the authors read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1.　Wu, J.; Bauer, M. Evaluating the Effects of Shadow Detection on QuickBird Image Classification and Spectroradiometric Restoration. *Remote Sens.* **2013**, *5*, 4450–4469. [CrossRef]

2.　Yang, J.; He, Y.; Caspersen, J. Fully constrained linear spectral unmixing based global shadow compensation for high resolution satellite imagery of urban areas. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *38*, 88–98. [CrossRef]

3.　Luus, F.P.S.; van den Bergh, F.; Maharaj, B.T.J. Adaptive Threshold-Based Shadow Masking for Across-Date Settlement Classification of Panchromatic QuickBird Images. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 1153–1157. [CrossRef]

4.　Liu, W.; Yamazaki, F. Object-Based Shadow Extraction and Correction of High-Resolution Optical Satellite Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2012**, *5*, 1296–1302. [CrossRef]

5.　Ng, K.P.; Yuan, Q.; Yan, L.; Sun, J. An Adaptive Weighted Tensor Completion Method for the Recovery of Remote Sensing Images with Missing Data. *IEEE Trans. Geosci. Remote Sens.* **2017**, *PP*, 1–15. [CrossRef]

6.　Li, J.; Yuan, Q.; Shen, H.; Zhang, L. Noise Removal from Hyperspectral Image with Joint Spectral–Spatial Distributed Sparse Representation. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 5425–5439. [CrossRef]

7.　Arbel, E.; Hel-Or, H. Shadow Removal Using Intensity Surfaces and Texture Anchor Points. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *33*, 1202–1216. [CrossRef] [PubMed]

8.　Liasis, G.; Stavrou, S. Satellite images analysis for shadow detection and building height estimation. *ISPRS J. Photogramm. Remote Sens.* **2016**, *119*, 437–450. [CrossRef]

9.　Liu, J.; Fang, T.; Li, D. Shadow Detection in Remotely Sensed Images Based on Self-Adaptive Feature Selection. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 5092–5103.

10.　Panchal, M.H.; Gamit, N.C. A Comprehensive Survey on Shadow Detection Techniques. In Proceedings of the International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), Chennai, India, 23–25 March 2016.

11.　Adeline, K.R.M.; Chen, M.; Briottet, X.; Pang, S.K.; Paparoditis, N. Shadow detection in very high spatial resolution aerial images: A comparative study. *ISPRS J. Photogramm. Remote Sens.* **2013**, *80*, 21–38. [CrossRef]

12.　Shahtahmassebi, A.; Yang, N.; Wang, K.; Moore, N.; Shen, Z. Review of shadow detection and de-shadowing methods in remote sensing. *Chin. Geogr. Sci.* **2013**, *23*, 403–420. [CrossRef]

13.　Makarau, A.; Richter, R.; Muller, R.; Reinartz, P. Adaptive Shadow Detection Using a Blackbody Radiator Model. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 2049–2059. [CrossRef]

14.　Wang, Q.; Yan, L. Anisotropic Scattering Shadow Compensation Method for Remote Sensing Image with Consideration of Terrain. In Proceedings of the ISPRS—International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 2016, Prague, Czech Republic, 12–19 July 2016; Volume XLI-B1, pp. 401–405.

15. Guo, R.; Dai, Q.; Hoiem, D. Paired Regions for Shadow Detection and Removal. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 2956–2967. [CrossRef] [PubMed]

16. Lorenzi, L.; Melgani, F.; Mercier, G. A Complete Processing Chain for Shadow Detection and Reconstruction in VHR Images. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 3440–3452. [CrossRef]

17. Tolt, G.; Shimoni, M.; Ahlberg, J. A shadow detection method for remote sensing images using VHR hyperspectral and LIDAR data. In Proceedings of the 2011 IEEE International Geoscience and Remote Sensing Symposium, Vancouver, BC, Canada, 24–29 July 2011; pp. 4423–4426.

18. Li, Y.; Gong, P.; Sasagawa, T. Integrated shadow removal based on photogrammetry and image analysis. *Int. J. Remote Sens.* **2005**, *26*, 3911–3929. [CrossRef]

19. Giles, P.T. Remote sensing and cast shadows in mountainous terrain. *Photogramm. Eng. Remote Sens.* **2001**, *67*, 833–840.

20. Stevens, M.R.; Pyeatt, L.D.; Houlton, D.J.; Goss, M.E. Locating shadows in aerial photographs using imprecise elevation data. In *Computer Science Technical Report CS-95-105*; Colorado State University: Fort Collins, CO, USA, 1995.

21. Khan, S.H.; Bennamoun, M.; Sohel, F.; Togneri, R. Automatic Shadow Detection and Removal from a Single Image. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *38*, 431–446. [CrossRef] [PubMed]

22. Guo, R.; Dai, Q.; Hoiem, D. Single-image shadow detection and removal using paired regions. In Proceedings of the 2011 IEEE Conference on Computer Vision & Pattern Recognition, Colorado Springs, CO, USA, 20–25 June 2011; pp. 2033–2040.

23. Su, N.; Zhang, Y.; Tian, S.; Yan, Y.; Miao, X. Shadow Detection and Removal for Occluded Object Information Recovery in Urban High-Resolution Panchromatic Satellite Images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2016**, *9*, 2568–2582. [CrossRef]

24. Song, H.; Huang, B.; Zhang, K. Shadow Detection and Reconstruction in High-Resolution Satellite Images via Morphological Filtering and Example-Based Learning. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 2545–2554. [CrossRef]

25. Chung, K.; Lin, Y.; Huang, Y. Efficient Shadow Detection of Color Aerial Images Based on Successive Thresholding Scheme. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 671–682. [CrossRef]

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

27. Otsu, N. A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. Man Cybern.* **1979**, *9*, 62–66. [CrossRef]

28. Ngo, T.T.; Collet, C.; Mazet, V. MRF and Dempster-Shafer theory for simultaneous shadow/vegetation detection on high resolution aerial color images. In Proceedings of the 2014 IEEE International Conference on Image Processing, Paris, France, 27–30 October 2014; pp. 5037–5041.

29. Li, H.; Zhang, L.; Shen, H. An Adaptive Nonlocal Regularized Shadow Removal Method for Aerial Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 106–120. [CrossRef]

30. Xiao, C.; She, R.; Xiao, D.; Ma, K.L. Fast Shadow Removal Using Adaptive Multi-Scale Illumination Transfer. *Comput. Graph. Forum* **2013**, *32*, 207–218. [CrossRef]

31. Aliakbarpour, H.; Prasath, V.; Dias, J. On Optimal Multi-Sensor Network Configuration for 3D Registration. *J. Sens. Actuator Netw.* **2015**, *4*, 293–314. [CrossRef]

32. Prasath, V.B.S.; Haddad, O. Radar shadow detection in synthetic aperture radar images using digital elevation model and projections. *J. Appl. Remote Sens.* **2014**, *8*, 83628. [CrossRef]

33. Li, J.; Hu, Q.; Ai, M. Joint Model and Observation Cues for Single-Image Shadow Detection. *Remote Sens.* **2016**, *8*, 484. [CrossRef]

34. Safdarinezhad, A.; Mokhtarzade, M.; Valadan Zoej, M. Shadow-Based Hierarchical Matching for the Automatic Registration of Airborne LiDAR Data and Space Imagery. *Remote Sens.* **2016**, *8*, 466. [CrossRef]

35. Levin, A.; Lischinski, D.; Weiss, Y. A Closed-Form Solution to Natural Image Matting. *IEEE Trans. Pattern Anal. Mach. Intell.* **2008**, *30*, 228–242. [CrossRef] [PubMed]

36. Levin, A.; Lischinski, D.; Weiss, Y. A Closed Form Solution to Natural Image Matting. In Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, New York, NY, USA, 17–22 June 2006; pp. 61–68.

37. Kosov, S.; Rottensteiner, F.; Heipke, C. 3D classification of crossroads from multiple aerial images using Markov Random Fields. In Proceedings of the ISPRS—International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 2012, Melbourne, Australia, 25 August–1 September 2012; Volume XXXIX-B3, pp. 1–4.

38. Cramer, M. The DGPF-Test on Digital Airborne Camera Evaluation Overview and Test Design. *Photogramm. Fernerkund. Geoinf.* **2010**, *2010*, 73–82. [CrossRef] [PubMed]

39. Elbakary, M.I.; Iftekharuddin, K.M. Shadow Detection of Man-Made Buildings in High-Resolution Panchromatic Satellite Images. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 5374–5386. [CrossRef]

40. Shen, X.; Li, Q.; Tian, Y.; Shen, L. An Uneven Illumination Correction Algorithm for Optical Remote Sensing Images Covered with Thin Clouds. *Remote Sens.* **2015**, *7*, 11848–11862. [CrossRef]