Article

A Radiative Transfer Model-Based Multi-Layered Regression Learning to Estimate Shadow Map in Hyperspectral Images

Usman A. Zahidi *,† [ID]**, Ayan Chatterjee** *,† [ID] **and Peter W. T. Yuen** [ID]

Centre for Electronic Warfare, Information and Cyber, Cranfield Defence and Security, Cranfield University, Defence Academy of The United Kingdom, Shrivenham SN6 8LA, UK

*   Correspondence: usmanzahidi@gmail.com (U.A.Z.); ayan@outlook.com (A.C.)

†   These authors contributed equally to this work.

check for updates

Abstract: The application of Empirical Line Method (ELM) for hyperspectral Atmospheric Compensation (AC) premises the underlying linear relationship between a material's reflectance and appearance. ELM solves the Radiative Transfer (RT) equation under specialized constraint by means of in-scene white and black calibration panels. The reflectance of material is invariant to illumination. Exploiting this property, we articulated a mathematical formulation based on the RT model to create cost functions relating variably illuminated regions within a scene. In this paper, we propose multi-layered regression learning-based recovery of radiance components, i.e., total ground-reflected radiance and path radiance from reflectance and radiance images of the scene. These decomposed components represent terms in the RT equation and enable us to relate variable illumination. Therefore, we assume that Hyperspectral Image (HSI) radiance of the scene is provided and AC can be processed on it, preferably with QUick Atmospheric Correction (QUAC) algorithm. QUAC is preferred because it does not account for surface models. The output from the proposed algorithm is an intermediate map of the scene on which our mathematically derived binary and multi-label threshold is applied to classify shadowed and non-shadowed regions. Results from a satellite and airborne NADIR imagery are shown in this paper. Ground truth (GT) is generated by ray-tracing on a LIDAR-based surface model in the form of contour data, of the scene. Comparison of our results with GT implies that our algorithm's binary classification shadow maps outperform other existing shadow detection algorithms in true positive, which is the detection of shadows when it is in ground truth. It also has the lowest false negative i.e., detecting non-shadowed region as shadowed, compared to existing algorithms.

Keywords: shadow mapping; empirical line method; hyperspectral imaging

1. Introduction

A high-fidelity hyperspectral imagery contains crucial spatial and spectral information of a given scene. Presence of shadows causes significant challenges for both satellite and airborne data analyses. Shadows cast by scene geometry or clouds cause hurdles in remote-sensing data analyses, including inaccurate atmospheric compensation, biased estimation of Normalized Difference Vegetation Index (NDVI), confusion in land cover classification, and anomalous detection of landcover variation. Therefore, shadows are a significant source of noise in Hyperspectral Image (HSI) data, and their detection is a vital pre-processing step in most analyses [1,2].

Over the years, various methods of shadow detection proposed are object-based shadow detection methods which classify clouds, their shadows, and non-shadowed regions by applying image segmentation at different bandwidth images of HSI imagery, e.g., [3,4], and color invariance-based shadow detection methods create RGB in invariant color space and exploit it for classification as in [5]. Some algorithms use band indices to detect shadows in an HSI image [6,7]. Another class of algorithms require an a priori Digital Surface Model (DSM) [8] or Terrestrial Laser Scanning data together with the HSI image to find shadows cast from scene geometry [9].

Beril et al. [10] proposed a color-invariant function for detecting buildings. Once buildings are detected, then they used the grayscale histogram of the image to detect shadows around the building using the Otsu algorithm [11], their algorithm is referred to as Beril's algorithm in the results. In another contribution, Teke et al. [12] proposed a false color space consisting of red, green and near infrared (NIR) bands. They dropped the blue color because it contains scattered light and removing it will increase the contrast between shadow and non-shadow regions, and will facilitate detection. They have named their algorithm the Land Use Land Cover classification method, or LULC in their code. Therefore, we will refer to their work as LULC algorithm in our analysis. Sevim et al. [13] modified the C_1,C_2,C_3 color space [14] to accommodate the NIR band, and supplemented it to become the C'_1,C'_2,C'_3,C'_4 color space. We refer to their work as RGBN algorithm. Gevers et al. [15] proposed color-invariance functions to separate shadow and non-shadow regions; their algorithm is referred to as Gevers' algorithm.

The approaches mentioned above are limited to use in particular bands, and may not use the complete hyperspectral data. Our algorithm may ideally use HSI data and can be down-scaled to multispectral imagery only in cases where data-acquisition sensor response is known. Spectral response is typically available for most Earth-observation satellites. Our algorithm does not address RGB images because QUAC [16] may not be applied to retrieve reflectance from RGB images. More importantly, our algorithm provides a mathematical foundation for shadow detection based on the RT model and highlights the sources of errors. Retrieval of radiance components by machine learning enables it to be extended for shadow compensation.

2. Radiative Transfer Model-Based Relationship between Shadowed and Non-Shadowed Regions

The scope of this research is within optical shadowing, and therefore, the subsequent discussion does not consider thermal radiance and shadowing and their relative terms in Radiative Transfer (RT) equations. This section is divided into two parts—the first presents a general description of the RT equation highlighting relevant parameters and elaborating the sources of errors and their impact on this work, and the second establishes the proposed general relationship between variably illuminated regions based on the RT equation.

2.1. Radiative Transfer Equation

The RT equation for at-sensor radiance within an optical spectrum is given as Equation (1), [17].

$$Ls(\theta_i, \theta_r, \phi, \rho, \Omega_i) =$$
$$f_r(\theta_i, \theta_r, \phi)E_s(\theta_i)cos(\theta_i)\tau_i(\theta_i)\tau_r(\theta_r)\rho + \tau_r(\theta_r)\rho \int_F \iint_{\Omega_i} f_r(\theta_i, \theta_r, \phi)cos(\theta_i)L_d^{\Omega_i}(\theta_i, \phi)d\Omega_i + L_p(\theta_i, \phi) \quad (1)$$

where Ls is total at-sensor radiance, f_r is the Bidirectional Reflection Distribution Function (BRDF) [18], E_s is exoatmospheric solar spectral irradiance, τ_i and τ_r are incident and reflected transmittance, L_p is

path radiance, F is the view factor of sky, and solid angle Ω_i is given as $d\Omega_i = sin\theta_i d\theta_i d\phi$. Equation (1) is rephrased as Equation (2) for brevity.

$$Ls = (L_r + L_d)\rho + L_p \tag{2}$$

where

$$L_r = f_r(\theta_i, \theta_r, \phi)E_s(\theta_i)cos(\theta_i)\tau_i(\theta_i)\tau_r(\theta_r); \tag{3}$$

$$L_d = \tau_r(\theta_r) \underbrace{\int_F}_{e_1} \iint_{\Omega_i} \underbrace{f_r(\theta_i, \theta_r, \phi)}_{e_2} cos(\theta_i)L_d^{\Omega_i}(\theta_i, \phi)d\Omega_i \tag{4}$$

The analytical form of Equation (1) is achieved by assuming Lambertian BRDF of material at the cost of estimation error, which we will briefly derive. Moreover, we will subsequently describe error sources e_1 *(sky-view factor error)* and e_2 *(BRDF error)* of Equation (4).

Lambertian BRDF yields diffuse flux, which is computed by integrating diffuse radiance over the hemisphere, which is related to Equation (5). A denotes the area on which flux is incident, E is the incident irradiance on area A, and is given as $E = \frac{\Phi}{A}$.

$$\Phi = \frac{\int_0^{2\pi} \int_0^{\pi/2} L_d A cos\theta_i r^2 sin\theta_i d\theta_i d\phi}{r^2} \tag{5a}$$

$$\Phi = L_d A \underbrace{\int_0^{2\pi} \int_0^{\pi/2} cos\theta_i sin\theta_i d\theta_i d\phi}_{\pi} \tag{5b}$$

$$E = L_d \pi \tag{5c}$$

The concept of spherical albedo of atmosphere and the property in Equation (5c) transforms Equation (1) into Equation (6). Further description is found in [19–21]. Equation (6) forms the basis of discussion in Section 2.2 where ELM calibration panels are described in terms of it.

$$Ls(\rho) = L_p + \frac{\tau E_g(0)\rho/\pi}{1 - s\rho} \tag{6}$$

where τ, $E_g(0)$, and s are the total ground-to-sensor transmittance, global flux on the ground for $\rho = 0$, and the spherical albedo of the atmosphere, respectively. τ is the sum of the direct and diffuse transmittance, i.e., $\tau = \tau_{dir} + \tau_{diff}$. Equation (6) shows that the effective global flux $E_g(\rho) = \frac{E_g(0)}{1 - s\rho}$ depends on the ground reflectance and spherical albedo [22].

2.1.1. BRDF Error

First, we briefly define Phong BRDF model to illustrate how Lambertian distribution is compared to other materials' BRDF. Phong BRDF is given as

$$f_r = \frac{\rho_d}{\pi} + \frac{\rho_s(n+1)cos^n\alpha}{2\pi cos\theta_i} \tag{7}$$

ρ_d is the diffuse reflection constant, n determines the angular divergence of the lobe, and ρ_s determines the peak value or "strength" of the lobe [23]. Figure 1 depicts BRDF distribution for different values of ρ_d, ρ_s, and n. Lambertian is a special case where $\rho_s = n = 0$, leaving only the first term of Equation (7). In Figure 1,

the second row shows BRDF of some real materials which approximate those in Figure 1. Rubber's BRDF is closer to our assumption.

Components of reflected light based on Phong reflection model are ambient, diffuse, and specular radiance, as shown in Figure 2. Ambient reflection is both in shadowed and non-shadowed regions [13]. Diffuse (Lambertian) is assumed in this work; therefore, the specular component is the actual source of error, e_1, of Equation (4). Ref. [13] assumes that the specular reflection can be ignored for most urban and rural areas because these images are usually matte; we also maintain this assumption in this work.
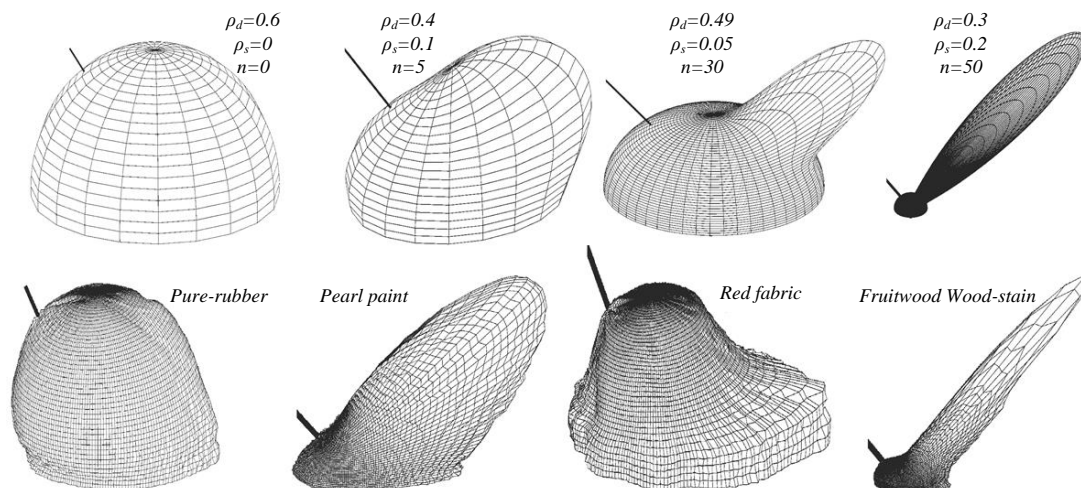
Figure 1. First row: Phong BRDF distribution for different values of ρ_d, ρ_s and n, refer. Second row: BRDF distribution for different materials, using visible spectrum [24,25].
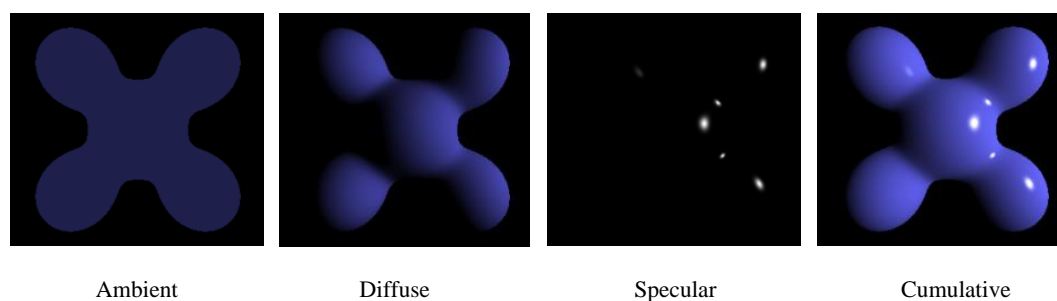

Figure 2. Phong reflection, components (Ambient, Diffuse, Specular) and their cumulative effect [26].

2.1.2. Sky-View Factor Error

3-D geometry in the scene tends to cause visual occlusion to the sky-view Line of Sight (LoS). Sky-view LoS is quantified as a sky-view factor and normalized between zero and one, where zero is complete occlusion and one is clear sky-view [27]. This occlusion is one of the causes of shadow casting (apart from clouds, partial solar eclipse, etc.). Therefore, the presence of this error in the estimation would enhance the discrimination between shadowed and non-shadowed regions; hence it facilitates detection of variable illumination.

2.2. Proposed Radiative Transfer Model-Based General Relationship between Variably Illuminated Regions

ELM represents the at-sensor radiance from in-scene white and black calibration panels. Radiance reflected from the white panel L_w and black panel L_b is deduced from Equation (6) and shown in Equations (8) and (9), respectively.

$$L_w = L_p + \frac{\tau E_g(0)/\pi}{1 - s\rho} \tag{8}$$

$$L_b = L_p \tag{9}$$

In Equation (8), L_w radiance includes both path radiance and total ground-reflected radiance, which are the first and second terms, respectively. We introduce two parameters α and β in Equation (10) that represent the total ground-reflected radiance and path radiance, respectively, and brings the convention to Equation (10), which is called the ELM equation.

$$\rho = \alpha Ls + \beta \tag{10}$$

where α and β are,

$$\alpha = 1/(L_w - L_b) \tag{11}$$

$$\beta = L_b/(L_w - L_b) \tag{12}$$

Reflectance ρ is independent of illumination conditions, and enables us to rephrase Equation (10) as Equation (13), emphasizing only parameters α and β.

$$\Rightarrow f(\alpha, \beta) = Ls = \frac{\rho - \beta}{\alpha} \tag{13}$$

We rewrite Equation (10) for shadowed (sub-scripted S) and non-shadowed (sub-scripted NS) regions of the scene as Equations (14) and (15), respectively.

$$\rho = \alpha_{NS} Ls_{NS}(\rho) + \beta_{NS} \tag{14}$$

$$\rho = \alpha_S Ls_S(\rho) + \beta_S \tag{15}$$

Equating Equations (14) and (15) we get Equation (16).

$$Ls_{NS} = \frac{\alpha_S Ls_S(\rho) + \beta_S - \beta_{NS}}{\alpha_{NS}} \tag{16a}$$

$$Ls_{NS} = \frac{\alpha_S Ls_S(\rho) + \Delta\beta}{\alpha_{NS}} \tag{16b}$$

where $\Delta\beta = \beta_S - \beta_{NS}$. Equation (16) is rephrased as Equation (17).

$$Ls_{NS} = \gamma Ls_S(\rho) + \delta \tag{17}$$

where

$$\gamma = \alpha_S/\alpha_{NS} \tag{18a}$$

$$\delta = \Delta\beta/\alpha_{NS} \tag{18b}$$

Equation (17) shows γ and δ that are two unknown parameters responsible for illumination variability between shaded and non-shaded regions. Ideally, if $\gamma = 1$ and $\delta = 0$, there is no variability in illumination across the scene.

3. Proposed Multi-Layered Regression Learning Algorithm

In the previous section, a general relationship between illumination under shadowed and non-shadowed region within an HSI image is established. Estimation of discriminant parameters α_{NS}, β_{NS}, α_S, and β_S is vital for good detection. We divide our learning algorithm into three phases: (i) regression learning; (ii) feature learning; and (iii) classification, as shown in Figure 3.
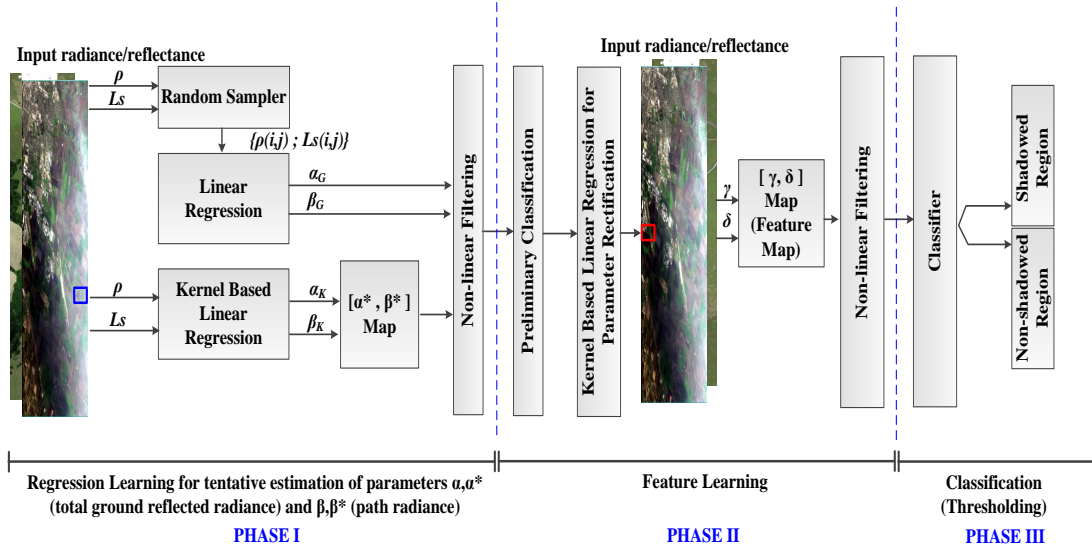

Figure 3. Design of proposed multi-layered regression learning algorithm is shown to have three phases: (i) regression learning; (ii) feature learning; and (iii) classification. Inputs are radiance and reflectance images that are randomly sampled to find neutral (both shadowed/non-shadowed regions) parameter estimates for α_G and β_G. A kernel of size 3 used for kernel-based linear regression parameters are α_K and β_K, representing a more localized and homogeneous (either shadowed or non-shadowed regions) estimate. In Phase II, more discriminating parameters are found in the second layer of regression learning, which rectifies parameters estimated by the previous phase. Two non-linear filter layers are shown. Eventually, the classifier layer segregates shadowed and non-shadowed regions.

During learning, Equation (10) becomes Equation (19), where $\tilde{\rho}$ is the approximate reflectance at any given search iteration, while reflectance computed from QUAC is the reference reflectance, referred in Equation (20). Therefore, the cost function $J(\alpha, \beta)$, to minimize is given in Equation (21), which is illustrated in Figure 4.

$$\tilde{\rho} = \alpha Ls + \beta \tag{19}$$

$$\rho_{QUAC} = \rho \tag{20}$$

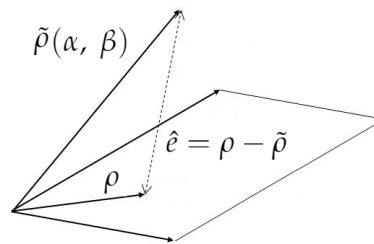$$J(\alpha, \beta) = \min_{\alpha, \beta} \{\rho - \tilde{\rho}\} \tag{21}$$

Figure 4. Regression learning reduces error $\hat{e}$ to find parameters α (total ground-reflected radiance) and β (path radiance).

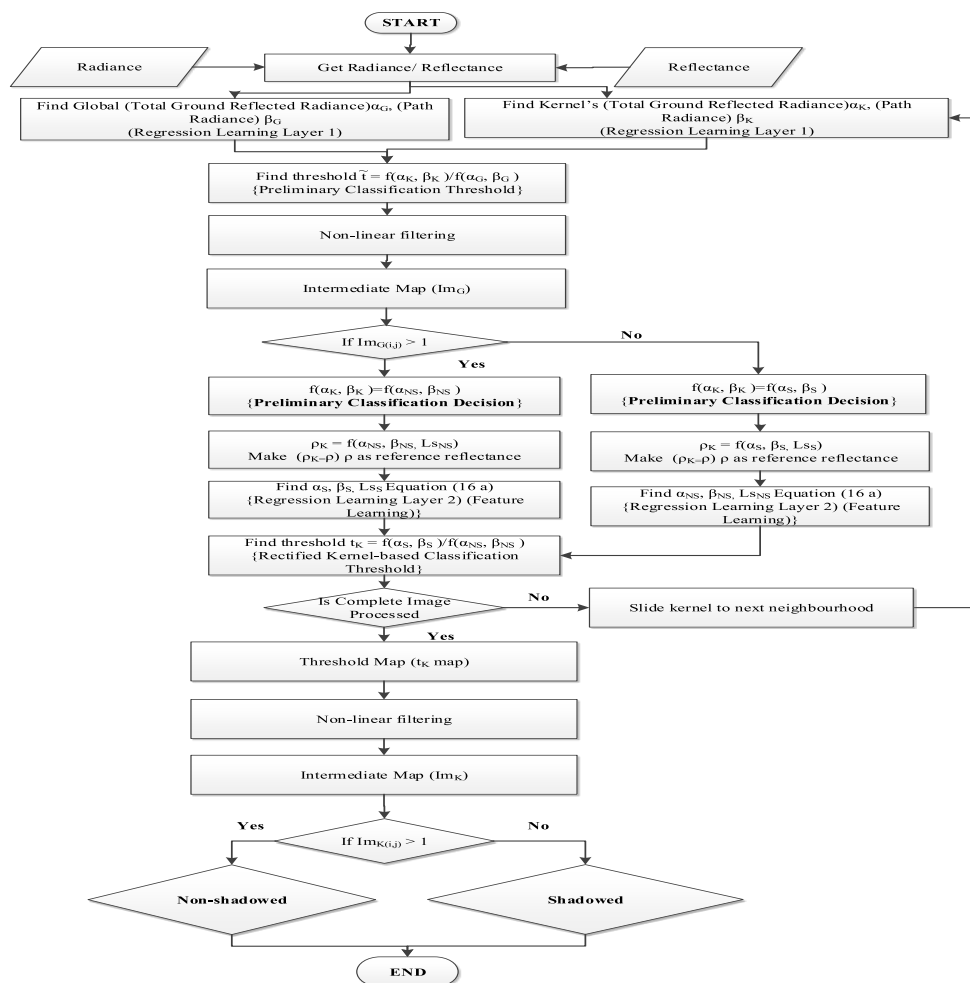A complete flowchart of multi-layered learning is shown in Figure 5.

Figure 5. Flowchart of the proposed multi-layered regression learning algorithm. Parameters α, β with subscript "$_G$" denote global, and "$_K$" denotes local (under the kernel). Moreover, the parameter subscript "$_{NS}$" denotes non-shadowed and "$_S$" denotes shadowed regions. ρ stands for reflectance, and t denotes threshold.

3.1. Regression Learning Phase

This phase is denoted as Phase I in Figure 3 and is further sub-divided into two steps:

- *Global Search*: Search for α_G and β_G estimates of Equation (19) on random samples drawn from the whole image given a global search across the scene, as shown in Figure 3.
- *Local Search*: Create a 3×3 kernel and search for parameters α_K and β_K in the kernel only i.e., localized search.
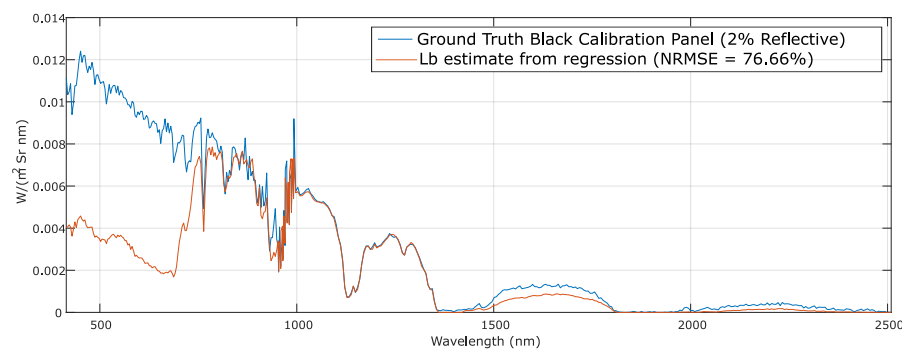
3.1.1. Global Search

Satellite/airborne images cover a larger landscape where the number of bands is more contiguous in HSI images. The high-resolution image has an immediate implication of an increase in both computing and memory requirements. To reduce these requirements, we introduced random sampling on the whole image. A random sampler selects several samples from the whole image and regression learning is performed on these samples to estimate Empirical Line Method (ELM) parameters. As the input image and samples contain neutral (both shadowed and non-shadowed) regions, parameters α_G and β_G also represent the same. Equation (6) for global search case is reformulated in Equation (22).

$$\rho_G = \alpha_G Ls_G + \beta_G \tag{22}$$

An estimate of L_w and L_b found in this phase is shown in Figure 6.
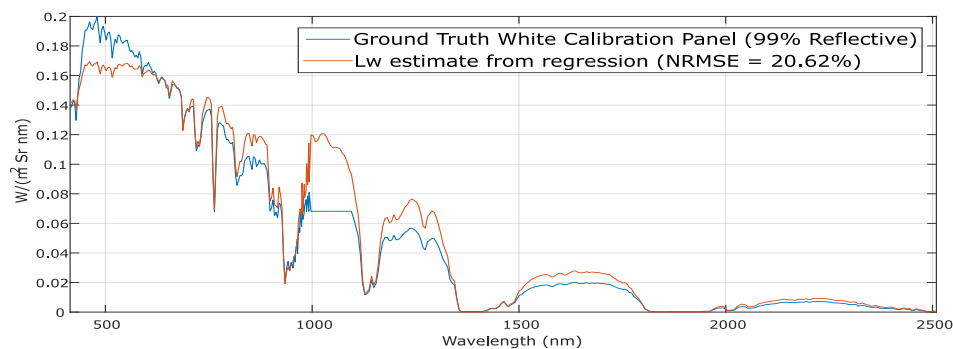
(**a**) White Panel Estimation

(**b**) Black Panel Estimation

Figure 6. White and black panel estimates from global search.

3.1.2. Local Search

In this part of regression learning, a 3×3 kernel is used. A smaller kernel provides a rationale for the assumption that pixels under the kernel are homogeneously illuminated. Correctness of this assumption is further reinforced for high-resolution images that possess lower ground-sampling distance (GSD). Due to learning on a sliding kernel, this search is more time-consuming than the global one. Outputs from this step are parameter (α_K, β_K) maps. An estimate of Lw_K and Lb_K for 465.611 nm is shown in Figure 7. In this case, we reformulate Equation (6) as Equation (23). Figure 8 shows the sliding kernel, and it is input and output parameters.

$$\rho_K = \alpha_K Ls_K + \beta_K \tag{23}$$
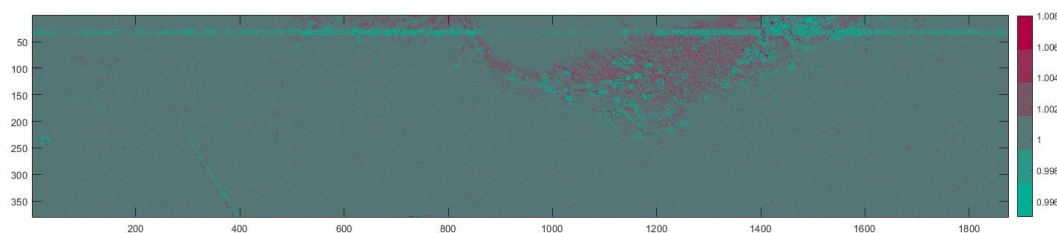
(**a**) 10×10 Kernel Sliding Window
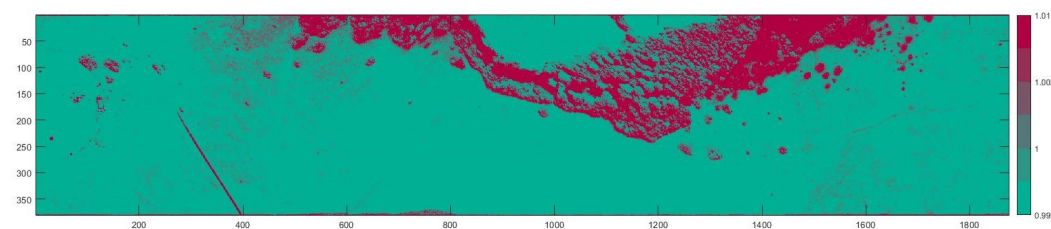
(**b**) 3×3 Kernel Sliding Window

Figure 7. Intermediate Map extracted by proposed method on coarse (10×10) and fine (3×3) sliding window ROI.
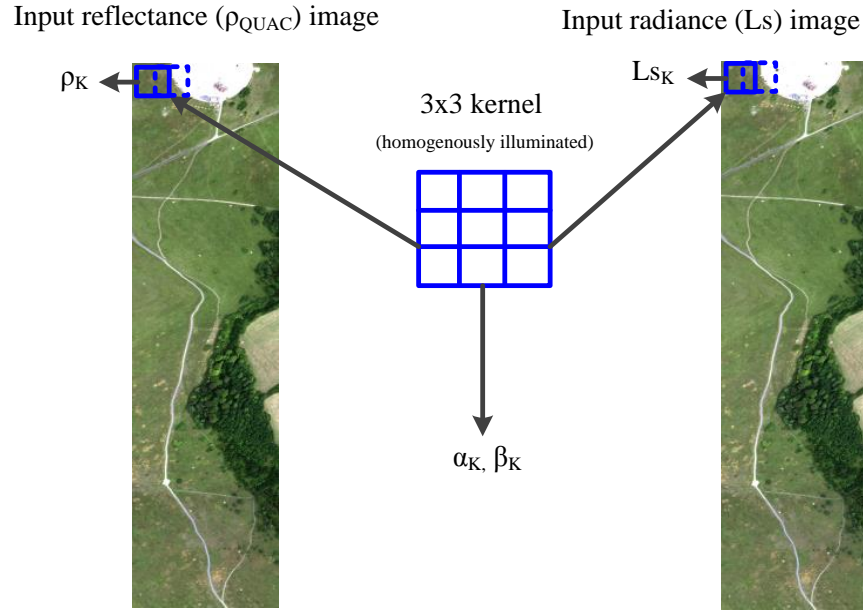
Figure 8. Kernel-based regression performs search for localized parameters α_K and β_K that are only within kernel.

3.2. Feature-Learning Phase

In the previous phase, we have found both global and local parameters, although we assume that since a smaller kernel has homogeneous illumination it is yet to be ascertained whether it is shadowed or non-shadowed. This phase will establish the discriminant function, first for tentative and then for the final classification. Components of feature-learning phase in Figure 3 are described in the subsequent sections.

3.2.1. Preliminary Classification

Global radiance Ls_G is estimated by inclusion of both shadowed and non-shadowed regions. Ls_K is however assumed to be either of them. A global and local version of Equation (13) is given as

$$f(\alpha_K, \beta_K) = Ls_K = \frac{\rho_K - \beta_K}{\alpha_K} \tag{24a}$$

$$f(\alpha_G, \beta_G) = Ls_G = \frac{\rho_G - \beta_G}{\alpha_G} \tag{24b}$$

In this phase, a ratio between $f(\alpha_K, \beta_K)$ and $f(\alpha_G, \beta_G)$ is calculated by Equation (25), which yields approximate threshold $\tilde{t}$.

$$g(\alpha_G, \beta_G, \alpha_K, \beta_K) = \frac{f(\alpha_K, \beta_K)}{f(\alpha_G, \beta_G)} = \tilde{t} \tag{25}$$

$$\tilde{t} = \begin{cases} > 1 & , \quad f(\alpha_K, \beta_K) = f(\alpha_{NS}, \beta_{NS}), \quad \text{(Non-shadowed)} \\ < 1 & , \quad f(\alpha_K, \beta_K) = f(\alpha_S, \beta_S), \quad \text{(Shadowed)} \end{cases} \tag{26}$$

The kernel region is assigned a shadow or non-shadow label based on Equation (26). If the value of t is greater than one, then the region under the kernel is more likely to be non-shadowed than otherwise. The above is under the intuitive assumption that the scene has more non-shadowed regions than shadowed ones. This provides a preliminary classification as shown in Figures 3 and 5, and leads us to either $(Ls_{NS}, \alpha_{NS}, \beta_{NS})$ or $(Ls_S, \alpha_S, \beta_S)$ of Equation (16a), as per our assumption. The selection of either of these parameter sets is shown as two potential flows in Figure 5.

In practice, $\tilde{t}$ is noisy data across the optical spectrum. Therefore, non-linear filtering is applied to it for smoothing and creating an intermediate map. As filtering is applied on both global and kernel thresholds $\tilde{t}$ and t_K, respectively, it is discussed later in Section 3.3.

3.2.2. Regression Learning for Parameter Rectification

To create a more reliable discriminant threshold function than that of Equation (26), we need to establish a relationship between our processed parameters to bring it in the shape of Equation (16). Global and local processing has provided us with a mathematical basis for tentative classification, which is plausible under the laws of physics. We suppose that the kernel region is labelled, by preliminary classification, as non-shadowed which implies that we found $(Ls_{NS}, \alpha_{NS}, \beta_{NS})$; however $(Ls_S, \alpha_S, \beta_S)$ of Equation (16a) are yet to be determined for the region under the kernel. We rewrite Equation (16a), for convenience.

$$Ls_{NS} = \frac{\alpha_S Ls_S(\rho) + \beta_S - \beta_{NS}}{\alpha_{NS}} \tag{27}$$

Ls_S, α_S, β_S are determined by another layer of regression learning. The cost function $Q(Ls_S, \alpha_S, \beta_S)$, to minimize is given in Equation (21), which is illustrated in Figure 9.

$$\tilde{\rho} = \alpha Ls_S + \beta_S \tag{28}$$

$$\rho = \alpha_{NS} Ls_{NS} + \beta_{NS} \tag{29}$$

$$Q(Ls_S, \alpha_S, \beta_S) = \min_{Ls_S, \alpha_S, \beta_S} \{\rho - \tilde{\rho}\} \tag{30}$$
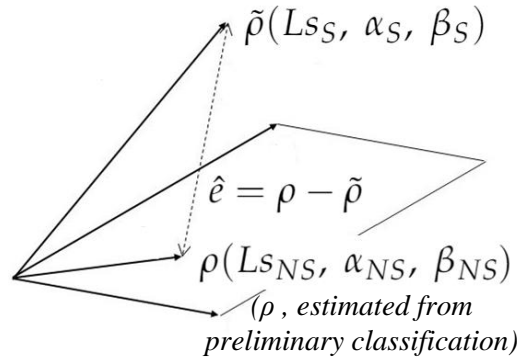

Figure 9. Regression learning reduces error $\hat{e}$ to find parameters α_S (total ground-reflected radiance in the shadow region) and β_S (path radiance in shadow region) and L_S (total radiance under shadow region).

This learning is performed by Algorithm 1. After this learning phase, we estimated all unknown parameters of Equation (27), therefore we may rewrite Equation (25) as Equations (32) and (31). Please note that this equation is only for the kernel region, which has the same reflectance ρ and has shadow and non-shadow parameters instead of global and local.

Algorithm 1 Gradient-descent algorithm for global/kernel-based search

---

1: Let HSI Radiance Image be "L" with "s" samples and "b" bands, and its reflectance estimated by Run

   QUAC to find reflectance R
2: Assign outputs L_w and L_b as two zeros vector of "b"
3: Let stepSize be 0.01 with a decay of 0.995
4: **while** $(\Delta L_w \leq 1 \times 10^{-10}$ and $\Delta L_b \leq 1 \times 10^{-10})$ **do**
5:    Select 2 pixels at random
6:    Estimate δL_w and δL_b for selected pixels by ELM equation
7:    $L_w = (1 - \text{stepSize}) \times L_w + \text{stepSize} \times \delta L_w$
8:    $L_b = (1 - \text{stepSize}) \times L_b + \text{stepSize} \times \delta L_b$
9:    stepSize = stepSize $\times$ decay

It is extremely important to note that if preliminary classification finds $(\alpha_K, \beta_K) = (\alpha_{NS}, \beta_{NS})$ then t takes the form of Equation (31). If it finds $(\alpha_K, \beta_K) = (\alpha_S, \beta_S)$ then t takes the form of Equation (32).

$$g(\alpha_{NS}, \beta_{NS}, \alpha_S, \beta_S) = \frac{f(\alpha_{NS}, \beta_{NS})}{f(\alpha_S, \beta_S)} = t \tag{31}$$

$$g(\alpha_S, \beta_S, \alpha_{NS}, \beta_{NS}) = \frac{f(\alpha_S, \beta_S)}{f(\alpha_{NS}, \beta_{NS})} = t \tag{32}$$

Replacing both numerator and denominator of Equations (31) and (32) by right-hand side of Equation (13) in their respective form, we get Equations (33) and (34).

$$t = \frac{\alpha_{NS}(\rho - \beta_S)}{\alpha_S(\rho - \beta_{NS})} \tag{33}$$

$$t = \frac{\alpha_S(\rho - \beta_{NS})}{\alpha_{NS}(\rho - \beta_S)} \tag{34}$$

Threshold t is a better estimate than $\tilde{t}$, because both parameters are computed for the region under the kernel. This process is termed as rectification in Figures 3 and 5. When the kernel slides through the image, it estimates t at each iteration. On completion, it creates an intermediate map for the whole image. As we discussed in Section 3.2.1, both $\tilde{t}$ and t are computed for all bands of HSI image and tend to get very noisy at some bands. This problem is tackled by filtering, which is covered in Section 3.3.

3.3. Filtering

Filtering is a supplementary but vital step for the performance of detection. This is the third and final layer of regression learning. Here we estimate an activation function for a non-linear filter based on Equation (35a). The cost function $J(x, t)$ finds the value of threshold scalar x which maximizes $h(x, t)$. The subscript k shows the band number of the HSI image with B bands. Threshold t is estimated from

Equations (33) and (34) and $\tilde{t}$ is found from (25), respectively. We applied this filter on our test dataset which reduces noise, as shown in Figure 10.

$$h(x,t) = \sum_{k=1}^{B} e^{-(x-t_k)} \tag{35a}$$

$$J(x,t) = \max_{x} \{h(x,t)\} \tag{35b}$$

The gradient-descent algorithm finds x in pseudo-code presented in Algorithm 2. Figure 10 shows output after filtering. Intermediate maps are created after the filtering process.
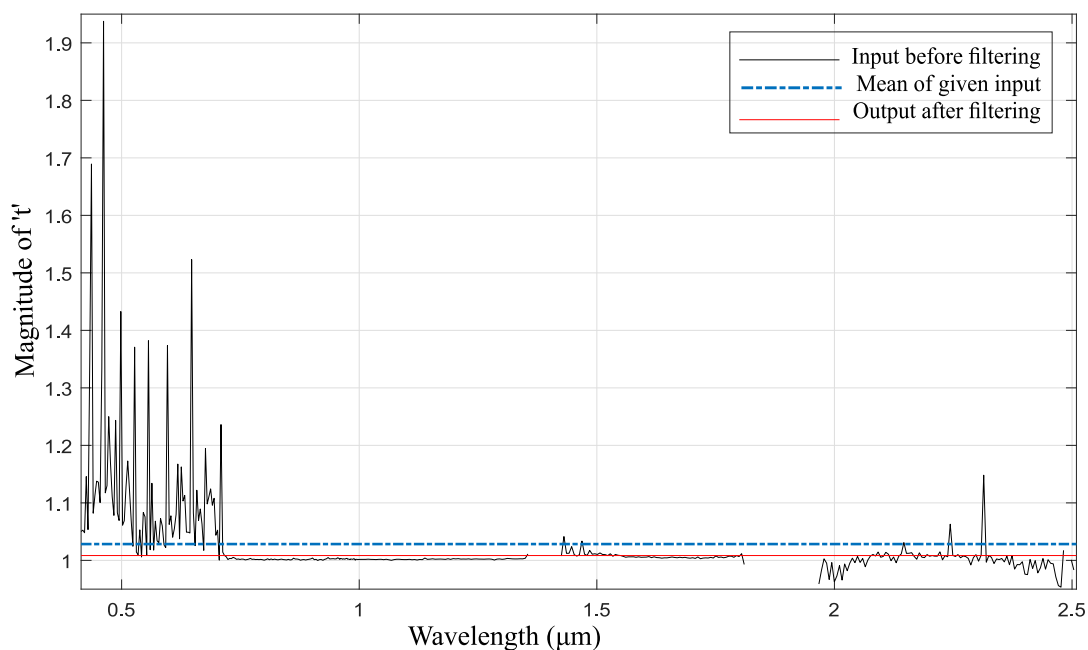

Figure 10. Effect of filtering on a given pixel.

Algorithm 2 Gradient-descent algorithm to estimate h(x, t)

1: For an n-dimensional input 't' and 'M' be the maximum number of iterations

2: Let the initial value of 'x' be mean(t)

3: Let output h_x be h(x, t)

4: Let stepSize be 0.01 with a decay of 0.995

5: **for** i = 1 to M **do**

6:      h_d = h((x-stepSize), t)

7:      h_i = h((x+stepSize), t)

8:      **if** $h_x < h_d$ **then**

9:          $h_x = h_d$

10:          x = x − stepSize

11:      **if** $h_x < h_i$ **then**

12:          $h_x = h_i$

13:          x = x + stepSize

14:      stepSize = stepSize × decay

3.4. Classification

The threshold map generated by Feature learning and filtering represents multiple levels of illumination. Therefore, we may generate a binary or multi-label classification from this intermediate map.

3.4.1. Binary Map

To create a binary map, our approach is similar to Equation (26) threshold function, shown in Equation (36) as:

$$t = \begin{cases} > 1 & , \quad \text{(Non-shadowed)} \\ < 1 & , \quad \text{(Shadowed)} \end{cases} \tag{36}$$

Figure 7b,c show the binary map for Selene HSI dataset using 10×10 and 3×3 kernel sizes, respectively.

3.4.2. Multi-Label Classification

In practice, any decrease in radiance caused by shadows may include soft and hard shadows, depending on the amount of direct radiance blockage and scattering, which causes multi-level illumination. In order to accommodate this case when a user-defined threshold is also not a priori, dual marching squares algorithm [28] is used to automate the process. By employing this algorithm, multiple levels of illumination are separated in different contour levels, as covered in Section 5.1.3. In this algorithm, one may define several levels to generate the contour map accordingly.

4. Experimental and Validation Data

4.1. Experimental Data

For experimental validation, we have used two real images. Our algorithm is tested on the Selene SCIH23 dataset [29] which was acquired by the Defence Science and Technology Laboratory (DSTL) covering 0.4 to 2.5 μm. It is a registered image which is separately taken from HySpex VNIR-1600 (160 bands) and SWIR-384 (288 bands) sensors mounted on the same airborne platform. This scene was acquired near Salisbury, UK, on 12 August 2014 BST 13:00:04. The registered image has a GSD of 70×70

cm, QUAC was applied using ENVI software for atmospheric compensation. The second image is taken from AVIRIS (224 bands), covering Modesto, California (Long 121°17′45.9″ N Lat 37°56′44.37″ W to 121°4′50.31″ N 37°51′55.67″ W) on 10 February 2015 BST 22:00:04. The latter is publicly available from https://aviris.jpl.nasa.gov/. RGB images of Selene and Modesto scenes are shown in Figure 11a,b, respectively.

(**a**)

(**b**)
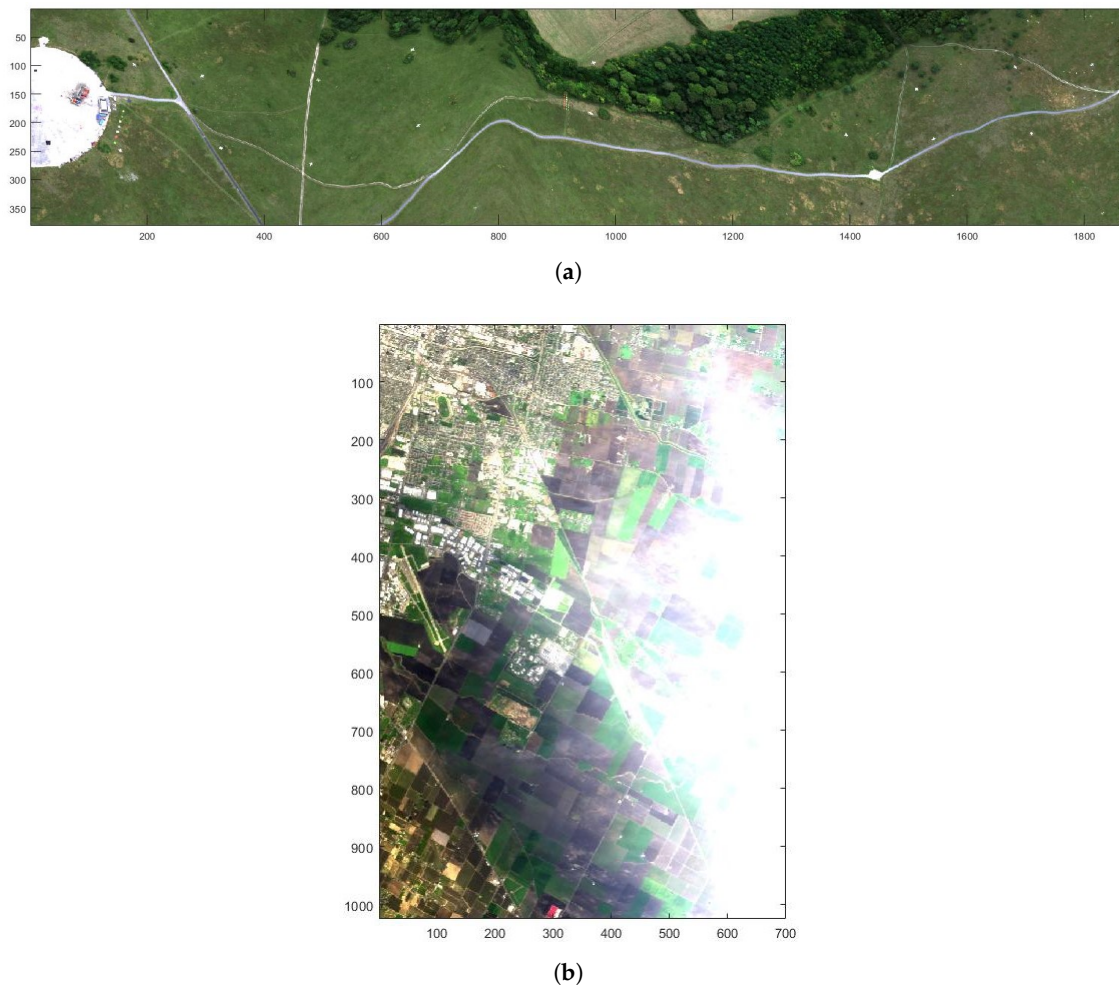
Figure 11. Experimental HSI imagery (**a**) Selene SCI H23 (0.4 μm–2.51 μm, 448 bands), UK; (**b**) AVIRIS Modesto (0.36 μm–2.49 μm, 224 bands), California, USA.

4.2. Validation Data

The Selene SCIH23 scene's terrain was mapped with a high-resolution LIDAR to create a contour-based DSM of the scene. This DSM is used in a ray-tracer to generate a shadow map, which is termed as ground truth for Selene results. Modesto does not have DSM data available. Therefore, classification performance is tested on visual perception only.

5. Results and Validation

In this section, we will present results of the Selene and Modesto scenes. Construction of GT for the Selene scene enables us to provide a more quantitative validation for it by means of confusion matrix,

while for Modesto, it is more qualitative and evaluated visually. For comparative analysis, we have considered both classical methods exploiting descriptor-based shadow detection from RGB input, as in Gevers [15], and Beril [10] (see [30] for source code), and methods that take in multispectral radiance input, as in RGBN [13] requiring RGB and a single NIR band, and false-color shadow detection method, LULC [12] (see [31] for source code), requiring five input bands within 0.3 μm to 2.5 μm.

5.1. Selene Scene

Firstly, the algorithm is executed on the Selene SCI H23 scene for both global and local search sub-phases of the regression learning phase. This scene has several white and black calibration panels planted, providing us with a GT for comparison of estimated global L_w and L_b.

5.1.1. Result of Regression Learning on Whole Image (Global Search)

The comparative result for global search sub-phase is shown in Figure 6. The Normalized Root Mean Square Error (NRMSE) for L_w is 20.62%, which shows that the algorithm can reconstruct the white panel with substantial accuracy. The deviation in the blue region shows under-estimation of scattering/sky radiance, while lower magnitude in NIR is due to over-estimation of the adjacency effect, which is primarily caused by the abundance of vegetation in the background scene. The NRMSE for L_b is 76.66%, which is higher than L_w, it shows that the adjacency effect is over-estimated, hence the estimated black panel looks similar to the vegetating signature. These errors were incurred due to QUAC reflectance, which is the reference for calculating the L_w and L_b in our algorithm.

5.1.2. Results of Kernel-Based Regression Learning (Local Search)

Figure 7a shows the GT shadow map generated from ray-tracer on DSM. Figure 7b,c show results of local search sub-phase. In the case of former, the algorithm was run with a coarse kernel (10 × 10) and the latter with a fine (3 × 3) kernel. Better separation of shadowed and non-shadowed regions is visible with 3 × 3 kernel compared to its coarse counterpart.

In case of the fine kernel, shadows due to bumps on the terrain are also captured which seems to be missing from the coarser one.

5.1.3. Results of Classification

- *Intermediate and Binary maps*

  Algorithms adopted for comparative analysis create an intermediate map and apply manual threshold as suggested by respective authors. We propose a cut-off value of 1 due to the ratio between shadow and non-shadow radiance. Intermediate maps of all algorithms for the Selene scene are shown in Figure 12. After applying respective threshold values, we get binary classification maps for the Selene scene as shown in Figure 13. The Selene scene has shadowed regions near trees and concrete fields. The concrete field has some buildings that are casting shadows there. The grass field has a few bumps and has some stand-alone tree distribution along the track.

  The binary map for the Selene scene GT is constructed by applying a threshold on the DSM contour data. Similarly, the binary map of all competitive algorithms is also constructed. A confusion matrix of each algorithm against GT is tabulated, which is shown in Table 1. Because the area of the non-shadowed region is significantly larger, covering 92.6377% of the scene, compared to the shadowed region, covering 7.3623% as estimated from the DSM binary map, we divide by the cardinality in the table to minimize bias in overall percentage accuracy.
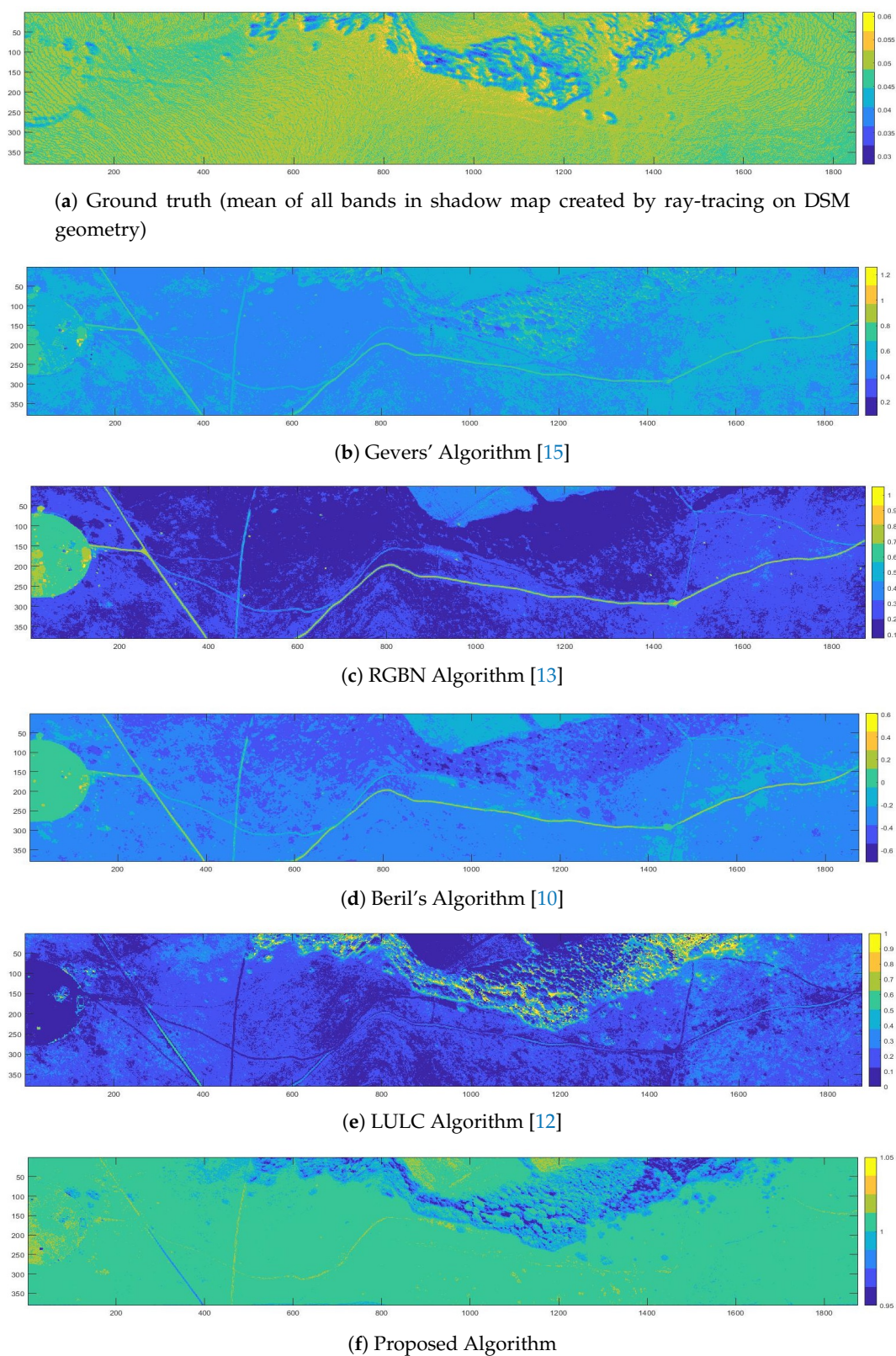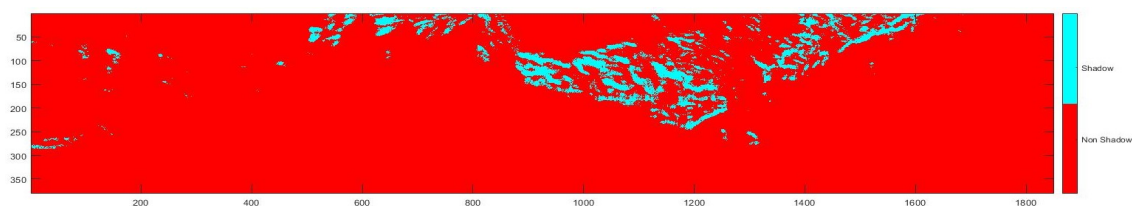
(**a**) Ground truth (mean of all bands in shadow map created by ray-tracing on DSM geometry)

(**b**) Gevers' Algorithm [15]

(**c**) RGBN Algorithm [13]

(**d**) Beril's Algorithm [10]

(**e**) LULC Algorithm [12]

(**f**) Proposed Algorithm

Figure 12. SELENE: Intermediate maps extracted by shadow detection methods before thresholding for a binary classification map.
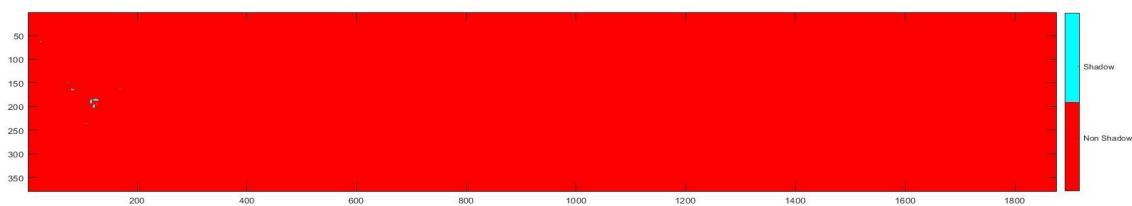
(**a**) Ground Truth: (Thresholding the Ground Truth map of Figure 12 (a) creates a binary map.)
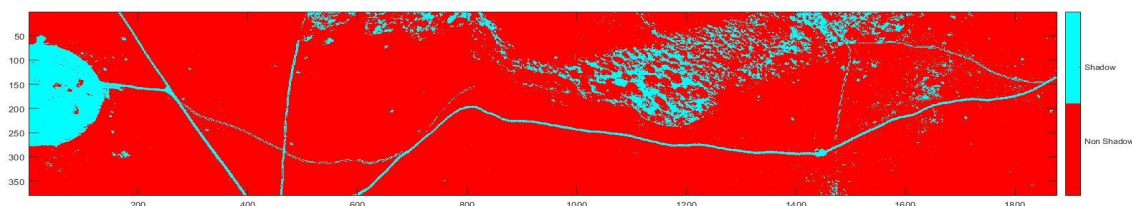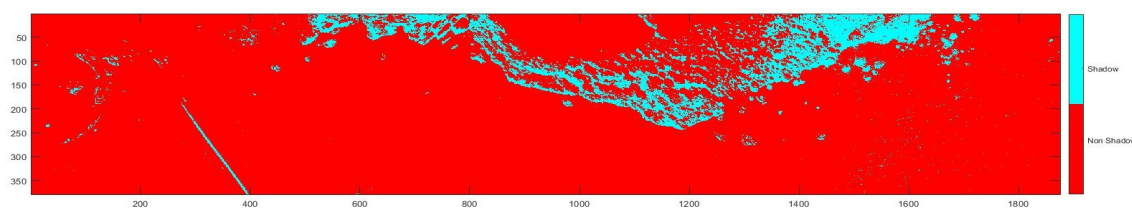

(**b**) Gevers: It detects roads and concrete field as shadows because they have lower reflectance.
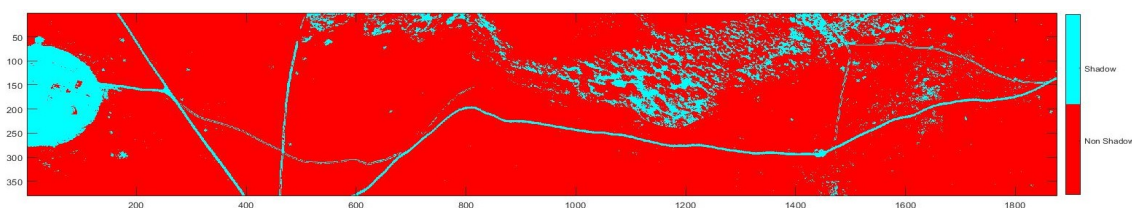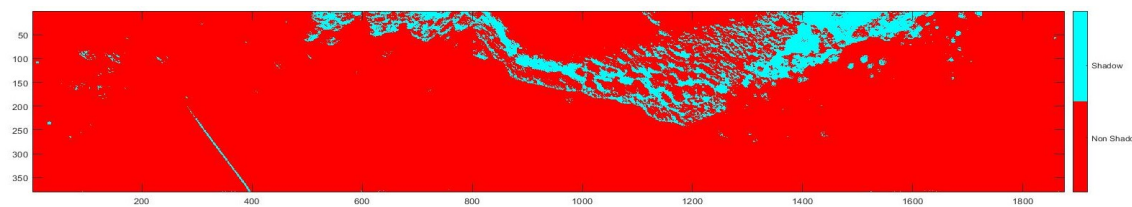

(**c**) RGBN: Similar to Gevers, it detects roads and concrete field as shadows.

(**d**) Beril: It only detects some building shadows on the concrete field.

(**e**) LULC: It detects shadows on both vegetation and concrete field regions, as expected. A portion of road is also detected as shadow.

(**f**) Proposed: Similar to LULC, it also detects shadows on both vegetation and concrete field regions. Similarly, a portion of road is also detected as shadow, but it is slightly thinner.

Figure 13. SELENE: Binary classification map separating shadow and non-shadow regions.

Table 1. Evaluation of accuracy of the methods in Figure 13 where true positive and true negative is successful detection of shadowed and non-shadowed areas respectively compared with ground truth (GT) DSM simulated map, and "*n*" represent cardinality.

Method	$\frac{n(\text{True Positive})}{n(\text{GT Shadow})} \times 100$	$\frac{n(\text{False Positive})}{n(\text{GT Shadow})} \times 100$	$\frac{n(\text{True Negative})}{n(\text{GT Non-Shadow})} \times 100$	$\frac{n(\text{False Negative})}{n(\text{GT Non-Shadow})} \times 100$
Gevers	28.9013%	112.1443%	89.7015%	5.5403%
RGBN	27.1437%	108.9442%	89.9558%	5.6801%
Beril	0%	**0.2418%**	**98.5948%**	7.8373%
LULC	**39.5298%**	72.6458%	92.8406%	4.6957%
Proposed	**40.4656%**	64.4731%	93.4901%	**4.6213%**

The average of true positive and true negative of Gevers, RGBN, Beril, LULC, and the proposed method are 59.3014%, 58.5497%, 49.2974%, 66.1852%, and 66.9779% respectively.

- *Multi-level Classification*

The DSM-based GT shadow map in Figure 14 shows contours; therefore, the shadow map from our algorithm is also converted to contour levels. Both datasets are generated to have four contour levels and are compared one to one. The first layer has the lowest accuracy, with 53.09%, which is in the non-shadowed region. The second layer has a partial shadowed region; a raise in the accuracy is observed in this case. As we step further to Layer 3 and Layer 4, which appear to capture shadowed regions, the accuracy increased to 69.6% and 88.2%, respectively.
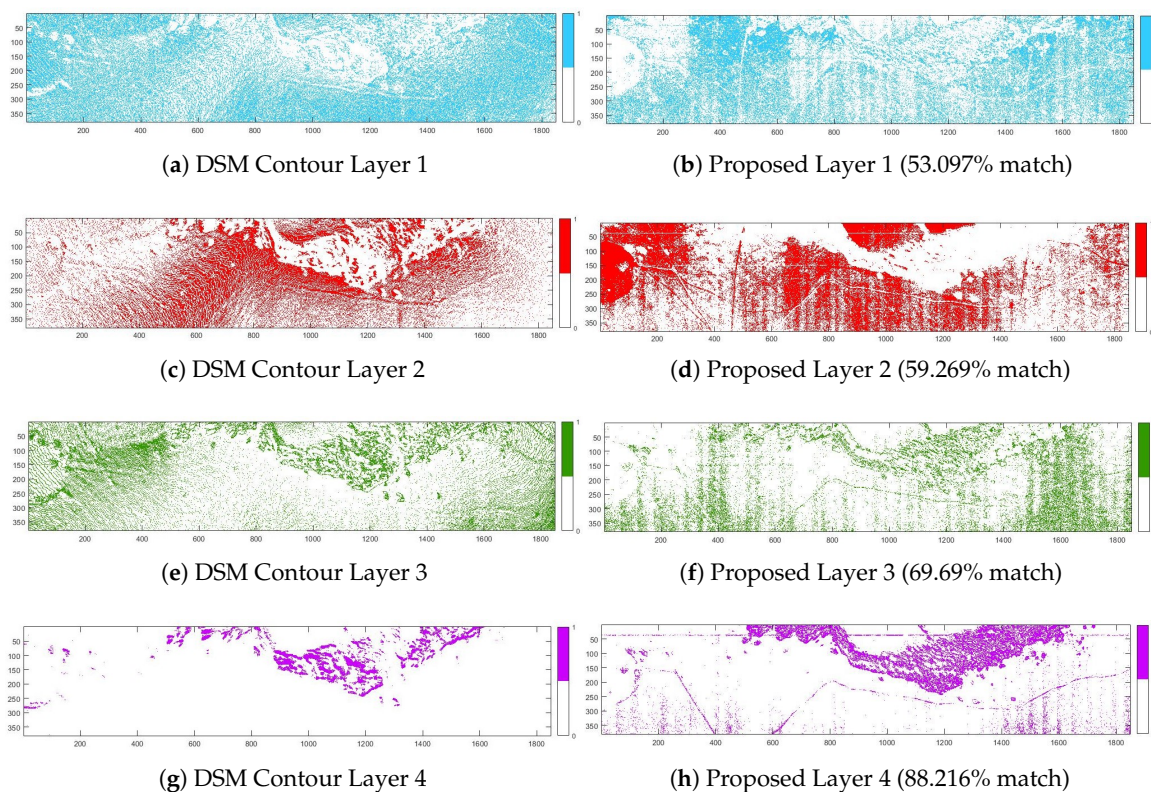

(**a**) DSM Contour Layer 1

(**b**) Proposed Layer 1 (53.097% match)

(**c**) DSM Contour Layer 2

(**d**) Proposed Layer 2 (59.269% match)

(**e**) DSM Contour Layer 3

(**f**) Proposed Layer 3 (69.69% match)

(**g**) DSM Contour Layer 4

(**h**) Proposed Layer 4 (88.216% match)

Figure 14. SELENE: Multi-level classification as contour layers 1 to 4, with decreasing order of illumination (1 = non-shadow, 4 = shadow).

5.2. Modesto Scene

The Modesto scene does not have DSM data. Therefore, ground truth for this scene is not available. The evaluation, in this case, is rather qualitative, based on visual perception. Intermediate maps of all algorithms of interest for the Modesto scene are shown in Figure 15. Their binary counterparts are in Figure 16. The proposed algorithm appears to classify the scene into three regions: (i) with clouds, (ii) with shadows, and (iii) the lit ground region. After the threshold is applied to the intermediate map, the scene is classified into shadowed and non-shadowed regions, as shown in Figure 16.
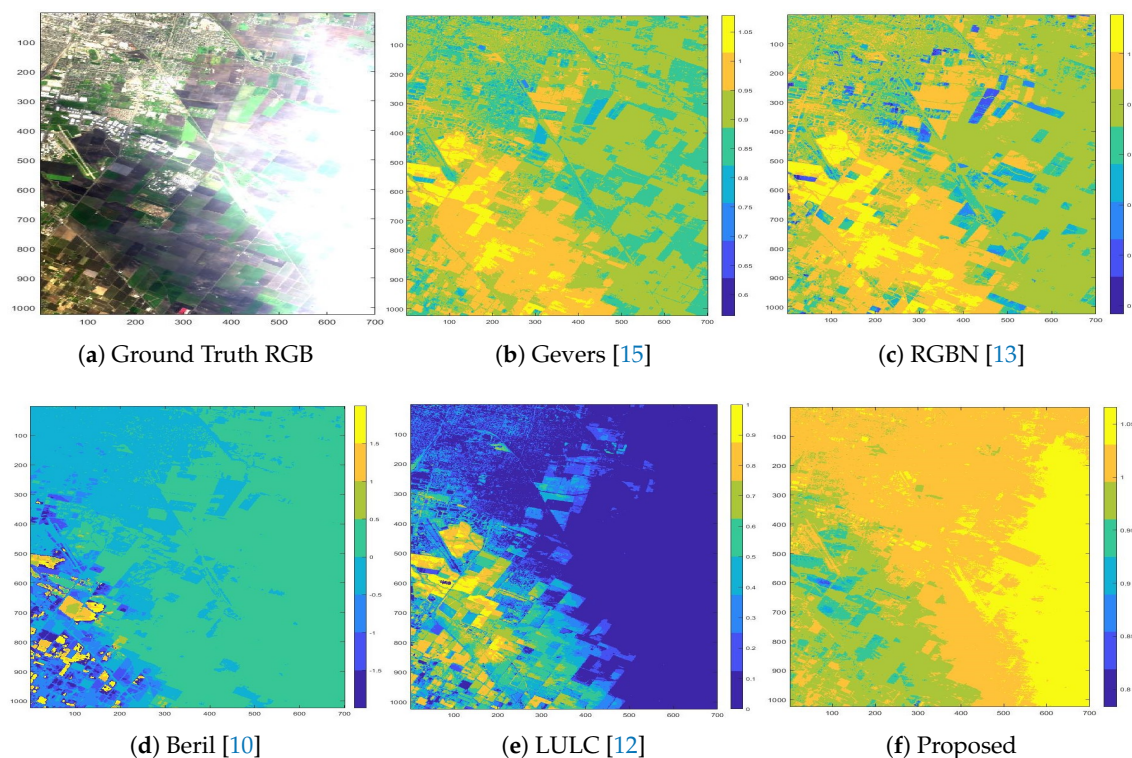

(**a**) Ground Truth RGB   (**b**) Gevers [15]   (**c**) RGBN [13]

(**d**) Beril [10]   (**e**) LULC [12]   (**f**) Proposed

Figure 15. MODESTO: Intermediate maps extracted by shadow detection methods before thresholding for a binary classification map.
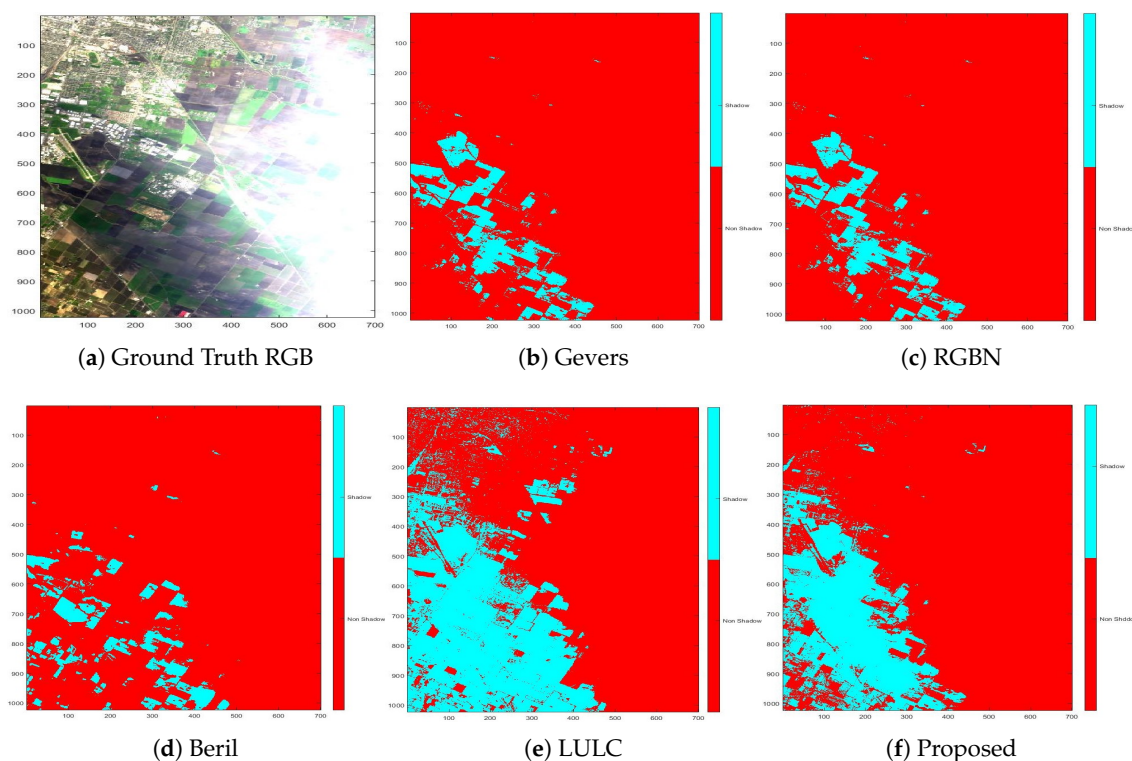
Figure 16. MODESTO: Binary classification map separating shadow and non-shadow regions.

6. Discussion

The Selene scene GT shadow map enables us to quantify the results of comparative algorithms in context. The proposed algorithm is shown to have 40.46% true positive, i.e., correct detection of a shadow region compared to other counterparts. In this case, the LULC algorithm is marginally inferior to our algorithm, yielding 39.52%. Beril, RGBN, and Gevers reach 0%, 27.14%, and 28.90%, respectively. In terms of false positive, i.e., detecting non-shadow as a shadow, our algorithm continues to perform better than other algorithms, reaching 64% compared to 72.64% of LULC, 112.14% of Gevers, and 108.94% of RGBN. Although Beril shows a lower value of 0.24%, it should be noted that it is the most biased performer, detecting 98% of the region as non-shadowed. Ignoring the Beril algorithm due to bias, our algorithm tops the true-negative detection as well, i.e., detecting non-shadowed region correctly. Our result is 93.49% compared to 92.8%, 89.9%, and 89.7% of LULC, RGBN, and Gevers algorithms, respectively. Finally, our algorithm is also the best performer in false negative, i.e., detecting non-shadow as a shadow, yielding only 4.62% compared 4.69%, 5.68%, and 5.54% of LULC, RGBN, and Gevers algorithms, respectively. We conclude that LULC is marginally inferior to our algorithm while others are completely outperformed.

For the Modesto scene, a qualitative evaluation was performed, and it looks like our algorithm performed reasonably well in this dataset.

In the case of multi-label classification of the Selene scene, GT and our algorithm intermediate maps were divided into four contour levels. Firstly, we can demonstrate the ability of this proposed approach to estimate shadows without any manual input from the user. Moreover, as Figure 14 manifests, the accuracy of the proposed method increased on contour levels from non-shadow top layer with 53% to shadowed bottom layer with 88% accuracy.

We believe that this direction of RT-based methods for shadow-map detection is more faithful than other intensity and band index-based methods. Moreover, it can be conveniently extended to shadow compensation work.

Abbreviations

The following abbreviations are used in this manuscript:

AC	Atmospheric Compensation
BRDF	Bidirectional Reflection Distribution Function
DSM	Digital Surface Model
ELM	Empirical Line Method
GSD	Ground-Sampling Distance
GT	Ground Truth
HSI	Hyperspectral Image/Imaging
LIDAR	Light Detection and Ranging
LoS	Line of Sight
LULC	Land Use Land Cover
MSI	Multispectral Image/Imaging
NDVI	Normalized Difference Vegetation Index
NIR	Near Infrared
NRMSE	Normalized Root Mean Square Error
NS	Non-Shadow
QUAC	QUick Atmospheric Correction
RGB	Red Green Blue
RGBN	Red Green Blue NIR
RT	Radiative Transfer
S	Shadow
SWIR	Short-Wave Infrared

References

1. Arvidson, T.; Gasch, J.; Goward, S.N. Landsat 7's long-term acquisition plan—An innovative approach to building a global imagery archive. *Remote Sens. Environ.* **2001**, *78*, 13–26. [CrossRef]
2. Irish, R.R. *Landsat 7 Automatic Cloud Cover Assessment*; International Society for Optics and Photonics: Orlando, FL, USA, 2000; pp. 348–355. [CrossRef]
3. Zhu, Z.; Woodcock, C.E. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sens. Environ.* **2012**, *118*, 83–94. [CrossRef]
4. Zhang, H.; Sun, K.; Wenzhuo, L. Object-Oriented Shadow Detection and Removal From Urban High-Resolution Remote Sensing Images. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 6972–6982. [CrossRef]

5. Salvador, E.; Cavallaro, A.; Ebrahimi, T. Shadow identification and classification using invariant color models. In Proceedings of the 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing (Cat. No. 01CH37221), Salt Lake City, UT, USA, 7–11 May 2001; Volume 3, pp. 1545–1548. [CrossRef]

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

7. Imai, N.N.; Tommaselli, A.M.G.; Berveglieri, A.; Moriya, E.A.S. Shadow Detection in Hyperspectral Images Acquired by UAV. *ISPRS-Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2019**, *4213*, 371–377. [CrossRef]

8. 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. [CrossRef]

9. Hartzell, P.; Glennie, C.; Khan, S. Terrestrial Hyperspectral Image Shadow Restoration through Lidar Fusion. *Remote Sens.* **2017**, *9*, 421. [CrossRef]

10. Sirmacek, B.; Unsalan, C. Damaged building detection in aerial images using shadow Information. In Proceedings of the 2009 4th International Conference on Recent Advances in Space Technologies, Istanbul, Turkey, 11–13 June 2009; pp. 249–252. [CrossRef]

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

12. Teke, M.; Başeski, E.; Ok, A.Ö.; Yüksel, B.; Şenaras, Ç. Multi-spectral False Color Shadow Detection. In *Photogrammetric Image Analysis*; Stilla, U., Rottensteiner, F., Mayer, H., Jutzi, B., Butenuth, M., Eds.; Springer: Berlin, Heidelberg, 2011; pp. 109–119.

13. Sevim, H.D.; Çetin, Y.Y.; Başkurt, D.Ö. A novel method to detect shadows on multispectral images. *Proc. SPIE* **2016**, *10004*, 100040A. [CrossRef]

14. Sarabandi, P.; Yamazaki, F.; Matsuoka, M.; Kiremidjian, A. Shadow detection and radiometric restoration in satellite high resolution images. In Proceedings of the 2004 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2004), Anchorage, AK, USA, 20–24 September 2004; Volume 6, pp. 3744–3747. [CrossRef]

15. Smeulders, A.W.M.; Bagdanov, A.D. Color Feature Detection. In *Color in Computer Vision*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2012; Chapter 13, pp. 187–220. [CrossRef]

16. Bernstein, L.S.; Jin, X.; Gregor, B.; Adler-Golden, S.M. Quick atmospheric correction code: Algorithm description and recent upgrades. *Opt. Eng.* **2012**, *51*, 111719. [CrossRef]

17. Schott, J.R. *Fundamentals of Polarimetric Remote Sensing*; SPIE Press: Bellingham, WA, USA, 2009. [CrossRef]

18. Nicodemus, F.E. Directional Reflectance and Emissivity of an Opaque Surface. *Appl. Opt.* **1965**, *4*, 767–775. [CrossRef]

19. Stamnes, K.; Tsay, S.C.; Wiscombe, W.; Jayaweera, K. Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media. *Appl. Opt.* **1988**, *27*, 2502–2509. [CrossRef] [PubMed]

20. Tanre, D.; Herman, M.; Deschamps, P.Y. Influence of the background contribution upon space measurements of ground reflectance. *Appl. Opt.* **1981**, *20*, 3676–3684. [CrossRef] [PubMed]

21. Kaufman, Y.J. The atmospheric effect on the separability of field classes measured from satellites. *Remote Sens. Environ.* **1985**, *18*, 21–34. [CrossRef]

22. Richter, R.; Bachmann, M.; Dorigo, W.; Muller, A. Influence of the Adjacency Effect on Ground Reflectance Measurements. *IEEE Geosci. Remote Sens. Lett.* **2006**, *3*, 565–569. [CrossRef]

23. Willers, C.J. *Electro-Optical System Analysis and Design: A Radiometry Perspective*; SPIE Press: Bellingham, WA, USA, 2013. [CrossRef]

24. Matusik, W.; Pfister, H.; Brand, M.; McMillan, L. A Data-Driven Reflectance Model. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2003.

25. Matusik, W.; Pfister, H.; Brand, M.; McMillan, L. Mitsubishi Electric Research Laboratory BRDF Database, Version 2, 2006. Available online: https://www.merl.com/brdf/ (accessed on 5 August 2019).

26. Smith, B. Illustration of the phong reflection model. *Wikipedia, the Free Encyclopedia*. 2006. Available online: https://en.wikipedia.org/wiki/Phong_reflection_model (accessed on 19 July 2019).

27.   Bernard, J.; Bocher, E.; Petit, G.; Palominos, S. Sky View Factor Calculation in Urban Context: Computational Performance and Accuracy Analysis of Two Open and Free GIS Tools. *Climate* **2018**, *6*, 60. [CrossRef]

28.   Gong, S.; Newman, T.S. Dual Marching Squares: Description and analysis. In Proceedings of the 2016 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI), Santa Fe, NM, USA, 6–8 March 2016; pp. 53–56. [CrossRef]

29.   Piper, J.; Clarke, D.; Oxford, W. A new dataset for analysis of hyperspectral target detection performance. In Proceedings of the HSI 2014, Hyperspectral Imaging and Applications Conference, Coventry, UK, 15–16 October 2014.

30.   Sirmacek, B. Source code for Beril's Algorithm. The MathWorks, Inc. 2016. Available online: https://www.mathworks.com/matlabcentral/fileexchange/56263-shadow-detection (accessed on 5 August 2019).

31.   Teke, M. Source code for LULC algorithm. GitHub, Inc. 2014. Available online: https://github.com/mustafateke/FalseColorShadowDetection/blob/master/False_Color_Shadow_Detection_w_LULC.m (accessed on 5 August 2019).