DEM Densification Using Perspective Shape From Shading Through Multispectral Imagery

Zhe Chen, Qianqing Qin, Liyu Lin, Qiong Liu, and Wenfeng Zhan

Abstract—Although photogrammetry has long been used as the primary method to produce digital elevation models (DEMs), there are still some locations that are not covered by multimage imagery because of various limitations. Moreover, traditional interpolation methods for producing denser DEMs often cause oversmoothing, particularly over rough terrain. To formulate a robust procedure that is capable of reconstructing the terrain surface using sparse ground control points in heterogeneous areas, this letter describes an improved spatial enhancement method combining perspective shape from shading (SFS) using a single optical image. This method is called SFS-based densification using multispectral information (SDMI). First, the image irradiance equation based on the perspective Lambertian model is formulated as a static Hamilton–Jacobi equation and is solved using a fast sweeping strategy. We then reconstruct the relative surface shape and apply an edge-preserved iterative interpolation method to generate a higher resolution DEM grid. Multispectral information is used to reveal the actual surface reflection properties, and the land surface is classified into several types to better estimate surface reflectance. Experiments indicate that SDMI is effective for the interpolation of a sparse grid DEM over heterogeneous terrain.

Index Terms—Digital elevation model (DEM), interpolation, multispectral image, perspective model, shape from shading (SFS).

I. INTRODUCTION

DIGITAL elevation models (DEMs) are important data resources and are extensively used in topographical analysis, surveying and mapping, 3-D visualization, and other domains. Photogrammetry, the main method to produce DEMs, is based on establishing the correspondence between image pairs; however, this method often encounters difficulties in extracting the corresponding pairs in textureless areas. Given this drawback, conventional interpolation methods such as inverse distance weighted interpolation (IDW), Kriging, Shepard, and spline algorithm are utilized to enhance the spatial resolution of DEMs from a sparse grid. All these methods construct new data points within the range of a discrete set of known points. Whereas, these methods oversimplify the surface in cases where important points are lost in sampling; this oversimplification particularly occurs in mountainous areas. To find a robust procedure for terrain surface reconstruction, Rajabi and Blais introduced the shape from shading (SFS) method into DEM densification using satellite images [1], [2].

The SFS method recovers the surface shape from the gradual variations in the shading of the input images based on the “image irradiance equation,” which states that the measured brightness of the image is proportional to the radiance at the corresponding point on the surface. Pioneered by Horn and Brooks [3], this method uses an orthographic camera model based on the Lambert surface assumption. Surveys such as [4] and [5] provide thorough overviews of solving the SFS problem. Since [6] proposed casting the SFS as an optimal control problem, a new category of propagation approaches was proposed that was based on the theory of viscosity solutions to the Hamilton–Jacobi (HJ) equations [7]–[9]. To solve the HJ equation numerically, a fast marching strategy was first used in [10] and [11]. Most of these aforementioned approaches assume that the Hamiltonian is convex and homogeneous. Recently, an iteration strategy has been developed in which the Godunov Hamiltonian was combined with the Gauss–Seidel iteration to reconstruct surfaces [12]–[15]. In addition to synthetic data, SFS has been also used on real-world surfaces [16]–[18].

To generate a high-resolution DEM by applying SFS techniques on a single optical image with higher spatial resolution than the DEM, Rajabi [1], [2] investigated the utilization of SFS in the spatial enhancement of DEMs, and promising results were generated. However, in his studies, i) only a grayscale image is utilized; ii) both the reflectance and projection models are oversimplified; hence, the albedo is considered to be a constant; and iii) the surface normal is estimated by a minimization approach, which sometimes causes an oversmoothing of the output. To overcome these problems, we propose an improved SFS-based densification algorithm using a multispectral image (SDMI). i) The grayscale image is replaced by a multispectral image. ii) Pixels of different land cover are separated to estimate the reflectance of land cover. iii) We reconstruct the surface under the perspective Lambertian model using the Lax–Friedrichs sweeping method because the propagation approach is considered the best to solve the SFS problem [11]. iv) We apply an iterative interpolation to obtain the accurate denser DEM grid.

II. GENERATION OF DEM USING THE SFS TECHNIQUE

A. Algorithm Steps

The block diagram of SDMI is depicted in Fig. 1:

I) De-noise the multispectral image by a median filter with a 5 × 5 square window and coregistrate data sources.

II) Extract homogenous areas, estimate the associated albedo, and reconstruct the surface normal by SFS.
III) Interpolate the sparse DEM with an edge-preserved iterative technique using those reconstructed surface normals.

The details for each step are elaborated below.

B. Algorithm Details

Coregistration of Multispectral Image and DEM Grid: The registration method consists of three steps: i) The four corners (upper left, upper right, lower left, and lower right) on the DEM grids are considered as control points, and the corresponding points on the satellite image can be found by their geographic coordinates. ii) Affine transformation is applied. iii) The image is transformed by means of this function. Image values in noninteger coordinates are estimated using a bilinear technique.

Extracting Homogenous Areas From the Image: Different land covers such as water, rock, and road have different reflectance properties, and each should have a unique albedo. To simplify the albedo estimation, we assume that each land cover has a standard spectral reflectance; classification is made prior to the reflectance estimation. The term “albedo” herein is not the real radiation reflected from the surface but is a composite parameter that includes factors such as the strength of illumination and the reflectivity of the surface.

First, pixels that represent each land cover feature are selected to identify each class. According to visual interpretations, we selected about 100 training samples from each class. Supervised classification using the Mahalanobis distance (M-distance) decision rule is then performed. For each class, the intensity value of the cluster center is considered as the albedo, and a “normalized irradiance map” (the brightness differences due to the topographical changes in pixel intensity) is obtained as

\[ I_{\text{new}} = I_{\text{old}} / \rho \]  

where \( I_{\text{new}} \) and \( I_{\text{old}} \) are the normalized irradiance and the image intensity (pixel brightness), respectively, and \( \rho \) is the albedo for each class in a certain spectral band.

SFS Using Lax–Friedrichs-Based Sweeping: In the following, the SFS theory is utilized to recover the surface shape. The image irradiance equation of a Lambertian surface under a distant point light source is defined as

\[ I(u, v) = N \cdot L \]

where \( I(u, v) \) is the normalized irradiance at the image point \((u, v)\) (under image coordinate) corresponding to the surface point \((x, y)\) (under world coordinate system), which is equal to \( I_{\text{new}} \): \( L = (\alpha, \beta, \gamma) \) is the illumination direction; \( N = (-p, -q, 1) \) is the surface normal at \((u, v)\). By the Legendre transform, (2) can be equivalently formulated as an HJ equation, written as

\[
\begin{align*}
H(u, v, \nabla z) &= I \sqrt{p^2 + q^2 + 1} + p\alpha + q\beta - \gamma \\
R(u, v) &= 0
\end{align*}
\]

where \( H \) is a Hamiltonian, \( \nabla z = (p, q) \), \( (\alpha, \beta, \gamma) \) is the normalized illumination direction and satisfies \( \alpha^2 + \beta^2 + \gamma^2 = 1 \), and \( R(u, v) \) is the depth value at a surface boundary. With the formulation shown in (3), we can apply the Lax–Friedrichs method to approximate the numerical solution for this partial differential equation (PDE) on the basis of finite differences. This method is described as the FTCS (forward in time, centered in space) scheme with an artificial viscosity term of 1/2.

Assuming that the grid size is one unit, \( Z \) can be updated as

\[ Z_{n+1}^{u,v} = \frac{1}{\sigma_u + \sigma_v} \left( R(u, v) - H(u, v, \nabla z) + \sigma_u \phi_u + \sigma_v \phi_v \right) \]

\[ p = \frac{Z_{u+1,v} - Z_{u,v-1}}{2}, \quad q = \frac{Z_{u,v+1} - Z_{u,v-1}}{2} \]

\[ \phi_u = \frac{Z_{u+1,v} + Z_{u,v-1}}{2}, \quad \phi_v = \frac{Z_{u,v+1} + Z_{u,v-1}}{2} \]  

Let

\[ \sigma_u = \max_{u,v} \left\{ \max \{|I+\pi|, |I-\pi|\} \right\} \]

\[ \sigma_v = \max_{u,v} \left\{ \max \{|I+\bar{\pi}|, |I-\bar{\pi}|\} \right\} \]  

First, the surface is initialized with known points, and \( z(u, v) \) is then updated by sweeping through the DEM grid in four directions. Readers who are interested in more details of the process are referred to [14].

DEM Interpolation Using SFS Results: After acquiring the surface normals, the SFS result is used to improve the spatial resolution of the DEM grid. To preserve the terrain edges during densification, a multilevel iteration algorithm is applied to generate a denser grid, as shown in Fig. 2. The first level is a DEM with the lowest spatial resolution, and then, the DEM grid is interpolated at twice the resolution of the former level [see Fig. 2(a)]. The procedure for each level of the DEM is performed in two steps, as shown in Fig. 2(b) and (c) [20].
Fig. 2. Interpolation processes. (a) Iteration from sparse DEM to denser grid. (b) The first step in each iteration using four diagonal pixels. (c) The second step of interpolation with four nearest neighbors.

I) Interpolate the center points using the four diagonal pixels. As illustrated in Fig. 2(b), light points with known elevation data are utilized to interpolate the unknown center point, shown as the dark pixel. The slope is estimated from the five relative heights of the SFS result by

\[
\text{slope} = \frac{\arctan \sqrt{f_x^2 + f_y^2} \cdot 180}{\pi}
\]

where \(f_x\) and \(f_y\) are the gradient in the \(x\)- and \(y\)-directions. Formulating the reflectance model as a function of heights (z), we have five nonlinear equations corresponding to the image intensity of each point with four known heights (the grid points) and one unknown height. These nonlinear equations are solved by a least-squares method.

II) Interpolate the points using the four diagonal pixels. The remaining gaps [e.g., the black triangle pixel in Fig. 2(c)] are filled complying with the same rule as Step i), using the four nearest neighbors (in horizontal and vertical directions). A denser DEM grid with finer spatial resolution is generated by combining Steps i) and ii).

III) Repeat Steps i) and ii) until the DEM grid has the same resolution as the satellite imagery.

III. STUDY AREA

The study area is located in XinJiang Province (Lon. and Lat: 81.6733°–84.6738° E, 40.7470°–42.7402° N). There are a number of ridges and valleys in this area, and the surface is mainly composed of clay, mudstone, and sandstone. The data utilized in the study include a composite image from Landsat ETM+ and the corresponding DEMs (see Fig. 3). The attributes of the ETM+ image are shown in Table I.

Fig. 3(a) shows a Landsat ETM+ RGB composite image (the seventh, fourth, and second bands for the R, G, and B channels, respectively; spatial resolution: 10 m). Typical minerals have two absorption peaks in the range of 1–2.5 \(\mu m\), and only a few minerals have an absorption peak under 1 \(\mu m\), a peak which is generally weak. Therefore, the characteristics of the near-infrared (NIR) spectrum are treated as the basis of discrimination between different rock and mineral types, and the different types of rock boundary are clearly shown in the composite image of short wave infrared, NIR, and Blue-Green bands. Fig. 3(b) shows the DEM grids with a resolution of 2.5 m. This DEM is downsampled to the same resolution as the image and is considered to be the reference to verify the accuracy of different methods.

I) The DEM data are further downsampled to 1/10 resolution of the image, resulting in a low-resolution DEM grid. The resultant grid is shown in Fig. 4.

![Fig. 3. Study area. (a) Landsat ETM+ image. (b) DEM grid (unit: meters).](image-url)
II) Bilinear, IDW, Kriging, Rajabi’s algorithm, and SDMI are implemented to produce the denser DEM.

III) All the results are compared with the reference to test the performance of the methods.

IV. RESULTS AND DISCUSSION

A. Estimated Albedo

Brightness differences due to the land cover changes in pixel intensity were first removed. The surface-type classification map is shown in Fig. 5; the reflectance was estimated by extracting the clustering centers for each class. The normalized albedo for each class in different spectral bands is shown in Table II. On the basis of this result, the normalized irradiance map was obtained by (1), and this map was used in the following SFS solution.

<table>
<thead>
<tr>
<th>Land cover type</th>
<th>Band 2</th>
<th>Band 4</th>
<th>Band 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clay</td>
<td>0.793</td>
<td>0.840</td>
<td>0.696</td>
</tr>
<tr>
<td>Mudstone</td>
<td>0.571</td>
<td>0.548</td>
<td>0.618</td>
</tr>
<tr>
<td>Sandstone</td>
<td>0.283</td>
<td>0.314</td>
<td>0.213</td>
</tr>
</tbody>
</table>

B. Densified DEM

The sparse DEM is then densified by SDMI (see Fig. 6), which resembles Fig. 3(b) and has more terrain detail than the sparse grid (see Fig. 4). Band 4 is used for this calculation here.

C. Intercomparison Among Methods

For further comparison among methods, bilinear, IDW, Kriging, and Rajabi’s algorithm were also implemented to produce the denser grid.

As shown in Fig. 7, we compare the contour maps generated by these methods. Bilinear, IDW, and Kriging algorithms suffer from the loss of terrain details in the mountainous areas, while the topographic configuration is faithfully reconstructed using SDMI. Topographic information would be lost in resampling if there are only sparse ground control points (GCPs), and traditional interpolation approaches can only be achieved using curve fitting or regression analysis based on insufficient points, which causes erroneous grid values. Although Kriging performs better than Bilinear and IDW (by considering both the distance and the degree of variation among known data points when estimating unknown points), details are still missing, as shown in the middle of Fig. 7(c). Moreover, Kriging causes a blocky-looking appearance in the output, and many elaborate parameters need to be understood to use the method well.

We notice that Fig. 7(d) captures more terrain variations in the middle mountainous area compared with the first three methods. However, there are erratic points within the contours (on the left of this subfigure), and several other areas lose significant terrain information (in the middle-left). From the accuracy assessment shown in Table III, both the maximum and minimum errors of Rajabi’s method are very high, although the corresponding absolute mean error is lower than that of Kriging. This is probably due to the biases caused by the
TABLE III
ACCURACY OF DIFFERENT INTERPOLATION METHODS

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Error (meters)</th>
<th>Max error</th>
<th>Min error</th>
<th>AME</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bilinear</td>
<td>97.7</td>
<td>-67.3</td>
<td>12.5</td>
<td></td>
</tr>
<tr>
<td>IDW</td>
<td>128.6</td>
<td>-113.8</td>
<td>23.3</td>
<td></td>
</tr>
<tr>
<td>Kriging</td>
<td>59.5</td>
<td>-59.9</td>
<td>23.3</td>
<td></td>
</tr>
<tr>
<td>Rajabi’s</td>
<td>88.3</td>
<td>-69.1</td>
<td>12.1</td>
<td></td>
</tr>
<tr>
<td>SDMI</td>
<td>37.1</td>
<td>-41.2</td>
<td>6.2</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 8. AE map of different algorithms, the color bar shows a certain column of the image. (a) Bilinear. (b) IDW. (c) Kriging. (d) Rajabi’s method. (e) SDMI.

constant albedo utilized in SFS; hence, the irradiance variations among different land covers are misconstrued.

Fig. 8 shows the absolute error (AE) map for the aforementioned methods. We observe that IDW has the highest error at each location, whereas the result using Kriging is better than that of Rajabi’s method. Among all the algorithms, SDMI performs the best (also see Table III).

Although advances have been made, problems remain with the SDMI method. In addition to the errors introduced in each step of the SDMI method, the spatial resolution of DEMs also significantly affects the interpolation accuracy. To investigate how the densification accuracy behaves under different spatial resolutions, a series of DEM grid data is generated; each of these is downsamled to half the resolution of the former and is then interpolated back to the same resolution as the original data. We do not include the error analysis and other experiments herein as a result of limited paper length, but these experiments show that i) all methods achieve high accuracy when the DEM and imagery have similar resolutions, ii) the interpolation error increases as the resolution of the DEM grid becomes lower, and iii) all methods fail when the GCPs are reduced to a certain extent (1/20 resolution of the image in our experiments). In our tests, SDMI always performs the best.

V. CONCLUSION

This study proposes a method of DEM densification using a single multispectral satellite image. Results indicate that the DEM densification using SFS better captures the features of real terrain. By using surface reflectance values, it becomes possible to apply the SFS to real-world images, and terrain can thus be reconstructed more accurately. Compared with bilinear, IDW, Kriging, and Rajabi’s algorithm, SDMI has a higher accuracy.

It should be mentioned that in very complex or densely vegetated areas, additional data (e.g., LIDAR and spectral reflection) are required to help the DEM densification.

Several assumptions of the SFS technique, such as the Lambertian hypothesis, still require further examination. Additionally, extending the perspective SFS to satellite images incorporating non-Lambertian surfaces is a potential direction for further improvements.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable comments, which were quite useful for improving this letter.

REFERENCES