

# Refinement of Digital Elevation Models from Shadowing Cues

James Hogan and William A. P. Smith  
Department of Computer Science, University of York, UK  
james@albanarts.com, wsmith@cs.york.ac.uk

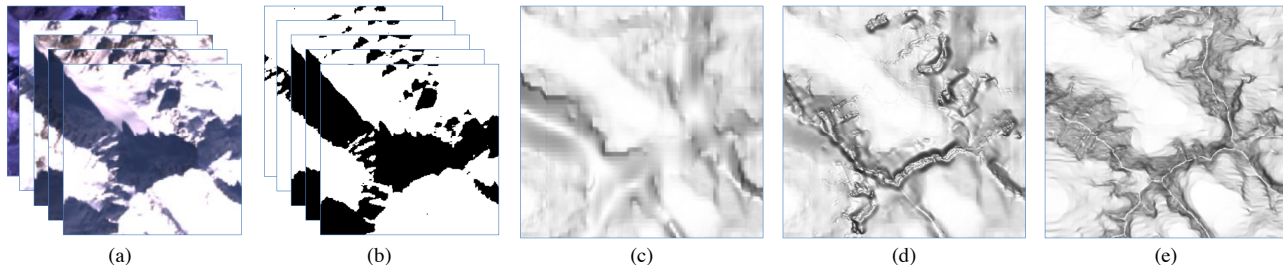


Figure 1: Overview of proposed refinement process: (a) 5 Landsat terrain images; (b) cast shadow maps estimated from (a); (c) interpolated SRTM data; (d) refinement of (c) using shadowing cues from (b); (e) ground truth.

## Abstract

*In this paper we derive formal constraints relating terrain elevation and observed cast shadows. We show how an optimisation framework can be used to refine surface estimates using shadowing constraints from one or more images. The method is particularly applicable to the digital elevation models produced by the Shuttle Radar Topography Mission (SRTM), which have an abundance of voids in mountainous areas where elevation data is missing. Cast shadow maps are detected automatically from multi-spectral satellite imagery using a simple heuristic which is reliable over varying types of surface cover. We show that the combination of our shadow segmentation and terrain correction methods can restore the structure of mountain ridges in interpolated SRTM voids using five satellite images, decreasing the RMS error by over 25%.*

## 1. Introduction

In 2000, NASA’s SRTM made available terrain elevation data of much of the mid latitudes of the earth [4]. Although this dataset represents the most complete high-resolution digital topographic database of Earth to date, the method of acquisition is affected by mountain and desert no-data areas, resulting in *voids* where elevation data is unavailable. Such void regions represent approximately 0.3% of the surveyed area [6] and typically occur in areas of very high relief. For example, voids occur in all summits over 8,000 metres, most summits over 7,000 metres, many Alpine and similar summits and ridges, and many gorges and canyons.

Completing these missing regions is a challenging problem because voids tend to occur in mountainous regions where the terrain contains much high frequency detail. Existing void-filling techniques have typically used interpolation, sometimes in conjunction with auxiliary Digital Elevation Maps (DEM) [12]. Such approaches suffer from two limitations: 1. Interpolation methods must make assumptions about the likely nature of the terrain within void regions, 2. the resolution of the auxiliary DEMs is often lower than the SRTM data, the result is that it only offers a very coarse guide to the terrain structure within voids.

In this paper, we show for the first time how to exploit additional information obtained from terrain imagery to intelligently guide the void-filling process. Namely, we exploit constraints suggested by shadowing, adjusting the terrain elevation data until it is consistent with the topography implied by the shadow regions. Cast shadows provide hard constraints about relative elevations, and shadow edges usually lie along the most interesting features of mountainous regions making them ideal cues for restoring those features. Our first contribution is to suggest a simple, yet robust and efficient method for segmenting multi-spectral satellite images into shadow/non-shadow regions on a per-pixel basis. Our second contribution poses correction of terrain elevation data as an optimisation problem which aims to satisfy shadowing constraints derived from one or more images. In contrast to existing shape-from-darkness techniques, our method requires only a small number of images (perhaps even only one). We present experimental results on SRTM data and show that our method can significantly improve the accuracy of interpolated void regions.

## 1.1. Related Work

Truly shadowed regions would reflect no light and would be easily identifiable as pixels with zero brightness. However, because of inter-reflections and ambient illumination, shadow regions in real images typically have non-zero brightness and are difficult to distinguish from objects with low albedo or which are darkly shaded. Lu and Drew [10] used illuminant invariance theory to define an illuminant discontinuity measure such that nearby pixels on opposite sides of a shadow boundary will have a high discontinuity. They used Markov random fields and graph cut optimisation to find the shadow regions. In multispectral images, Ranson and Daugtry [11] showed that as shadows increased, red and near-infrared reflectance decreased but the normalised difference between the two increased.

Extracting object shape from shadow maps of the same scene under varying illumination is known as shape-from-darkness. Kender and Smith [8] were the first to consider this problem and presented a solution for the one-dimensional case. Daum and Dudek [2] extended this work to enable reconstruction of a 3D surface from 64 shadow maps. They used constraints on the upper and lower bounds of pixel heights, applying them iteratively and in an order chosen carefully to maximise convergence and the propagation of information through the image. Such approaches require a large number of images since only shadow edges convey useful information. In contrast, traditional photometric stereo requires only 3 or 4 images and uses the information at every pixel. Chandraker *et al.* [1] combined shape-from-darkness with photometric stereo by using shadowing constraints to constrain surface integration.

Voids exist in the SRTM data where no reliable elevation measurements are available. These can be caused by steep terrain that is facing away from the radar (shadowing), or facing toward the radar (foreshortening or layover) [4]. It can also be caused by surfaces which don't reflect enough energy back towards the sensor to make an image, such as water which is highly specular or sand which absorbs radar signals. We are primarily concerned with void-filling in mountainous terrain where generic interpolation techniques such as bilinear interpolation produce extremely poor results on anything but the smallest voids. Therefore, state of the art approaches have combined interpolation with additional sources of information. Grohman *et al.* [5] use a delta surface fill method which replaces voids with values from fill data adjusted to match the SRTM boundary values. The results more closely match the original SRTM data while maintaining the shape of the fill data. Link *et al.* [9] improved interpolation of SRTM voids using Landsat TM/ETM+ imagery and GTOPO30 data. They manually labelled valleys in the image data. The segments of valley that were void were then linearly interpolated, before interpolating the rest of the void based on the GTOPO30 data.

## 2. Preliminaries

We begin by providing formal definitions of the effects of shadowing and show how these lead to constraints on terrain geometry.

### 2.1. Pointwise and Transition Shadow Constraints

Shadows can be divided into two classes: *self shadows* (where the surface faces away from the light source) and *cast shadows* (where the surface faces towards the light source but where another part of the surface blocks the light from reaching it). Self shadows are simpler to model, being dependent only the local surface normal direction. Cast shadows on the other hand depend on the global geometry of the surface. The two are related in the sense that every cast shadow has a corresponding self shadow on the surface that is obscuring the light.

We can state formally a number of shadowing rules which apply to a single point on a surface and transitions between adjacent points. We do so by defining various properties which pixels may exhibit and then stating the formal relation between points exhibiting combinations of such properties. For each property  $Y$ , the set  $Y_D$  contains all pixels which exhibited the property  $Y$  in image dataset  $D$ .

A pixel  $p$  exhibits the property  $l$  if  $p$  is lit, or more specifically if the path from  $p$  to the light is unobstructed. The pixels exhibiting this property can be recorded in a shadow-map obtained by shadow segmentation or by ray-tracing using the reference DEM and light direction. Stated formally, the lit property is as follows:

$$p \in l_D \Leftrightarrow \forall n \in \mathbb{R}^+ \left( alt_{H^*} \left( \vec{X}_p + n \cdot \vec{L}_D \right) > 0 \right), \quad (1)$$

where  $alt_h \left( \vec{V} \right)$  is the altitude of the position  $\vec{V}$  directly above the terrain  $h$ ,  $\vec{X}_p$  is the position of  $p$  and  $\vec{L}_D$  is the direction of the light source in  $D$ .

A related property  $f$  is exhibited by a pixel  $p$  if it is facing the light source. This can be determined by the dot product of the light direction vector  $\vec{L}_D$  with the surface normal vector  $\vec{N}_p$  of the terrain at  $p$ :

$$p \in f_D \Leftrightarrow \vec{L}_D \cdot \vec{N}_p > 0 \quad (2)$$

These two properties are not the same. Any pixel  $p \notin f_D$  will obscure light from reaching another pixel  $q \in f_D$ . In this case both pixels are unlit ( $p \notin l_D$  and  $q \notin l_D$ ), even though  $q$  does face the light source, i.e.  $q \in f_D$  (see Figure 2).

The fact that  $p \in f_D$  can be determined directly from the terrain, and  $p \in l_D$  can be determined directly from a shadow-map motivates finding constraints which relate the two properties. All pixels must be facing the light in order to receive any light, so it is clear that  $l_D \subset f_D$ :

$$p \in l_D \Rightarrow p \in f_D. \quad (3)$$

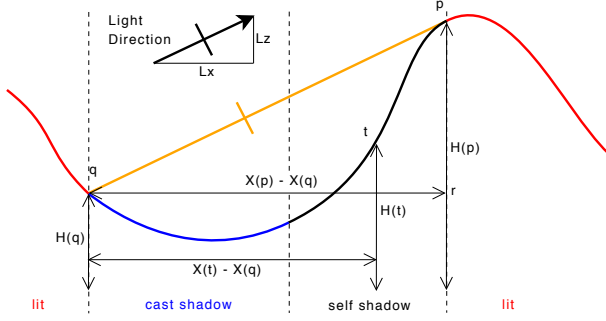


Figure 2: Lighting transitions and shadow geometry.

A pixel  $p$  that is not lit is cast shadowed ( $p \in c_D$ ) if it faces the light source:

$$p \in c_D \Leftrightarrow p \notin l_D \wedge p \in f_D, \quad (4)$$

and is self shadowed ( $p \in s_D$ ) if it faces away from the light source:

$$p \in s_D \Leftrightarrow p \notin f_D. \quad (5)$$

The valid property combinations for a pixel are shown in Table 1.

We can now state a number of constraints regarding the transition between combinations of pixel properties as we move away from the light source over the terrain. If  $q$  is the pixel adjacent to  $p$  and further away from the light source, the possible transitions are as follows:

- l→s Shadow entrance at the highest part of the shadow line.
- l→c It is impossible to enter a cast shadow without first passing through a self shadow.
- s→l It is impossible to leave a self shadow into sunlight without first passing through cast shadow.
- s→c Passing from self shadow to cast shadow is necessary.
- c→l Shadow exit at the lowest part of the shadow line.
- c→s Passing from cast shadow to self shadow may occur if a smaller hill is in the shadow of a larger hill.

Table 2 summarises which transitions are valid. We assume that the terrain is a height map with no downward facing surface normals. This implies there are no caves or arches, which would make the l → c transition valid. We also assume that there is sufficient resolution that there is no more than one transition between adjacent pixels. If this assumption were broken then both invalid transitions could occur between the pixels.

	Facing ( $p \in f_D$ )	Non-facing ( $p \notin f_D$ )
<b>Lit</b> ( $p \in l_D$ )	Lit ( $p \in l_D$ )	Invalid
<b>Unlit</b> ( $p \notin l_D$ )	Cast shadow ( $p \in c_D$ )	Self shadow ( $p \in s_D$ )

Table 1: Valid pixel property combinations

	$q \in l_D$	$q \in s_D$	$q \in c_D$
$p \in l_D$	✓	✓	✗
$p \in s_D$	✗	✓	✓
$p \in c_D$	✓	✓	✓

Table 2: Valid pixel property transitions

## 2.2. Geometric Shadow Constraints

We can use the formal shadow constraints given above to derive geometric constraints relating terrain elevation to position in the image plane (with reference to Figure 2). The most important geometric constraint is that of the vertical separation of the ends of the shadow. The shadow entrance  $p$  and the shadow exit  $q$  are on opposite boundaries of a shadow so that every pixel between them is in shadow. The straight line between  $p$  and  $q$  shall be referred to as the *shadow ceiling*. It is the boundary between light and shadow therefore it is inclined at the same gradient as the light direction:

$$\frac{H(p) - H(q)}{X(p) - X(q)} = \frac{L_z}{L_x}, \quad (6)$$

where  $H(x)$  is the elevation of the terrain at pixel  $x$ ,  $X(x)$  is the horizontal distance in the light direction of pixel  $x$  and  $L_x, L_z$  are horizontal and vertical components of the light direction vector.

Any pixel  $t$  between  $p$  and  $q$  is in shadow by the definition of  $p$  and  $q$ . Therefore  $t$  is restricted to lie below the shadow ceiling or it would cease to be in shadow. In other words the gradient of the line from  $q$  to  $t$  cannot exceed the gradient of the shadow ceiling. This is the second geometric shadow constraint:

$$\frac{H(t) - H(q)}{X(t) - X(q)} < \frac{L_z}{L_x}. \quad (7)$$

It is also the case that the tangent of the terrain at  $p$  is parallel to the light direction which can be used as another constraint:

$$H'(p, \vec{L}_{xy}) = \frac{L_z}{L_x}, \quad (8)$$

where  $H'(p, \vec{V})$  is the gradient of the terrain  $H$  at  $p$  in the direction  $\vec{V}$ .

Finally, at  $p$  the terrain must be downward sloping so as to form a peak instead of a valley. This is handled by the shadow ceiling limit and the constraint that a lit pixel must be facing the light.

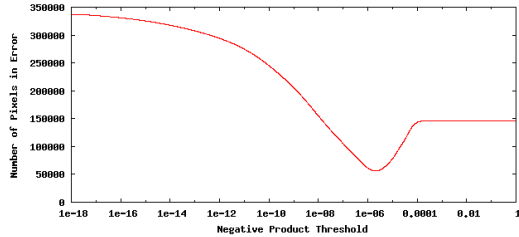


Figure 3: Optimising the negative product threshold.

### 3. Shadow Segmentation

State-of-the-art shadow segmentation techniques are based on an illuminant discontinuity measure [10]. However, to use such techniques the illuminant colour must be known. For tristimulus images, in which log-chromaticity space is two dimensional, this is straightforward, requiring only the estimation of a single angular value. However for multispectral images, obtaining a stable estimate is less well constrained. Further, such approaches do not deal well with pixels that are saturated. Many images in the SRTM dataset contain large numbers of saturated pixels, particularly in alpine, snow covered regions, and hence in practice such algorithms do not perform well. For this reason, we develop a very simple approach for shadow classification which is both efficient and sufficiently robust that the results can be used for shape-from-darkness.

Our shadow classification is based on the following simple heuristic. We assume that all materials visible in the satellite imagery reflect light (i.e. have a high albedo) in at least one of the frequency bands measured by the camera. Where our assumption holds, pixels which are dark in all bands can only be so because they are receiving much less light and are hence in shadow. On the other hand, pixels which are bright in one or more bands must be receiving light and are considered lit. An object with low albedo over all measured wavelengths would appear dark in every band and may be incorrectly labelled as shadowed by our method.

Our classifier takes as input a set of pixels  $p \in P$ , each with intensities  $p_1 \dots p_k$  corresponding to the  $k$  frequency bands, and the location  $X_p$  of each pixel. The result of the classification  $s(p) \in \{S, L\}$  indicates whether the pixel is lit. A straightforward implementation of our heuristic is obtained by the product of image negative values over all frequency bands (which we term the *negative product*):

$$f(p) = \prod_{i=1}^k (1 - p_i)^{\eta_i}. \quad (9)$$

A pixel  $p$  which is bright in any one band (and therefore has a low image negative value in that band) will result in a low value, i.e.  $f(p) \approx 0$ . On the other hand, a pixel which is

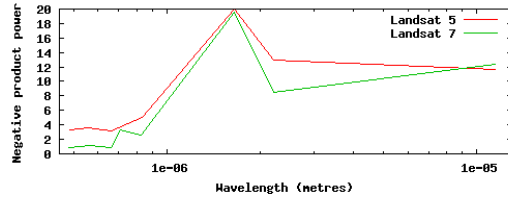


Figure 4: Negative product powers plotted against the mean wavelength sensitivity of the corresponding bands.

dark in all bands (and therefore has image negative values of approximately 1 in all bands) will result in a high value, i.e.  $f(p) \approx 1$ . It has been shown [11] that low image negative values in certain bands correlate more strongly with shadowing than others. This is particularly true of the infrared bands. We can exploit this idea by including certain bands more times than others or, more generally, raising each element of the product to a value greater than 1. These  $\eta_i$  values form a vector of parameters which, when well chosen, can vastly improve the shadow classification accuracy.

In order to classify pixels, we threshold the negative product:

$$s(p, t) = \begin{cases} L & \text{if } f(p) < t \\ S & \text{otherwise} \end{cases}, \quad (10)$$

where  $t$  is chosen to optimise the classification performance. This optimisation is achieved by computing the classification error between estimated and ground truth shadow maps over a set of representative data. An example is shown in Figure 3.

#### 3.1. Optimising Negative Product Parameters

The shadow classification error can be reduced by optimising the parameter  $\eta_i$  for each of the frequency bands. Moreover, these parameters turn out to generalise well to novel data and hence need computing only once. The error associated with a vector of negative product parameters,  $\eta = [\eta_1 \eta_2 \dots \eta_k]^T$ , is the classification error with optimal choice of threshold. We use gradient descent to optimise the parameters, using a finite difference approximation of the gradient of  $\eta$ . For training data we use a large area of terrain but sub-sample in order to cover a range of terrain types without prohibitive computation requirements.

Figure 4 shows two typical distributions of negative product parameter values plotted as a function of the wavelengths to which the bands are sensitive. The parameters were obtained by optimising for a Landsat 5 and Landsat 7 dataset. The important observation here is that different datasets result in a largely similar distribution of optimal parameter values. Also note that the infrared bands are given higher weight than those in the visible range, suggesting

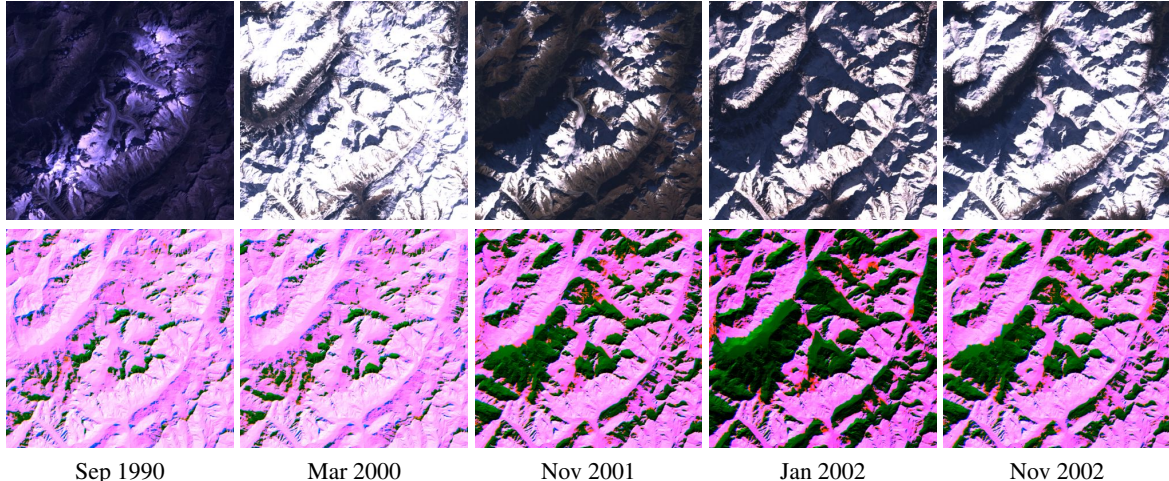


Figure 5: Shadow segmentation results. Top row: RGB channels of input multispectral images, bottom row: segmentation results (see text).

that brightness in an infrared image is more highly correlated with being lit.

### 3.2. Shadow Segmentation Evaluation

We evaluate our multispectral shadow segmentation algorithm on a large area of terrain in the Alps (the geographical rectangle with corners  $46^{\circ}3'N$   $6^{\circ}45'E$  to  $45^{\circ}46'N$   $7^{\circ}8'E$ ). It is an area of very rugged terrain with many cast shadows. The first row of Figure 5 shows the RGB bands for 5 sets of terrain imagery (the first image is from the Landsat 5 satellite, the remainder are from Landsat 7). The second row shows the segmentation results. These have the reference shadow map in the red channel, Lambertian shaded terrain in the green channel, and our estimated shadow segmentation map in the blue channel. The reference shadow map is calculated by ray-tracing using the reference DEM. Shadows which are correctly classified appear green, while false positives (shadow in segmentation, not in reference) appear blue and false negatives (shadow in reference, not in segmentation) appear orange. Corresponding quantitative results are shown in Table 3.

Our simple heuristic leads to a segmentation that is of a good quality (less than 10% error) and matches the shape of the shadows well. Errors that do occur either lie along the shadow boundary (caused by shadow penumbra) or where the incident light is close to grazing (where slight errors in the reference DEM cause large changes in the shadow map and the classification error is hence unreliable).

### 4. Elevation Correction

Our aim is to refine an initial terrain  $H_0$  (a set of pixel elevations) to obtain an improved terrain  $H$  which approximates as closely as possible the true terrain  $H^*$ .  $H$  must

Date	Shadow	Error	Shadows unclassified	Non-shadows classified
Sep 1990	10.84%	7.04%	48.11%	2.05%
Mar 2000	11.83%	5.92%	35.17%	1.99%
Nov 2001	33.96%	7.42%	9.24%	6.49%
Jan 2002	52.82%	7.76%	5.83%	9.91%
Nov 2002	35.76%	8.53%	11.12%	7.11%

Table 3: Shadow Segmentation Results.

satisfy the constraints imposed by which pixels are lit. We assume that the illumination direction vectors are known. This is a reasonable assumption in terrain imagery where the direction of the sun can be calculated from the image timestamp and location. We denote the terrain elevation at a pixel  $p$  according to terrain map  $H$  as  $H(p)$ .

Our approach is to exploit a number constraints inferred from the image data in an optimisation framework. For each of  $n$  images, our algorithm takes as input: shadow-maps  $S = \{S_1 \dots S_n\}$  and light direction vectors  $L = \{\vec{L}_1 \dots \vec{L}_n\}$ . We also provide per-pixel confidence values  $R(p)$  which quantify measurement uncertainty. For the SRTM data these are simply binary values indicating whether the terrain is void or non-void. We construct an objective function which seeks to satisfy shadowing and regularisation constraints. The error induced by the  $i$ th constraint at pixel  $p$  is denoted  $\varepsilon_i(p, H, R, S, L)$ . The objective function is a weighted combination of these costs:

$$\varepsilon(H_i, R, S, L) = \sum_{p \in P} \sum_i w_i \cdot \varepsilon_i(p, H_i, R, S, L). \quad (11)$$

We define some cost functions on a per image basis, denoted

$\varepsilon_i^{\text{im}}(p, H, R, S, \vec{L})$ . In this case, the total cost is defined as the sum over all images:

$$\varepsilon_i(p, H, R, S, L) = \sum_j \varepsilon_i^{\text{im}}(p, H, R, S_j, \vec{L}_j). \quad (12)$$

We obtain the terrain estimate at iteration  $i$ ,  $H_i$ , from the previous estimate  $H_{i-1}$  using gradient descent. We approximate the gradient of the objective function using finite differences computed by perturbing each terrain vertex. For reasons of efficiency and numerical stability, these perturbations are applied to a region around each vertex, computed using a radial function with linear fall-off.

In the remaining subsections we describe how each cost function is derived from the shadowing constraints described in Section 2.

#### 4.1. Lit pixels face light

This function penalises pixels which are lit but are not facing the light source. Computing the angle between the surface normal and light source direction for each image is costly as the surface normals will need recomputing every time the terrain elevation is adjusted. To avoid this, we compute the error directly in terms of the surface gradients:

$$\text{facing}(p, H, \vec{L}) = \frac{L_z}{\|\vec{L}_{xy}\|} - \frac{H(\vec{X}_p) - H(\vec{X}_p + \vec{L}_{xy})}{2\|\vec{L}_{xy}\|}. \quad (13)$$

This calculation becomes unstable when the light source is close to directly overhead. However, since such a lighting configuration will result in no shadowing we assume that such an image would not be used. We define the error caused by violations of this constraint as:

$$\varepsilon_1^{\text{im}}(p, H, R, S, \vec{L}) = \begin{cases} 0 & \text{if } S(p) \neq \text{lit} \vee \text{facing}(p, H, \vec{L}) \geq 0 \\ \text{facing}(p, H, \vec{L})^2 & \text{otherwise} \end{cases} \quad (14)$$

#### 4.2. Shadow entrance and exit elevation

The first cost related to the shadowing constraints ensures the correct vertical separation between corresponding shadow transition pixels. We precompute a number of quantities for each pixel in the shadow-map:

- $\text{ent}(S, p)$  - the last pixel in shadow at the entrance of the current shadow in the light direction.
- $\text{exit}(S, p)$  - the last pixel in shadow at the exit of the current shadow in the negative light direction.
- $\text{entDist}(S, p)$  - the horizontal distance between  $p$  and  $\text{ent}(S, p)$ .
- $\text{exitDist}(S, p)$  - the horizontal distance between  $p$  and  $\text{exit}(S, p)$ .

To compute these values, we trace from each shadowed pixel toward and away from the light source until a lit pixel is reached. If the shadow-map edge is reached and it is also in shadow then the position of the transition cannot be known. The computational complexity of this calculation is  $O(whd)$  where  $w$ ,  $h$  and  $d$  are respectively the width and height of the shadow-map, and the maximum shadow depth. It is possible to optimise this process by caching partial results from each pixel but in practice the computation time is not significant.

If  $\text{ent}(S, p) = p$  then  $p$  is a shadow entrance pixel. The error caused by such a pixel deviating from its calculated elevation is given by:

$$\varepsilon_2^{\text{im}}(p, H, R, S, \vec{L})^{\frac{1}{2}} = H(\text{exit}(S, p)) + \text{exitDist}(S, p) \frac{L_z}{\|\vec{L}_{xy}\|} - H(p). \quad (15)$$

Similarly if  $\text{exit}(S, p) = p$  then  $p$  is a shadow exit pixel and the following cost is applicable:

$$\varepsilon_3^{\text{im}}(p, H, R, S, \vec{L})^{\frac{1}{2}} = H(\text{ent}(S, p)) - \text{entDist}(S, p) \frac{L_z}{\|\vec{L}_{xy}\|} - H(p). \quad (16)$$

#### 4.3. Shadow ceiling limit

This cost function uses data from the shadow entrance and exit elevation costs. The elevation of the shadow ceiling at a pixel is found by linearly interpolating the elevations of the transitions, weighting using the distances to each:

$$\text{shadCeil}(p, H, S) = \frac{H(\text{exit}(S, p)) \cdot \text{entDist}(S, p) + H(\text{ent}(S, p)) \cdot \text{exitDist}(S, p)}{\text{entDist}(S, p) + \text{exitDist}(S, p)} \quad (17)$$

This can then be used to construct a cost function which penalises pixels which lie above the shadow ceiling:

$$\varepsilon_4^{\text{im}}(p, H, R, S, \vec{L}) = \max(0, H(p) - \text{shadCeil}(p, H, S))^2. \quad (18)$$

#### 4.4. Shadow entrance slope

The tangent of the terrain at the shadow entrance should be parallel to the light source. As for the lit pixel computation, this can be calculated efficiently in terms of surface gradients. If  $\text{ent}(S, p) = p$  then  $p$  is a shadow entrance pixel and the following cost is applicable:

$$\varepsilon_5^{\text{im}}(p, H, R, S, \vec{L}) = \text{facing}(p, H, \vec{L})^2. \quad (19)$$

#### 4.5. Terrain prior

We use a confidence measure to encourage pixels to stay at the same elevation if that elevation is reliable. This is particularly important when used in combination with shadow

exit elevation constraints. When the shadow exit elevation is reliable the terrain prior prevents it from dropping which forces the shadow entrance elevation to increase towards its true value. We define the terrain prior as proportional to the confidence in the elevation and the square of the elevation change:

$$\varepsilon_6(p, H, R, S, L) = R(p) \cdot (H(p) - H_0(p))^2. \quad (20)$$

For our dataset we use binary reliability values indicating whether the terrain is void or non-void. However, this constraint is flexible and could make use of a continuous reliability value, for example related to measurement uncertainty.

#### 4.6. Smoothness

To encourage a smooth solution and to ensure that elevation changes are diffused throughout the terrain, we impose a smoothness error term:

$$\varepsilon_7(p, H, R, S, L) = (\partial_{xx}H(p))^2 + (\partial_{yy}H(p))^2. \quad (21)$$

Derivatives are approximated numerically.

#### 4.7. Shadow entrance curvature

We also impose higher order constraints on the shadow entrance. Moving along the surface in the direction of the light source, the terrain should be convex at the shadow entrance, i.e. a hill rather than a valley. Non-convex shadow entrances are penalised as follows:

$$\varepsilon_8^{\text{im}}(p, H, R, S, \vec{L}) = \max \left( 0, \frac{H(\vec{X}_p - \vec{L}_{xy}) + H(\vec{X}_p + \vec{L}_{xy})}{2} - H(p) \right)^2, \quad (22)$$

where  $\vec{X}_p$  is the location of pixel  $p$ .

### 5. Experimental Evaluation

For terrain imagery we use satellite images from the Landsat 5 and Landsat 7 programmes. Although the satellites are in sun-synchronous orbits, Landsat 7 sun elevations in the Alps vary between approximately  $18^\circ$  and  $62^\circ$ , and sun azimuths vary between approximately  $134^\circ$  and  $163^\circ$ . For our initial terrain elevation estimate, we use data from the CGIAR-CSI-SRTM [7] dataset which is currently the best regarded global void-filled SRTM dataset and uses interpolation methods and auxiliary DEMs to fill void regions. For our reference dataset, we use de Ferranti’s ViewFinder Panorama DEMs [3]. For our region of interest, these DEMs are produced from detailed topographic maps with considerably higher resolution and accuracy than SRTM.

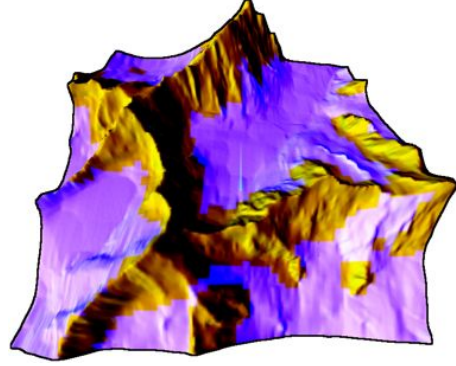


Figure 6: Test region reference DEM. Yellow is SRTM void. Lilac is non-void. North is right. French/Swiss/Italian border is the summit of Mont Dolent towards the lower left.

For our experimental evaluation, we use a region of rugged alpine terrain at the head of the Argentiere glacier (see Figure 6). It is an area where CGIAR-CSI-SRTM has major discrepancies with the reference DEM. 43.76% of the SRTMv2 data in this area are voids.

We compute shadow maps from the input images in the top row of Figure 7 using the method described in Section 3. The weights associated with each constraint are chosen based on observation as follows:  $w_{1..8} = \langle 255, 10, 1, 10, 1, 10, 2.5, 1000 \rangle$ . Our algorithm refines the initial terrain shown in the bottom left panel of Figure 7 to obtain the terrain shown in the bottom middle panel. The reference terrain is shown bottom right. It is clear visually that much of the fine detail and ridge structure has been restored to the initial overly smooth terrain. Our algorithm converges in approximately 500 iterations reducing the RMSE from 115.79m to 86.82m. The spatial distribution of errors is shown in Figure 8. It is clear that much of the improvement is concentrated along the ridge lines and summits (see Figure 9).

### 6. Conclusions

We have presented a method for refining terrain elevation estimates using shadow maps estimated from multispectral satellite imagery. Our approach is flexible and we intend in future work to explore incorporating shading constraints to improve the recovery of shape outside of shadow edges. Using only 5 images with a limited range of sun directions, our algorithm is able to substantially improve the terrain elevation estimates, adding fine detail and restoring ridge structure. The complete source code of our method is available under the terms of the GNU General Public License Version 2 from: <http://repo.or.cz/w/tecorrec.git>.

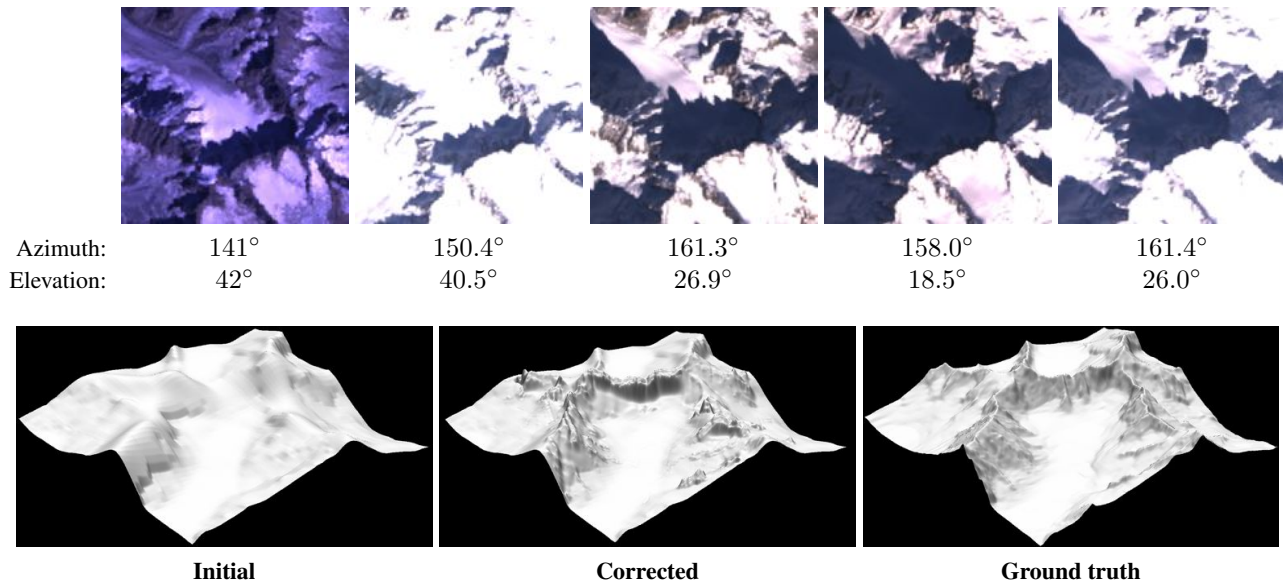


Figure 7: Top row: Input images and sun directions. Bottom row: initial, refined and reference terrains.

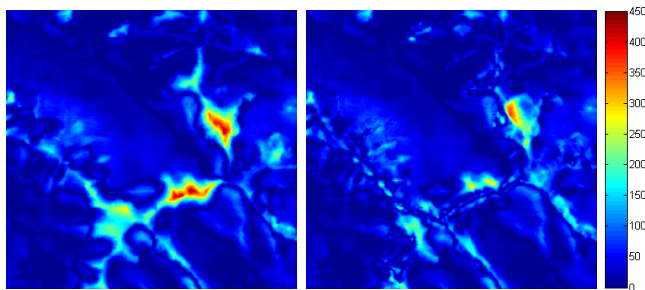


Figure 8: Absolute elevation error (m) of initial SRTM terrain (left) and corrected terrain (right).

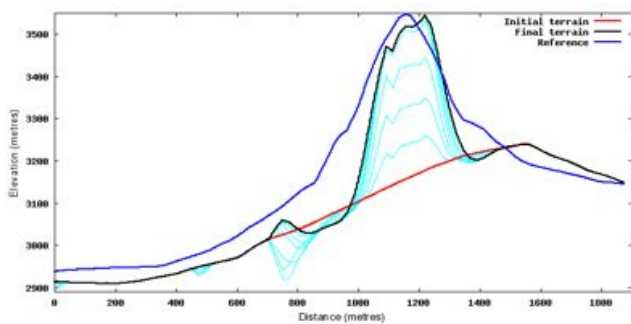


Figure 9: Ridge cross section showing initial, intermediate, converged and reference profiles.

## References

- [1] M. Chandraker, S. Agarwal, and D. Kriegman. Shadowcuts: Photometric stereo with shadows. In *Proc. CVPR*, 2007.
- [2] M. Daum and G. Dudek. On 3-D surface reconstruction using shape from shadows. In *Proc. CVPR*, pages 461–468, 1998.
- [3] J. de Ferranti. Digital elevation data, 2008. <http://www.viewfinderpanoramas.org/dem3.html>.
- [4] T. G. Farr, P. A. Rosen, E. Caro, R. Crippen, R. Duren, S. Hensley, M. Kobrick, M. Paller, E. Rodriguez, L. Roth, D. Seal, S. Shaffer, J. Shimada, J. Umland, M. Werner, M. Oskin, D. Burbank, and D. Alsdorf. The shuttle radar topography mission. *Rev. Geophys.*, 45:RG2004, 2007.
- [5] G. Grohman, G. Kroenung, and J. Strebeck. Filling srtm voids: the delta surface fill method. *Photogrammetric Engineering and Remote Sensing*, 72(1):213–216, 2006.
- [6] O. Hall, G. Falorni, and R. L. Bras. Characterization and quantification of data voids in the shuttle radar topography mission data. *IEEE Geosci. Remote Sens. Lett.*, 2(2):177–181, 2005.
- [7] A. Jarvis, H. Reuter, A. Nelson, and E. Guevara. Hole-filled SRTM for the globe version 4, 2008. <http://srtm.csi.cgiar.org>.
- [8] J. R. Kender and E. Smith. Shape from darkness: Deriving surface information from dynamic shadows. In *Proc. AAAI*, pages 664–669, 1986.
- [9] F. Ling, Q. W. Zhang, and C. Wang. Filling voids of srtm with landsat sensor imagery in rugged terrain. *Int. J. Remote Sens.*, 28(2):465–471, 2007.
- [10] C. Lu and M. Drew. Shadow segmentation and shadow-free chromaticity via markov random fields. In *Proc. Color Imaging Conference*, 2005.
- [11] K. Ranson and C. Daughtry. Scene shadow effects on multispectral response. *IEEE Trans. Geosci. and Remote Sens.*, 25(4):502–509, 1987.
- [12] H. I. Reuter, A. Nelson, and A. Jarvis. An evaluation of void-filling interpolation methods for SRTM data. *Int. J. Geogr. Inf. Sci.*, 21(9):983–1008, 2007.