Error Detection and DEM Fusion Using Self-Consistency*

Howard Schultz, Edward M. Riseman, Frank R. Stolle
Computer Science Department
University of Massachusetts
Amherst, MA 01003
USA
{hschultz,riseman,stolle}@cs.umass.edu

Dong-Min Woo
School of Electrical Engineering
Myongji University
South Korea
dmwoo@wh.myongji.ac.kr

Abstract

The ability to efficiently and robustly recover accurate 3D terrain models from sets of stereoscopic images is important to many civilian and military applications. Our long-term goal is to develop an automatic, multi-image 3D reconstruction algorithm that can be applied to these domains. To develop an effective and practical terrain modeling system, methods must be found for detecting unreliable elevations in digital elevation maps (DEMs), and for fusing several DEMs from multiple sources into an accurate and reliable result.

This paper focuses on two key factors for generating robust 3D terrain models, (1) the ability to detect unreliable elevations estimates, and (2) to fuse the reliable elevations into a single optimal terrain model. The techniques discussed in this paper are based on the concept of using self-consistency to identify potentially unreliable points. We apply the self-consistency methodology to both the two-image and multi-image scenarios. We demonstrate that the recently developed concept of self-consistency can be effectively employed to determine the reliability of values in a DEM. Estimates with a reliability below an error threshold can be excluded from further processing. We test the effectiveness of the methodology, as well as the relationship between error rate and scene geometry by processing both real and photo-realistic simulations.

1 Introduction

There is a substantial body of research in two-image stereo reconstruction [3][8], and a smaller but growing body of work on multi-view stereo[1][4]. These methods have been loosely divided into two categories, feature matching and texture matching. Independent of these classifications 3D reconstruction algorithms can be divided into image space methods[2][8] in which the matching occurs without regard to the physical characteristics of the surface; and object-space matching in which the images must be consistent with the geometry and physical properties of the surface[1][4]. Each of these techniques share a common goal of identifying elements (features [5][6][11], patterns [8][12]) across two or more views of a scene. Unfortunately, there are a variety of mechanisms that will cause any of these techniques to fail locally – a flawed imaging mechanism (e.g., dropouts, dust, thermal noise), poorly modeled optical properties (e.g., using a Lambertian reflectance model for asphalt or water), occlusions (across varying image viewpoints), and/or problematic texture patterns (e.g., repetitive patterns, linear features that align with epipolar lines).

Because it is impossible to take into account all of the problems, a robust 3D reconstruction system should

contain a mechanism for identifying and removing blunders. The techniques discussed in this paper are based on the concept of using self-consistency measures, first introduced by Leclerc, Luong and Fua [10], to identify unreliable points in a distribution. The main focus of their work was to obtain a quality measure for correspondence algorithms without relying on any ground truth. Their algorithm obtained a probability distribution by counting the number of corresponding image points for each object point that is consistent with the viewing geometry within a specified error limit. In a closely related application [9], they extend their work to detect changes in terrain by applying the concept of self-consistency to elevations. We extend the idea of this work to detect unreliable elements in a digital elevation map (DEM) generated from stereoscopic image pairs, and to fuse multiple DEMs. In our case, multiple pairs of images of the same site are processed using a hierarchical texture matching system.

In our application domain of environmental monitoring, aerial images are used to produce highly accurate DEM which are used to produce geo-referenced maps of biomass, ground cover classes, etc. Our input data are sequences of overlapping digitized aerial photographs (or more recently, high-quality digital video). Our 3D terrain modeling system produces a single, accurate geo-referenced DEM from multiple images (typically 3-12). We are utilizing an automatic stereo reconstruction algorithm that employs a hierarchical, texture matching scheme to generate DEMs from pairs of images that must be fused to form a large mosaic. The fusion process serves several purposes, including improving the accuracy by averaging redundant elevation estimates, detecting and removing outliers, and estimating the geospatial uncertainty. The key to a reliable fusion process is the use of self-consistency measures to identify and remove unreliable elevation estimates.

Of prime interest to us is the automatic detection of matching errors related to variables of the scene geometry. In particular we are interested in reconstructing the shape of the terrain from images which may have been taken from widely varying viewpoints. As the separation between the camera positions increases, elevation estimates become more precise [8]. At the same time, widely separated viewpoints result in a substantial perspective distortion between images, and increases the likelihood of encountering occlusions, which increase the chances of generating false matches. Even when viewing a surface from nadir, local variations in surface slope may result in highly distorted surface elements. In addition, aerial survey lenses often have a wide field-of-view [13] (typically 90°), which will increase the perspective distortion near the image edges. This tradeoff between increased accuracy versus increased error rate must be taken into account when designing a terrain modeling system.

This problem will be mitigated when multiple images are processed from different viewpoints. If a surface patch is highly distorted or obscured in one view, it may be clearly visible in another. Thus, as the number of views increase it is more likely that a correct match can be found for any particular surface patch, and it becomes imperative to detect unreliable elevation estimates, particularly when processing images taken at oblique viewing angles or when fusing multiple DEMs.

The methodology described in this paper for detecting errors in DEMs by self-consistency can be applied to both the two-image and multi-image scenarios. When only two images are available, two DEMs can be generated by reversing the roles of the reference and target images in the matching process. We will demonstrate that self-consistency measures can be effective employed to determination of the reliability of values in a DEM, which can then be excluded from further analysis. We will test the effectiveness of the methodology as well as the relationship between error rate and scene geometry by processing multiple overlapping views of a real terrain, as well as synthetic images of a realistic, 3D terrain model.

2 Photo-realistic simulation

Any comprehensive analysis and evaluation of a dense array of elevations estimates generated from images requires a dense array of ground truth. The typical use of a few ground control points spread out over an entire scene is simply not sufficient to compute meaningful statistics. Unfortunately, even with a technologically advanced system measurement system, such as an airborne scanning laser range mapping system or IFSAR (interferrometric synthetic radar), a high-resolution dense array of elevation spread out over several square kilometers would require an enormous effort.

To facilitate the analysis of the techniques described in this paper, we develop a method of generating pseudo ground truth through photo-realistic synthetic images. The process begins with a digital elevation map (DEM) and an ortho-image, which may come from an independent source (e.g., DFED or USGS digital ortho-quads), or which may have been generated from aerial images using a 3D reconstruction algorithm. We begin by treating the DEM and ortho-image as if they were precise, error-free representations of the terrain. Next, a ray tracing program is used to generate photo-realistic, synthetic images of the terrain from any viewpoint. The synthetic views then serve the function of real images.

Clearly pseudo ground truth and synthetic images are not a completely satisfactory substitute for real data. Nevertheless, the method does provide means for generating otherwise unobtainable ground truth samples, and furthermore, it generates expected types of errors due

to perspective distortion, and occlusion that will occur in real data. The images used to generate the pseudo ground truth were extracted from six 9 inch × 9 inch aerial images of a desert area near 29 Palms California[1]. Four of the images were taken so that a section of the terrain was visible in each image. This set of four overlapping images (one of which is shown in Figure 1) is the basis of the data presented in this paper.

The pseudo ground truth were generated from two of the four overlapping images. The choice of which two images to use was arbitrary. This DEM will undoubtedly contain some anomalies (deviations from the actual physical world). Because these anomalies become part of the pseudo ground truth and are correctly manifested in the synthetic images, they will not affect the quality of our experiments. The pseudo ground truth covered an area of 157.5m × 368.4m, with a ground sampling distance of approximately 0.35m.

To test the validity of the simulation procedure, we synthesized the images that were not used to generate the pseudo ground truth DEM, and then compared the synthesized images to the real ones. In other words, starting with four overlapping images, labeled A, B, C, D, we generated a DEM from images A and B (denoted by Z_{AB}). We then synthesized two images labeled C′ and D′ which had the same camera parameters as images C and D. A simple analysis of visual inspection, as well image differencing showed a remarkable similarity between the real and synthesized images. This has reasonably convinced us that the DEMs generated from the photo-realistic simulated data will be valuable in test the self-consistency algorithms. Figure 1 show one of the four original images, a small 400×400 pixel region of the original image, and the corresponding region of the synthetic image.

3 Self-consistency

At the heart of our method is an expectation of consistency between DEMs when the computation is accurate. In addition, we rely on the observation that for many 3D reconstruction algorithms [7] two DEMs can be generated for a single image pair. We consider the class of image matching algorithms that generate a dense array of disparities $\mathbf{D}_{AB}(i,j)$ such that the pixels (i,j) in image A and the pixels $(i+\mathbf{D}_{AB}(i,j), j)$ in image B are projections of the same surface elements. In this notation image A is the reference image and image B is the target image. By reversing the roles of images A and B we can generate a second disparity map $\mathbf{D}_{BA}(i,j)$ such that the pixels (i,j) in image B and the pixels $(i+\mathbf{D}_{BA}(i,j), j)$ in image A are projections of the same surface elements. Because of the

---

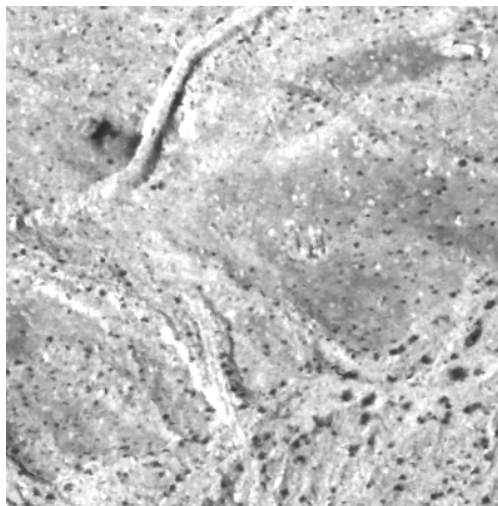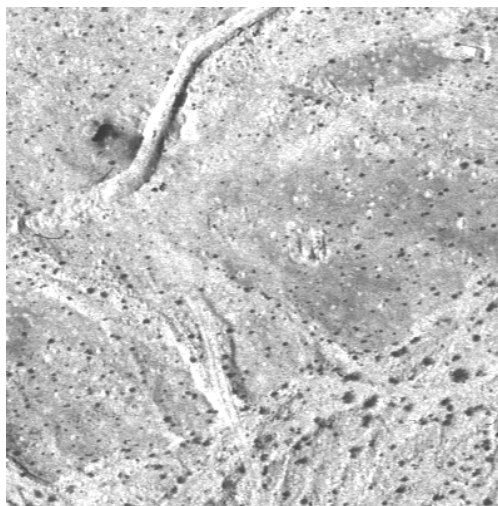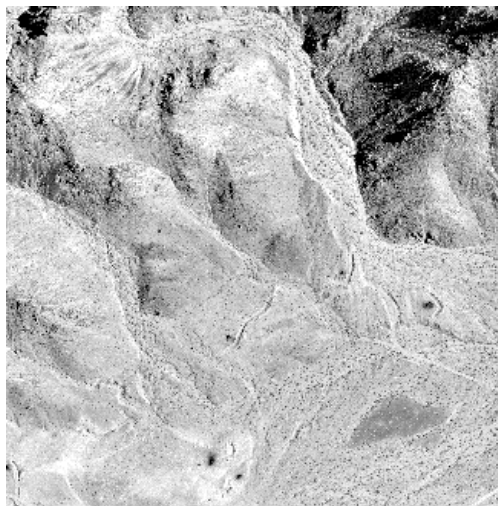[1] This data set was make available by the Army Topographic Engineering Laboratory

Figure 1. One of the original images (top); a small 400 x 400 pixel region taken from the original image (middle); the same region taken from the photo-realistic synthetic image (bottom).

nonlinear and adaptive manner in which disparity maps are usually computed, it is generally the case that $\mathbf{D}_{AB}$ and $\mathbf{D}_{BA}$ do not produce the same set of elevations.

Our analysis is based on testing the consistency of a pair of DEMs generated from a pair of overlapping images. By comparing reference-target duality in DEMs, inconsistency becomes a means of detecting errors. We denote the elevations recovered from images A and B by Z_{AB}, where the first subscript is the label of the reference image and the second subscript is the label of the target image. Without loss of generality, Z_{AB} may be written as the sum of the actual surface Z^* and a geospatial error term δ_{AB}. Thus, for any two overlapping images A and B, the two recovered surface models are

$$Z_{AB} = Z^* + \delta_{AB}$$
$$Z_{BA} = Z^* + \delta_{BA}$$

Taking the difference of the two computed elevation maps gives an expression that is independent of the surface structure.

$$Z_{AB} - Z_{BA} = \delta_{AB} - \delta_{BA} \tag{1}$$

The left hand side of Equation 1, which is very similar to self-consistency described in [10], depends only on images A and B. The right hand side is the difference between the geospatial errors, which requires ground truth to evaluate. Currently, we are developing a comprehensive model that expresses the conditional probability density function (PDF) of the geospatial errors to self-consistency distribution. As a first step, we will examine methods based on statistics derived from the distribution self-consistency.

We begin by assuming that the distribution of geospatial error measures δ_{AB} and δ_{BA} are comprised of two distinct populations:

1. Errors associated with correct correspondences, which we refer to as *inliers*. These errors result from small uncertainties in camera orientation, digitization, numerical roundoff, etc. Inliers are modeled by a zero-mean, normally distributed random process with a standard deviation (s.d.) in object space that corresponds to approximately one pixel registration error in image space.

2. *Outliers* which arise from false matches. These errors have a very broad distribution which may span hundreds of pixels in image space with correspondingly large errors in object space.

A simple test for separating the two populations can be devised by taking the s.d. of both sides of Equation 1,

$$\sigma(Z_{AB} - Z_{BA}) = \sigma(\delta_{AB} - \delta_{BA}) \tag{2}$$

This formula may be applied locally (e.g., within an $n \times m$ window), globally, or conditionally (e.g., as a function of incidence angle). Assuming that the distribution of δ_{AB} and δ_{BA} are identical, Equation (2) can be rewritten as

$$\sigma(Z_{AB} - Z_{BA}) = A\sigma(\delta) \tag{3}$$

where A is a function, that depends on several factors, including the local slope of the terrain, the scene geometry and the processing scheme. In the simplest case, when δ_{AB} and δ_{BA} are samples from an uncorrelated, normally distributed random processes, the function A reduces to a constant equal to $\sqrt{2}$, and Equation 3 can be rewritten as

$$\sigma(Z_{AB} - Z_{BA}) = \sqrt{2}\sigma(\delta)$$

Although this simple case cannot be applied to most situations, it is important to observe that large values of the geospatial set-consistency are likely to be associated with corresponding large geospatial errors, which in turn are associated with false matches. In other words, we expect that the distribution of the object space self-consistency measure will mimic the geospatial error distribution.

A method of determining the s.d. of a population containing a mixture of inliers and outliers was developed that reduces the influence of outliers. This was necessary because the inlier population is expected to have a zero mean and a s.d. that is fraction of a meter, whereas the outlier population is expected to have values on the order of several meters. A few outliers, therefore, will drastically affect the computation of the s.d. of the entire population. To minimize the influence of the tails of the distribution, which we expect to be dominated by outliers, standard deviations were computed by fitting the histogram to a Gaussian plus a constant of the form

$$h_{max} \exp-\left[\frac{(z_i - z_0)^2}{2\sigma^2}\right] + h_0 \tag{4}$$

where z_i are the histogram values, and the peak value (h_{max}), the floor (h_0), the mean (z_0), and standard deviation (σ) are parameters of the fit. If the population was normally distributed, there would be no need to compute σ using Equation 4. Instead σ could be found by evaluating the standard computational formula

$$\sigma = \frac{1}{n-1}\sqrt{\sum_i (z_i - z_0)^2}$$

With outliers present in the population, the curve fitting technique reduces the influence of non-Gaussian distributions in the tails.

The outlier threshold for the analysis presented in this paper was set to 2σ. It must be understood that whatever method is used to determine a threshold, the resultant reliability cannot be considered exact. Some of the values marked as reliable may not be, and vice versa, for two reasons:

- In general, the self-consistency distribution is not isomorphic with the geospatial error distribution. As pointed out above, a more comprehensive formulation is need to precisely predict the statistics of the geospatial errors from the self-consistency distribution (planned future work).

- It is possible, although very rare, that a false match will produce the same results when the reference and target images are reversed. In fact, we were not able to find any errors of this type in our data.

Although we do not offer an objective method for setting the outlier threshold, one can easily determine if the threshold is set too low or too high by checking the number of rejected points. Furthermore, we observed that the fusion results are fairly insensitive to the outlier threshold, and we were able to achieved similar results with σ set to values between 2 and 4. We chose a conservative threshold (i.e., a low value) to ensure that the DEM computed for the real data had as few unreliable elevations contributed to the fused DEM as possible. The 2σ cutoff threshold was selected because it resulted in a complete recovery for the real data, although some DEM elements in the simulated data set did not have any reliable points (note: these points could be filled in by interpolating the fused DEM).

To test the procedures we compared the s.d. of the self-consistency $(Z_{AB} - Z_{BA})$, the percent of detected inliers, the s.d. of the geospatial errors $(Z_{AB} - Z^{*})$, $(Z_{BA} - Z^{*})$, and $(\frac{1}{2} \cdot (Z_{AB} + Z_{BA}) - Z^{*})$ as a function of various off-nadir view angles θ for cameras A and B. The s.d. of the geospatial errors were computed by evaluating the standard computational formula for all elevation estimates less than the 2σ threshold.

The results are summarized in Table 1, and a plot of the histogram and the fitted distribution curve for $b/h = 0.277$ are shown in Figure 2. For the data summarized in Table 1, h_{max} was at least two order of magnitude greater than the outlier floor h_0, and the total extent of the histogram tails was approximately ± 60m, which was much greater than width of the fitted distribution σ. The real data were all taken with the camera looking approximately straight down. For the simulated data, however, we were able to place the cameras in any position. This allowed us to study the affects of viewing geometry on the performance of the algorithm.

Inspection of Table 1 shows that the trend is for the standard deviation of the geospatial errors to decrease with increasing base-to-height ratio (b/h) and for the number of outliers to increase with increasing b/h. As mentioned in Section 1, this is the expected trend based on standard stereo geometry [7]. The last column shows the standard deviation of the geospatial error measured relative to the average of the two retrieved DEMs.
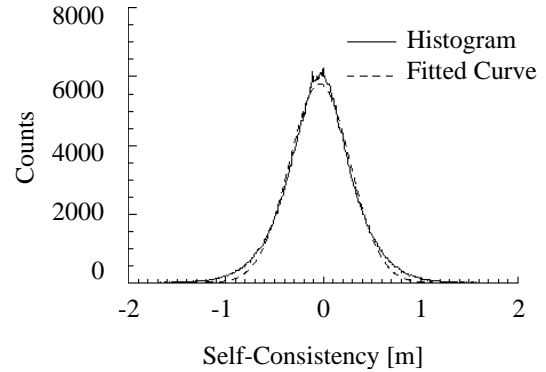

Figure 2. Self-consistency distribution for image pair 3

4 Fusing multiple DEMs

The method of using self-consistency measures to detect unreliable elevation estimates can applied to the problem of fusing multiple generated from several overlapping images. The basic principle is to identify outliers (unreliable elevations estimates) and then compute a weighted average of all reliable points. We employed the following two-pass algorithm:

- Given a set of n images where n is greater than 2, compute the set of order dependent DEMs and self-consistency distributions. For the four overlapping images discussed above (labeled A,B,C,D), 12 DEMs $(Z_{AB}, Z_{BA}, Z_{AC}, Z_{CA}, \ldots, Z_{CD}, Z_{DC})$ and six self-consistency distributions $(Z_{AB}-Z_{BA})$, $(Z_{AC}-Z_{CA})$, ..., $(Z_{CD}-Z_{DC})$ can be computed.

- For each DEM pair compute an outlier threshold by (i) fit the formula in Equation 4 to the histogram of the self-consistency distribution, (ii) set the threshold to a predetermined multiple of the s.d. of the fitted function, and (iii) assign summation weights (planned future work).

- Mark as unreliable all elements in the DEM with a self-consistency measure greater than the threshold.

Table 1. Self-consistency and geospatial error statistics as a function of the base-to-height ratio (*b/h*) and the angle between the optic axis and vertical (θ) for images A and B. $\sigma(Z_{AB}-Z_{BA})$ is the standard deviation of the self-consistency distribution; % inliers is the percent of all elevations where $|Z_{AB}\text{-}Z_{BA}| < 2\sigma(Z_{AB}-Z_{BA})$; the last three columns are the standard deviations of the difference between the pseudo ground truth Z^*, and Z_{AB}, Z_{BA} and average of Z_{AB} and Z_{BA}. All standard deviations are in meters.

B/h	θ_A	θ_B	$\sigma(Z_{AB}-Z_{BA})$	% Inliers 2σ cutoff	$\sigma(Z^*-Z_{AB})$	$\sigma(Z^*-Z_{BA})$	$\sigma\left(Z^*-\bar{Z}\right)$
0.277	0°	15°	0.451189	91.90	0.332601	0.244706	0.213685
0.293	15°	30°	0.486813	92.50	0.344480	0.330056	0.260698
0.575	15°	-15°	0.311553	91.36	0.163137	0.213822	0.131443
0.868	-15°	30°	0.203503	89.40	0.157535	0.194275	0.152326
1.230	30°	-30°	0.167713	84.24	0.155302	0.188295	0.155993

- For each element in the DEM grid, compute the average of reliable elevation estimates $\bar{Z}$.

- An optional step would be to compute an interpolated value for all elements without any reliable elevations. This option was not in the examples presented in this paper.

- Re-apply the threshold test to the difference between the fused DEM $\bar{Z}$ and the unreliable elements. If any of the differences are less than the outlier threshold, re-label the point as reliable. This step is necessary because some of the unreliable matches occur in $\mathbf{D}_{AB}$ and not $\mathbf{D}_{BA}$ (or visa versa).

- Update the fused DEM by including the newly relabeled elevation estimates.

This fusion algorithm was applied to two data sets – the set of four overlapping images, and the photo-realistic simulated data (see Figure 1). For the simulated data we were able compare the results to an absolute reference. The results generated from the simulated data are shown in Figure 3, and the results from processing the real data are shown in Figure 4. The simulated data covered a significantly narrower region than the real data. This reduction in the size occurred because the image matching algorithm cannot process data along the boarder of the image; and the synthesized images are computed only for regions where a DEM exists.

For the simulated data set, 99.54% of the DEM elements had at least one reliable point (the missing elements show up in Figure 3 as small black flakes), and the overall accuracy, measured by the standard deviation of the difference between the pseudo ground truth and the fused DEM (i.e., $\sigma(Z^* - \bar{Z})$) was 0.169m (the ground sampling distance was 0.35).

Our methodology involves automatic determination of a reliable set of elevation estimates for each value in each DEM. Therefore, there will be a varying number of contributions to the final elevation at each point in the $2k \times 2k$ fused DEM. This is graphically displayed by gray value in Fig. 4, which shows in gray value the number of elevation estimates used (varying from 0 to a maximum of 12. About 98% of the points had 10 or more reliable values contributing, and only 215 points of the 4 million had 3 or less values (157 with no estimates). Most of these cases were isolated, and median filtering would achieve reasonable estimates for many of them.

Of course, for the real data we were not able to compute the accuracy of the retrieval. However, we did find that the algorithm found at least one reliable element in 99.99% of the scene. In addition, there were no apparent anomalies in the fused DEM. That is, the appearance of the ortho-image seemed consistent with the terrain everywhere in the scene. Our qualitative inspection consisted of checking the relationship between objects such as rocks and ditches to their shadows, and checking to see if streams (or in this case dry streambeds) flow down hill everywhere

5 Conclusions

We successfully applied the notion of self-consistency to the specific task of improving the generation of DEMs using dual symmetric stereo processing on pairs of images from a subset of overlapping aerial images. The photo-realistic simulation results shows an ability to detect almost all outliers by using a conservative threshold to mark reliable elevation estimates in dual DEMs from a pair of images. There is a significant decrease in the residual errors of the model compared to individual DEMs generated without regard to self-consistency. Experimental results are in good correlation with theoretical predictions. The strategy for fusion of the multiple DEMs can be further improved and this will be investigated in the future.

There is a significant computational overhead associated with this technique. It can be justified in cases where a very accurate DEM is important, such as for off-
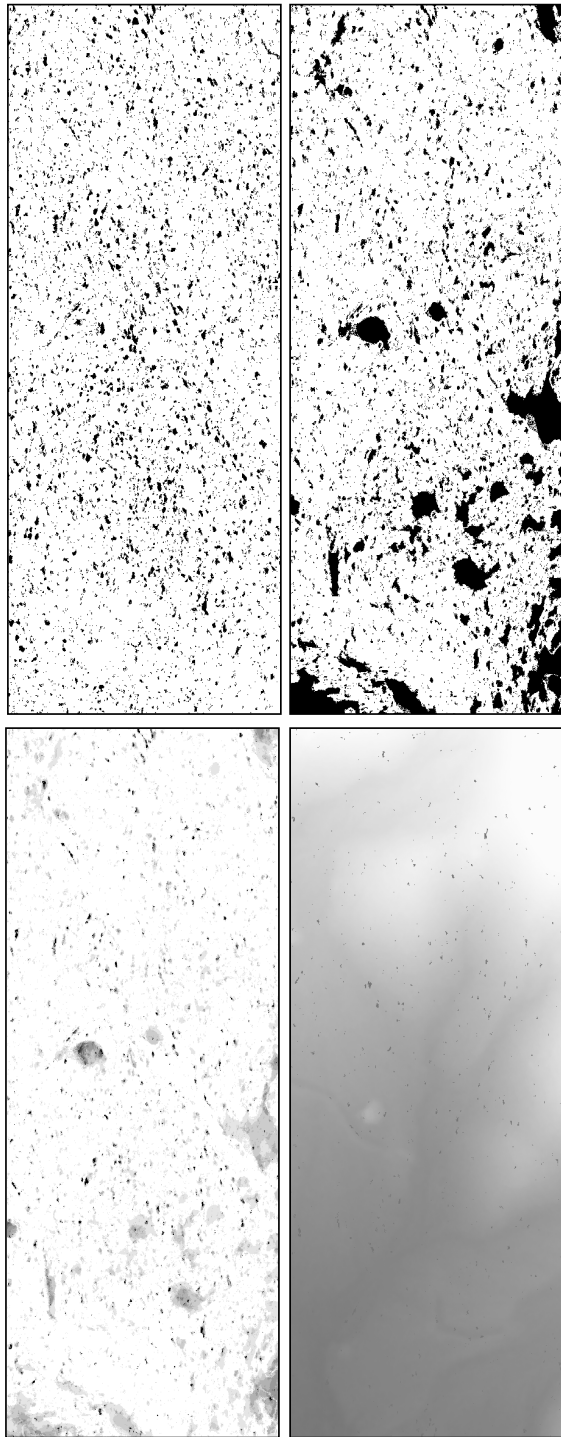
Figure 3. The inlier mask for $\theta_A=0°$ and $\theta_B=15°$ (top-left); inlier mask for $\theta_A=-30°$ and $\theta_B=+30°$ (top-right); an image of the number of samples used to compute the fused DEM, black corresponds to $n=0$ and white corresponds to $n=10$; the fused DEM, the black flecks are elements with no elevation estimates ($n=0$).

road autonomous driving. The work presented is the first step in a more comprehensive attempt to apply the notions of geospatial uncertainty and self-consistency in order to generate near-optimal elevation models. The focus of our ongoing work will be on extending the technique to investigating the value of self-consistency measures among real and across synthetic imagery. This will require a better model for the propagation of errors and weighting the contributions from the different sources.

From our experimental results, we conjecture that most of the benefit of the fusion process derives from a relatively small subset of the data. In the future, we will conduct studies that will give one the ability to select, from the available images, the optimal/minimal subset that conforms to a given error tolerance.

References

[1] Agouris, Peggy and T. Schenk, Automated Aerotriangulation Using Multiple Image Multipoint Matching, Photogrammetric Engineering and Remote Sensing, Vol. LXII, No. 6, June 1996, pp. 703-710.

[2] Ayache, N., and B. Faverjon, "Efficient Registration of Stereo Images by Matching Graph Description of Edge Segments," *Int'l J. Computer Vision,* pp. 107-131, 1987.

[3] Aschwanden, P. and W. Guggenbuehl, "Experimental Results From a Comparative Study on Correlation-Type Registration Algorithms," *Robust Computer Vision,* Foerstner and Ruwiedel, eds., pp. 268-289, Wichmann, 1993.

[4] Fua, P. and Y.G. Leclerc, "Taking Advantage of Image-Based and Geometry-Based Constraints to Recover 3-D Surfaces," *Computer Vision and Image Understanding,* vol. 64, no. 1, pp. 111-127, 1996.

[5] Grimson, W.E.L., "Computational Experiments With a Feature Based Stereo Algorithm," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 7, no. 1, pp. 17-34, January, 1985.

[6] Hoff, W., and N. Ahuja, "Surface From Stereo: Integrating Feature Matching, Disparity Estimation and Contour Detection," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 11, pp. 121-136, 1989.

[7] Horn, Berthold K. P., Robot Vision, MIT Press, Cambridge, MA.., 1986.

[8] Kanade, T. and M. Okutomi, "A Stereo Matching Algorithm With an Adaptive Window: Theory and Experiment," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 16, no. 9, pp. 920-932, Sept. 1994.

[9] Leclerc, Y.G., Q.T. Luong, and P. Fua, "A Framework for Detecting Changes in Terrain," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 20, no. 11, pp. 1143-1160, November 1998.

[10]   Leclerc, Y,G., Q.T. Luong, et al., "Self-consistency: A novel approach to characterizing the accuracy and reliability of point correspondence algorithms," *DARPA Image Understanding Workshop*, Monterey, CA, Morgan Kauffman, 1998.

[11]   Medioni, G. and R. Nevatia, "Segment-Based Stereo Matching," *Computer Vision, Graphics, and Image Processing,* vol. 31, pp. 2-18, 1985.

[12]   Schultz, H., "Terrain Reconstruction from Widely Separated Images", Proc. SPIE, Volume 2486, pp. 113-123, Orlando, FL, April, 1995.

[13]   Slama, Chester C. (Editor), Manual of Photogrammetry, 4ed., American Society of Photogrammetry, Falls Church, VA., 1980.

[14]   Witkin, A., D. Terzopoulos, and M. Kass, "Signal Matching Through Scale Space," *Int'l   J. Computer Vision,* pp. 133-144, 1987.
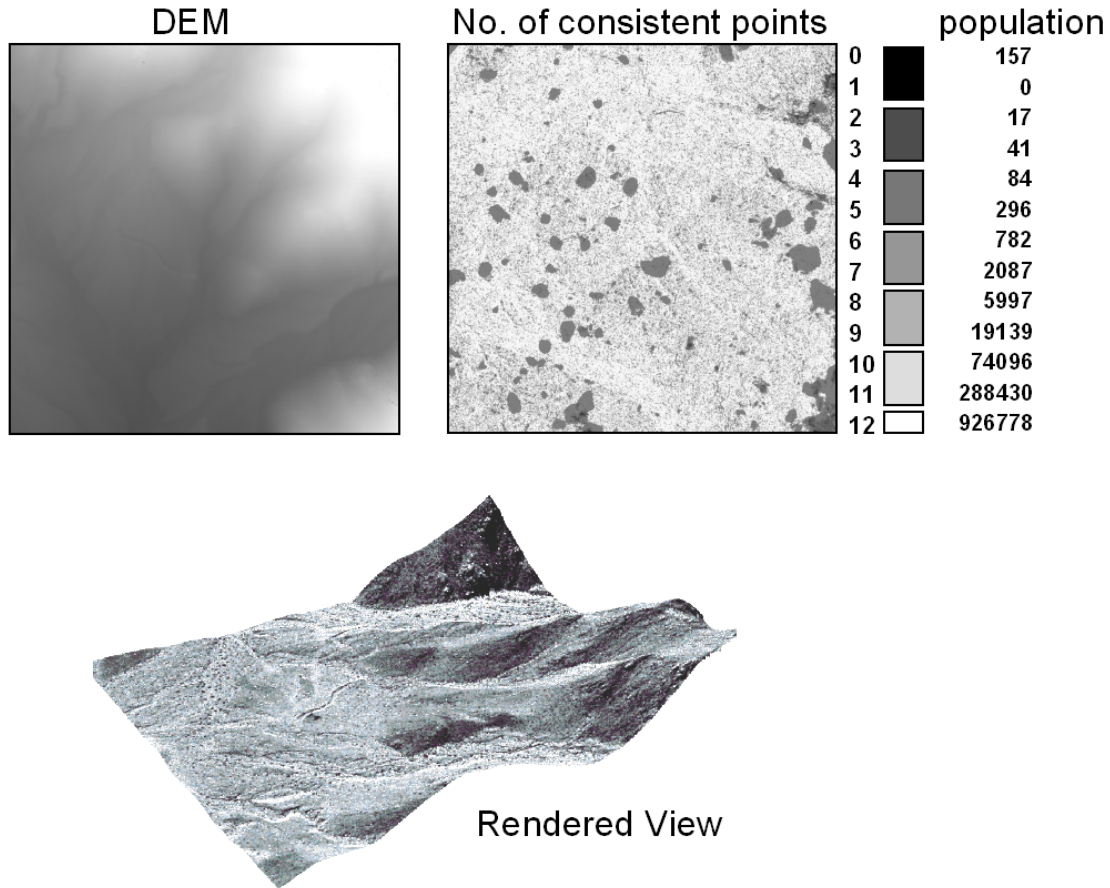
Figure 4. DEM fusion from the four real images. The DEM (top-left), a map of the number of consistent elevation estimates (top-right), the ortho-image draped over the DEM (bottom).