

TOWARDS A SEAMLESS ROBUST DIGITAL SURFACE MODEL (DSM)

B. Ameri, N. Goldstein, H. Wehn, A. Moshkovitz, H. Zwick

MacDonald Dettwiler and Associates Ltd. (MDA)
13800 Commerce Parkway, Richmond, BC. Canada
Tel: (604) 231-2308, Fax: (604) 278-2117

bameri@mda.ca, norman@mda.ca, hw@mda.ca, amoshkov@mda.ca, hhz@mda.ca

Commission III

KEY WORDS: Multiple image matching, DEM/DTM, Graph theory, Sensor calibration, Digital cameras

ABSTRACT:

FIFEDOM (Frequent-Image-Frames Enhanced Digital Ortho-rectified Mapping) is an ongoing, multidisciplinary research project aiming to provide low-cost, information-rich digital mapping through the development of an intelligent digital airborne image acquisition strategy, and novel image processing techniques. In particular, FIFEDOM is designed to exploit the high degree of overlapping in digital aerial imagery to generate accurate map products. FIFEDOM utilizes high-overlap digital image acquisition to provide multiple-look-angle reflectance for each pixel in a scene. The FIFEDOM project concentrates on four technology areas: Image Acquisition System and Sensor Calibration; Radiometric Balancing; Digital Surface Model (DSM) generation; and Bi-directional Reflectance Distribution Function (BRDF) model extraction.

This paper emphasizes the FIFEDOM DSM module. It introduces an innovative, breakthrough methodology for automated DSM generation. This is achieved through utilization of multi-view, multi-frame highly-overlapped digital images. This capability significantly improves the result significantly by generating a very dense, high-quality, reliable Digital Surface Model.

1. INTRODUCTION

Since the 1970s, an increasing amount of research work has been reported aiming to replace the manual time consuming image mensuration process with a semi/fully automated operation (Helava 1976, Förstner 1982, Ackermann 1984). This investigation leads to a series of more complex and successful commercially available systems, for e.g. automatic Digital Elevation Model (DEM) generation or Automatic Aerial Triangulation (AAT). In spite of a high degree of automation in those systems, a significant amount of work is still required for post-processing and editing of extracted e.g. elevation data especially in problematic areas such as forest or built-up areas. This is due to a number of reasons such as occlusion, height discontinuities, repetitive patterns, shadows, and lack of texture, to name a few.

The proposed work is an attempt to overcome some of these issues especially in built-up areas where the user community would like more frequent updates if the cost could be contained. The importance of this project to the geomatic industry is based on the reality that digital airborne cameras are now practical for mapping. Modern CCD performance and cost characteristics make this a cost-effective approach, and multiple digital image pairs are not expensive compared to traditional aerial photography.

The paper introduces an innovative methodology for automated DSM generation. This is achieved through utilization of multi-view, multi-frame highly overlapped digital images. Existing tools are based on stereo image matching techniques, where only two overlapping images are processed. This leads to unsatisfactory results (e.g. in built-up areas, where occlusion is a major issue). The method proposed by FIFEDOM overcomes

this problem, if not completely, then at least partially, by applying novel multiple image matching techniques. In this manner, if an area/point is occluded in one image, it is still likely to be seen in other overlapping images, captured from different view angles, and it is possible to reconstruct a 3D representation of the occluded terrain surface. This capability improves the result significantly by generating a very dense, high-quality, reliable Digital Surface Model (DSM). The resulting DSM is subsequently (beyond the scope of this paper) used to create Digital Elevation Models (DEMs) and other related image products (e.g., true ortho-rectified images).

2. OVERVIEW

The FIFEDOM project employs digital RGB imagery from the Nikon D1 camera, 12 bits per band. The camera has been mounted on an airplane by the Range and Bearing Environmental Resource Mapping Corporation, who also acquire the project's data sets, and supply the GPS/INS positioning for each frame of imagery.

Before reaching the DSM module, a flat field radiometric calibration is applied to the imagery. The flat field is a multiplicative mask that was previously derived from a white light calibration procedure.

Also, geometrically, the camera was self-calibrated for the given camera settings. The tie points are automatically generated using area correlation without manual intervention, using the remotely sensed imagery that is routinely collected. However, for calibration, a small set of flight lines are needed that are orthogonal to the main set of flight lines, to supply a sufficiently geometrically rich data set for extracting the camera calibration

parameters. These consist of standard radial and tangential distortion, on top of an affine transformation between the focal and image planes. Calibration is not required for every data set due to the inherent stability of the camera.

The DSM module is described in the next section. After the DSM construction, (near true) ortho mosaicking is performed, as well as using the ortho image frames to explore the BRDF qualities of the data set. For a typical FIFEDOM data set, there are about 26 distinct frames, which view a region from different zenith and azimuth angles. Moreover, since the FIFEDOM data have 12-bit radiometric resolution, even small changes in view-angle induce a detectable change in image brightness. FIFEDOM explores a wide range of the issues associated with BRDF-based land classification, but that discussion is beyond the scope of this paper.

3. DSM

This section describes the workflow of the FIFEDOM DSM module and discusses its major components and their interrelationships. The process is performed based on a hierarchical coarse-to-fine strategy, where the final output is a seamless DSM that covers the entire area. The process consists of four fundamental steps as follows:

1. Image Pyramid Generation – images are down sampled to different resolution layers.
2. Point Feature Extraction – a list of interesting points is extracted individually for each of the images.
3. Multiple Points Matching – the extracted points in the different images are matched.
4. Multiple Forward Intersection – match points consisting of two or more points enables reconstruction of the 3D location of the object point on the terrain.

3.1 Initial Search Space

One of the fundamental tasks in the proposed DSM generation algorithm is to identify and to measure conjugate (homologous) points in two or more overlapping images, a process known as feature/image matching. There are many interesting aspects to image matching, such as determination of initial values and suitable approximations, the selection of exact correspondence from multiple solutions, outliers detection and removal, and mathematical modeling of terrain surface, to name a few. In this section we focus on the constraints which restrict the space of possible matching solutions, such as setting bounds on the search area, which also begins the image matching process rather close to the true solution.

The FIFEDOM camera is a frame camera with a central perspective projection. Central perspective projection provides a very powerful constraint, namely that of epipolar geometry. Given two images and a 3D point in object space, the epipolar geometry is defined as a plane containing the object point and the two projection centers of both images. If the relative orientation of two images is known, for a given point in one image the epipolar line in the other image can be computed, and the corresponding point must lie on this epipolar line. Thus, the image-matching problem is reduced from a two- to a one-dimensional task. In order to facilitate matching along epipolar lines in a multiple image environment such as FIFEDOM, the epipolar geometry should be constructed on-the-fly using the relative orientation parameters of the images. The epipolar

geometry constraint is applied in the process of Multiple Points Matching (see Section 3.3), where the process is searching for the candidate homologous points within a search area along the epipolar lines. This reduces ambiguity and computational cost.

To further reduce the initial search space in the matching process a hierarchical coarse-to-fine strategy is used. In this process images are represented in a variety of resolutions, leading to an image pyramid. The images are organized from coarse-to-fine pixel resolution, and results achieved on one resolution are considered as approximations for the next finer level. A coarser resolution is equivalent to a smaller image scale, and a larger pixel size. Thus, the ratio between the flying height and the terrain height increases as the resolution decreases, and local disturbances such as occlusions become less of a problem (Hepkie 1996). In this work, a low-pass Gaussian kernel is used to generate the image pyramid of FIFEDOM images.

3.2 Point Feature Extraction

In the previous section we discussed some simple but powerful methods to reduce the search space. Now, the issue is the selection of appropriate matching primitives. In fact, the distinction between different matching primitives is probably the major difference between the various matching algorithms. This is true because this selection influences the whole process of the matching. These primitives fall into two broad categories, *windows* composed of gray values (area based matching), and *features* extracted in each image a priori (feature based matching). We selected a feature-based matching approach in this work, where the Förstner operator is used to extract salient points as matching primitives in each image individually prior to the matching process (Förstner 1986). The advantage of the point selection is obvious. It leads to a significant information reduction, as we have only to deal with a set of points, not all pixels in the images.

Each extracted point is characterized by a set of attributes. These attributes are the key elements in the success of the upcoming process of Multiple Point Matching, as they are input parameters to a *weight function* (Equation 1), which actually establishes the similarity measures between the candidate homologous points in the overlapping images (see Section 3.3 for details).

3.3 Multiple Point Matching

The objective of the multiple point matching is to determine precise locations of homologous points from n images ($n \geq 2$). All the points should contribute simultaneously to the solution to exploit a major advantage that digital image matching offers. This results in a higher redundancy for the matching problem and thus a greater reliability is achieved for the results (Baltasavias 1991). The following two sections discuss the overall concept of the multiple point matching problem and the potential solution to this problem.

Multiple Point Matching Problem: Figure 1 illustrates the multiple point matching problem. We have five overlapping images, all of which contain the same building. Point p_1 in Image 1 is a point of interest (the upper right corner of the building). The multiple point matching problem in this case is to find the set of points in all the other images that best match point p_1 in image 1. Each of the Images 2 to 5 has a set of candidate points that potentially match point p_1 . These are the

red and green points (the blue points are candidates for other points, but not for point p_1). The Image 2 has two candidate points, and Images 3, 4, and 5 each have three candidate points. It is impossible to have a candidate point from the same image as point p_1 .

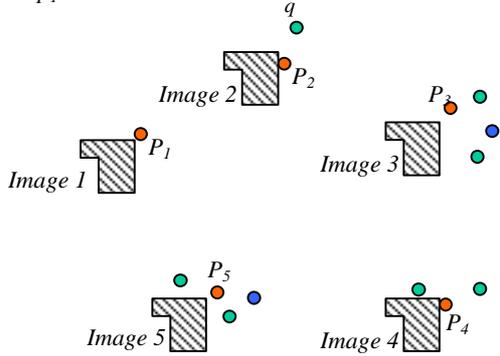


Figure 1: Multiple point matching problem

By simply looking at the figure, we see that points $p_2, p_3, p_4,$ and p_5 (drawn in red) are the set of points that best match point p_1 in Image 1. But how do we determine algorithmically that this is the optimal solution?

3.3.1 N-partite maximum-weight clique problem: The first step in solving the problem is to associate a weight, or a measure of similarity, between p_1 and each candidate point for p_1 in the other images. The similarity weight is computed based on the weight function f_{ij} :

$$w_{ij} = f_{ij}(s_1, s_2, s_3, \dots) \quad \begin{matrix} i = 1, 2, \dots, n \\ j = 1, 2, \dots, n \\ i \neq j \end{matrix} \quad (1)$$

Where w_{ij} is the correspondence weight, s_1, s_2, s_3, \dots are the similarity measures between the candidate homologous points such as normalized cross correlation, and n is the number of overlapping images that contain candidate points for p_1 .

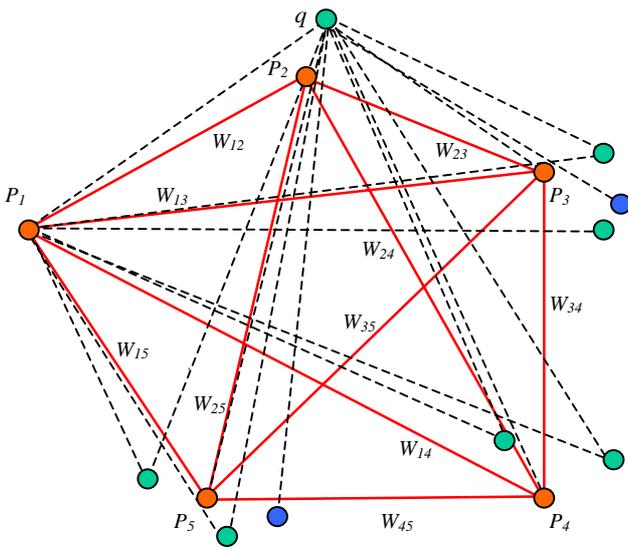


Figure 2: n-partite undirected weighted graph

Once we have generated the weights for every pair of candidate points, we then construct an undirected weighted n-partite graph $G=(V, E)$ where V is the set of all distinct candidate points and E is generated by considering the weight of all the points previously matched. Figure 2 shows the resulting partially-drawn graph. All the edges connecting candidate points are drawn in for point p and q , along with their weights. There would certainly be more edges connecting candidate points for points $p_2, p_3, p_4,$ and p_5 , but the graph would be far too crowded. Once this graph G is constructed, we solve the problem of multiple points matching by finding a set of vertices, which form a *maximum weight clique*.

Thus, we can reduce the multiple point matching problem to the maximum-weight clique problem. The edges that connect the maximum-weight-clique are drawn with solid lines in Figure 2. All other lines are dashed. A similar approach is reported by Tsingas (1994) to solve the matching problem in the digital point transfer process.

3.4 Multiple Forward Intersection

The idea is that the 3D position of the matching points in object space (terrain/model co-ordinates) should be determined from the simultaneous contribution of the 2D co-ordinates of the homologous points in n ($n \geq 2$) images. This can be obtained directly by spatial forward intersection of corresponding space rays as the result of a least squares solution, as depicted in Figure 3. The mapping relation between the point in 3D object space and its perspective projection in 2D image space is represented by the classical collinearity equations (Ameri 2000). The required mapping parameters are derived through a bundle adjustment process in earlier phases, which form the refined model parameters and are input to the DSM module.

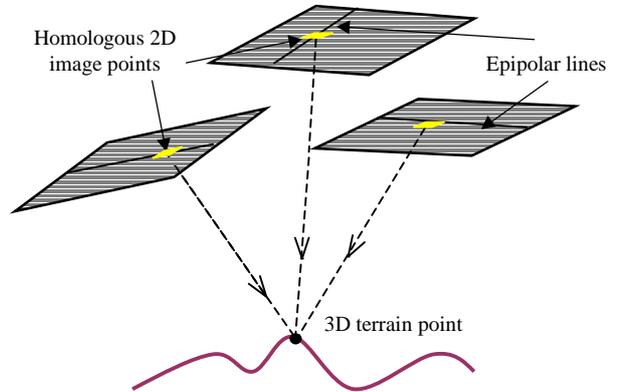


Figure 3: Multiple Forward Intersection

The cloud of 3D points generated in each pyramid level n is used as a coarse approximation of terrain heights for the next level $n+1$. Therefore a simple terrain modeling process is used to filter out the erroneous points and convert the cloud of 3D points into a possibly outliers-free TIN data structure, and from that to a GRID structure.

Figure 4 illustrates an example of the generated digital surface model. The data set consists of seven overlapping images acquired in three different flight lines, 1500m above the ground with 60% overlap in both cross- and along-track direction. It covers an area of 900m x 700m of the terrain.

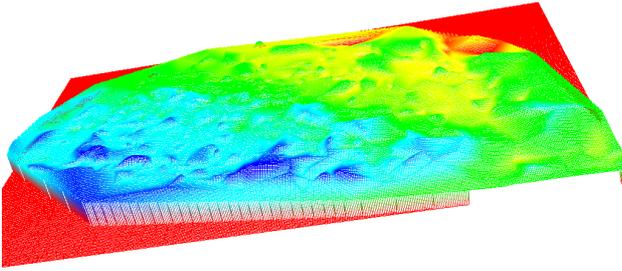


Figure 4: 3D perspective view of generated DSM

Since the 3D position of the points are derived based on a LS solution, the estimated variances (σ_x^2 , σ_y^2 , σ_z^2) of the terrain point co-ordinates are used as the quality measures to detect and remove the potential outliers from the result.

Planimetric threshold $t_p = 0.5 \times \text{GSD} = 0.6\text{m}$ Altimetric threshold $t_h = 1.5 \times t_p = 0.9\text{m}$ Total # of 3D points = 12530; Total # of potential outliers = 1736						
# of intersecting rays	2	3	4	5	6	7
# of reconstructed 3D points	999	7322	2653	918	377	181
# of potential outliers / % of potential outliers	822/ 82%	853/ 12%	99/ 2%	2/ 0%	0/ 0%	0/ 0%

Table 1: Relative number of potential outliers vanishes with an increasing number of images.

Table 1 summarizes some of the statistical measures regarding the generated DSM. Note that the number of outliers is approaching zero as the number of images incorporating into the solution increases. This preliminary result is in agreement with the initial motivation of the FIFEDOM project, that high information content and the great wealth of information in highly-overlapped image frames improves the resulting geospatial production significantly.

4. CONCLUSION

We have introduced a new method for the reliable and efficient reconstruction of a digital surface model through utilization of multiview, multiframe, highly-overlapped digital images. The results presented are intermediate results and thus some of the processes are only implemented in a simplified form such as terrain surface modeling and outliers detection. These processes will be fully implemented in the next stage of this study, which subsequently improves the result and efficiency of the proposed method.

5. REFERENCES

Ackerman, F. (1984), Digital image correlation: performance and potential application in photogrammetry, *Photogrammetric Record*, Vol. 11, No. 64, pp. 429-439.

Ameri, B. (2000), Feature Based Model Verification (FBMV): a new concept for hypothesis validation in building reconstruction, *IAPRS* Vol. 33, part B3, Amsterdam, The Netherlands

Baltsavias, E. (1991), *Multiphoto geometrically constrained matching*, Dissertation, Institut für Geodäsie und Photogrammetrie, Mitteilungen der ETH Zürich (49).

Förstner, W. (1982), *On the geometric precision of digital correlation*, *IAPRS* Vol. 24, No. 3, pp. 176-189.

Förstner, W. (1986), A feature based correspondence algorithm for image matching, *IAPRS*, Vol. 26, No. 3/3, pp. 150-166.

Heipke, C. (1996), Overview of image matching techniques, *OEEPE Workshop on the Application of Digital Photogrammetric Workstations*, OEEPE Official Publications No. 33, 173-189

Helava, U. V. (1976), Digital correlation in photogrammetric instruments, *IAP*, Vol. 21, No. 3.

Schenk, T. (1999), *Digital Photogrammetry*, Vol. 1, TerraScience, OH, USA.

Tsingas, V. (1994), A graph-theoretical approach for multiple feature matching and its application on digital point transfer, *IAPRS* Vol. 30, No.3, pp. 865-871.

6. ACKNOWLEDGEMENTS

The authors express their gratitude to Precarn Incorporated for partial funding support of this project, and to Triathlon Ltd. for helping to define the DSM requirements and continued evaluation of the results. We acknowledge the Range and Bearing Environmental Resource Mapping Corporation for acquisition of the test data.