

HIGH RESOLUTION DIGITAL SURFACE MODEL (DSM) GENERATION USING MULTI-VIEW MULTI-FRAME DIGITAL AIRBORNE IMAGES

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

MacDonald Dettwiler and Associates (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 IV

KEY WORDS: Multiple image matching, DSM, Graph theory, Sensor calibration, Digital cameras

ABSTRACT:

FIFEDOM (Frequent-Image-Frames Enhanced Digital Ortho-rectified Mapping) is a multidisciplinary ongoing research project aiming to provide low-cost and high information content digital mapping by developing an intelligent digital airborne image acquisition strategy, and novel image processing techniques. In particular, FIFEDOM is designed to exploit the high degree of overlapping of the 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, and breakthrough methodology for automated DSM generation. This is achieved through utilization of multi-view, multi-frame highly overlapped digital images. This capability improves the result significantly by generating a very dense, high quality, and reliable Digital Surface Model. The paper also discusses sensor calibration and radiometric-balancing, as they are prerequisite steps for the successful completion of the DSM task.

1. INTRODUCTION

Since the seventies, 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. automated DEM generation or AAT. In spite of a high degree of automation in those systems, still a significant amount of work is 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, and reliable Digital Surface Model (DSM). The resulting DSM is subsequently (beyond the scope of this paper) used to create Digital Elevation Models (DEM) and other related image products e.g., true ortho-rectified images.

The paper also briefly discusses other FIFEDOM technology areas, especially sensor calibration and radiometric-balancing, as they are prerequisite steps for the successful completion of the DSM task.

2. OVERVIEW

Figure 1 illustrates a context diagram of the FIFEDOM data processing system. The inputs are D1 12-bit per band raw colour imagery, an ascii file containing GPS/INS data that accompanies the acquired imagery and is generated by the acquisition system's support software, and a product specification determining the processing flow and which of the outputs to generate. The outputs are the camera's radiometric and geometric calibration results (section 3.3), BRDF results (section 5), radiometrically balanced image frames (section 4) and DSM results (section 6).

Figure 2 illustrates the relationship between the DSM module and other FIFEDOM components. The module receives the refined model parameters, the single channel of radiometrically balanced images, and a coarse DSM as input and delivers a TIN/GRID dense DSM as output.

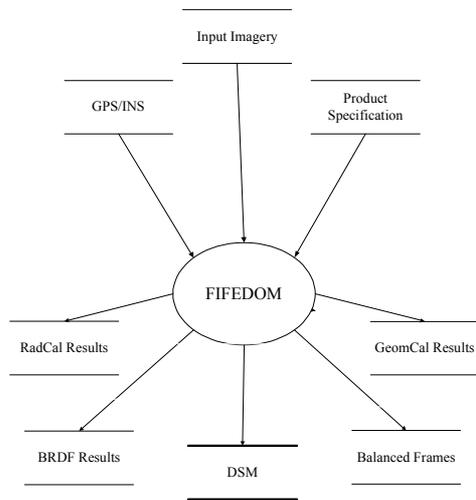


Figure 1: FIFEDOM system context diagram

The refined model parameters are estimated in a post bundle adjustment (BND) process where the initial values are provided using the onboard GPS/INS system during the flight mission. The radiometrically-balanced images are the output of Radiometric Balancing (RadBal) module of FIFEDOM where, using sparse tie points (TIPs), the corrective gains and offsets of the individual images are estimated and applied.

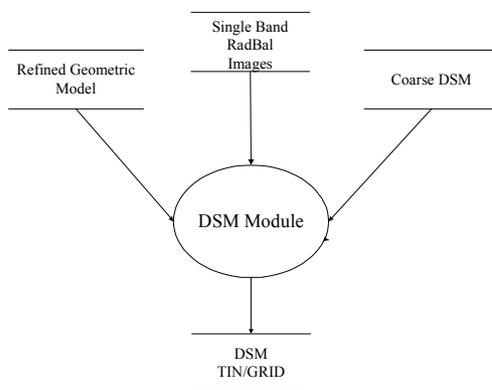


Figure 2: DSM module external interfaces

This process is discussed in more detail in section 4. In addition, the sparse TIPs are used to generate a coarse DSM representing an approximation of terrain surface, which is used as a starting point for DSM generation. The balancing is applied in one band only (green band) and is sufficient for 3D-point extraction, but not necessarily to form a globally balanced and calibrated mosaic of the data set. The actual and a more sophisticated RadBal process is applied to all 3 bands in a later stage, where the output of the DSM module is used as an input to that stage.

The following sections contain a detailed description of the above processes.

3. FIFEDOM ACQUISITION SYSTEM

This section describes how imagery from the D1 camera is prepared, as input to the DSM, BRDF and Radiometric Balancing undertakings. There are two steps involved:

1. Extracting imagery from the camera.
2. Applying calibration (geometric and radiometric)

The following sections elaborate on these tasks.

3.1 Acquisition

One factor in choosing the Nikon D1 is its rugged construction. The camera has been mounted on a twin piston engine, Piper Navajo Chieftain airplane, based in Victoria, BC 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.

The camera is mounted so that the long dimension is along track. The acquisitions are executed to obtain 90% overlap in the along track direction, and 60% side lap. This ensures that each point on the ground will be in at least 20 different image frames with 20 different view directions.

3.2 Camera

The Nikon D1 camera lens has a nominal focal length of 14mm and a single focal plane, with the CCD R, G and B pixels interspersed in a standard Bayer pattern (Figure 3). The focal plane has 1324 rows by 2012 columns, half the pixels are green, one quarter are red and one quarter are blue. A thin film low-pass filter covers the entire focal plane, to reduce the aliasing effects within each band.

3.2.1 Image format/preparation: The camera is able to generate various image formats – jpeg, tiff and raw. The format used in the FIFEDOM project is the raw format, as this is simply a dump of the focal plane to a disk file, without any balancing or special effects being applied. The raw imagery allows us to do a stable radiometric calibration (that is valid for a fixed set of camera settings).

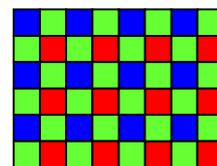


Figure 3: Bayer pattern of CCD on the D1 focal plane.

The raw image is in a proprietary format, which Nikon calls NEF, which is actually (non-standard) TIFF. The FIFEDOM project derives several imagery formats from the NEF file: Two of these are:

1. *Bilinearly Interpolated.* This is the workhorse format for generating DSMs and orthophoto mosaics. Each band is bilinearly resampled to the grid of the focal plane, resulting in the three bands being geometrically registered with each

other. The resulting imagery loses a 1-pixel border around the focal plane, and has dimensions 1322x2010.

2. *Separated RGB.* In order to provide the project with imagery that is unsullied by the bilinear interpolator, we are also able to extract the three bands from the NEF file, and place them into their own separate image files. No two bands are registered, but all the pixels are unaltered from the NEF file.

3.3 Calibration

An important goal of the FIFEDOM project is to command accurate placement and interpretation of the imagery. For this, the geometric and radiometric calibrations are essential. Another goal is that the calibration procedures be easily repeatable, so that it is convenient to

1. Verify the stability/accuracy of the calibration, and
2. Try different camera settings.

To a large extent, our geometric calibration procedures are easily repeatable. We are currently experimenting with various radiometric calibration approaches.

3.3.1 Geometric: The geometric calibration is divided into the two phases of camera interior and camera alignment calibration, which are discussed in more detail, below.

Camera Interior: The highly overlapped (low parallax) imagery allows tiepoints to be found completely automatically, simply with image correlation, and with no need of an Earth elevation model. We assume a standard distortion model for the camera (radial and tangential distortions), and use the tie points to constrain the distortion parameters. One caveat is that the focal plane tends to be parallel to the ground, so that a small number of GCPs are also used, to remove the focal length ambiguity. We have achieved 0.1 pixel accuracy in this way.

Camera Alignment: The GPS/INS describe the pose (position and orientation) of the GPS/INS frame, but not necessarily the pose of the camera. To this end, after the camera interior calibration has been applied, the tiepoints are used again, this time to calculate the pose of the camera relative to the GPS/INS frame. This increases the accuracy of the relative positioning of the frames in subsequent data sets, and improves the robustness and correctness of the tie point generation.

3.3.2 Radiometric

The goal of radiometric calibration is to calculate the radiance at the camera for each pixel of an image frame. This is one step in the chain of calculating the ground reflectance of each pixel, which is essential for BRDF investigations and for the absolute comparison of the acquired data across different data sets and even different sensors. The process can be divided into two parts, *relative* and *absolute*.

Relative calibration imposes a flat field on each band, i.e. the radiance values are known up to an (unknown) scale factor that is the same for every pixel of the band. Absolute calibration determines the gains (1 per band) that convert the relative calibration to true radiance units. We are experimenting with various approaches:

1. Calibration panels,
2. Overcast sky shots (with and without a diffusing opal glass filter), and

3. Integrating sphere.

In light of the criteria of correctness, convenience and cost, we are currently evaluating these approaches, but are, for now, using the diffuse sky shots to calculate the relative calibration, which is a set of gains for each band; as depicted in Figure 4.



Figure 4: Typical image of 1/gains for relative calibration

4. RADIOMETRIC BALANCING

There are two main goals for radiometric balancing in the FIFEDOM project:

1. Supply a single band of balanced imagery for DSM generation. The green band is used, as it has the highest spatial resolution of the 3 bands.
2. Produce seamless RGB mosaics that are pleasing to the eye, while maintaining, as much as possible, the proper spectral properties of the imagery.

The radiometric balancing approach is empirical, to remove residual errors from the 3 radiometric processes that precede the radiometric balancing step. These 3 processes are:

1. Radiometric calibration: Resulting in camera radiance.
2. Atmospheric correction: Resulting in ground reflectance for the pixel-specific sun and view angles, and
3. BRDF removal: Resulting in ground reflectance for a sun angle and a view angle that are the same across all pixels and images in the data.

Moreover, it is not always the case that all 3 of these processes will be applied to a data set. In such a situation, the radiometric balancing must still result in a balanced set of images, albeit not necessarily tied to engineering units.

The radiometric balancing approach is to lay a coarse grid on each image, with a gain/offset at each grid location. The tie points are used to constrain the gains and offsets, which are then applied to each image. At non-grid locations, the gain/offsets are obtained by bilinear interpolation from those at the grid locations.

5. BRDF

FIFEDOM data sets allow a user to view a given region on the ground from many different view angles. For example, for a typical FIFEDOM data set, there are about 26 distinct frames, which view a region from different zenith and azimuth angles. Thus, since the Nikon D1 is an RGB camera, this corresponds to an image with 78 “bands”. This large number of bands can be compared with hyperspectral images. Moreover, since the FIFEDOM data have 12-bit radiometric resolution, even small changes in view-angle dependent image brightness can be detected.

This great wealth of information about the view angle dependent reflectance of the ground can be fitted to a bi-directional reflectance distribution function (BRDF) model. Since this BRDF model for an object is characteristic of that object, it can serve as the basis for a scene classification.

FIFEDOM explores a wide range of the issues associated with BRDF-based land classification. For example, for forestry applications it is hoped that physical parameters of the forest canopy such as LAI, crown closure, tree geometry etc. can be derived from the FIFEDOM data.

6. 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 based on similarity measures and N-partite graph solution.
4. Multiple Forward Intersection – match points consisting of 2 or more points enables to reconstruct the 3D location of the object point on the terrain. Reconstruction of the object point and rejection of false points is achieved through a Least Square solution.

6.1 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 modelling 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 6.3), where the process is searching for the candidate homologous points within a search area along the

epipolar lines. This constraint is fundamental in reducing ambiguity and computational cost. Even if only approximate values for the parameters of relative orientation are known, the epipolar constraint should be used in order to restrict the search space in the direction perpendicular to the base line.

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 organised 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.

6.2 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 grey 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 great information reduction, as we only have to deal with a set of points and not with 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 6.3 for details).

Note that the interest operator should be carried out on each resolution level of the image pyramid separately, since points can vanish or be displaced from one level to the next due to the low pass filtering which is inherently present when decreasing the resolution. The lists of selected points and their associated attributes for every image are passed to the next module to establish the correspondence relationships between homologous points in overlapping images.

6.3 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 (Baltsavias 1991). The following two sections discuss the

overall concept of the multiple point matching problem and the potential solution to this problem.

6.3.1 Multiple Point Matching Problem: Figure 5 illustrates the multiple point matching problem. We have 5 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 have a set of candidate points which 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 3 candidate points. It is impossible to have a candidate point from the same image as point p_1 .

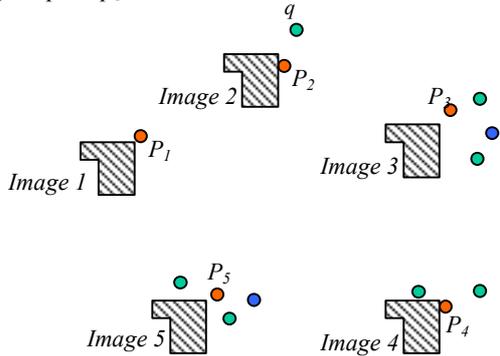


Figure 5: Multiple point matching problem

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?

6.3.2 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 .

Once we have generated the weights for every pair of candidates 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. The Figure 6 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 the Figure 6. All other lines are dashed. A similar approach is reported by Tsingas (1994) to solve the matching problem in the digital point transfer process.

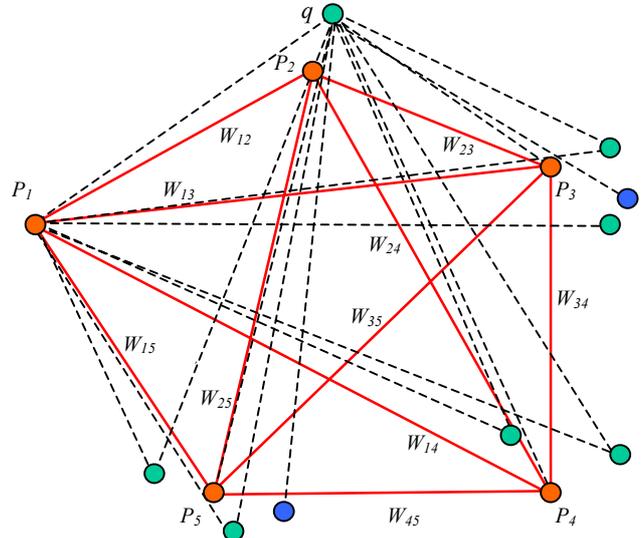


Figure 6: n-partite undirected weighted graph

6.4 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 7. 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.

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 modelling process, for the time being, is advised in order 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.

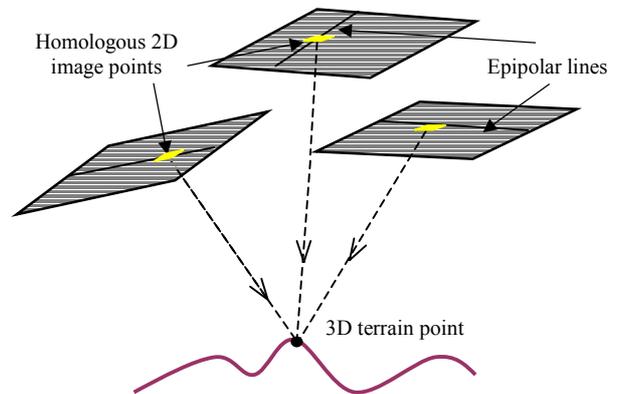


Figure 7: Multiple Forward Intersection

A more robust approach to terrain surface modelling with finite elements technique is reported by Ebner et al. (1980), where surface curvatures are introduced as constraints in order to regularize the model surface in areas where 3D measured points are missing.

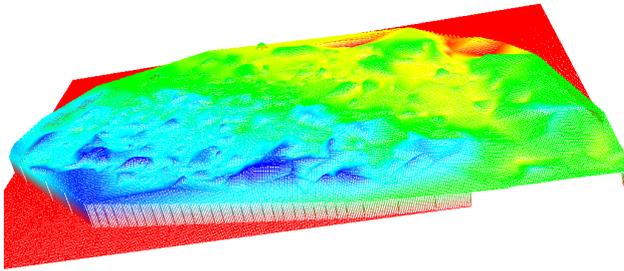


Figure 8: 3D perspective view of generated DSM

Figure 8 illustrates an example of the digital surface model generated based on the proposed method. The data set consists of 7 overlapping images acquired in 3 different flight lines, 1500m above the ground with 60% overlap in both cross and along track direction. It covers an area of 900mx700m on the terrain.

Due to the fact that a reliable ground truth information was not available at the time of writing this paper, we were unable to perform a true analysis on the quality of the resulting DSM. However, since the 3D positions 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 (currently) detect and remove the potential outliers from the result.

Planimetric threshold $t_p = 0.5 \times \text{GSD} = 0.5\text{m}$ Altimetric threshold $t_h = 1.5 \times t_p = 0.75\text{m}$ Total # of 3D points = 9822; Total # of outliers = 1305						
# of intersection rays	2	3	4	5	6	7
# of reconstructed 3D points	898	5580	2364	830	150	0
% of detected outliers	19.6	18.7	3.7	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 great wealth of information in highly overlapped image frames improves the resulting geo-spatial production significantly. The output of this component is in fact an input to other process such as Radiometric Balancing or an additional filtering process to generate the DEM product, which are outside the scope of this paper.

7. 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 modelling 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.

8. 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).
- Ebner, H., B. Hoffmann-Wellenhof, P. Reiß, F. Steidler, (1980), *HIFI* - A minicomputer program package for height interpolation by finite elements, *IAPRS* Vol. 23, part B4, pp. 202-215.
- 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.

9. ACKNOWLEDGEMENTS

The authors express their gratitude to Precarn Incorporated for partial funding support of this project. Also 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.