NEW AUTOMATED PROCEDURES
FOR CREATING LARGE SCALE DIGITAL ORTHOPHOTOGRAPHY IN URBAN AREAS

John A. Thorpe
Chairman and Chief Technical Officer

Wolfgang Schickler
Manager Advanced Technologies

Analytical Surveys, Inc., Colorado Springs, CO, USA
e-mail: wolfgang@anlt.com

Commision IV, Working Group 4

KEY WORDS: Automation, Digital Elevation Models, Orthoimages, Production

ABSTRACT

With the increasing demand for high-resolution ortho imagery of urban areas in the U.S. comes a need for faster and less expensive production methods. At ASI's Advanced Technologies Division (ATD) a completely new production system is being developed, using precise flight management systems, airborne GPS,an inertial measurement unit, automatic camera exposure, precision 60% sidelap photography, automatic film scanning, automatic interior orientation, automatic aerial triangulation, automatic DTM capture with minimal editing, orthophoto production and final image mosaicking into a seamless database. Although more powerful computers and extensive mass storage devices are required to handle the greatly increased computational load, the resulting decrease in total operator-hours and production schedule is considerable.

1 INTRODUCTION

1.1 Motivation

Automatic orthophoto production requires several components; a scanner and a plotter for digital input and output, a digital elevation model (DTM) and information about the interior and exterior orientation of the images to be rectified.

Up to now, most of the components need human interaction to be performed. The most time consuming parts are the aerial triangulation[1] and the data capturing for the DTM generation. Besides others, the automation of these two components are the major objectives of the current ATD research.

This paper documents the progress we have made so far. An overview of the whole system shows the combination of the individual components. First results of the development and test of some of the components are presented.

In this paper the results are given for 1:10,000 photography, a scale commonly used in U.S. for digital orthophotography of urban areas.

1.2 Overview

The objective when doing the image capturing is a GPS-INS guided flight mission resulting in very precisely positioned 60% lateral and 60% longitudinal overlapping imagery. Figure 1 shows this configuration.

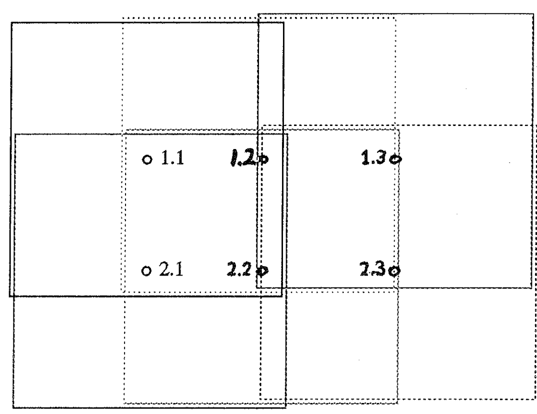

Figure 1: Configuration of 60% lateral and 60% longitudinal overlapping imagery

The main advantages of this flight configuration are as follows:

1. The Airborne GPS guided flight guarantees very precisely pre-planned exposure points. The processing of the GPS-INS navigation data captured during the flight results in very good approximate exterior orientation values for each image. These approximate values are accurate enough to directly apply an automatic aerial triangulation.
2. The 60/60 photography enables us to get a view from four different perspectives of the ground in between buildings, respectively four different perspectives of the ground. This is essential when doing automatic DTM generation in urban areas. This is a very promising approach in separating built-up objects from the ground.
3. With 60/60 photography, only the central 40% by 40% (plus a small overlap) is used for the final orthophotography. The resulting imagery is very high in quality due to minimal "building lean" (radial displacement) and excellent matching of radiometric values.
4. Note that although nearly twice as many images must be captured, the resulting savings from scanning directly the negative film (no diapositives are needed), automatic interior orientation, aerial triangulation and semi-automatic DTM capturing more than compensate

[1] in the case the density of available control points is not high enough to perform individual spatial resections for each image

for the additional flying cost.

The whole process of the automated digital orthophoto production process consists of the following nine steps:

1. **Pre-Flight planning**
   The flight plan is very carefully developed, with the required coordinates of each photograph pre-programmed and entered in on-board computer systems. As the aircraft flies along the desired flight path, the airborne GPS (AGPS) automatically triggers the camera when it is closest to the pre-selected coordinates. All relevant data for each exposure, including time, exposure number, preliminary GPS coordinates etc. are electronically printed on each negative.
2. **Photo mission**
   In a dramatic departure from conventional 20% or 30% sidelap photography, the photography is flown with 60% forward and side overlap, resulting in the photo centers falling at the corners of almost perfect squares.
3. **Scanning and image preprocessing**
   Instead of making diapositives or paper prints, we directly digitize the original negative film, using a **Vexcel** scanner. In future this will be done using a scanner with automatic film feed and batch processing capabilities. The whole image is scanned at a resolution of $15\mu m$, and an image pyramid for the further processes is calculated.
4. **Automatic interior orientation**
   The interior orientation is done fully automatically, using the software package AIO [cf. POTH Z., SCHICKLER W. 96], which has been developed by Bonn University in cooperation with Zeiss company. We have an agreement with Zeiss to use this software for our own internal purposes. The AIO automatically determines how the image was oriented in the scanner by locating an asymmetric feature. This has to be done only once per film roll scanning process. The accuracy of the AIO is a tenth of a pixel for a single fiducial measurement, and the sigma naught of the transformation estimation is about 0.2 to 0.3 pixels.
5. **Initialization**
   The initialization is necessary for the fully automatic aerial triangulation. In this part approximate tie point areas are found and the position of paneled control points in all the images are automatically measured.
6. **Automatic aerial triangulation**
   The aerial triangulation is done automatically using INPHO's MATCH-AT [KRZYSTEK P. 96] software package followed by a manual checking of the results and editing if necessary.
7. **Automatic DTM Generation**
   For the DTM generation, we developed an extension to INPHO's MATCH-T [KRZYSTEK ET AL. 91] software package. This enables us to produce high precision DTM's in urban areas while reducing the necessity of the manual or interactive part. Though this process works mainly automatically there is still a need for manual checking and possibly editing.
8. **DTM editing**
   After the DTM has been created automatically, it is very important to check that there are no gross errors. This must be manually done in a stereo mode, with tools available to modify the DTM where necessary.
   At ASI we have developed stereo viewing and editing techniques using our own proprietary production software.
9. **Digital differential rectification and mosaicking**
   At ASI we use proprietary ORBIT (Ortho Rectification By Image Transformation) software to make the final orthophotos. The only changes being made for the new production system are conversion to the Silicon Graphics platform and conversion to accommodate the new geo-Tiff image format which is used in our AIO, MATCH-AT, MATCH-T and the stereo editing software.

The steps which lead from the film roll to an automatically derived DTM (steps 3 to 7) are done mainly fully automatically in batch processing for a whole block of aerial images or a subset of a block. The most innovative parts are described in more detail in the following sections.

2 Initialization

ASI, in collaboration with INPHO GmbH of Stuttgart, has developed the initialization phase of MATCH-AT. This is the process needed to find the image coordinates of

a) surveyed control points, and
b) bundle adjustment of tie point areas,
   (these are approximate values in the sense of a search space, for the automatic aerial triangulation)

in every photograph on which they fall.

Both steps are done automatically, and will be described in the next subsections.

2.1 Finding tie point areas

This procedure is critical to the efficiency of the automatic aerial triangulation adjustment. There are three critical elements:

- Coordinates of the camera lens, obtained from AGPS to an accuracy of about 100 m.
- Orientation of the of the camera, obtained from a MOTIONPACK inertial measurement unit to an accuracy of about 0.5 degrees.
- An existing digital elevation model. In the U.S. we are fortunate in having the USGS DTM covering most of the country, to an accuracy of 7 m vertical.

To obtain the image coordinates of the tie points, the classical three von Gruber points of each photograph are projected down to the DTM, and then projected back to all the other photographs. After a check is made for proximity to other tie points, a selection is made automatically of the best final tie point position on the ground, and then points are re-projected into all the relevant images. As a control process, polygons containing water and dense forest (where these are available from existing maps), where projected tie points fall in the water or forest (or on steep slopes identified from the DTM) they are weighted accordingly after checking if they were slightly moved they could fall outside the polygons or on flatter ground.

When precise 60/60 photography is used, almost all the tie point fall on nine different images, so the final solution is extremely robust. When any of the three critical elements described above is missing, the best possible approximations are used, like digitizing the photo centers from the flight plan, assuming zero camera angles, and assuming flat ground.

2.2 Locating control points

Also part of the initialization is an automatic control point measurement, where this is possible.

When panels are used to mark the control points, a pattern matching technique can be used to locate the control point in the images. There are different approaches to solve this problem. With the availability of very good approximate orientation values for the exterior orientation, the search space for the control points in the image is relatively small. The above described approximate values result in a search area which is approximately 100 by 100 pixels. Also the approximate size of the panel in the image and the panel orientation relative to north up to 10^o is known.

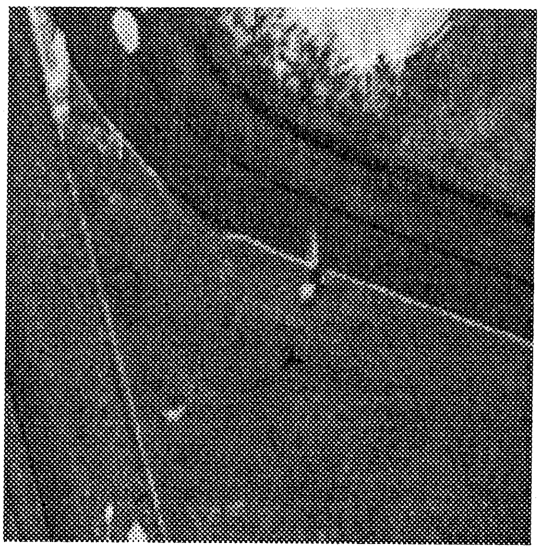

Figure 2: Image showing a paneled control point

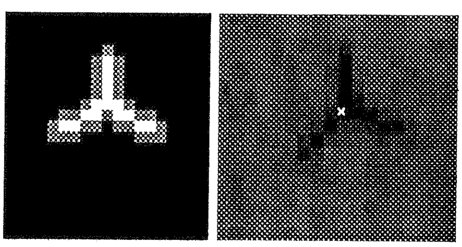

Figure 3: Pattern of the control point (left) and result of the automatic control point localization (right)

This enables us to use control point patterns and a cross correlation technique to estimate the precise position of the control point within an accuracy of a tenth of a pixel, which takes less than a second per control point on an SGI Indigo2. Figures 2 and 3 show the result of a control point localization.

In the case only natural control points, such as fence corners, small bushes or line intersections are available, the approximate values can be used to automatically pop up all the image sections where a certain control point falls. This speeds up the manual control point measurement dramatically, and also prevents misinterpretations.

3 Automatic Aerotriangulation

INPHO's MATCH-AT software is based on the same matching algorithm (FÖRSTNER W. 86; FÖRSTNER W., GÜLCH E. 87) as the MATCH-T automatic DTM correlation software. The initialization phase shows the computer where to look for homologous image features (feature points). Starting at a low resolution (usually the 6-th pyramid level), and working down in an iterative process to the highest resolution, planned in our case to be 30 μm. As the final sigma naught is of the order of one tenth of a pixel, or 3 μm, this resolution is quite sufficient for our needs. In fact the use of 60 μm scans may will be sufficient for creating digital orthophotos (but not contours), but this remains to be seen.

At each tie point location a cluster of 20 - 30 points are matched at each pyramid level, and the outliers are detected using a robust estimation technique similar to PAT-B.

The result of MATCH-AT is very accurate exterior orientation elements of each photograph.

4 Automatic DTM Generation

The task here is to generate automatically a high precision DTM in urban areas describing the topographical surface of the ground. Topographical surface, in this sense, means the real ground surface without objects like buildings and trees. We use the commercial software package MATCH-T [2] which we apply to six different stereo models from six different perspectives showing the same area on the ground, due to the special flight configuration.

In the next sections we first discuss the problems of the automatic DTM generation in urban areas, then a new approach for generating high precision DTM's in urban area is presented, and finally we introduce our procedure for generating a DTM automatically for a whole block of images.

4.1 DTM generation in urban areas, a challenge for each automatic system

The basic strategy of MATCH-T is described in KRZYSTEK 91. The principle is to use a hierarchical matching strategy to find homologous image features (feature points). Starting with a low resolution (top of the image pyramid), and a plane or other external prior information as an approximation of the ground surface, homologous feature points are used to compute their 3-D coordinates and from these, a refined DTM. This is done for each pyramid level using the DTM of the previous step as an approximation until the highest resolution (bottom of the image pyramid) is reached. MATCH-T uses a regularization (ref. TERZOPOULOS 86) technique when deriving the DTM from the matched raw 3-D points (further called a *3-D point cloud*). This technique is designed to treat measured points on trees or buildings as outliers and to eliminate them. This works well on small scale imagery, but not as well in large scales.

There are several reasons why using automatic DTM correlation directly is **not** appropriate for our purposes:

[2]Copyright: INPHO GmbH, Stuttgart

- When applying MATCH-T to large scale and high resolution imagery, the percentage of points matched on roof tops or trees is very high, resulting in a surface description which includes these objects, especially in very densely built-up areas.
- The regularization, even if using local adaptive techniques (cf. WEIDNER 94), has a tendency to smooth the surface and to round off sharp edges. This is a big disadvantage, especially when using this kind of surface description for building detection.
- The strategy of MATCH-T is based on matching only two images. Our approach is to use four images showing the same ground area. To our knowledge there is no commercial DTM generation package available that uses simultaneously more than two images, except for combining the final DTM's, which is not an appropriate approach to use the information.

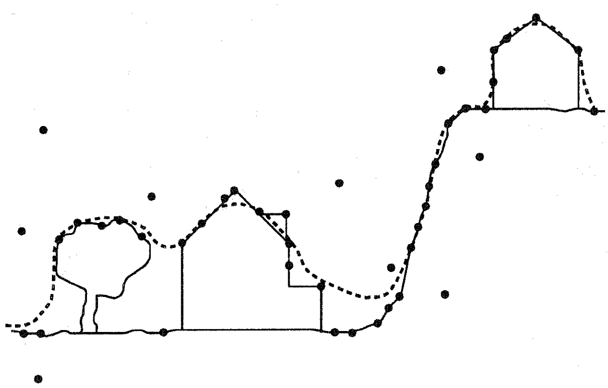

Figure 4: Basic principle of automatic DTM generation

Figure 4 shows the principle of deriving a DTM from a set of 3-D points computed by feature matching and the effect of the regularization. The solid line represents the real surface, the dots show the 3-D point cloud including some false matches, and the dashed line the surface description derived using a regularization technique.

4.2 Using morphologic filtering to determine the topographic surface

As shown in the last section, applying regularization techniques directly to the point cloud is not an appropriate approach to derive a description of the topographic ground surface.

Therefore, we first use a special filtering technique to eliminate points matched on buildings, and use the filtered 3-D point cloud to estimate the regular DTM grid. This filtering is applied to the combination of all six 3-D point clouds derived from imagery showing the same ground area. The basic idea behind the filtering is the fact that buildings or trees are higher than the surrounding topographic surface within a certain region.

A similar filtering technique is used by FÖRSTNER, WEIDNER 95 for purposes of building extraction from high resolution DTM's. They apply a morphological filter on a regular DTM grid to derive an approximation of the topographic surface without buildings. The difference in our approach is that we apply the filtering directly to an irregular set of 3-D points, and we added some robust techniques to prevent from falling into local minima due to outliers i.e. false matches.

The principle of the filter we use is from the field of mathematical grey scale morphology (cf. HARALICK ET AL. 87) called *opening*.

This generally is an *erosion*

$$\underline{z} = z \ominus w \tag{1}$$

followed by an *dilation*

$$\overline{\underline{z}} = \underline{z} \oplus w \tag{2}$$

both using the same structural element $w(x, y)$ having constant z values. The structural element generally may have an arbitrary shape depending on the application.

In order to apply this theory to the irregularly distributed 3-D point cloud $z(x, y)$, to eliminate points on top of buildings or trees, we first generate a regular grid $h(x, y)$ with a fixed grid size δ. One has to make sure that the structural element $\mathcal{W}$ is not entirely contained in a building outline. On the other hand, it should be small enough to keep small hills in the data set. We choose a squared window of the size of $200ft$. With that, the opening can be performed by simply placing the structural element on each grid point and applying a robust minimum filter

$$\underline{h} = rob\ inf\left\{ z(x,y)|x,y \in \mathcal{W} \right\} \tag{3}$$

followed by a maximum filter

$$\overline{\underline{h}} = sup\left\{ \underline{h}(x,y)|x,y \in \mathcal{W} \right\} \tag{4}$$

As the 3-D point cloud contains a certain percentage of outliers, it is not appropriate to use the absolute minimum for the erosion (minimum filtering). This would cause gross errors in $\overline{\underline{h}}$. Therefore, we use a generalization of the median, the so called *k-th sorted element* (k = 50% is equivalent to the median). The robust property of the generalized median is very useful to eliminate outliers i.e. points that are below the ground surface. From numerical investigations, we found k = 3% appropriate for the erosion.

The opening results in a regular grid $\overline{\underline{h}}(x, y)$, which is a discrete approximation of the topographic surface. Using this surface description, each point in the 3-D point cloud can be classified as being:

above the topographic surface,
(Mainly points on top of buildings and trees, but also some outliers)

on the topographic surface or

below the topographic surface
(Mainly outliers due to false matches)

by applying the following thresholding scheme

$$z(x,y) := \begin{cases} \textbf{above} & : \left(z(x,y) - \overline{\underline{h}}\right) & > & \Delta h \\ \textbf{on} & : \left\| z(x,y) - \overline{\underline{h}} \right\| & \leq & \Delta h \\ \textbf{below} & : \left(z(x,y) - \overline{\underline{h}}\right) & < & -\Delta h \end{cases} \tag{5}$$

Figure 5: Example of filtered 3-D point cloud superimposed on a stereo model. Black crosses mark points that are removed, white crosses are kept. Remember: The point cloud is generated using imagery from six different models, therefore some points may look erroneous.

The threshold parameter Δh can be derived from prior knowledge about the height of buildings taking the accuracy of the 3-D point cloud into account. We use $5ft$. The 3-D points classified as being **on** the topographic surface can now be used to estimate a high precision DTM, especially in urban areas. This represents the real topographic surface much better than a DTM estimated using all 3-D points from the point cloud directly.

Figure 5 illustrates an example of the above described filtering. The image scale was $1:10,000$, the image is scanned at $60\mu m$. Black crosses are classified as being **above** or **below**, and white crosses are classified as being **on** the topographic surface.

One can see that the 3-D points classified as being **above** are mostly on top of buildings or on trees. It is remarkable that the outline of buildings and even structures of roof constructions are recognizable in the classified 3-D point cloud. Fine structures of roofs may be recognizable in a higher resolution imagery, say $15\mu m$ at the same image scale. It seems to be very promising to use additional structural and radiometric information about buildings to separate points on trees from those on buildings and to extract the 3-D structure of the buildings using the 3-D point cloud.

4.3 DTM generation in a production environment

The above described technique has been embedded in a production environment, so that the DTM generation, works fully automatically in batch mode for a whole block of images, the number of images being limited only the available disk space. The whole procedure of automatically generating a DTM for a whole block of images consists of the following steps:

1. **Automatic model setup.**
   Within this first step the overlapping area of all the images is determined by projecting the image boundaries onto the ground using a coarse DTM. Each overlapping area, which fulfills a criterion based on the base length of the two images, is declared as a *model area* and the necessary project files for running MATCH-T are automatically generated.
2. **Generation of the 3-D point cloud using MATCH-T**
   For each model found in 1) the 3-D point cloud is generated in batch mode. Only the model area is processed which saves disk space and computation time. The result of this step is a 3-D point cloud for the whole project area which is, except for the project border, generated from four different perspectives, respectively six different model combinations including the diagonals (which only occurs in 60/60 geometry).
3. **Point cloud filtering**
   From the 3-D point cloud, covering the whole project area, so called *DTM units* are extracted and processed by the morphologic filter. The result of this step is a 3-D point cloud which ideally includes only those points on the ground, all points matched on top of trees or buildings are eliminated.

4. **DTM estimation**
We use a module of the MATCH-T software package, which is basically a robust estimation, to generate a regular DTM grid from the filtered point cloud. This is done for each of the overlapping DTM units separately. The final step is the averaging of the overlapping areas, which results in a DTM for the whole project area.

4.4 Results of first tests using the morphological filter

The above described automatic process for DTM generation has been tested on a data set which was flown with 60/60 geometry, in terms of run time, disk space usage and accuracy.

Table 1 summarizes the result of run times on each single step of the DTM generation for a DTM shape, 92 by 92 mm in plate system. Note: Components marked with * have do be done six times separately for each of the six different combination of four images. The times shown in the table are therefore six times a "normal" MATCH-T run.

Run time DTM processing per 6 x 1/4 image		
Resolution	$30\mu m$ [min:sec]	$60\mu m$ [min:sec]
Normalization *	14:12	3:42
Image Pyramid *	5:42	1:30
Feature pyramid *	8:12	1:54
Point cloud generation (matching) *	15:24	5:48
Point cloud filtering	2:55	2:43
Final DTM estimation	6:17	6:03
Total	52:42	21:40

Table 1:

Table 2 shows the disk space usage for MATCH-T run on a single DTM shape model, and for the final (all four perspectives combined) DTM generation. Most of the data has to be kept only temporarily. Only the final DTM and the normalized images are needed for the final DTM editing.

Disk space usage		
Resolution	$30\mu m$ $[MB]$	$60\mu m$ $[MB]$
DTM processing per 1/4 image		
DTM shape	18	5
Pyramid of the normalized DTM shape	23	6.6
Feature pyramid	10	2.6
DTM processing per 6 x 1/4 image		
Point cloud (MATCH-T)	9	3.8
Final DTM (binary, 24 ft grid)	0.1	0.1

Table 2:

Table 3 summarizes the accuracy of different DTM's and the 3-D point cloud. The reference for this investigation was a conventionally measured surface description that represents the "real ground", namely a dense set of 3-D mass points and break lines which were originally measured for accurate contour interpolation. The table shows the RMS value $\left(\sqrt{\sum(z_i - z_j)^2/(n-1)}\right)$ of the different data compared to this reference. The difference in the RMS of the DTM computed with and without the morphological filtering, confirms the necessity of the filtering, especially in urban areas.

Accuracy	RMS $[ft]$
USGS DTM (level 1)	14.9
MATCH-T DTM without filtering ($60\mu m$)	10.7
Filtered 3-D point cloud ($60\mu m$)	3.4
Filtered 3-D point cloud ($30\mu m$)	3.7
Final DTM using filtered 3-D point cloud ($60\mu m$)	3.4

Table 3:

5 Final Conclusions of first results

At the time of writing the system had not been fully run from beginning to end. However, the preliminary results of all the different phases, particularly the AIO, the automatic measurement of paneled control points, the filtering of the 3-D point cloud, and the stereo editing system, have all been extremely gratifying. The system is expected to go into limited production this summer.

REFERENCES

[FÖRSTNER W., GÜLCH E. 87] A Fast Operator for Detecting and Precise Location of Distinct Points, Corners and Centers of Circular Features. Proceedings ISPRS Intercommission Workshop, Interlaken 1987, pp. 281-305

[FÖRSTNER W. 89] A Feature Based Correspondence Algorithm for Image Matching and Least Squares Matching. International Archives of Photogrammetry and Remote Sensing, Vol. 26, Part 3, Rovaniemi, 1989

[FÖRSTNER W., WEIDNER U. 95] Towards Automatic Building Extraction from High Resolution Digital Elevation Models. ISPRS Journal 50(4) pp. 38-49, 1995

[HARALICK R.M., STERNBERG S.R., ZHUANG, X. 87] Image Analysis Using Mathematical Morphology. *IEEE T-PAMY*, Vol 9, pp. 523 - 550, 1987.

[KRZYSTEK P. 91] Fully Automatic Measurement of Digital Elevation Models. In: *Proceedings of the 43rd Photogrammetric Week, Stuttgart*, pp 203 -214, 1991

[KRZYSTEK P., HEUCHEL T, HIRT U. 96] An Integral Approach To Automatic Aerial Triangulation and Automatic DEM Generation. XVIII ISPRS, Commission III, Vienna, 1996

[POTH Z., SCHICKLER W. 96] The Automatic Interior Orientation and its Daily use, XVIII ISPRS, Commission III, Vienna, 1996

[WEIDNER U. 94] Parameterfree Information-Preserving Surface Rectoration. In: Eklundh J.-O. (Ed.), *Computer Vision -ECCV 94, Vol.11 Proceedings*, pp. 218 - 224, 1994