Automatic Processes in Photogrammetry

Tamas Jancso

Institute of Geoinformatics, Alba Regia Technical Faculty, Óbuda University
Pirosalma u. 1-3, Székesfehérvár, H-8000, Hungary
jancso.tamas@amk.uni-obuda.hu

Abstract— **Digital Photogrammetric Workstations (DPWs) and especially some software packages serving the processing of UAV images offer automatic algorithms to accelerate the gaining of end products like digital terrain or surface models or orthophotos. To achieve the final steps in the image evaluation, the orientation process, particularly the exterior orientation requires manual measurements of control points. The measurements of tie points necessary for building models from stereo image pairs or finding tie points for strips of images needed for aerial triangulation require automatic image matching. These algorithms are based mainly on finding the same terrain points, lines or other objects on different images using the area-based image matching, semi-global image matching or other object-based image matching techniques. The paper summarizes the recent status and the perspectives of these image processing solutions.**

Keywords— **DPW, UAV, photogrammetry, image matching**

I. Introduction

In digital photogrammetry the orientation and the evaluation process are supported by automatic processes. Fig. 1 summarizes the possible steps for automation.

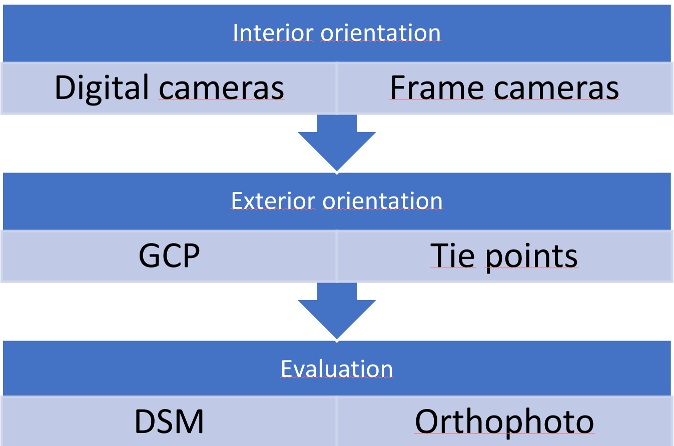


Figure 1. Possible steps for automation

During the interior orientation we build the geometric model of the camera including the distortion model of the projected image. For digital cameras, the interior orientation is an easy step since we need to input only the sensor data including the calibrating focal length, the coordinates of the principle point, the pixel size, the image size (width and height n pixels) and the parameters describing the image distortion.

For frame cameras, this step is more complicated since the photo coordinate system should be determined by the fiducial marks. The measurement of fiducial marks can be automated based on an area-based image matching technique.

The goal of the exterior orientation is the determination of the coordinates of the exposure centre and the calculation of the rotation angles of the camera relatively to the ground coordinate system for each photo. If the IMU/INS (Inertial Navigation System) together with the GNSS equipment is capable to measure the exterior orientation elements with high accuracy during the flight then the exterior orientation doesn't need ground control points (GCPs), which means the direct georeferencing.

The image matching algorithms are developing very fast and they become so sophisticated that the DSM (Digital Surface Model) and orthophoto production can be fully automatic.

All these aspects helped the spread of UAV (Unmanned Aerial Vehicle) technology as well.

II. Flight Logs and Metadata

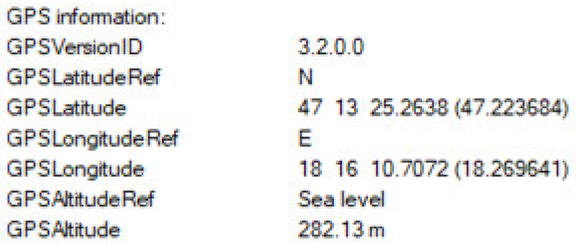
The IMU together with the GNSS equipment logs the coordinates of the exposure centre and the yaw, pitch, roll rotation angles. The GNSS data can be used directly in the exterior orientation process, but the yaw, pitch, roll angles should be converted to phi, omega and kappa angles used in photogrammetry. The main difference between the two sets of angles is:

- Yaw, pitch, roll angles define the rotation of the camera body with respect to its navigation coordinate system.
- Omega, phi, kappa angles define the rotation between the image coordinate system and a ground coordinate system used in photogrammetry.

The conversion is more less straightforward, but we should consider some aspects [1]:

- the meridian convergence,
- the interpretations of angular range and sign of rotations,
- misalignments ex, ey, ez between INS and camera.

Each image contains several metadata including the time of exposure, the focal length, the aperture value, the GPS coordinates, the yaw, pitch, roll angles. These data can be used as approximate values for the precise interior and exterior orientation. Fig 2 shows a fragment of image metadata captured by an UAV camera.



GPS information:	
GPSVersionID	3.2.0.0
GPSLatitudeRef	N
GPSLatitude	47 13 25.2638 (47.223684)
GPSLongitudeRef	E
GPSLongitude	18 16 10.7072 (18.269641)
GPSAltitudeRef	Sea level
GPSAltitude	282.13 m

Figure 2. GPS information block in image metadata

Based on the calculated S_{avg} and pixel size Δ_p, we can get the GSD value, in other words the ground resolution by the following equation:

III. Orientation

The whole orientation process is divided into two steps: interior orientation and exterior orientation.

The task of the interior orientation is well defined and can be automatized. In case of frame cameras., according to the method of least squares, we try to solve the task of interior orientation with adjustment, like all orientation tasks, by introducing more measurements than minimally necessary. In case of affine transformation, we need to measure at least four fiducial marks. The system of equations based on equations (1) is linear for the unknowns and can be solved directly [3], [4].

$$\begin{aligned} X &= a_0 + a_1 x + a_2 y \\ Y &= b_0 + b_1 x + b_2 y \end{aligned} \qquad (1)$$

Notation in (1):
X, Y : photo coordinates
x, y : image (pixel) coordinates
$a_0, a_1, a_2, b_0, b_1, b_2$: transformation parameters

The measurement of fiducial marks can be automatic with the help of the cross correlation between the stored image of the fiducial mark as a pattern and the image segment containing the fiducial mark. This area-based image matching is useful not only for the automatic measurement of fiducial marks, but it can be used also for the measurement of common points (tie points, ground control points for the exterior orientation and DSM points for orthophoto production).

A. *Area-based image matching*

Most programs apply the raster correlation pattern matrix and use the cross-correlation formula (2) [3].

$$\rho = \frac{\sum_{r=1}^{R}\sum_{c=1}^{C}(g_1(r,c)-\mu_1)(g_2(r,c)-\mu_2)}{\sqrt{\sum_{r=1}^{R}\sum_{c=1}^{C}(g_1(r,c)-\mu_1)^2 \sum_{r=1}^{R}\sum_{c=1}^{C}(g_2(r,c)-\mu_2)^2}} \qquad (2)$$

Where:
g_1 - the greyscale value of the pixel in the target area
g_2 - The greyscale value of the pixel in the search area
r, c - row, column index
μ_1, μ_2 - Average grey values in the target and search areas
R, C - number of rows and columns in the sample area

The value of the correlation coefficient ρ can vary from -1 to 1, where 1 represents the maximum fitting. If the value of the correlation coefficient reaches a predetermined threshold (typically 0.7), then the fitting is accepted.

Point measurement with autocorrelation does not give a 100% reliable result, so if it is possible and supported by the evaluation program. It is worthwhile to create a correlation image, where the weakly correlated places can be seen, and we know the areas where more ex-post control may be needed. The other main source of problems is the automated evaluation of homogeneous areas (water, arable land, forest, sandy areas, large roads, airports). The correlation procedure is ambiguous at these locations, and it can result errors in elevation data. The masking procedure is commonly used to prevent errors of this nature, i.e., these areas are excluded from the correlation procedure. This method is effective when there are contiguous areas and their height can be specified after a small number of height values. Another way to reduce this type of error is to incorporate subsequent filtering and smoothing procedures.

In general, it can be stated that the subsequent manual verification of the measurements made in this way and the performance of additional measurements are inevitable. When producing terrain models, the result can be dramatically improved if we already have altitude values to describe important terrain details (e.g. mountain peak, road network, embankments) before performing the correlation process.

The method outlined here can be used

- for automated measurement of fiducial marks,
- for measurement of Gruber points for relative orientation,
- for measurement of any field points,
- for automated measurement of DSM points.

B. *Feature-based image matching*

In the area-based image matching we search for the correspondence of common points. We can extend this approach if we are looking for the matching of features like points, lines, curves, or polygons where each feature is described also by a feature descriptor. In this case the matching procedure is done only after the identification of features and feature descriptors in each image. The separation of features and feature descriptors is done by different image processing techniques. Commonly used feature detectors are [4]:

- Förstner operator,
- Sobel operator,
- Laplace operator,
- SUSAN (smallest univalue segment assimilating nucleus) operator,
- FAST (features from accelerated segment test) operator,

- SIFT (scale invariant feature transform) operator,
- Affine invariant SIFT operator,
- SURF (Speed-Up Robust Features) operator.

C. Semi-global image matching

The autocorrelation method described in Section III.*A* finds the greatest fit between image details by comparing square or rectangular areas when searching for related points. Such a rectangular area behaves like a low-pass filter, blurring the sharp boundaries of depth differences, resulting in an edge-free elevation model. The smaller the image matrix used for pattern matching, the smaller the blurring of the edges. Ideally, the search matrix would be the size of one pixel, but this is not a feasible solution due to ambiguity. A pixel alone carries too little information to decide a clear fit. The resulting altitude model will not be consistent, so additional information would be needed, e.g. to constraint of "smoothness". Another problem with cross-correlation based on pattern fitting is that we assume there are no height differences within the correlation pattern matrix, or at least only to a negligible extent. In this case, the pattern matching will fail or leading to the determination of an incorrect height value in the worst case.

In general, when we try to fit the left image $I_1$to the right image I_2 with maximum reliability, we are actually looking for the minimum of the energy function (3) [2]:

$$E(D) = \sum_p \left(C(p, d_p) + \sum_{q \in N_p} P\left[\left| d_p - d_q \right| \geq 1 \right] \right) \quad (3)$$

The first term of the function (3) sums the cost C of fitting for the given pixel p, where d_p represents the parallax between the images I_1 and I_2. Because the fitting is done pixel by pixel, only the intensity values $I_1(p)$ and $I_2(q)$ are compared to calculate the cost. In a simpler solution, the absolute difference in intensity values is considered as a cost. In the formula, on the other hand, the second term represents a kind of penalty P for those pixels q where, taking into account the neighborhood N_p of the pixel p, the parallax difference is greater than one pixel. That is, the term of penalty P can result in mismatch even if the difference in intensity values between $I_1(p)$ and $I_2(q)$ is zero, but the parallax is too large relative to the surrounding pixels. This also means that this concept may give preference to point pairs where the difference in intensity values may be larger, but the parallax difference does not exceed 1 pixel. The resulting d_p parallaxes are stored separately in a matrix D with equal extent to the extent of the I_1 image, which can be displayed as a depth image with appropriate coding. In the depth image, different shades represent different parallaxes or depth ranges derived from them. The member P, which penalizes a sudden change in height, is bound to the pixel p in the formula, which in turn is related to the neighbouring pixels q, which are also related to their own neighbors. As a result, each pixel is related to each pixel while looking for the minimum of the function $E(D)$. This method is called global image matching, which is difficult to implement in practice. The main reason is that the computation time increases exponentially as the image size increases and above a certain image size, the practical application of the method does not pay off.

The semi-global image matching (SGM) method is not based on the whole image to reduce the computation time, but only on a path taken in 4 or 8 directions around the pixel p. Paths are chosen radially and symmetrically around the pixel p (Fig. 3).

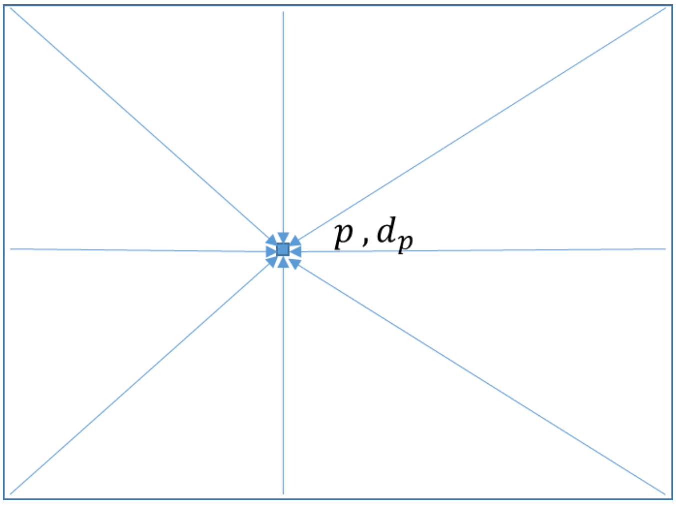


Figure 3. 8-way path to select the optimal d_p parallax

The cost calculation is performed separately for each path according to the energy function (3) and the parallax d_p belonging to the lowest cost is stored, from which the complete depth image can be generated (Fig. 4).

Figure 4. Applying SGM on a pair of images (source: http://lunokhod.org)

With this method, we can only process stereo image pairs. Depth maps from multiple stereo image pairs, on the other hand, can be integrated into a common model.

IV. DSM AND ORTHOPHOTOS

DSM data acquisition from a stereo image pair can be further automated during digital evaluation by various stereo-correlation methods, which automate stereoscopic orientation and thus allow the extraction of ground point coordinates.

A common feature of correlation methods is that the search for homologous point pairs is done using image patterns (correlation image matrices). The image patterns can be compared in the original image coordinate system or - after generating the normal stereogram from the original image pair - in the normal-case images, and finally the comparison can be realized in the raster space of the orthophoto, in which case the orthophoto is generated. In addition to the correlation image matrix, the comparison can be done independently or in addition to shape and topology recognition. Then the alignment of the related points in the image (line, polyline) is preceded by the selection according to the interest operators, which further increases the robustness of the correlation.

V. Conclusions

As a summary we can say that a fully automatic evaluation of stereoscopic images pairs or block of images is not possible in every case. The image matching techniques are working well only in those cases if the following conditions are fulfilled [4]:

- GCPs are marked with same signs (size, colour, form is unique and same for every GCP).
- There are no occlusions in the corresponding images and the common points are visible.
- There are no ambiguous object structures or transparent surfaces, which can result several candidates for common points or features.
- For regions with poor texture, the solution is stable or non-sensitive with respect to minor disturbances in the image (noise).
- Intensities in all images cover the same spectral regions.
- Constant illumination, atmospheric effects, and media interfaces for the period of image acquisition.
- Stable object surface over the period of image acquisition.
- Macroscopically smooth object surface.
- Opaque object surface.
- Largely diffuse reflection off the surface.
- Known approximate values for orientation data (image overlap) and object data (geometric and radiometric parameters).

We can meet some conditions in the above list but not all ones. Therefore, the fully automatic orientation and evaluation of images remains a task. On the other hand, we can see a very progressive development in this area thanks to the development of UAV software solutions where the demand and the requirement for urge solutions for fully automation are coming from the end users.

References

[1] M. Bäumker and F.J. Heimes: New Calibration and Computing Method for Direct Georeferencing of Image and Scanner Data Using the Position and Angular Data of an Hybrid Inertial Navigation System, Proceedings of OEEPE Workshop on Integrated Sensor Orientation, 2002

[2] Hirschmüller H.: Semi-Global Matching – Motivation, Developments and Applications, in: Dieter Fritsch (Ed.): Photogrammetric Week'11, Stuttgart, Wichmann Verlag, Berlin, ISBN 978-3-87907-507-2, pp. 173.183., 2011

[3] Kraus K.: Photogrammetry, Geometry from Images and Laser Scans, Walter de Gruyter GmbH & Co. KG, Berlin, 2. kiadás, ISBN 978-3-11-019007-6, 2007

[4] T. Luhmann, S. Robson, S. Kyle and I. Harley: Close Range Photogrammetry, Whittles Publishing, ISBN 1-870325-50-8, pp. 284., 2006