Automatic Digital Surface Model (DSM) generation procedure from images acquired by Unmanned Aerial Systems (UASs)

Andrea LINGUA, Prof., Politecnico di Torino, Italy, andrea.lingua@polito.it
Davide MARENCHINO, PhD, Politecnico di Torino, Italy, davide.marenchino@polito.it
Francesco NEX, PhD student, Politecnico di Torino, Italy, francesco.nex@polito.it

Abstract: In the last few years the use of Unmanned Aerial Systems (UAS) has became an alternative way of performing photogrammetric flights. These instruments are less expensive and safer than traditional flights in several applications such as in circumscribed area surveys and in emergency management. Nevertheless, their image acquisition is still far from that of flights performed by manned planes: their dimensions and their lack of weight never allow them to fly over the previously set course and, consequently, their images are often affected by large rotations and small overlaps. In these conditions, traditional algorithms (already implemented in commercial software) are not able to orientate images or generate reliable Digital Generation Models (DSM). For these reasons, the research has focused on improving performances of these two topics in UAS applications. The procedure for the DSM generation set up by the research group at the Politecnico di Torino is proposed in this paper. This approach takes advantage of both Computer Vision and multi-image matching algorithms for the extraction of points and edges from images and their matching for a dense DSM generation. It will be shown how this procedure allows reliable results to be reached in terms of extracted points, location accuracy and quickness in the results production in difficult geometric taking conditions.

Keywords: DSM extraction, multi-image matching, filtering, UAS, ortho-photo

1. Introduction

The word “UAS” (Unmanned Aerial Systems) is well known in the research community, and in the last few years it has also been used more and more in the ordinary language. Scientific reviews, newspapers and magazines have published many articles on the potentiality and the applications of these systems; and many different types of UASs exist, each with different capacities that respond to different user needs. The Geomatics research group at the Politecnico di Torino, in cooperation with ITHACA (Information Technology for Humanitarian Assistance, Cooperation and Action) and the Department of Aerospace Engineering (DIASP) at the Politecnico di Torino, has developed a mini-UAS devoted to emergency management of environmental disasters. Satellite images are usually unable to quickly acquire data relative to a catastrophic event, therefore, in situ missions have to be foreseen. For this purpose a UAS, called “Pelican”, has been developed. “Pelican” is a low-cost mini-UAS equipped with photogrammetric sensors which is capable of autonomous navigation (GPS/IMU) and automatic digital image acquisition (characterized by a suitable geometric and radiometric quality). The aeronautical specifications of the platform, the availability of a navigation system for autonomous flight (autopilot with GPS/IMU), and the features of the payload installed onboard (RICOH GR, high resolution digital camera), allow this system to be employed for photogrammetric surveys. The UAS can follow flight courses in a completely automatic way (excluding take-off and landing), according to the flight plan specifications. Therefore, the “Pelican” UAS is able to perform photogrammetric surveys in remote and disaster-affected areas [1], where it is not possible to carry out traditional photogrammetric flights. Nevertheless, the performed flights are far from traditional photogrammetric flights; the small dimensions of the UAS and the reduced weight do not allow images to be obtained near to
normal case. For this reason, the commercial solution for image orientation, DSM extraction and orthophoto production were not sufficient to achieve reliable results. A new ad hoc process was necessary to furnish maps in all operative conditions in a quick way.

In the following sections, the basic algorithm that was set up by the Research Group at the Politecnico di Torino in order to solve this task will be described. This approach takes advantage of both the Computer Vision and the Photogrammetric application field. Computer Vision algorithms were adapted in the image orientation procedure and multi-image matching techniques were used in the DSM generation.

The developed algorithm is an ongoing process that still has to be improved. In the following, one of the first tests and the validation results will be shown in detail.

2. Algorithm workflow

Over the last 15 years, photogrammetric techniques have significantly improved. The availability of digital images and the swift development of computer science have allowed the photogrammetric process to be increased. Several commercial software programmes have been developed by photogrammetry industries. Nevertheless, these solutions only achieve reliable results when traditional flights and normal image acquisitions have to be performed. For this reason, it was necessary to produce an ad hoc solution for UAS application. The research developed by the Politecnico di Torino Geomatics Research Group, in this application field, has involved a photogrammetric process that can be divided into several steps, as shown in Fig. 1:

- **Image acquisition.** The “Pelican” UAS carries out image acquisition in an automatic way. The operator designs the flight plan and uploads it into a Ground Control Station (GCS), which remotely controls the aerial platform. Therefore, it is possible to perform flights with a programmed image acquisition, which can be monitored in real time by an operator through the GCS. At the end of the flight, the images and the position/attitude data are downloaded, in order to be processed in the following steps.

- **Image pre-processing.** The image pre-processing techniques allow the image quality to be enhanced when it has been taken in non-ideal conditions. In photogrammetry, these algorithms are frequently used to enhance the image quality, to reduce the noise effects and to increase the image contrast. A Wallis filter [2] was implemented in the studied algorithm. It can, in fact, work both as a low pass and a high pass filter. The goal of the Wallis filter is to force the mean and the standard deviation of an image to fit some given values [3]. It is widely used to enhance image
A. Lingua, D. Marenchino, F. Nex
Automatic Digital Surface Model (DSM) generation procedure from images acquired by Unmanned Aerial Systems (UASs)

Texture patterns, and to provide good radiometric information for point and region detectors. In the proposed approach, the Wallis filter algorithm was tested on aerial images and an optimized set of parameters for the image enhancement was determined in order to achieve the most stable result.

- **Image orientation.** A good orientation allows 3D points to be matched in the object space and to perform the subsequent automatic algorithms. The mathematical computation of the external orientation of a set of images is performed through a Bundle Block Adjustment (BBA). This computation is made possible through the integration of image information (automatic tie-point extraction), the ground control points (achieved by a survey), and position/attitude information (furnished by autopilot GPS/IMU systems).

  The *automatic tie-point extraction* is performed using the SIFT (Scale Invariant Feature Transform) operator. This is a region detector(descriptor [4] which allows us to extract features that are invariant to image scaling and rotation and partially invariant to changes in illumination and 3D camera viewpoints (affine transformation). A vector of dimension 128, called descriptor is associated to each feature. This allows a feature matching between image pairs, which satisfies the geometrical previously mentioned transformation. The SIFT operator allows a great number of points to be matched, in any taking condition. Nevertheless, a robust LMS relative orientation must be performed in order to detect mismatches. The extracted points are useful in the BBA as tie-points and they furnish an approximate DSM in the multi-image matching algorithm, as will be explained in the following.

  The autopilot MP2128g installed on the UAS allows information of the *position and the attitude of the plane* to be obtained during the flight. Several tests have been performed in order to evaluate the suitability of this system in a Direct Geo-referencing; the achieved results have shown that this instrument does not supply sufficiently accurate attitude data. Nevertheless, these row exterior orientation parameters can be employed as approximate values in the BBA. For this reason, a survey is still required in order to determine the ground control points and, then, to collimate them on the images.

  Until now, the BBA has been performed using the commercial and academic software that are already available. The experimental tests have been carried out with the Leica LPS software. An automatic tool for the importing of the tie-points extracted by SIFT into Leica LPS software and for the production of the approximate DSM, using these tie-points and the external orientation parameters, has been created.

- **DSM extraction.** The DSM extraction is made possible through the matching of points in two or more images. A multi-image matching approach, called Multi-Image Geometrically Constraint Cross-Correlation (MIGC3), has been implemented. This approach is an area-based matching algorithm, founded on the Cross-Correlation (CC) technique. The MIGC3 procedure is based on the concept of multi-image matching guided from the object space, therefore any number of images can be matched simultaneously and the epipolar constraints is integrated implicitly. Together with an adaptive determination of the correlation parameter, it has the ability to reduce the problems caused by surface discontinuities, occlusions and repetitive structures, and to produce dense and reliable point matching results [5]. The method works as follows: one of the stereoscopic distortion-free image (all the images are previously undistorted) is chosen as the reference image, and the others serve as the search images. The feature extraction is carried out in the reference image with the Forstner operator [6]. In this way, a great number of points is extracted over the entire image. The image matching procedure of each point of interest is then performed. An image point \(p_0\) is projected onto the object space, using the collinearity equations (Fig. ). The approximate DSM obtained by the SIFT operator feature extractor allows a first 3D position of the point \(P_0\) to be defined. Since the approximate height \(Z_0\) may not be accurate, the correct position of \(p_0\) in the object space should lie between \(P_{\text{max}}\) and \(P_{\text{min}}\), with height values of \(Z_0-\Delta Z\) and \(Z_0+\Delta Z\) respectively, along the image ray \(cp_0P_0\). If these two points are back-projected onto the search images, the
bounds of the epipolar lines are defined (yellow line in Fig. 2). Therefore, the research of the corresponding image points \( p_1, p_2 \) is restricted to the image points which lie along these lines.

Fig. 2. Image matching with the MIGC3 technique. Geometric constraints and window warping procedure.

After the application of the geometric constraints, the research of the homologous points along the epipolar segment is performed using the similarity constraint. In order to define the homologous points, the NCC (Normalized Cross-Correlation) value is computed between the image windows and the other image. Compared to the traditional Cross-Correlation method, the NCC function implemented in this approach is computed with respect to the height value in the object space. The NCC value between two corresponding windows in the reference image \( I_0 \) and in the \( i^{th} \) search image \( I_i \) is definite as:

\[
NCC_i(p_0, Z) = \frac{\sum_{x,y \in W} (I_0(x_0,y_0) - \bar{I}_0)(I_i(x(Z),y(Z)) - \bar{I}_i)}{\sqrt{\sum_{x,y \in W} (I_0(x_0,y_0) - \bar{I}_0)^2} \cdot \sqrt{\sum_{x,y \in W} (I_i(x(Z),y(Z)) - \bar{I}_i)^2}}
\]

where, \( W \) defines the correlation window around point \( p_0 \) in the reference and search images; furthermore, \( I_0(x_0,y_0) \) represents the intensity value of each pixel within the reference image. \( I_0(x(Z), y(Z)) \) denotes the radiometric values of the corresponding pixels in the search image \( I \) at a determined height \( Z \), which varies in the range \([Z_0 - \Delta Z, Z_0 + \Delta Z]\).

Unlike the traditional NCC technique, the correlation windows of the search images are not defined as a constant square shape, but they are computed as a back-projection of the reference image window from the object space to each search image. This process is named correlation window warping procedure and it exploits the approximate DSM and the image orientation parameters. Each point in the reference image window is projected onto the approximate DSM, and a piece of surface patch in the object space is therefore obtained. Then, by means of a back-projection, the corresponding image windows in the search images are determined. Consequently, a square correlation window in the reference image may appear as a general quadrangle window in the search image (red quadrangles in Figure 2).

The NCC functions of all individual stereo pairs are defined in a unique framework, which can be mathematically expressed by the SNCC function:
The correct match is not evaluated by the single NCC function analysis, but through a unique function that integrates the similarity constraints supplied by each stereo pair. The basis algorithm has been improved by an adaptive approach that allow us to automatically set some important parameters that optimize the MIGC3 performances and reduce the ambiguity problems, such as the repetitive texture patterns, the occlusions and the surface discontinuities.

The results achieved by the MIGC3 algorithm still have some gross errors. The elimination of these blunders is carried out through a robust filtering algorithm called the Self-tuning Standard deviation Median Filter (S2MF) which has been implemented by the authors. S2MF is a distance-function filter method, which quickly processes 3D point clouds extracted using photogrammetric procedures. It produces filtered DSM with a nearly negligible rate of residual errors. The algorithm is based on the median filter, which is widely used for the elimination of outliers in DDSMs (Dense Digital Surface Models) produced by Lidar techniques [6]. In the case of DSMs extracted by automatic photogrammetric procedures, the lack of dense 3D point clouds causes the failure of the median filter; therefore, the development of an auto adaptive approach, in relation to the density of the point cloud, is required. S2MF locates outliers in several steps.

A first global filter, which deletes the outliers extracted by GC3 with a huge altimetric error, has been realized. The outliers are easily removed with a median filter applied to the whole dataset. The points that do not respect the distance function (that is to say that are outside the interval value) are classified as gross errors and are removed.

\[ \text{med}(Z_k) - 4\sigma(Z_k, \text{med}) \leq Z_k \leq \text{med}(Z_k) + 4\sigma(Z_k, \text{med}) \]  

where \( \text{med}(Z_k) \) defines the median operator applied to the set of \( Z_k \) values, while \( \sigma(Z_k, \text{med}) \) is the standard deviation of the same values, computed using the median instead of the mean value. Factor 4 was experimentally defined.

After the elimination of the largest errors, a second filter operation is carried out. The planimetric DSM area is split into \( n_b \) bins and each point is classified, according to its position, in the respective bin, which is denoted by a row and column index \((i,j)\). The bin size is a fundamental parameter which affects the performances of the filter. On one hand, small bins allow a local analysis of the object to be performed, and increase the reliability of the method; however, it is likely that some bins do not have enough points for a statistical analysis. On the other hand, if bins of large dimensions are used, each of them have hundreds of points but an ineffective smoothing can be achieved. For this reason, the square bin size \( a_b \) is defined in relation to the density of the points in the object space:

\[ a_b = w_f \cdot \left( \frac{W_x \cdot W_y}{m} \cdot \text{GSD} \right) \]  

where \( W_x \) and \( W_y \) are the analysed area dimensions [pixel] on the reference image in \( x, y \) directions; \( m \) is the number of points extracted by Forstner; \( w_f \) is the window side factor (defined by the user), which denotes how many times the bin side is higher than the medium linear density (pixel/m) of the point cloud. The median, \( \text{med}_{i,j} \) and the standard deviation, \( \sigma_{i,j} \) of the heights are computed for each bin. The median is a robust estimator of the mean value of a random variable; therefore, it is not sensitive to the presence of outliers in the data set. The standard deviation of the data, which is computed with respect to the median, is instead sensitive to the measures affected by...
gross errors. Thus, the distance function, which compares the elevations of each point with the altimetry trend of the surrounding bins, is also sensitive to the gross errors. In order to avoid these problems, a self-tuning approach for the computation of the standard deviation of each bin has been developed. Given a bin with \( s \) points, the standard deviation \( \sigma_i \) of the whole data set of z-values with respect to the median value is computed. Then, the minimum and maximum z-values are removed and a new value of standard deviation \( \sigma_2 \) is estimated. The process is iterated until the difference of the standard deviations between two adjacent iterations is smaller than the threshold \( \Delta_{st} \), defined by the user, as shown in the following the equation: \( \sigma_{iter-1} - \sigma_{iter} \leq \Delta_{st} \).

The process allows the points that are likely affected by gross errors to be removed from the computation of the standard deviation with an automated robust approach. In the example in Fig. 3 the process stops at the fourth iteration, where the difference between the final standard deviation and the previous one, is negligible. The outliers, however, are not eliminated by the data set during this step, but they are ruled out in the self-tuning standard deviation computation.

![Figure 3. Self-tuning computation of the standard deviation of the point elevations of a bin](image)

The median \( \text{med}_{ij} \) and the self-tuning standard deviation \( \sigma_{ij} \) of each bin are the two parameters necessary for the definition of the distance function. For each point, given its coordinates \((X_k, Y_k, Z_k)\), the planimetric position is computed in the bin reference system as \((i+di, j+dj)\). Hence, the median and standard deviation to apply to this point are computed considering the median and standard deviation values of the 4 nearest bins, by means of a bilinear interpolation:

\[
\text{med}_{i+di,j+dj} = (1 - di) \text{med}_{i,j} + didj \text{med}_{i+1,j+1} + (1 - dj)(1 - di) \text{med}_{i,j} + di(1 - dj) \text{med}_{i+1,j} \tag{5a}
\]

\[
\sigma_{i+di,j+dj} = (1 - di) \sigma_{i,j} + didj \sigma_{i+1,j+1} + (1 - dj)(1 - di) \sigma_{i,j} + di(1 - dj) \sigma_{i+1,j} \tag{5b}
\]

Therefore, a new distance function for each point can be expressed as:

\[
\text{med}_{i+di,j+dj} - n\sigma_{i+di,j+dj} < Z_k < \text{med}_{i+di,j+dj} + n\sigma_{i+di,j+dj} \tag{6}
\]

where \( n \) is a multiplicative factor of the standard deviation, which defines the width of the confidence interval. This distance function is applied for each point of the DSM in order to identify the outliers. Thus, the 3D point cloud is triangulated with the Delaunay algorithm, and the final TIN is ready for the visual inspection and the manual editing.

The results achieved by MIGC3 and S2MF on the images taken by the Z/I Imaging DMC are shown in Fig. 4 as an example.
Solid True Ortho-photo. In the recent years, the researches concerning Geomatics, have dealt with the integration of digital images and 3D models automatically extracted with Lidar or photogrammetric techniques. In particular, the Geomatics research group at the Politecnico di Torino, has developed and implemented an innovative product, called Solid True OrthoPhoto (STOP) [7] which combines the high radiometric resolution of orthophotos with the 3D information of DSMs. STOP is the ideal cartographic support for emergency operations in damaged areas. The radiometric and the 3D spatial data allow basic geometric information to be extracted, which can be easily utilized for risk assessment and emergency interventions. STOP allows 3 colour values of the true colour image (RGB) extracted from the orthophoto and memorized in 3 bytes and 1 height value derived from the DDTM and memorized as an integer (2 or 4 bytes) to be recorded for each of its pixels.

STOP has been produced by SIRio, a commercial software that has been implemented by SIR, a Politecnico di Torino Spin off.

3. Photogrammetric survey

In the 2008, photogrammetric test was carried out on the archaeological area of Augusta Bagiennorum. The ruins of the Roman city of Augusta Bagiennorum are situated a few kilometres away from the actual city of Bene Vagienna (Figure 5). Due to its strategic position in the middle of the Tanaro valley, Augusta Bagiennorum became an important centre which was exploited by the Romans for agricultural purposes and urban development. The construction phases are referable to the early Christian epoch (V-VI sec. D.C.) and high medieval period (VII-VIII D.C.).

The automatic acquisition of the theatre area provided several images of the ruins area. Three images were suitable for a photogrammetric processing of the Theatre. These were acquired at different flight altitudes (from 60 to 70 m) and in non-normal conditions (Fig. 0.).
The extraction of the homologous points by means of SIFT allowed the approximate DSM of the area to be produced, with more than 1000 points. Once the images had been oriented, the points to be matched were extracted by the Forstner operator in the area of the Theatre. The reference image in this procedure is the central one.

![Fig. 6. Points in the reference image extracted by Forstner operator](image)

The triple stereoscopic coverage of the theatre ruins and the good approximation of the initial DSM allowed quite good results to be obtained with MIGC3 and S2MF. Nevertheless, some outliers were not detected and filtered in the open field and in the area of the trees (Fig. 7). This was probably due to the lack of a multi-stereoscopic coverage. The correct matches rate (50.9%) in fact refers to the points that were extracted in the ruins of the theatre, where three images were available. In the other zones, especially in the field to the south of the ruins, the correct matches rate was dramatically reduced.

![MIGC3-S2MF Theatre 386239](image)

<table>
<thead>
<tr>
<th>MIGC3-S2MF Theatre 386239</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area (pixels)</td>
</tr>
<tr>
<td>Feat. Extracted</td>
</tr>
<tr>
<td>Feat. Matched</td>
</tr>
<tr>
<td>No match</td>
</tr>
<tr>
<td>Multi-match</td>
</tr>
<tr>
<td>% match</td>
</tr>
<tr>
<td>% no match</td>
</tr>
<tr>
<td>% multi match</td>
</tr>
<tr>
<td>Outliers</td>
</tr>
<tr>
<td>Final 3D points</td>
</tr>
<tr>
<td>% Outliers</td>
</tr>
<tr>
<td>% Final 3D points</td>
</tr>
<tr>
<td>Density (m²/point)</td>
</tr>
<tr>
<td>Plan. Step (m)</td>
</tr>
</tbody>
</table>

![Fig. 7. Report of the MIGC3 – S2MF algorithms of the theatre area (a). Filtered TIN of the theatre (b).](image)

The final TIN was regularized with the Kriging technique (step= 0.10 m), then the STOP of the area was carried out (pixel size= 0.03 m), as shown in Fig. 8.
The accuracy of the orthophoto was checked according to Italian regulations. The pixel size (0.03 m) is suitable for the production of a STOP with a nominal scale in the range 1:500 – 1:1000 range. As shown in Table 1, the CE95(EN) is lower than the tolerance limit (T\textsubscript{EN}) of the 1:500 map scale for ordinary orthophotos.

<table>
<thead>
<tr>
<th>TERRAIN</th>
<th>(\sigma\text{CP,EN}(m))</th>
<th>0.02</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(CE95\text{CP}(m))</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>(RMSE\text{OP}(m))</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>(CE95\text{OP}(m))</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>(CE95\text{EN}(m))</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>(T\text{EN 1:500 quick})</td>
<td>0.26</td>
</tr>
<tr>
<td></td>
<td>(T\text{EN 1:500 ord.})</td>
<td>0.17</td>
</tr>
</tbody>
</table>

4. Conclusion and future developments

The first results achieved by the UAS of the research group at the Politecnico di Torino have been presented. It has been shown how this application is now able to quickly acquire data relative to a catastrophic event and quickly produce cartography that is useful for the mapping activities.

The algorithms for the automatic extraction of DSMs require a multi-image acquisition, with the aim of exploiting the redundancy of the similarity constraint. Consequently, the flight plans have to be planned according to a multi-image approach. In the presented photogrammetric flight, the rate of correct homologous points is good, when the triple stereoscopic overlap is assured, because it allows the computation of the SNCC function and an improvement in the matching performances. In addition, the multi-image matching algorithm is very sensitive to camera calibration and the problems due to the calibration of the camera can reduce the capacity of the SIFT-LMS and MIGC3 techniques. If the internal orientation parameters are not correct, the homologous points will in fact have residual parallaxes, and the search along the epipolar line could not provide good correlations.

S2MF is a quick and reliable method, which assures good performances for the filter operations on DSMs automatically extracted by photogrammetric techniques. The algorithm does not require a priori knowledge of a set of error-free points; it works very quickly, and it is able to discard gross errors in a completely automatic way. The performances of S2MF depend above all
on the quality of the DSM row. Furthermore, the geometrical characteristics of the object affect the efficiency of the algorithm. S2MF is sensitive to localized discontinuities, because it works under the hypothesis of planar terrain within a single bin; in this way, the located features above the bare earth are removed. Nevertheless, the smoothing effects of S2MF vary over a wide range, which depends on the threshold parameters.

The photogrammetric tests on the archaeological ruins of Augusta Bagiennorum have underlined the possibility of using aerial surveys with UAS for mapping purposes, even in the cultural heritage application field. A semi-automatic approach led to the production of orthophotos with the information content and the nominal scale map suitable for archaeological analysis (1:500 map scales).

The UAS project is still an ongoing work. The presented results have to be ameliorated by future developments. The developed UAS still does not allow a completely automated photogrammetric flight to be performed: the assembled GPS/IMU is not sufficient to allow a direct geo-referencing. For this reason, the research group at the Politecnico di Torino is currently working on how to improve the performances of the GPS/IMU system.

The DSM extraction will be improved by extracting edges and increasing the point density through a grid point matching: the final result will be a dense DSM. A new S2MF, which defines its threshold automatically, is at present under development.

5. References