Photogrammetric Record, 17(98): 343–353 (October 2001)

AN APPROACH TO SINGLE IMAGE AUTOMATIC ORIENTATION AND POINT DETERMINATION BY USING ORTHO-IMAGES AND A DTM

Jeffrey Shan (jshan@ecn.purdue.edu)
Purdue University, Indiana, USA

Abstract

This paper describes an approach to single image automatic orientation and point determination by using current ortho-images and a DTM, and the experience gained in its implementation. The procedure proposed automatically extracts and matches feature points in evenly distributed patches on aerial images and ortho-images. A large number of image measurements (up to several thousand) are obtained in this process and are included in a robust space resection to determine the orientation parameters of the aerial image. For point determination with a single image, a method is formalised which integrates the DTM interpolation into the space resection so that the 3D ground coordinates of the image points can be determined in a unified mathematical model. Tests and analyses of this method show that the large number of automatic image measurements relieves the requirement for complicated and precise feature extraction and matching methods. The ground points obtained from single image intersection have an accuracy of approximately 1 pixel in planimetry, which fulfils the requirement for ortho-image updating. The elevation accuracy is mainly dependent on the quality of the current DTM and the interpolation method applied to it.

Keywords: DTM, image orientation, ortho-image, point determination

Introduction

Research from government and commercial organisations shows that the ortho-image is becoming an indispensable component in geographical information systems (GISs) and its cycle for updating is shortening to 4 to 5 years or less (Baltsavias, 1993; Knabenschuh, 1999). Therefore, an efficient method is required which can best utilise existing database information, such as ortho-images and digital terrain models (DTMs), to update the current GIS database.

Exterior image orientation is the first step in this process. One of the main concerns in automatic processing is its accuracy (Ackermann, 1996). In order to achieve a high accuracy, the proposed approach is based on the extraction of feature points from evenly distributed patches on a newly acquired single aerial image and on existing ortho-images. An image matching process is then carried out to associate the extracted feature points in the aerial image with those in the existing ortho-images. In this way, thousands of image measurements are obtained. The planimetric ground coordinates of the aerial image measurements are taken directly from their corresponding points in the ortho-images, while their elevations are obtained by means of interpolation of existing DTM data. The orientation of the aerial image is then determined by a space resection involving this large number of image measurements and the corresponding ground coordinates of the image points. Robust estimation can be introduced to further detect and eliminate mismatched measurements to ensure the required quality for the estimation of the image orientation parameters.

Once the image orientation is established, the ground positions of any additional image measurements in the aerial image can be determined by using the existing DTM data together with the calculated image orientation. The proposed method for the single image point determination integrates the DTM data into the space resection so that a unified solution is obtained. In addition, the proposed method is used to evaluate the performance of the orientation approach described above. In this way it provides an accuracy estimation for the updated ortho-image. The Task B data provided in the test organised by the European Organization for Experimental Photogrammetric Research (OEEPE) on "Automatic Orientation of Aerial Images on Database Information" has been used for this study (Höhle, 1999). The entire approach, including feature extraction and matching as well as the space resection, is implemented by using Java as a stand-alone program.

The remainder of this paper consists of four sections. The next section briefly formulates and describes the methods for feature extraction and matching for the subsequent space resection. Common interest operators and a least squares image matching (LSM) algorithm are presented. In the following section, the approach to point determination with a single image is described, whereby DTM interpolation is integrated into the space resection calculation. Results for feature extraction, correspondence and matching, single image orientation and point determination with the OEEPE test data are presented and analysed before the entire approach is summarised in the final section, with concluding remarks.

FEATURE EXTRACTION AND MATCHING

The proposed approach is based on image feature extraction and matching between an aerial image and its corresponding ortho-images. The feature extraction allows a large number of image measurements to be automatically obtained so that a reliable calculation of the exterior orientation of the image can be assured by using a space resection. The successful matching of the aerial image with the corresponding ortho-images provides the planimetric ground locations for the image measurements. The elevations of the image points can be obtained by subsequent DTM interpolation calculations. Space resection is thereafter conducted by including all successfully matched image points and their three-dimensional coordinates.

The proposed approach starts by selecting a number of image patches in several standard locations in the aerial image. As the aerial image to be oriented has a smaller scale than the existing ortho-images, it covers an area of four adjacent ortho-images. Therefore, the patches are interactively selected from the aerial image and the four ortho-images. In order to reduce the size of the images to be handled, the selected image patches are stored in separate files for the subsequent processing of feature extraction and matching. In this way, the search space, memory requirement and amount of computation are greatly reduced in later processes. In order to ensure a high reliability for the automatic process, each patch contains at least 512 × 512 pixels, and nine evenly distributed patches are selected on the aerial image. Corresponding patches on the aerial and ortho-images undergo the successive steps of feature extraction, feature correspondence, feature matching and precision LSM.

Fig. 1 illustrates the workflow of the proposed approach. Feature extraction is conducted on the selected aerial image patches and the corresponding ortho-image patches using the Moravec operator and the Förstner operator (Förstner and Gülch, 1987), respectively. Their formulae are summarised in Tables I and II. Experience at Purdue University shows that the two operators do not cause a significant difference in the image orientation calculation, which indicates that there is no requirement for the use of sophisticated feature extraction techniques. It also demonstrates that the large number of feature points extracted can compensate for the effect of any potential weaknesses in the feature extraction algorithms on the subsequent photogrammetric reductions. The number of extracted feature points in each patch can reach up to a few thousand. Such a large number of measurements will substantially ensure the quality of image matching and orientation computation.

After feature points have been extracted, feature correspondence is performed between them by using cross correlation within a bounded search window in the

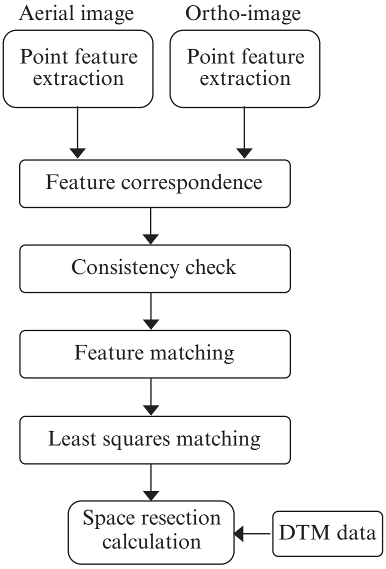

FIG. 1. Flowchart of image orientation with ortho-image and DTM.

TABLE I. Feature extraction with the Moravec operator.

Step	*Calculation*	*Formula*
1	Grey value differences in four directions	v1 = abs(NW-SE), v2 = abs(N-S) v3 = abs(NE-SW), v4 = abs(E-W)
2	Interest value	$V = \min(v1, v2, v3, v4)$
3	Feature point	Location where V is local maximum

TABLE II. Feature extraction with the Förstner operator.

Step	*Calculation*	*Formula*
1	Normal matrix	$\mathbf{N} = \begin{bmatrix} \sum g_r^2 & \sum g_r g_c \\ \sum g_r g_c & \sum g_c^2 \end{bmatrix}$
2	Major axis of feature ellipse	$W = \dfrac{1}{tr(\mathbf{N}^{-1})} = \dfrac{\det(\mathbf{N})}{tr(\mathbf{N})}$
3	Roundness of feature ellipse	$q = \dfrac{4\det(\mathbf{N})}{(tr\mathbf{N})^2}$
4	Interest value	$W^*(r, c) = \begin{cases} W(r, c) & q > q_0,\ W > W_0 \\ 0 & else \end{cases}$
5	Feature point	Location where $W^*(r, c)$ is local maximum

g_r and g_c are the grey gradients in row (r) and column (c) directions respectively.
q_0 and W_0 are the given thresholds for q and W, respectively.

corresponding aerial and ortho-image patches. In order to avoid false matching, local topological and geometrical constraints are introduced in this step. Experience shows that these constraints are necessary for obtaining reliable correspondence results. Once the feature correspondence is established, cross correlation and least squares matching (Ackermann, 1984) are applied to further refine the results. The formulation of least squares matching is given as

$$g_1(x_1, y_1) = t_0 + t_1 g_2(x_2, y_2) \tag{1}$$

where (x_1, y_1) and (x_2, y_2) are the image coordinates, and g_1 and g_2 are the grey values, in the corresponding aerial and ortho-image windows, respectively, and t_0 and t_1 are the radiometric transformation parameters between these two windows. The geometric transformation between these two windows is described by

$$\begin{aligned} x_2 &= a_0 + a_1 x_1 + a_2 y_1 \\ y_2 &= b_0 + b_1 x_1 + b_2 y_2. \end{aligned} \tag{2}$$

Together with the radiometric parameters t_0 and t_1, the geometric parameters a_0, ... b_2 are determined by the least squares computation within the two corresponding windows (Ackermann, 1984).

Once the precision matching has been undertaken, the planimetric ground coordinates of the image measurements on the aerial image can be directly determined by scaling and translating the coordinates of their corresponding points on the ortho-images, whilst their elevations are calculated by means of a local bilinear interpolation of the existing DTM. The exterior orientation parameters are determined by a space resection with the successfully matched feature points obtained in the preceding steps. Since large numbers of matched points are involved in the calculation, robust estimation can be introduced in this step to detect and eliminate remaining false matches so that the quality of the automatic orientation is further ensured.

SINGLE IMAGE POINT DETERMINATION

When the exterior orientation parameters have been obtained, the ground position corresponding to any image measurements can be determined by using the existing DTM data. This calculation is one of the fundamental problems in single image mapping and database revision. For this purpose, a general algorithm for point determination using a single image has been developed, which integrates the DTM interpolation with the space resection. A brief derivation of this algorithm is given below.

The well-known transformation which leads to the collinearity equation for a given point is written as

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \lambda \mathbf{R} \begin{pmatrix} x \\ y \\ -f \end{pmatrix} + \begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix} \tag{3}$$

where (X, Y, Z) and (X_c, Y_c, Z_c) are the ground coordinates of the point and the camera perspective centre respectively, (x,y) are the image coordinates of the corresponding ground point, f is the principal distance of the camera, $\mathbf{R}$ is the 3×3 orthogonal rotational matrix whose elements are determined by orientation angles of the aerial image, and λ is the scale factor related to that point. Let

$$\begin{pmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{pmatrix} = \mathbf{R} \begin{pmatrix} x \\ y \\ -f \end{pmatrix}. \tag{4}$$

Eliminating the coefficient λ, (3) can be rewritten as

$$\begin{aligned} \bar{x}(Z - Z_c) - \bar{z}(X - X_c) &= 0 \\ \bar{y}(Z - Z_c) - \bar{z}(Y - Y_c) &= 0. \end{aligned} \tag{5}$$

The given DTM can be expressed in general as

$$Z = F(X,\ Y) \tag{6}$$

where F denotes the elevation interpolation function, which is chosen in a bilinear form throughout the computation in this paper.

Linearisation of (5) and (6) is needed for iterative calculation, because these two equations are cross dependent when being used to determine the planimetric coordinates as well as the elevation.

With the notation (X^0, Y^0, Z^0) for the approximate coordinates of the ground point and (dX, dY, dZ) for their corrections, the linearised form of (5) and (6) can be written as

$$
\begin{aligned}
\bar{z}dX - \bar{x}dZ &= \bar{x}(Z^0 - Z_c) - \bar{z}(X^0 - X_c) \\
\bar{z}dY - \bar{y}dZ &= \bar{y}(Z^0 - Z_c) - \bar{z}(Y^0 - Y_c) \\
\frac{\partial F}{\partial Z}dX + \frac{\partial F}{\partial Y}dY - dZ &= Z^0 - F(X^0, Y^0). \qquad (7)
\end{aligned}
$$

Given the DTM as expressed in general in (6) and the image orientation parameters determined earlier based on the automatic matching output, for each image point its 3D ground coordinates can be calculated with (7). The calculation is iterative. Initially, the Z coordinate of the image point is given as the average elevation of the area covered by the image, whereas the initial coordinates X^0, Y^0 are then calculated with (5). The interpolated elevation $F(X^0, Y^0)$ is obtained by using (6) with the currently approximate X^0, Y^0 coordinates. The calculated corrections from (7) are added to the approximate coordinates for the use of subsequent iterations. The iteration stops when the coordinate corrections dX, dY and dZ are smaller than the given tolerance. In this way, the 3D ground location of the image point is determined.

RESULTS AND ANALYSES

This section presents the test results of feature extraction and matching, image orientation and single image point determination, as well as analyses based on these test results. The test has been conducted with the Task B data provided by OEEPE for the "Test on Automatic Orientation of Aerial Images on Database Information". Details concerning the data and the test are given by Höhle (1999).

Feature Extraction and Matching

Table III lists the number of image points obtained in automatic feature extraction and image matching. As can be seen, on average each 512 × 512 patch has more than 1000 feature points extracted. In total, more than 10 000 feature points are extracted in the nine patches from the new aerial image and the existing orthoimages. Through feature correspondence and matching steps, singular feature points that appear only in one of the two corresponding image patches, or can not reach a certain reliable match level, are screened out. More than one in ten of the extracted feature points are carried over to the final orientation calculation. Experience at Purdue University shows that the large number of measurements makes an essential contribution to the high quality of the subsequent image orientation and point determination steps, and relieves the requirement for a complex feature extraction and matching technique.

Orientation Parameters

Exterior orientation parameters are determined by space resection with robust estimation. Different thresholds (1·8 to 2·0 σ_0, where σ_0 is the standard deviation

TABLE III. Numbers of feature points, correspondence points and matched points.

	Feature extraction		*Correspondence and matching*	
Patch number	*Aerial image*	*Ortho-image*	*Correspondence*	*Least squares matching*
1	957	1037	340	206
2	1039	1237	285	195
3	1072	2287	360	240
4	1992	1542	506	320
5	1335	1468	502	312
6	3749	4022	218	156
7	1252	1460	260	160
8	966	2043	157	95
9	924	1093	223	155
Total	**13286**	**16189**	**2851**	**1840**

of image measurements) are used to eliminate blunders remaining from the previous automatic feature extraction and matching steps. It is shown that different thresholds have only minor effects on the final orientation parameters, although the orientation angles show a small instability relative to the various thresholds. The first row of Table IV gives the average values of the orientation parameters obtained from the space resection calculation. The second row of Table IV gives the estimated standard deviation for the orientation parameters at $\sigma_0 = 31{\cdot}9\,\mu\mathrm{m}$ obtained from the covariance matrix in the space resection. A calculation based on the image scale of 1:27 000 shows that the planimetric precision of the camera position is within 10 μm or ⅓ pixel (1 pixel ≈ 30 μm or 0·8 m on the ground), while the precision of Z_c is approximately 0·002% of the flying height (4100 m above the ground). Multiplying the standard deviations for φ and ω by the flying height indicates that these values are compatible with the precision of the camera positions. Such a high precision is only made possible by the high redundancy in the space resection from the large number of measurements and the good geometric configuration of the wide-angle photography. The standard deviation σ_0 for the image coordinates reflects the precision of feature extraction and image matching, and errors caused by, for example, image scanning, camera calibration and DTM interpolation.

TABLE IV. Data for exterior orientation parameters.

	Camera orientation (°)			*Camera position (m)*		
Measurement	φ	ω	κ	X_c	Y_c	Z_c
Value	−1·298	−0·093	88·394	46·65	604·87	4202·19
Standard deviation, σ	0·0032	0·0025	0·0012	0·26	0·23	0·09
Difference, Δ	0·0447	−0·0293	−0·0150	−4·32	2·07	0·47

Standard deviation of image measurements, $\sigma_0 = 31{\cdot}9\,\mu\mathrm{m}$.
Δ is the difference between the orientation parameters obtained from manual and automatic measurements.

In order to evaluate the accuracy of the proposed approach, 97 evenly distributed conjugate points were manually measured on the aerial image and on the ortho-images. The image coordinates from the ortho-images were scaled by the ground distance of each pixel and translated to obtain the planimetric coordinates on the ground. Elevations for those ground points were then obtained by DTM interpolation. These derived 3D ground coordinates were used to carry out the space resection to calculate the orientation parameters of the aerial image. As Table IV demonstrates, errors in image measurements mainly cause the planimetric differences for the camera position. Since the manual measurements are obtained in mono-observation mode, it is reasonable to believe that they cause a large percentage of the differences in the orientation parameters. However, their influence on the ground point determination is far less significant than on the camera position, as is shown below.

Single Image Point Determination

The ground positions of image points on the aerial image are determined by using the approach described in the previous section. Two sets of the exterior orientation parameters of the aerial image obtained from manual and automatic measurements respectively have been used to carry out the single image point determination for a group of 25 image points, whose coordinates on the aerial image were provided by OEEPE (Höhle, 1999). The comparison of the ground coordinates obtained from the two sets of orientation parameters shows the influence of orientation parameters on the single image point determination. The first row in Table V lists the root mean square error (rmse) of the ground coordinates at these 25 points obtained with the two sets of orientation parameters.

A first analysis of Table V reveals that the proposed automatic approach obtains virtually the same elevation as the manual measurements. This result suggests that in the single image point determination, the elevation accuracy for a ground point depends mainly on the DTM interpolation rather than the space resection. The difference for planimetric coordinates is of the order of 1 pixel (0·8 m on the ground), even when a large percentage of matched points is screened out in the space resection. Since the manual measurement is performed on the screen in mono mode, space resection conducted with those measurements will be no better than that from automatic precision matching. Thus, as a conservative estimate, when the two resection methods have the same effect on the rmse as shown in Table V, an estimation for the standard deviation of the planimetric coordinates determined by using automatic matched measurement is $\sim 0{\cdot}8/\sqrt{2} = 0{\cdot}57\,\text{m}$ (or 0·7 pixel).

TABLE V. Rmse at the ground points determined.

Comparison methods	Number of points for comparison	Root mean square error (m) X	Y	Z
Based on two sets of orientation parameters	25	0·782	0·825	0·055
Based on manually measured check points	97	1·079	1·137	0·199

In order to obtain a comprehensive and reliable understanding of the proposed approach given in this paper, another comparison has been conducted by using all the manual image measurements as check points. After the exterior orientation parameters had been obtained by automatic image matching, single image intersection was performed for all 97 check points to calculate their ground coordinates, which were then compared with the "ground truth" (the coordinates derived from measurements on the ortho-images). The rmse is listed in the second row of Table V. Since the measurement is carried out manually on the screen in mono mode, its precision is expected to be no better than 1 pixel and several blunders have been detected and deleted. Bearing this in mind, the rmse for planimetric coordinates listed in the second row of Table V in fact has two error components, one caused by space resection ($\sigma_{\text{resection}}$) and one by measurements of the check points themselves ($\sigma_{\text{check points}}$), namely

$$\sigma^2_{X,Y} = \sigma^2_{\text{resection}} + \sigma^2_{\text{check points}} \tag{8}$$

where $\sigma_{X,Y} = \sigma_X = \sigma_Y$. Taking the average of rmse^2 for X and Y in Table V as an estimate for $\sigma^2_{X,Y}$ and counting $\sigma^2_{\text{check points}}$ as 0·8 m (1 pixel), the actual effect of space resection on planimetric coordinates on the ground is then estimated as

$$\sigma^2_{\text{resection}} = (1{\cdot}079^2 + 1{\cdot}137^2)/2 - 0{\cdot}8^2 = 0{\cdot}588$$

or

$$\sigma_{\text{resection}} = 0{\cdot}77\,\text{m}\ (\sim 1\ \text{pixel}).$$

As the elevation of ground points is virtually determined by means of DTM interpolation, its accuracy σ_Z mainly depends on terrain slope and the planimetric accuracy. By using (6), it can be deduced that

$$\sigma_Z = \sqrt{(F^2_X + F^2_Y)}\ \sigma_{X,Y} \tag{9}$$

where F_X and F_Y are the partial derivatives of $Z = F(X, Y)$ with respect to X and Y, namely the terrain slopes in the X and Y directions respectively. For the flat areas covered by this test image, the slope angle is less than 5°, and therefore the elevation accuracy is estimated to be better than $1{\cdot}414\ \tan 5^\circ \sqrt{1{\cdot}23} = 0{\cdot}14\,\text{m}$ or 0·003% of the flying height, which is very close to the value estimated in Table IV.

CONCLUDING REMARKS

The approach to single image automatic orientation and point determination using ortho-images and a DTM presented in this paper is effective and reliable and can reach the ideal precision potential of the image measurements. Up to several thousand feature points extracted from evenly distributed image patches can ensure the quality of the space resection for sensor orientation. Such an approach has less critical requirements on the precision and use of complicated feature extraction and matching methods. Different feature extraction and matching methods essentially lead to almost the same image orientation results in practice, provided that large numbers of image measurements are involved. For point determination with a single image, the proposed method integrates the DTM into a space resection computation so that a unified solution is achieved. It is shown that the elevation accuracy of point determination with a single image is essentially dependent on the quality of the DTM and its interpolation, whereas the planimetric accuracy can reach ~1 pixel with automatic image measurements. This value is also the accuracy for ortho-image updating. Further research will be focused on automatically matching linear features between image and database and using them, together with point features, to automate exterior orientation, change detection and updating procedures.

Acknowledgements

The work described in this paper was conducted with the Task B test data provided by OEEPE. The paper is based in part on "Special contribution B2 to Task B" by the author, published as an addendum to the report by J. Höhle in *OEEPE Official Publication* 36 (Shan, 1999). The assistance of Dr. Höhle and the permission of OEEPE to publish these results are gratefully acknowledged.

REFERENCES

Ackermann, F., 1984. High precision digital image correlation. *Proceedings of 39th Photogrammetric Week, Stuttgart. Schriftenreihe* 9 (Institut für Photogrammetrie der Universität Stuttgart): 231–243.

Ackermann, F., 1996. Some considerations about feature matching for the automatic generation of digital elevation models. *OEEPE Official Publication* 33: 231–240.

Baltsavias, E. P., 1993. Integration of ortho-images in GIS. *State-of-the-Art Mapping*. SPIE, 1943: 314–324.

Förstner, W. and Gülch, E., 1987. A fast operator for detection and precise location of distinct points, corners and circular features. *Proceedings of ISPRS Intercommission Conference on Fast Processing of Photogrammetric Data*, Interlaken, Switzerland. 437 pages: 281–305.

Höhle, J., 1999. Automatic orientation of aerial images on database information. *OEEPE Official Publication* 36: 71–117.

Knabenschuh, M., 1999. Requirements of mapping organisations for automatic production of orthoimages and updated maps. *Ibid.*: 167–173.

Shan, J., 1999. Special contribution B2 to Task B. *Ibid.*: 151–157.

Résumé

On présente dans cet article une méthode pour orienter automatiquement une image isolée et en déterminer des points géoréférencés, en utilisant un MNT et des ortho-images existantes. On fait part également de l'expérience acquise au cours de la mise en œuvre de cette méthode. Le processus en question extrait et apparie automatiquement les points de diverses détails situés dans des zones régulièrement réparties sur des images aériennes et des ortho-images. On obtient ainsi un nombre considérable de déterminations sur ces images (allant jusqu'à plusieurs milliers) et on les injecte dans un relèvement dans l'espace bien conformé de façon à déterminer les paramètres d'orientation de l'image aérienne. Lorsqu'il s'agit de déterminer des points sur une image isolée, on a instruit une méthode qui intègre l'interpolation dans le MNT et le relèvement dans l'espace, de sorte que les coordonnées-terrain en 3D des points de cette image sont alors obtenue à partir d'un modèle mathématique unique. Les essais réalisés et leur analyse ont montré que ce nombre considérable de mesures automatiques sur l'image allège les conditions requises par les méthodes précises mais sophistiquées *d'appariement et d'extraction des éléments. Les points au sol, obtenus à partir d'une image isolée ainsi traitée, ont une précision planimétrique d'environ 1 pixel, ce qui suffit pour effectuer la mise à jour d'une ortho-image. La précision altimétrique dépend essentiellement de la qualité du MNT disponible et de la méthode d'interpolation qui lui est appliquée.*

Zusammenfassung

Dieser Beitrag beschreibt einen Ansatz für die automatische Orientierung von Einzelbildern und der Punktbestimmung mit Hilfe von existierenden Orthophotos und einem DGM und die damit verbundenen Erfahrungen bei der Implementierung. Die vorgeschlagene Methode extrahiert automatisch Punkte in gleichmäßig verteilten Bildausschnitten in Luftbild und Orthophotos und ordnet sie zu. In diesem Prozess wird eine große Zahl (mehrere Tausend) von Punkten generiert, die in eine robuste Triangulierung zur Bestimmung der Orientierungsparameter des Luftbildes einfließen. Zur Punktbestimmung in einem Einzelbild wird eine Methode entwickelt, die die DGM Interpolation in die räumliche Triangulierung integriert, so dass die 3D Objektkoordinaten der Bildpunkte in einem einheitlichen mathematischen Modell bestimmt werden. Untersuchungen und Analysen dieses Verfahrens zeigen, dass durch die hohe Anzahl automatisch gemessener Bildpunkte die Anforderungen an eine hochgenaue Merkmalsextraktion und Zuordnungsmethoden reduziert werden. Die Bodenpunkte, die durch Einzelbildauswertung abgeleitet werden, haben eine Lagegenauigkeit von ca. 1 Pixel, was die Anforderungen einer Orthophotofortführung erfüllt. Die Höhengenauigkeit ist vor allem von der Qualität des aktuellen DGMs und der angewandten Interpolationsmethode abhängig.