Theoretical Analysis of the Iterative Photogrammetric Method to Determining Ground Coordinates from Photo Coordinates and a DEM

Yongwei Sheng

Abstract

It is necessary to determine ground coordinates from a single aerial image and a digital elevation model (DEM). A widely used method in photogrammetry is to iteratively calculate the coordinates based on the inverse collinearity equations. The iterative photogrammetric method may be divergent when the terrain surface becomes complicated. However, there is a lack of theoretical analysis of this method. This paper theoretically analyzes the convergence condition and the convergence speed of the method, and validates the theory using a simulated surface containing various slope conditions. The elevation angle θ of the view ray and the inclination angle α of the profile, intersected by the terrain surface and the vertical view plane, play a critical role in the method. The necessary and sufficient condition for convergence is $\alpha < \theta$. The convergence speed is quantified as a function of α, θ, the convergence threshold ΔT and the offset ΔZ_0 of initial elevation estimate.

Introduction

It is often necessary in photogrammetry to convert coordinates between ground and photo (or image) coordinate systems. Suppose camera properties, location (X_s, Y_s, Z_s) and attitude (ω, ϕ, κ) are known, and the lens distortions are negligible. The relationship between the 2D image coordinates (x, y) and the 3D ground coordinates (X, Y, Z) is quantitatively described by the well-known collinearity equations, Equation 1 and their inverse forms, Equation 2.

$$\begin{cases} x = x_0 - f \cdot \dfrac{r_{11}(X - X_s) + r_{12}(Y - Y_s) + r_{13}(Z - Z_s)}{r_{31}(X - X_s) + r_{32}(Y - Y_s) + r_{33}(Z - Z_s)} \\ y = y_0 - f \cdot \dfrac{r_{21}(X - X_s) + r_{22}(Y - Y_s) + r_{23}(Z - Z_s)}{r_{31}(X - X_s) + r_{32}(Y - Y_s) + r_{33}(Z - Z_s)} \end{cases} \quad (1)$$

$$\begin{cases} X = X_s + (Z - Z_s) \cdot \dfrac{r_{11}(x - x_0) + r_{21}(y - y_0) + r_{31}(-f)}{r_{13}(x - x_0) + r_{23}(y - y_0) + r_{33}(-f)} \\ Y = Y_s + (Z - Z_s) \cdot \dfrac{r_{12}(x - x_0) + r_{22}(y - y_0) + r_{32}(-f)}{r_{13}(x - x_0) + r_{23}(y - y_0) + r_{33}(-f)} \end{cases} \quad (2)$$

where, f is the camera focal length; x_0 and y_0 are the image coordinates of the photo principal point; and finally, r_{11}, r_{12}, . . . , and r_{33} are the elements of the rotation matrix which is determined by ω, ϕ, and κ.

It is a straightforward process to determine photo coordinates from ground coordinates using Equation 1, and the projected coordinates (x, y) on the image plane of a point (X, Y, Z) in the object space can be solely determined. The reverse process, determining the ground coordinates (X, Y, Z) from photo coordinates (x, y) using Equation 2, however is much more complicated since the two equations are not sufficient for three unknowns (i.e., X, Y, and Z). Additional information is therefore needed. 3D ground coordinates can be obtained in photogrammetry from a stereo pair of photos through the process of space intersection, or from a single image and a digital elevation model (DEM) through the process of mono-plotting.

Determining the position in the 3D object space of a feature in an aerial image with the assistance of a DEM is known as the single-ray backprojection problem (Mikhail *et al.*, 2001). This is considered as the main process in digital mono-plotting (Makarovic, 1973; Radwan and Makarovic, 1980; Baltsavias, 1996) and the core of direct rectification in orthophoto generation (Konecny, 1979; Chen and Lee, 1993; Mueller *et al.*, 2001; Li, 2002; Sheng *et al.*, 2003).

The iterative photogrammetric (IP) method based on the inverse collinearity equations has been widely used in photogrammetry to solve the backprojection problem (Masry and McLaren, 1979; Radwan and Makarovic, 1980; Chen and Lee, 1993; Tudor and Sugarbaker, 1993; Doytsher and Hall, 1995; Zhang and Zhang, 1996). From Equation 2, X and Y can be determined if the elevation Z is provided. However, elevations in a DEM array are indexed by column (i.e., the X coordinate) and row (i.e., the Y coordinate) in the object space, so the elevation of a cell stays unknown until its column and row in the DEM array become available. Therefore, an iterative process is needed.

As shown in Figure 1, the following procedures are generally used in the IP method to determine the ground coordinates for pixel **P**:

(1) Using a given initial elevation Z_0, determine the X and Y coordinates of the initial projected point $\mathbf{A_1}$ using Equation 2.

Department of Environmental Resources and Forest Engineering, SUNY College of Environmental Science and Forestry, 312 Bray Hall, Syracuse, NY 13210 (ysheng@esf.edu).

Photogrammetric Engineering & Remote Sensing
Vol. 71, No. 7, July 2005, pp. 863–871.

0099-1112/05/7107–0863/$3.00/0

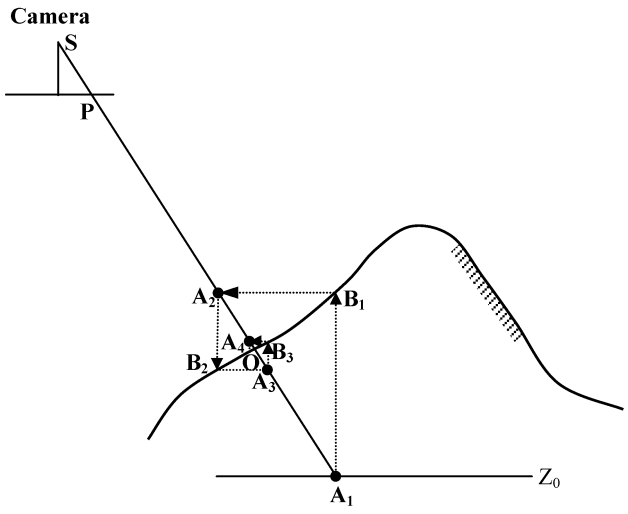


Figure 1. Iterative photogrammetric method. The ground coordinates of pixel **P** are the coordinates of the intersection point **O** between the view ray and the terrain surface, and the location of **O** is solved iteratively. Each iteration produces a projected point $\mathbf{A_i}$ ($i = 1, 2, \ldots, n$), and the sequence of $\mathbf{A_1}$, $\mathbf{A_2}$, $\mathbf{A_3}$, $\mathbf{A_4}$, . . . , $\mathbf{A_n}$ approaches **O** gradually if the sequence is convergent.

(2) From $\mathbf{A_1}$ vertically trace to surface point $\mathbf{B_1}$, and interpolate its Z coordinate Z_1 from the DEM using $\mathbf{A_1}$'s X and Y coordinates.
(3) Update the elevation Z using $\mathbf{B_1}$'s Z coordinate (i.e., $Z = Z_1$), and repeat Step 1 and Step 2 to determine subsequent projected points $\mathbf{A_2}$, $\mathbf{A_3}$, $\mathbf{A_4}$, . . . , $\mathbf{A_n}$ until convergence.
(4) Output the coordinates of $\mathbf{A_n}$ as the ground coordinates of pixel **P**.

In the above iterative process, one iteration is considered to include the two steps of determining the Z value (from either the initial elevation estimate or the DEM) and calculating X and Y coordinates using Equation 2. Each iteration produces one projected point $\mathbf{A_i}$ (i = 1,2, . . .) on the view ray. As the iteration continues, $\mathbf{A_1}$, $\mathbf{A_2}$, $\mathbf{A_3}$, $\mathbf{A_4}$, . . . , $\mathbf{A_n}$ form a sequence of projected points. If the last two projected points are close enough, then the iteration is considered convergent, and the coordinates of the last projected point are output as the ground coordinates of the pixel. Various distance definitions between the last two projected points may be used as the convergence measure in the iterative method. Tudor and Sugarbaker (1993) and Doytsher and Hall (1995) used the vertical distance, whereas Radwan and Makarovic (1980) and Mueller *et al.* (2001) took the planimetric distance. This paper uses the Euclidean distance as the convergence measure and defines the iteration error as the Euclidean distance between the last two projected points. The criterion for convergence is that the iteration error becomes less than a pre-defined convergence threshold ΔT.

The IP method described above is computationally efficient compared to iterative and non-iterative ray-tracing methods (Mikhail *et al.*, 2001; Börner *et al.*, 2001), and can produce precise coordinates when the surface is simple (Sheng, 2004). However, several problems associated with the IP method have been recognized. Zhang and Zhang (1996) and Sheng (2004) pointed out the IP method may be divergent when the terrain surface becomes complicated. In practice, the iteration is considered divergent if it does not converge after many iterations (i.e., a large number m). Jauregui *et al.* (2002) and Sheng (2004) illustrated that this method is problematic when occlusions are present in aerial imagery. Radwan and Makarovic (1980) qualitatively briefed that the convergence speed (i.e., the number of iterations needed for convergence) depends on the local terrain slope, the accuracy of the initial elevation estimate, the pixel's location in the photo, and the required accuracy. However, this iterative method has not been theoretically studied. The condition for convergence is largely unclear; the convergence speed is quantitatively unknown. This paper provides a theoretical analysis of the IP method, discussing the condition for convergence and quantitatively estimating the theoretical convergence speed of this method.

Theoretical Analysis of the Iterative Method

Being an iterative algorithm, the IP method suffers from the divergence problem. It is rather difficult to directly analyze convergence on general surface. To make the discussion easier, this section first analyzes the convergence condition and the convergence speed of this method on planar surface, and finally discusses convergence on general surface.

The Condition for Convergence

Convergence on planar terrain surface is analyzed from the viewpoint of geometry. The terrain profile illustrated in Figure 1 actually is the intersection between the terrain surface and the vertical view plane defined by the view ray and the vertical line through the camera lens center. Figure 2a shows the geometry for a planar surface whose slope angle is β. The ground coordinates of pixel **P** are the coordinates of the intersection point **O** between the terrain surface and the view ray **SO**, emitted from the camera lens center **S** at an elevation angle of θ. **OV**, with an inclination angle of α, is the profile intersected between the view plane **S-T-W** and the terrain surface. The profile in the view plane is illustrated in Figure 2b and Figure 2c for fore-viewed and back-viewed slopes, respectively. Suppose Z_0 is the initial elevation estimate, which offsets by ΔZ_0 from the true value; the distance between **O** and the first projected point $\mathbf{A_1}$ is $d^{(1)} = OA_1 = d = |\Delta Z_0|/\sin\theta$. As the iteration continues, $\mathbf{A_2}$, $\mathbf{A_3}$, and $\mathbf{A_4}$ are projected on the view ray using Equation 2. The distance between **O** and these points are $d^{(2)} = OA_2 = d \cdot (tg\ \alpha/tg\ \theta)$, $d^{(3)} = OA_3 = d \cdot (tg\ \alpha/tg\ \theta)^2$, and $d^{(4)} = OA_4 = d \cdot (tg\ \alpha/tg\ \theta)^3$. In a general form, the distance between **O** and the n^{th} (n = 1,2,3, . . .) projected point is

$$d^{(n)} = d \cdot (tg\,\alpha/tg\,\theta)^{(n-1)}; \tag{3}$$

$\{d^{(n)}\}$ forms a geometric sequence with a ratio of $tg\ \alpha/tg\ \theta$. If $\alpha < \theta$, then $d^{(n+1)} < d^{(n)}$ for any n. Examining the limit of $d^{(n)}$, we have

$$\lim_{n\to\infty} d^{(n)} = \begin{cases} 0 & \text{if } \alpha < \theta \\ d & \text{if } \alpha = \theta \\ \infty & \text{otherwise} \end{cases}.$$

If $\lim_{n\to\infty} d^{(n)} = 0$, then these projected points approach **O** gradually, and the sequence $\{d^{(n)}\}$ is convergent. Therefore, the necessary and sufficient condition for convergence is $\alpha < \theta$.

The value of $d^{(n)}$ is the distance between **O** and the n^{th} projected point, but the coordinates of **O**, which the iterative method intends to determine, stay unknown for $d^{(n)}$ calculation. Thus, $d^{(n)}$ cannot be used as the convergence measure. In practice, the distance $\Delta D^{(n)}$ between two successive projected points $\mathbf{A_{n-1}}$ and $\mathbf{A_n}$, is used instead in the IP method as the convergence measure.

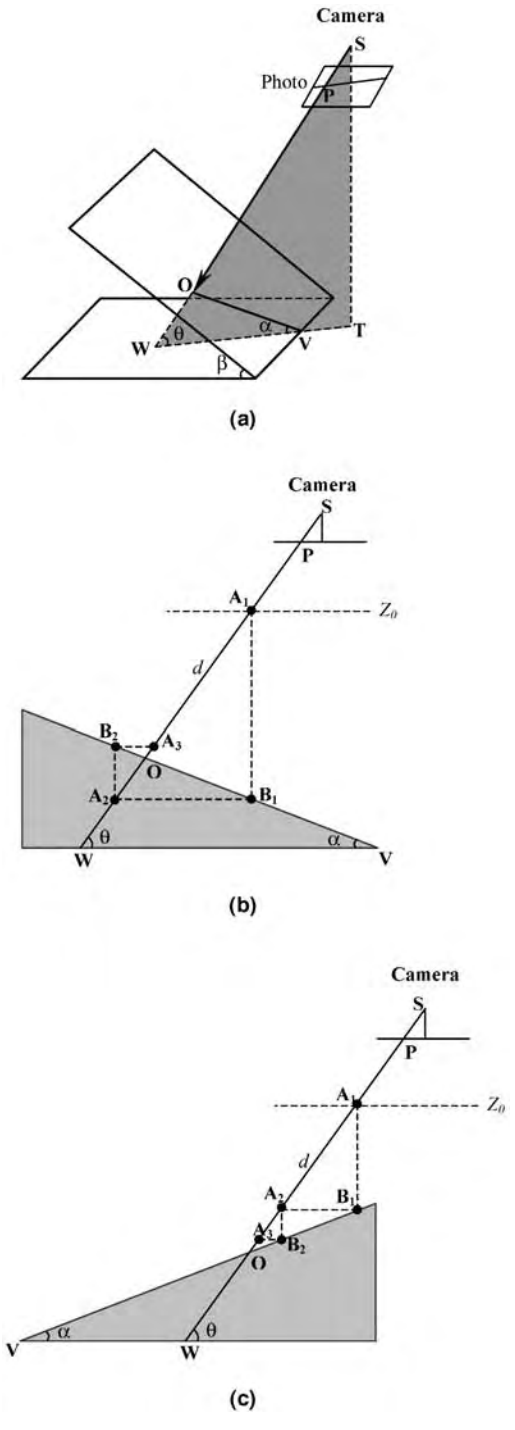


Figure 2. Geometry and convergence on planar surface. (a) Geometry of the view ray and the surface. Figure 1 shows the profile intersected between the terrain surface and the shaded vertical view plane **S**-**T**-**W**, defined by the view ray **SW** and the vertical line **ST** through the camera focal point **S**. Point **O** can be solved by analyzing the geometry in plane **S**-**T**-**W**. (b) Profile of a fore-viewed slope in plane **S**-**T**-**W**. Note that projected points $\mathbf{A_1}$, $\mathbf{A_2}$, $\mathbf{A_3}$, $\mathbf{A_4}$, . . . , $\mathbf{A_n}$ approach point **O** from both sides. (c) Profile of a back-viewed slope in plane **S**-**T**-**W**. Note that $\mathbf{A_1}$, $\mathbf{A_2}$, $\mathbf{A_3}$, $\mathbf{A_4}$, . . . , $\mathbf{A_n}$ approach point **O** from only one side.

The sequence of projected points gradually approaches point **O** when the iteration is convergent. The iteration converges in a different way on fore-viewed and back-viewed slopes, and this is intuitive in geometry. The iterates (i.e., the projected points) approach point **O** from both sides on fore-viewed slopes as illustrated in Figure 2b, and the convergence is called oscillatory in numerical analysis (Jacques and Judd, 1987). On the contrary, the iterates approach point **O** from only one side on back-viewed slopes (Figure 2c), and the convergence is monotonic. Thus, $\Delta D^{(n)}$ has to be calculated differently for fore-viewed and back-viewed slopes as follows:

$$\Delta D^{(n)} = \begin{cases} OA_{n-1} + OA_n = d^{(n-1)} + d^{(n)} & \text{fore-viewed slopes} \\ OA_{n-1} + OA_n = d^{(n-1)} - d^{(n)} & \text{back-viewed slopes} \end{cases}, n = 2,3,4,....$$

Introducing Equation 3 and after some rearrangement, we have

$$\Delta D^{(n)} = \begin{cases} d^{(n)} \cdot (tg\,\theta/tg\,\alpha + 1) = d \cdot (tg\,\alpha/tg\,\theta)^{(n-1)} \cdot (tg\,\theta/tg\,\alpha + 1) & \text{fore-viewed slopes} \\ d^{(n)} \cdot (tg\,\theta/tg\,\alpha - 1) = d \cdot (tg\,\alpha/tg\,\theta)^{(n-1)} \cdot (tg\,\theta/tg\,\alpha - 1) & \text{back-viewed slopes} \end{cases}. \quad (4)$$

In the extreme cases of $\theta = 90°$ or $\alpha = 0°$, $\Delta D^{(n)} = \begin{cases} d & \text{if } n = 2 \\ 0 & \text{if } n > 2 \end{cases}$.

Like $\{d^{(n)}\}$, $\{\Delta D^{(n)}\}$ is also a geometric sequence with a ratio of $tg\ \alpha/tg\ \theta$. Similarly,

$$\lim_{n\to\infty} \Delta D^{(n)} = \begin{cases} 0 & \text{if } \alpha < \theta \\ 2d \text{ (fore-viewed slopes)} \\ 0 \quad \text{(back-viewed slopes)} & \text{if } \alpha = \theta \\ \infty & \text{otherwise} \end{cases}$$

Thus, the general condition for convergence of $\{\Delta D^{(n)}\}$ is also $\alpha < \theta$. Both $\{d^{(n)}\}$ and $\{\Delta D^{(n)}\}$ are linearly convergent to zero, if and only if, $\alpha < \theta$. If $\{\Delta D^{(n)}\}$ converges to zero, then $\{d^{(n)}\}$ does, and vice versa. In summary, the theoretical necessary and sufficient condition for convergence of the IP method on planar surface is $\alpha < \theta$.

The Speed of Convergence

When $\Delta D^{(n)}$ becomes less than the pre-defined convergence threshold ΔT, then the iteration stops and is considered convergent at the n^{th} iteration. Let $\Delta D^{(n)} < \Delta T$, and from Equation 4, we have

$$|\Delta Z_0|/\sin\theta \cdot (tg\,\alpha/tg\,\theta)^{(n-1)} \cdot (tg\,\theta/tg\,\alpha \pm 1) < \Delta T$$

where the "+" and "−" signs are taken for fore-viewed and back-viewed slopes, respectively.

Solving for n, and taking extreme cases into consideration, the convergence speed is

$$n = \begin{cases} 3 & \text{if } \theta = 90° \text{ or } \alpha = 0° \\ 2 & \text{if } |\Delta Z_0| < \Delta T \cdot \sin\theta/(1 \pm tg\alpha/tg\theta) \\ \left[\dfrac{\ln\left(\dfrac{|\Delta Z_0| \cdot (tg\,\theta/tg\,\alpha \pm 1)}{\Delta T \cdot \sin\theta}\right)}{\ln(tg\,\theta/tg\,\alpha)}\right] + 1 & \text{if } \alpha < \theta \\ \text{Divergent} & \text{otherwise} \end{cases} \quad (5)$$

where $\ln(x)$ is the logarithmic function, and $[x]$ rounds x to the next integer.

This is the formula for the convergence speed of the IP method, and it can be used to predict the number of iterations needed for convergence. The IP method needs at least two iterations to converge since two or more projected points are needed to compute the convergence measure. For a nadir view ray (i.e., $\theta = 90°$), the iterative method converges at the third iteration no matter how large α is or how big ΔZ_0 is. The method also converges at the third iteration for a horizontal surface (i.e., $\alpha = 0°$), not depending on θ and ΔZ_0. When the initial elevation estimate is selected close enough to the true value, i.e., $|\Delta Z_0| < \Delta T \cdot \sin\theta/(1 \pm tg\,\alpha/tg\,\theta)$, the iterative method stops at the second iteration no matter $\alpha < \theta$ or not. α and θ are crucial to the convergence speed. Figure 3 shows the convergence speed as a function of α and θ for both fore-viewed and back-viewed slopes. When $\alpha > \theta$, n is either infinite or a negative number, and the iteration is divergent. When $\alpha < \theta$, n is a positive number, and the $\{\Delta D^{(n)}\}$ sequence converges after n iterations. If α is much smaller than θ, the iteration converges rapidly; as α approaches θ, the convergence slows down sharply. It is obvious from Equation 5 that the iteration converges slower on a fore-viewed slope than on a back-viewed slope with the same α. In addition, the magnitudes of ΔT and ΔZ_0 cancel out in Equation 5. For example, the setting of $\Delta Z_0 = 10$ m and $\Delta T = 1$ m is equivalent (in terms of convergence speed) to the setting of $\Delta Z_0 = 1$ m offset and $\Delta T = 0.1$ m. They need the same number of iterations to convergence.

There are four factors affecting the convergence speed in Equation 5: the profile inclination angle α, the view angle θ, the convergence threshold ΔT, and the offset ΔZ_0 of the initial elevation estimate Z_0. The first two factors are not selectable, and the third one is also fixed to the accuracy to be achieved. Only ΔZ_0 can be specified by selecting the initial elevation Z_0. The selection of Z_0 is crucial to the iterative method. The selection should be based on DEM properties, and the chosen Z_0 should be representative to the most popular elevation values in the DEM. Mean and median values of vicinity DEM elevation values are suitable candidates of Z_0.

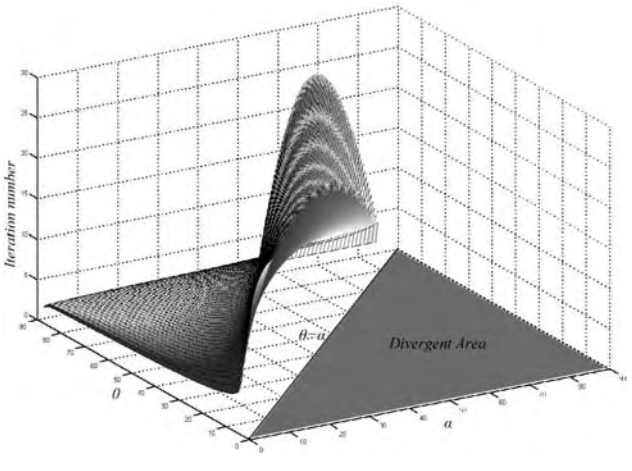


Figure 3. Convergence speed as function of α and θ. This plot shows the convergence speed for both fore-viewed slopes (the wire-framed surface) and back-viewed slopes (the shaded surface) using $\Delta T = 0.1$ m and $\Delta Z_0 = 10.0$ m. The iteration converges slower on fore-viewed slopes than on back-viewed slopes with the same α.

Convergence on General Surface

The angle α is not a constant for general surface, and this makes convergence analysis rather complicated. Figure 4a and Figure 4b show the situations of fore-viewed and back-viewed general slopes, respectively. Let the sequence of projected points on the view ray be $\mathbf{A_1}$, $\mathbf{A_2}$, . . . , and $\mathbf{A_n}$. Let $\mathbf{B_1}$ be the first traced surface point, whose elevation offset be h from the initial elevation estimate, and the subsequent traced points on the surface be $\mathbf{B_2}$, $\mathbf{B_3}$, . . . , and $\mathbf{B_n}$. Let the α angles between two successive traced points be α_1 for $\mathbf{B_1B_2}$, α_2 for $\mathbf{B_2B_3}$, α_3 for $\mathbf{B_3B_4}$, . . . , and α_{n-1} for $\mathbf{B_{n-1}B_n}$. Based on the geometry in Figure 4a and Figure 4b, the distance between two successive projected points can be derived as follows:

$$
\begin{aligned}
A_1A_2 &= A_1B_1/\sin\theta = |h|/\sin\theta,\\
A_2A_3 &= A_2B_2/\sin\theta = A_2B_1 \cdot tg\,\alpha_1/\sin\theta = A_1A_2 \cdot (tg\,\alpha_1/tg\,\theta),\\
A_3A_4 &= A_3B_3/\sin\theta = A_3B_2 \cdot tg\,\alpha_2/\sin\theta\\
&= A_2A_3 \cdot (tg\,\alpha_2/tg\,\theta), \text{ and so on.}
\end{aligned}
$$

In a general form,

$$
\begin{aligned}
A_{n-1}A_n &= A_{n-2}A_{n-1} \cdot (tg\,\alpha_{n-2}/tg\,\theta)\\
&= |h|/\sin\theta \cdot \prod_{i=3}^{n} (tg\,\alpha_{i-2}/tg\,\theta),\ n = 3,4,\ldots. \qquad (6)
\end{aligned}
$$

A_1A_2, A_2A_3, . . . , $A_{n-1}A_n$ form a sequence $\{\Delta G^{(n)}\}$, which is a general form of $\{\Delta D^{(n)}\}$ for planar surface (see the Appendix for a proof).

From Equation 6, one can see that $\Delta G^{(n)} < \Delta G^{(n-1)}$ if $\alpha_{n-2} < \theta$. If all the α angles are less than θ, then $\lim_{n\to\infty} \Delta G^{(n)} = 0$. Thus, the sufficient condition for convergence on general surface is that all the α angles are less than θ. It should be pointed out that this is not a necessary condition since the iteration may be still convergent even if some of these α angles are greater than θ.

It is rather difficult to analytically estimate the convergence speed on general surface. For a given convergence threshold ΔT, the number of necessary iterations is the smallest integer n satisfying $\Delta G^{(n)} < \Delta T$. That is,

$$
|h|/\sin\theta \cdot \prod_{i=3}^{n} (tg\,\alpha_{i-2}/tg\,\theta) < \Delta T. \qquad (7)
$$

Since too many α angles are involved, it is hardly possible to analytically solve n from Equation 7 as we did for planar surface. If the α angles along the profile do not change much in the neighborhood of $|\Delta Z_0|/tg\,\theta$, the general surface can be well analyzed using the theory developed from planar surface.

When the terrain surface becomes complicated, occlusions may occur in aerial imagery and cause problems to the iterative method. As illustrated in Figure 4c, the view ray intersects with the surface at more than one points, i.e., **A**, **B**, and **C**, and the true intersection point should be **A** since it is the closest point to the camera. However, the iterative algorithm may not always pick up the correct point, and this to a large extent depends on the initial elevation estimate. If the selected initial elevation is less than the elevation of point **B**, then the algorithm most likely finds **C** as the intersection point, which is wrong.

Experiments

The above theoretical analysis needs to be tested with various slope situations. Three pyramids (i.e., planar surfaces)

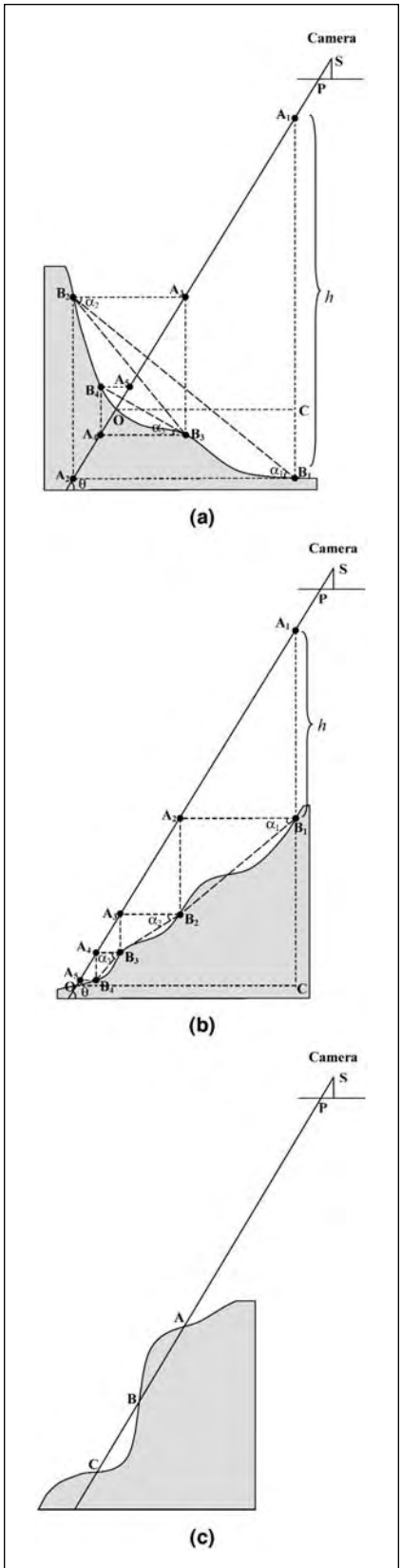


Figure 4. Convergence on general surface. (a) Convergence on fore-viewed general surface. (b) Convergence on back-viewed general surface. (c) Occlusions affecting the IP method.

and three *sinc* surfaces (i.e., general surfaces) of various settings were placed on a flat background plane (i.e., a horizontal surface) with an elevation of 25 m to compose a DEM (Figure 5a) for testing in this study. Pyramids 1, 2, and 3 with the slope angle of 30°, 40°, and 50° were placed at the upper-right, middle-left and lower-right portions of the DEM, respectively. A *sinc* surface is defined as $Z = p \cdot sinc(X,Y) = p \cdot sin(\sqrt{X^2 + Y^2})/\sqrt{X^2 + Y^2}$, where Z is the vertical coordinate, X and Y are the horizontal coordinates, and p is the height of the central peak of the surface. In addition, the number of peaks (q) in the surface can be specified. Smaller p and q produce a less relieved surface. A 30 m tall-3-peaked *sinc* surface 1, a 10 m-2-peaked *sinc* surface 2, and a 20 m-2-peaked *sinc* surface 3 were placed on the upper-left, middle-right, and lower-left portions of the DEM, respectively. The composed DEM covers a 120 m × 170 m area with a cell size of 0.24 m. The ground coordinates of the DEM's lower-left corner are (4307.00 m, 3951.00 m). The minimum, mean, median, and maximum values of the DEM are 18.48 m, 27.40 m, 25.00 m and 76.45 m, respectively.

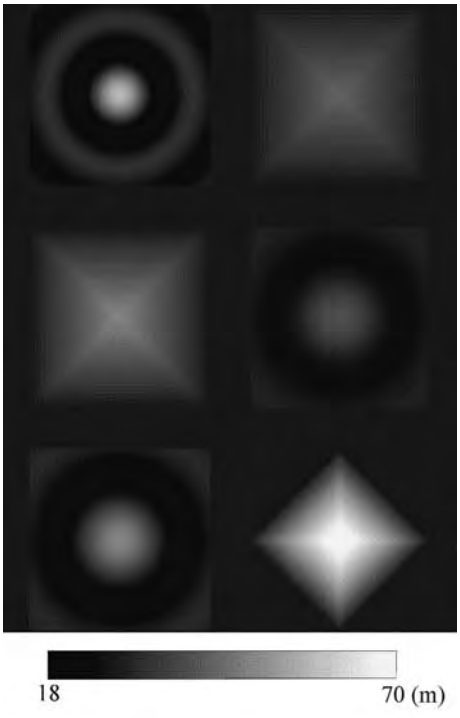
An aerial view image (Figure 5b) of the test area was simulated using ray-tracing techniques (Börner *et al.*, 2001) from the DEM with the Sun illuminating from the lower-right direction. The camera (f = 152.888 mm) with attitudes of (0.5015°, 0.0028°, 0.3228°) was placed at (4576.40 m, 4003.50 m, 404.70 m) to the right of the test area. The reflectance ρ of image pixels was modeled based on the incident angle θ_i between the sunlight ray and the surface normal, and the exit angle θ_e between the view ray and the surface normal (Woodham and Gray, 1987).

$$\rho = \frac{\tau + 1}{2\pi} \cdot \cos\theta_i^{\tau} \cdot \cos\theta_e^{\tau - 1} \tag{8}$$

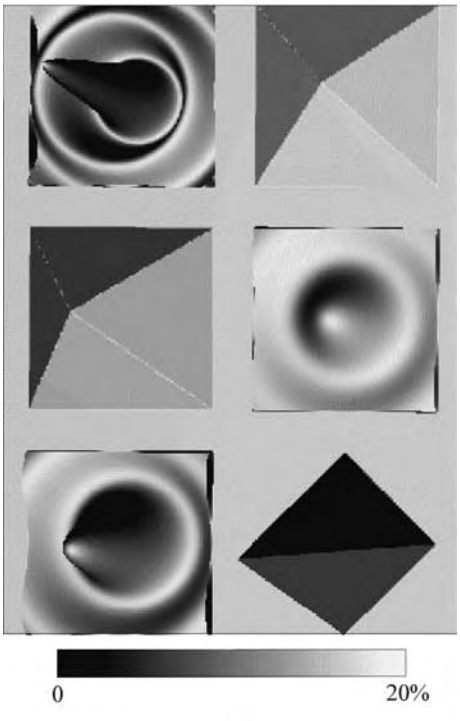
where τ is an adjusting coefficient ranging between 0 and 1, and is set to 0.8 in this paper.

Using the median value (i.e., 25.00 m) of the DEM as the initial elevation estimate, one-tenth of pixel size (i.e., 0.024 m) as the convergence threshold ΔT, and limiting the maximum iteration number to 50 (i.e., m = 50), the IP method computes the ground coordinates for all pixels in the image. The results are shown in Figure 6. Gradual slopes such as the horizontal background, Pyramids 1 and 2, and *sinc* surface 2 turn out to be fine in the produced Z coordinate map (Figure 6a). However, the method fails to produce coordinates on some steep slopes (the black areas marked as **A**, **B**, **C**, and **D**) in Figure 6a. Figure 6b shows the number of iterations for each pixel. The background plane in dark gray converges at the second iteration since its elevation happens to equal Z_0. The black color shows the divergent areas where the iteration does not converge within 50 iterations, and corresponds to the problematic black areas in the Z coordinate map (Figure 6a). The bright areas around these divergent areas, though are convergent, converge very slow with a large number of iterations. In addition, the results also show that the IP method is prone to occlusions. It is noticeable that Pyramid 3 loses its tip (marked by **E**) on both maps (Figure 6a and Figure 6b). Pyramid 3 in the aerial view image (Figure 5b) occludes the background plane to its left side. Since the initial elevation estimate (Z_0 = 25 m) is selected close to the elevation of the background plane, the IP method tends to pick up the occluded background points, leading to the missing tip of Pyramid 3 in Figure 6a and Figure 6b.

The convergence theory of the IP method developed for planar surface is applied to various surface conditions contained in this synthetic data set. α and θ play a critical role in convergence of the method. The derived α and θ

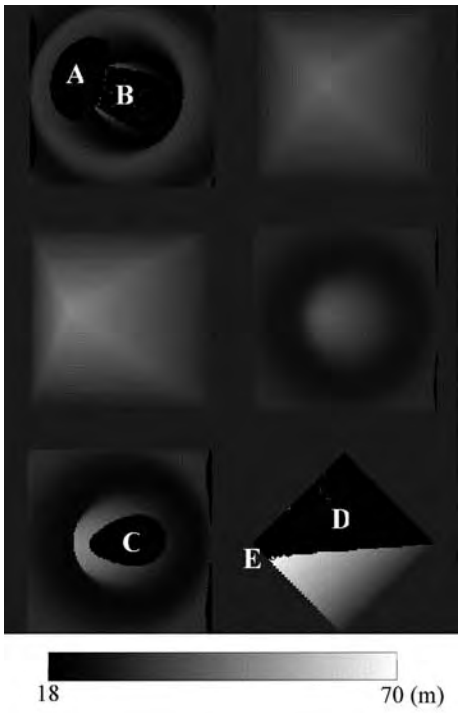


(a)

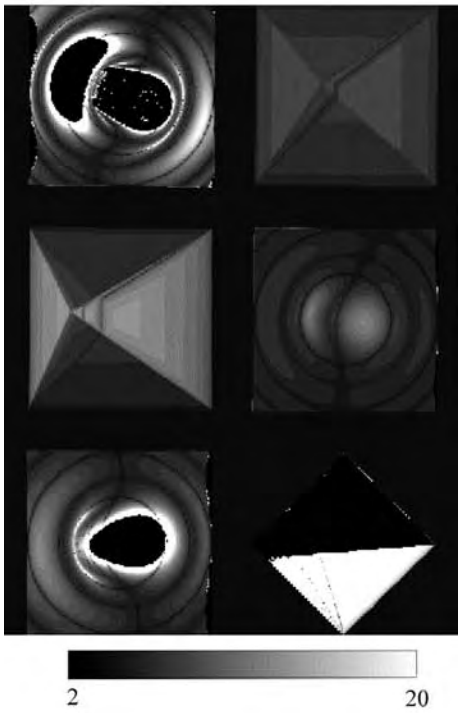


(b)

Figure 5. Composed DEM and simulated aerial image. (a) Composed DEM. (b) Simulated aerial image.



(a)



(b)

Figure 6. Outputs of iterative method. (a) *Z* coordinate map. Divergent areas are marked by **A**, **B**, **C**, and **D**. **E** marks the missing tip of Pyramid 3 due to occlusions. (b) Iteration number map.

maps are shown in Figure 7a and Figure 7b, respectively. Since α is greater than θ around the central peaks of *sinc* surfaces 1 and 3 and on the upper-right slope of Pyramid 3, they are divergent areas in Figure 7c detected using the $\alpha \geq \theta$ criterion. These modeled divergent areas in general agree with the real divergent areas (black patches marked by **A**, **B**,

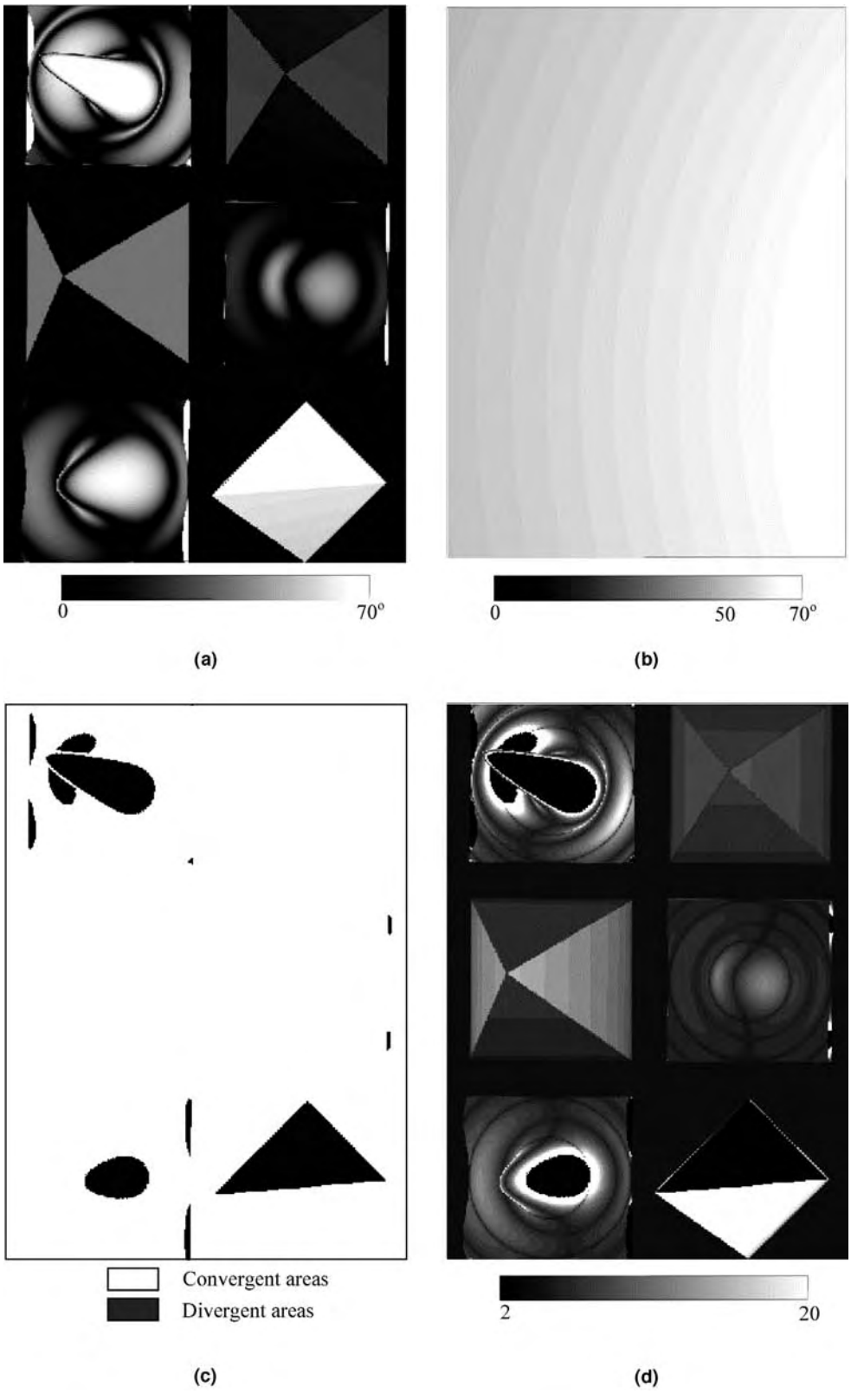


Figure 7. Analysis of convergence condition and convergence speed. (a) α map. (b) θ map. (c) Modeled divergent areas. (d) Predicted iteration number map.

C, and **D**) in Figure 6a. Figure 7d shows the iteration number map predicted from the α and θ maps using Equation 5. This map closely resembles the real iteration number map in Figure 6b with minor difference appearing around the edges where α changes abruptly.

Discussion

There has been little discussion about the performance of the IP method, and some properties of the method were not properly (or not precisely at least) understood. This theoretical analysis can help to clarify some misunderstandings about this method. Chen and Lee (1993) stated that the iteration slows down when the surface normal has a direction similar to the view ray, however this is not the case. For example, the IP method converges very fast on a nearly horizontal surface around the photo principal point, though the view ray direction is similar to the surface normal under this situation. Radwan and Makarovic (1980) included terrain slope as one of the factors affecting the performance of the iterative method. This research shows that α is critical to the convergence speed of the IP method, and α is not the slope angle β of the surface. Rather than being a surface property, α is the inclination angle of the profile intersected between the terrain surface and the vertical view plane: when under any situation, $\alpha \leq \beta$. As the slope aspect getting perpendicular to the view ray azimuth, α can be much smaller than β.

Though it is difficult to directly establish the convergence theory for general surface, the convergence theory developed for planar surface works quite well even for general surface as long as the surface can be locally represented by planar surfaces. In the experiments with various slope conditions in this study, the results show general agreements between the divergent areas and the theoretically detected extent, and similarities between the real iteration number and the theoretical estimate. Minor differences are due to the abrupt changes of α. These similarities support the investigated theory on convergence of the IP method.

Conclusions

This paper investigates the iterative photogrammetric method to the single-ray backprojection problem from a theoretical viewpoint, and tests the theory using a synthetic surface containing a variety of slope conditions. The following can be concluded from this research:

1. For general surfaces, the sufficient condition for convergence is that all the α angles are smaller than the elevation angle θ of the view ray. α is the inclination angle of the line segment formed by the two successive traced points on the surface.
2. It is difficult to analytically study the performance of the IP method on general surfaces. If the α angles do not change much in the vicinity, general surfaces can be analyzed using the theory developed for planar surface. The subsequent findings are drawn from the theoretical analysis on planar surface.
3. The values of α and θ play a critical role in the IP method. The necessary and sufficient condition for convergence is $\alpha < \theta$. When $\alpha \geq \theta$, the iteration is divergent.
4. The convergence speed depends on the inclination angle α, the view angle θ, the convergence threshold ΔT, and the offset ΔZ_0 of initial elevation estimate. If the iterative method converges at the nth iteration, then n is estimated as

$$n = \begin{cases} 3 & \text{if } \theta = 90^\circ \text{ or } \alpha = 0^\circ \\ 2 & \text{if } |\Delta Z_0| < \Delta T \cdot \sin\theta/(1 \pm tg\alpha/tg\theta) \\ \left[\dfrac{\ln\left(\dfrac{|\Delta Z_0| \cdot (tg\theta/tg\alpha \pm 1)}{\Delta T \cdot \sin\theta}\right)}{\ln(tg\theta/tg\alpha)} \right] + 1 & \text{if } \alpha < \theta \\ \text{Divergent} & \text{otherwise} \end{cases}$$

5. The method performs differently on fore-viewed slopes and back-viewed slopes. In the above formula, the "+" and "−" signs are taken for fore-viewed and back-viewed slopes, respectively. The iterative method converges slower on fore-viewed slopes than on back-viewed slopes with the same settings.
6. The method is prone to occlusions, and the initial elevation selection becomes critically important when occlusions are present. If the initial elevation is not properly estimated, the algorithm may mistakenly pick up the intersection point on the occluded side of the surface.

Appendix

$\{\Delta D^{(n)}\}$ is a special case of $\{\Delta G^{(n)}\}$ with constant α.

[Proof]

For planar surface, $\alpha_1 = \alpha_2 = \alpha_3 = \ldots = \alpha_{n-2} = \alpha$, and Equation 6 is simplified to

$$\Delta G^{(n)} = A_{n-1}A_n = |h|/\sin\theta \cdot (tg\,\alpha/tg\,\theta)^{(n-2)} \tag{9}$$

We can obtain the relationship between $|h|$ and d for planar surface from the geometry of fore-viewed slopes and back-viewed slopes.

For fore-viewed slopes (Figure 4a),

$$\begin{aligned} |h| = A_1B_1 &= A_1C + B_1C = A_1O \cdot \sin\theta + A_1O \cdot \cos\theta \cdot tg\,\alpha \\ &= d \cdot \sin\theta \cdot (1 + tg\,\alpha/tg\,\theta) \end{aligned} \tag{10a}$$

For back-viewed slopes (Figure 4b),

$$\begin{aligned} |h| = A_1B_1 &= A_1C - B_1C = A_1O \cdot \sin\theta - A_1O \cdot \cos\theta \cdot tg\,\alpha \\ &= d \cdot \sin\theta \cdot (1 - tg\,\alpha/tg\,\theta) \end{aligned} \tag{10b}$$

Taking Equations 10a and 10b into Equation 9, we get

$$\Delta G^{(n)} = \Delta D^{(n)} = \begin{cases} d \cdot (tg\,\alpha/tg\,\theta)^{(n-1)} \cdot (tg\,\theta/tg\,\alpha + 1) & \text{fore-viewed slopes} \\ d \cdot (tg\,\alpha/tg\,\theta)^{(n-1)} \cdot (tg\,\theta/tg\,\alpha - 1) & \text{back-viewed slopes} \end{cases}$$

Thus, $\{\Delta D^{(n)}\}$ is a special case of $\{\Delta G^{(n)}\}$ with constant α.

References

Baltsavias, E.P., 1996. Digital ortho-images – a powerful tool for the extraction of spatial- and geo-information, *ISPRS Journal of Photogrammetry and Remote Sensing*, 51(1):63–77.

Börner, A., L. Wiest, P. Keller, R. Reulke, R. Richter, M. Schaepman, and D. Schläpfer, 2001. SENSOR: a tool for the simulation of hyperspectral remote sensing systems, *ISPRS Journal of Photogrammetry and Remote Sensing*, 55(5–6):299–312.

Chen, L.C., and L.H. Lee, 1993. Rigorous generation of digital orthophotos from SPOT images, *Photogrammetric Engineering & Remote Sensing*, 59(3):655–661.

Doytsher, Y, and J. Hall, 1995. Fortran programs for co-ordinate resection using an oblique photograph and high-resolution DEM, *Computers & Geosciences*, 21(7):895–905.

Jacques, I., and C. Judd, 1987. *Numerical Analysis*, Chapman and Hall, New York, 326 p.

Jauregui, M., J. Vilchez, and L. Chacón, 2002. A procedure for map updating using digital mono-plotting, *Computers & Geosciences*, 28(4):513–523.

Konecny, G., 1979. Methods and possibilities for digital differential rectification, *Photogrammetric Engineering & Remote Sensing*, 45(6):727–734.

Li, Z., 2002. Orthoimage generation and measurement from single images, *Geographical Data Acquisition* (Y.Q. Chen and Y.C. Lee, editors), Springer-Verlag, New York, pp. 173–192.

Makarovic, B., 1973. Digital mono-plotters, *ITC Journal*, 1973(4):583–599.

Masry, S.E., and R.A. McLaren, 1979. Digital map revision, *Photogrammetric Engineering & Remote Sensing*, 45(2):193–200.

Mikhail, E.M., J.S. Bethel, and J.C. McGlone, 2001. *Introduction to Modern Photogrammetry*, John Wiley & Sons, New York, 479 p.

Mueller, R., M. Lehner, and R. Mueller, 2001. Verification of digital elevation models from MOMS-2P data, *Online Proceedings of ISPRS Joint Workshop on High Resolution Mapping from Space 2001, 19–21 September, Hannover, Germany*, URL: http://www.ipi.uni-hannover.de/html/publikationen/2001/isprs_cd.pdf (last date accessed: 13 April 2005).

Radwan, M.M., and B. Makarovic, 1980. Digital mono-plotting system – improvements and tests, *ITC Journal*, 1980(3): 511–534.

Sheng, Y., 2004. Comparative evaluation of iterative and non-iterative methods to ground coordinate determination from single aerial images, *Computers & Geosciences*, 30(3): 267–279.

Sheng, Y., P. Gong, and G.S. Biging, 2003. True Orthoimage Production for Forested Areas from Large-Scale Aerial Photographs, *Photogrammetric Engineering & Remote Sensing*, 69(3):259–266.

Tudor, G.S., and L.J. Sugarbaker, 1993. GIS orthographic digitizing of aerial photographs by terrain modeling, *Photogrammetric Engineering & Remote Sensing*, 59(4):499–503.

Woodham, R.J., and M.H. Gray, 1987. An analytic method for radiometric correction of satellite multispectral scanner data, *IEEE Transactions on Geoscience and Remote Sensing*, 25(3):258–271.

Zhang, Z., and J. Zhang, 1996. *Digital Photogrammetry*, Press of Wuhan Surveying and Technology University, Wuhan, China, in Chinese, 289 p.

(Received 01 December 2003, accepted 15 March 2004, revised 09 April 2004)