3D Reconstruction Methods Based on the Rational Function Model

C. Vincent Tao and Yong Hu

Abstract
The rational function model (RFM) is an alternative sensor model allowing users to perform photogrammetric processing. The RFM has been used as a replacement sensor model in some commercial photogrammetric systems due to its capability of maintaining the accuracy of the physical sensor models and its generic characteristic of supporting sensor-independent photogrammetric processing. With RFM parameters provided, end users are able to perform photogrammetric processing including ortho-rectification, 3D reconstruction, and DEM generation with an absence of the physical sensor model. In this research, we investigate two methods for RFM-based 3D reconstruction, the inverse RFM method and the forward RFM method. Detailed derivations of the algorithmic procedure are described. The emphasis is placed on the comparison of these two reconstruction methods. Experimental results show that the forward RFM can achieve a better reconstruction accuracy. Finally, real Ikonos stereo pairs were employed to verify the applicability and the performance of the reconstruction method.

Introduction
A sensor model relates 3D object point positions to their corresponding 2D image positions. It describes the geometric relationships between the image space and the object space. A well-designed sensor model ensures that 3D reconstruction (or stereo intersection) and ortho-rectification products generated from imagery are accurate.

Physical sensor models and generalized sensor models are the two broad categories of sensor models used (McGlone, 1996). The rational function model (RFM) is essentially a generalized sensor model. Use of the RFM to ‘‘replace’’ the rigorous physical sensor models has been in practice for a decade (Paderes et al., 1989; Greve et al., 1992). Described in the OpenGIS document (OCC, 1999), there are three main replacement sensor models, the grid interpolation model, the RFM, and the universal real-time sensor model. These models are all generalized, i.e., the model parameters do not carry physical meanings of the imaging process. The primary purpose of the use of ‘‘replacement sensor models’’ is their capabilities of sensor independence, high fitting accuracy, and real-time calculation (Madani, 1999; Dowman and Dolloff, 2000; Tao and Hu, 2001a).

The replacement sensor model should be accurate and robust enough so that it can be used, with no distinguishable loss of accuracy, for photogrammetric processing, e.g., ortho-rectification, 3D reconstruction, DEM generation, etc. The name ‘‘replacement sensor model’’ is sometimes confusing. From an end-user perspective, with the replacement sensor model provided, the user can perform photogrammetric processing without needing to know the rigorous physical sensor model, the sensor type, and the physical imaging process. However, to be able to replace the rigorous physical sensor model, the physical sensor model is often used for the determination of some unknown coefficients in the replacement sensor model.

The RFM has gained considerable interest recently in the photogrammetry and remote sensing community, mainly due to the fact that some satellite data vendors, for example, Space Imaging have adopted the RFM as a replacement sensor model for image exploitation. The rational function coefficients (RFCs) of the RFM instead of the physical sensor parameters are provided to end users for photogrammetric processing. Such a strategy may help keep confidential the information about the sensors because it is difficult to derive the physical sensor parameters from the RFM. On the other hand, the RFM facilitates the applications of high-resolution satellite imagery due to its simplicity and generality. It was reported in Grodecki (2001a) that the Ikonos rational model differs by no more than 0.04 pixel from the physical sensor model, with the RMS error below 0.01 pixel.

The RFM was initially used in the U.S. military intelligence community. Therefore, there have been few publications available to researchers, developers, and users until the past two years (Madani, 1999; Dowman and Dolloff, 2000; Yang 2000; Tao and Hu, 2001b; Tao and Hu, 2001c). The least-squares solution to the nonlinear RFM was derived and described in Tao and Hu (2001b). The accuracy assessment of the use of RFM for replacing the rigorous sensor models is provided in Dowman and Dolloff (2000), Yang (2000), and Tao and Hu (2001b; 2001c).

The RFM-based 3D reconstruction has been implemented in some softcopy photogrammetric software packages (Paderes et al., 1989; Greve et al., 1992; Madani, 1999) but without disclosures of the details regarding their methods. Yang (2000) described an RFM-based iterative procedure to compute the object point coordinates from a stereo pair. In his method, an inverse form of the RFM, where the planimetric coordinates are represented as rational functions of the image coordinates and the ground elevations, is used to establish the 3D reconstruction. The method was validated using both aerial and SPOT stereo pairs. 3D reconstruction using the forward form of the RFM was examined by Di et al. (2001) and Tao and Hu (2000). Tao and Hu (2000) developed a web-based demonstration system based on the RFM for the general photogrammetric community.

1. The term Rational Polynomial Camera (RPC) model used by Space Imaging is the same as the RFM used in this context.

C.V. Tao is with the York Geospatial Information and Communication Technology Lab, York University, 4700 Keele Street, Toronto, Ontario M3J 1P3, Canada (tao@yorku.ca).

Y. Hu is with the Department of Geomatics Engineering, The University of Calgary, 2500 University Drive, NW, Calgary, Alberta T2N 1N4, Canada (yhu@ucalgary.ca).

Photogrammetric Engineering & Remote Sensing
0099-1112/02/6807-705$3.00/0
© 2002 American Society for Photogrammetry and Remote Sensing
on the forward RFM reconstruction method. Di et al. (2001) tested both the upward (forward) and the downward (inverse) RFM for 3D shoreline mapping.

In this paper, we offer a detailed description of two 3D reconstruction methods, with an emphasis on their comparison. The two methods are compared using aerial photography data. In order to verify the applicability and the performance of the method, stereo Ikonos pairs with RFCs supplied by Space Imaging were used to evaluate the absolute and relative accuracies of the reconstructed object points.

### Rational Function Models (RFM)

**Forward and Inverse RFMs**

In the RFM, image pixel coordinates \((r, c)\) are expressed as the ratios of polynomials of object point coordinates \((X, Y, Z)\). The two image coordinates and three object point coordinates are each offset and scaled to fit the range from \(-1.0\) to \(+1.0\) over an image or image section in order to minimize the introduction of errors during the computation (NIMA, 2000). A detailed description of this normalization process can be found at the OpenGIS Consortium website (OGC, 1999). For the ground-to-image transformation, the defined ratios of polynomials have the **forward form** (Greve et al., 1992; OGC, 1999): i.e.,

\[
\begin{align*}
    r &= p_1(X, Y, Z) = p_1(X, Y, Z) \\
    c &= p_3(X, Y, Z) = p_3(X, Y, Z)
\end{align*}
\]

where \(r\) and \(c\) are the normalized row and column index of pixels in the image, respectively, and \(X, Y, \) and \(Z\) are the normalized coordinate values of points in object space. For the third-order case, the numerators and denominators in Equation 1 are 20-term polynomials: i.e.,

\[
p = \sum_{i=0}^{m_1} \sum_{j=0}^{m_2} \sum_{k=0}^{m_3} a_{ijk} X^i Y^j Z^k
\]

where the \(a_{ijk}\) are polynomial coefficients, called rational function coefficients (RFCs), and \(m_1, m_2,\) and \(m_3\) are maximum powers of ground coordinates and are typically limited to 3. The order of the terms is different from that used in the RFCs provided by Space Imaging. However, it is trivial and differs in the literature (Greve et al., 1992; Madani, 1999; Yang, 2000; Grodecki, 2001a; Tao and Hu, 2001b).

The forward form, Equation 1, defines the ground-to-image transformation. For image-to-ground transformation, an **inverse form** can be used (Yang, 2000): i.e.,

\[
\begin{align*}
    X &= \frac{p_5(r, c, Z)}{p_6(r, c, Z)} \\
    Y &= \frac{p_7(r, c, Z)}{p_8(r, c, Z)}
\end{align*}
\]

**Rational Function-Based Ortho-Rectification and 3D Reconstruction**

RFM-based ortho-rectification is straightforward. Either the forward form or the inverse form of the RFM can be employed. It results in two different rectification approaches: direct rectification, i.e., from the original image space \((r, c)\) to the object space \((X, Y, Z)\); and indirect rectification, i.e., from the original image space \((r, c)\) to the object space \((X, Y, Z)\). The advantages and disadvantages of each approach together with the resampling methods can be found in Novak (1992).

After solving the RFM for each image, 3D reconstruction can be performed using the corresponding points in a stereo pair by grouping the transformation equations (the forward form is used herein): i.e.,

\[
\begin{align*}
    r_i &= \frac{p_1(X, Y, Z)}{p_2(X, Y, Z)} \\
    c_i &= \frac{p_3(X, Y, Z)}{p_4(X, Y, Z)}
\end{align*}
\]

For multiple (e.g., three or more) images with overlapped area, two more transformation equations for each additional image would be added. As shown in Figure 1, compared to the conventional stereo intersection, there is no actual intersection of the light rays occurring at the object point. Therefore, we use the term “3D reconstruction” instead of “stereo intersection” throughout the paper, and we used dotted lines in Figure 1 to represent the concept of the “virtual intersection.”
3-D Reconstruction with the Inverse RFM

After the RFs of Equation 2 are solved for in each image, the 3D object point coordinates can be iteratively calculated using the conjugate image points in a stereo pair.

Algorithm Derivation

Applying a Taylor expansion of X and Y towards the input variable Z in Equation 2, we have the first-order approximations

\[
X \approx \hat{X} + \frac{\partial X}{\partial Z} \cdot \Delta Z \\
Y \approx \hat{Y} + \frac{\partial Y}{\partial Z} \cdot \Delta Z
\]

where

\[
\frac{\partial X}{\partial Z} = \frac{\partial p_5}{\partial Z} \cdot p_6 + \frac{\partial p_6}{\partial Z} \cdot p_5 - \frac{5}{6} \cdot \frac{\partial p_6}{\partial Z}
\]

\[
\frac{\partial Y}{\partial Z} = \frac{\partial p_7}{\partial Z} \cdot p_8 + \frac{\partial p_8}{\partial Z} \cdot p_7 - \frac{7}{8} \cdot \frac{\partial p_8}{\partial Z}
\]

\[
\frac{\partial p_6}{\partial Z} = a_1 + a_1 c + a_2 r + 2a_1 r + a_{12} c Z + 2a_{12} r Z + 3a_{12} c^2 + 3a_{12} r^2 + 3a_{12} c Z^2
\]

and \(\hat{X}\) and \(\hat{Y}\) are estimated by substituting some approximate values of \(r, c,\) and \(Z\) into Equation 2.

Given a pair of conjugate image points \((r_i, c_i)\) and \((r_j, c_j)\) and a value of \(Z\), we have

\[
X \approx \hat{X}_i + \frac{\partial X_i}{\partial Z} \cdot \Delta Z, \quad Y \approx \hat{Y}_i + \frac{\partial Y_i}{\partial Z} \cdot \Delta Z \\
X \approx \hat{X}_j + \frac{\partial X_j}{\partial Z} \cdot \Delta Z, \quad Y \approx \hat{Y}_j + \frac{\partial Y_j}{\partial Z} \cdot \Delta Z
\]

Eliminating \(X\) and \(Y\) from above equations, we have the error equations

\[
\begin{bmatrix}
\frac{\partial X_i}{\partial Z} & \frac{\partial X_j}{\partial Z} \\
\frac{\partial Y_i}{\partial Z} & \frac{\partial Y_j}{\partial Z}
\end{bmatrix}
\begin{bmatrix}
\Delta Z
\end{bmatrix}
= \begin{bmatrix}
\hat{X}_i - \hat{X}_j \\
\hat{Y}_i - \hat{Y}_j
\end{bmatrix}
\]

Then the least-squares solution to \(\Delta Z\) is

\[
\Delta Z = \left( \frac{\partial X_i}{\partial Z} - \frac{\partial X_j}{\partial Z} \right) \cdot w_x \left( \frac{\partial Y_i}{\partial Z} - \frac{\partial Y_j}{\partial Z} \right) + \left( \frac{\partial Y_i}{\partial Z} - \frac{\partial Y_j}{\partial Z} \right) \cdot w_y
\]

where \(w_x\) and \(w_y\) are weights for \(X\) and \(Y\).

Yang (2000) proposed an alternative correction with the form

\[
\Delta Z = \left( \frac{\partial X_i}{\partial Z} - \frac{\partial X_j}{\partial Z} \right) \cdot w_x \left( \frac{\partial Y_i}{\partial Z} - \frac{\partial Y_j}{\partial Z} \right) + w_y \left( \frac{\partial Y_i}{\partial Z} - \frac{\partial Y_j}{\partial Z} \right) \cdot \left( \frac{\partial X_i}{\partial Z} - \frac{\partial X_j}{\partial Z} \right)
\]

Reconstruction Procedure

Now we can sketch the procedure for computing the object point coordinates from a pair of conjugate points \((r_i, c_i)\) and \((r_j, c_j)\) in the image.

1. Find an initial approximate value for elevation \(Z\). This can often be specified as the median value of the elevation range (e.g., 0 for the normalized elevation range \([-1, 1]\)).
2. Calculate the correction \(\Delta Z\) using Equation 4, and then add \(\Delta Z\) to \(Z\).
3. Repeat Step 2, and update \(Z\) each time with \(\Delta Z\), until the conjugate image points in a stereo pair are solved for in each image, the 3D object point coordinates can be iteratively calculated using the third-order inverse form is used with coordinate normalization. The graph shows excellent convergence, with the correction value decreasing by many orders of magnitude in just three iterations. No further refinement can be regularly obtained with more iterations.

Therefore, the specified threshold can be strict. In our experiments, three iterations were always sufficient to ensure convergence (a threshold of \(1.0 \times 10^{-9}\) meters in \(Z\) was used in our experiments.)
testing). We have also found that the convergence of the iterative procedure is not dependent on the initial approximate value of $Z$ as long as it falls within the elevation range. For this reason, we start the reconstruction with the initial value of $Z$ set to the median elevation range.

It should be noted that the object point coordinates for the two images of a stereo pair should be un-normalized or be normalized using the same offset and scale values in the same object coordinate system. If the two images of a stereo pair are normalized separately using different offset and scale values for the object point coordinates, the computation equations should be modified accordingly. The object point coordinates separately normalized for the two images of a stereo pair can also be re-normalized to be in the same object coordinate system, and then the RFCs would be re-solved.

The inverse RFM reconstruction described above may not be able to obtain the best solution because it allows only one explicit least-squares solution for $Z$, and discrepancies may occur in the $X$ and $Y$ directions. As we will observe in the next section, the forward RFM allows for a simultaneous least-squares adjustment for all three object point coordinates. We will show that a better solution can be expected by treating the result of the inverse RFM as the initial approximation for the forward RFM reconstruction.

### 3D Reconstruction with the Forward RFM

After the RFCs of the forward RFM from Equation 1 are solved for, the 3D object position can be iteratively reconstructed from its corresponding image points. As we know, Space Imaging provides the values of RFCs and the normalization parameters for the forward form of the RFM. Let $X_0$, $Y_0$, and $Z_0$ be the un-normalized coordinate values of points in object space. The normalization of the ground coordinates is computed using the following equations (OpenGIS Consortium, 1999):

$$\begin{align*}
X &= X_u - X_o, & Y &= Y_u - Y_o, & Z &= Z_u - Z_o
\end{align*}$$

where $X_o$, $Y_o$, and $Z_o$ are offset values for three ground coordinates, and $X_u$, $Y_u$, and $Z_u$ are scale values for three ground coordinates. Moreover, the left and right images of a stereo pair are usually normalized separately using different offset and scale values for the ground coordinates. Therefore, the normalization parameters should be introduced so that those original, instead of separately normalized, object point coordinates are used in the adjusting equations for the images.

#### Algorithm Derivation

Similar to the previous section, we get first-order approximations by applying a Taylor expansion of $r$ and $c$ towards the three input variables $X$, $Y$, and $Z$ in Equation 1. Thus, when considering the normalization parameters, the four error equations for two corresponding image points $(r_i, c_i)$ and $(r_j, c_j)$ become

$$\begin{align*}
\begin{bmatrix} v_{i1} \\ v_{i2} \\ v_{i3} \\ v_{i4} \end{bmatrix} &= \begin{bmatrix} \frac{\partial r}{\partial Z} & \frac{\partial r}{\partial Y} & \frac{\partial r}{\partial X} \\ \frac{\partial c}{\partial Z} & \frac{\partial c}{\partial Y} & \frac{\partial c}{\partial X} \\ \frac{\partial r}{\partial Z} & \frac{\partial r}{\partial Y} & \frac{\partial r}{\partial X} \\ \frac{\partial c}{\partial Z} & \frac{\partial c}{\partial Y} & \frac{\partial c}{\partial X} \end{bmatrix} \begin{bmatrix} \Delta Z_u \\ \Delta Y_u \end{bmatrix} \\
\begin{bmatrix} Z_u \\ Y_u \end{bmatrix} &= \begin{bmatrix} r_{i1} - r_i \\ r_{i2} - r_j \\ r_{i3} - r_k \\ r_{i4} - r_l \end{bmatrix}
\end{align*}$$

Thus, the initial approximate values of object point coordinates $X_u$, $Y_u$, and $Z_u$ can be obtained by solving Equation 7 using the least-squares adjustment. When the coordinates are normalized, it is found that for pushbroom imagery (e.g., IKONOS and SPOT) the coefficient values of the constant and first-order terms in both the numerator and the denominator are larger by many orders of magnitude than those of the second- and third-order terms. Therefore, this method is suitable for spaceborne pushbroom-type imagery.

For frame camera imagery, the coefficients of the constant and first-order terms in both the numerator and the denominator do not dominate when compared with those of the second- and third-order terms. As a result, the initial values obtained using this method may often result in divergence of the correction computations with Equation 6. Therefore, for frame camera imagery, we use the median values of the three object coordinate ranges to start the reconstruction, for examples, three zeros for normalized coordinate ranges $[-1, 1]$. Where

$$\begin{align*}
\frac{\partial r}{\partial Z} &= \frac{\partial r}{\partial Y} = \frac{\partial r}{\partial X} = \frac{\partial c}{\partial Z} = \frac{\partial c}{\partial Y} = \frac{\partial c}{\partial X} = 0 \\
\frac{\partial r}{\partial Z} &= \frac{\partial r}{\partial Y} = \frac{\partial r}{\partial X} = \frac{\partial c}{\partial Z} = \frac{\partial c}{\partial Y} = \frac{\partial c}{\partial X} = 0
\end{align*}$$

and the remaining partial derivatives are similarly calculated. The least-squares solution is

$$x = [\Delta Z_u \, \Delta Y_u \, \Delta X_u]^T = (A^TW)^{-1} A^T W l$$

where $W$ is the weight matrix for the image points.

#### Determination of the Initial Approximate Values

The remaining problem is that initial approximate values of object point coordinates $X_u$, $Y_u$, and $Z_u$ should be used to start the computation. One method to obtain these initial values is to solve the RFM, using only the first-order terms and omitting the second- and third-order terms, i.e., Equation 1 is reduced to be

$$\begin{align*}
r &= a_0 + a_1 Z_u + a_2 Y_u + a_3 X_u \\
b &= b_0 + b_1 Z_u + b_2 Y_u + b_3 X_u \\
c &= c_0 + c_1 Z_u + c_2 Y_u + c_3 X_u \\
d &= d_0 + d_1 Z_u + d_2 Y_u + d_3 X_u
\end{align*}$$

and the remaining symbols are similarly determined.

Then the four error equations for two corresponding image points are

$$\begin{align*}
\begin{bmatrix} v_{i1} \\ v_{i2} \\ v_{i3} \\ v_{i4} \end{bmatrix} &= \begin{bmatrix} a_1 - r_{i1} & a_2 - r_{i2} & a_3 - r_{i3} & a_0 - r_{i4} \\ c_1 - c_{i1} & c_2 - c_{i2} & c_3 - c_{i3} & c_0 - c_{i4} \\ a_1 - r_{j1} & a_2 - r_{j2} & a_3 - r_{j3} & a_0 - r_{j4} \\ c_1 - c_{j1} & c_2 - c_{j2} & c_3 - c_{j3} & c_0 - c_{j4} \end{bmatrix}
\end{align*}$$

Thus, the initial approximate values of object point coordinates $X_u$, $Y_u$, and $Z_u$ can be obtained by solving Equation 7 using the least-squares adjustment. When the coordinates are normalized, it is found that for pushbroom imagery (e.g., IKONOS and SPOT) the coefficient values of the constant and first-order terms in both the numerator and the denominator are larger by many orders of magnitude than those of the second- and third-order terms. Therefore, this method is suitable for spaceborne pushbroom-type imagery.

For frame camera imagery, the coefficients of the constant and first-order terms in both the numerator and the denominator do not dominate when compared with those of the second- and third-order terms. As a result, the initial values obtained using this method may often result in divergence of the correction computations with Equation 6. Therefore, for frame camera imagery, we use the median values of the three object coordinate ranges to start the reconstruction, for examples, three zeros for normalized coordinate ranges $[-1, 1]$. Where
Test Results and Evaluation

Aerial Photograph Data Test

The two methods were tested using an aerial photography stereo pair provided by ERDAS Inc. The original stereo pair at a scale of 1:40,000 was taken over the Colorado Springs, Colorado area, and both photos were scanned at 100 μm per pixel. The overlap between the two images was about 68 percent. The scanned size was 2313 by 2309 pixels, and the ground pixel size was about 4.5 meters. The relief range was from 1847 meters to 2205 meters. A photogrammetric bundle block adjustment with OrthoBASE was done by ERDAS, and the rigorous collinearity equations with orientation parameters for both images were obtained. The average standard deviations after adjustment in object space were \( (m_x, m_y, m_z) = (1.7008, 2.1577, 0.2957) \) meters at five control points, and \( (m_x, m_y, m_z) = (4.2964, 0.7726, 3.8165) \) meters at one checkpoint. In the overlapping area of the stereo pair, 7499 conjugate points in left and right images were available, and the 3D coordinates of the corresponding 7499 object points were intersected using the rigorous collinearity equations. Figure 3 shows a 3D view of the terrain as well as the distribution of these object points on the terrain.

Accuracy of RFM Fitting

To solve for the RFCs for each image, the terrain-independent approach was used. A 3D control grid and a check grid in object space, as well as their corresponding image grids were generated using the rigorous collinearity equations. The image grid contained 11 by 11 points across the full extent of each image. The 3D control grid contained five terrain layers, each with 11 by 11 points, and the 3D check grid contained ten terrain layers, each with 20 by 20 points. The layers covered the full range of terrain relief. The unknown RFCs in Equation 1 and Equation 2 were determined, respectively, using the image grid points and the 3D object grid points. For the inverse RFM, the accuracy of the solution was checked in the object space, while, for the forward RFM, the accuracy was compared against the check grid in the image.

Tables 1 and 2 list the RFM accuracy results at the check grid for the left and right images with, respectively, the inverse and forward forms of the RFM. Both RMS errors and maximum absolute errors are given. Only the results calculated by the third-order RFM are provided. In Table 1, the notion of \( p6 = p6 \) means that the same denominator is used for \( X \) and \( Y \) in Equation 2. In the case of \( p6 = p6 = 1 \), the RFM becomes a regular third-order polynomial form. For comparison purposes, three cases, \( p6 = p6 \), \( p6 = p6 = 1 \), and \( p6 = p6 = 1 \), are all provided for the inverse RFM (Table 1) and for the forward RFM (Table 2).

It is found that both the inverse and the forward RFMs provide very high fitting accuracy to the collinearity equation.

<table>
<thead>
<tr>
<th>Case</th>
<th>Normalization</th>
<th></th>
<th>Un-normalization</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( X )</td>
<td>( Y )</td>
<td>( X )</td>
<td>( Y )</td>
</tr>
<tr>
<td>Left Image</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( p6 \neq p6 )</td>
<td>( 5.29 \times 10^{-09} )</td>
<td>( 2.19 \times 10^{-09} )</td>
<td>( 2.85 \times 10^{-03} )</td>
<td>( 6.50 \times 10^{-03} )</td>
</tr>
<tr>
<td>( p6 = p6 )</td>
<td>( 6.55 \times 10^{-07} )</td>
<td>( 1.04 \times 10^{-07} )</td>
<td>( 6.48 \times 10^{-03} )</td>
<td>( 1.92 \times 10^{-03} )</td>
</tr>
<tr>
<td>( p6 = p6 = 1 )</td>
<td>( 4.71 \times 10^{-06} )</td>
<td>( 3.59 \times 10^{-06} )</td>
<td>( 4.13 \times 10^{-03} )</td>
<td>( 3.48 \times 10^{-03} )</td>
</tr>
<tr>
<td>Right Image</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( p6 \neq p6 )</td>
<td>( 1.54 \times 10^{-05} )</td>
<td>( 1.97 \times 10^{-05} )</td>
<td>( 2.07 \times 10^{-02} )</td>
<td>( 1.89 \times 10^{-02} )</td>
</tr>
<tr>
<td>( p6 = p6 )</td>
<td>( 3.16 \times 10^{-04} )</td>
<td>( 2.14 \times 10^{-04} )</td>
<td>( 1.06 \times 10^{-01} )</td>
<td>( 7.60 \times 10^{-01} )</td>
</tr>
<tr>
<td>( p6 = p6 = 1 )</td>
<td>( 1.90 \times 10^{-03} )</td>
<td>( 1.81 \times 10^{-03} )</td>
<td>( 1.85 \times 10^{-03} )</td>
<td>( 1.88 \times 10^{-03} )</td>
</tr>
<tr>
<td>( p6 = p6 = 1 )</td>
<td>( 6.62 \times 10^{-02} )</td>
<td>( 1.05 \times 10^{-02} )</td>
<td>( 1.29 \times 10^{-01} )</td>
<td>( 1.01 \times 10^{-01} )</td>
</tr>
<tr>
<td></td>
<td>( 1.53 \times 10^{-01} )</td>
<td>( 1.07 \times 10^{-01} )</td>
<td>( 1.09 \times 10^{-03} )</td>
<td>( 1.76 \times 10^{-03} )</td>
</tr>
<tr>
<td></td>
<td>( 9.01 \times 10^{-01} )</td>
<td>( 4.55 \times 10^{-01} )</td>
<td>( 7.33 \times 10^{-03} )</td>
<td>( 4.11 \times 10^{-03} )</td>
</tr>
<tr>
<td></td>
<td>( 9.74 \times 10^{-03} )</td>
<td>( 6.80 \times 10^{-03} )</td>
<td>( 1.45 \times 10^{-04} )</td>
<td>( 6.84 \times 10^{-03} )</td>
</tr>
<tr>
<td></td>
<td>( 5.56 \times 10^{-04} )</td>
<td>( 4.57 \times 10^{-04} )</td>
<td>( 5.93 \times 10^{-04} )</td>
<td>( 4.56 \times 10^{-04} )</td>
</tr>
</tbody>
</table>

Figure 3. 3D view of the test data and check points.

Another method to obtain the approximate values of object point coordinates is to perform a 3D reconstruction using the inverse RFM form, described in the previous section. The values obtained will then be used as initial approximations for the forward reconstruction.

Reconstruction Procedure

Now we sketch the procedure that can be used to compute the object point coordinates from a pair of corresponding points \((r_x, c_x)\) and \((r_y, c_y)\) in the image.

1. Determine the initial approximate values for the object point coordinates \(X_o, Y_o, Z_o\) by solving Equation 7, by specifying the median values of the three object coordinate ranges, or by the reconstruction results from the inverse RFM, depending on the type of imagery.
2. Calculate the corrections \(\Delta X_o, \Delta Y_o, \Delta Z_o\) by computing Equation 6, then add them to \(X_o, Y_o, Z_o\).
3. Repeat Step 2 until the specified maximum number of iterations (e.g., ten) has been reached or \(\Delta X_o, \Delta Y_o, \Delta Z_o\) all converge.

In our experiments, the above procedure always converged when appropriate initial values were given. When the initial approximate values for \(X_o, Y_o, Z_o\) are obtained by solving Equation 7 or set to be the median values of the ground coordinate ranges, eight iterations are usually enough to converge (a threshold of \(1.0 \times 10^{-12}\) meters was used in our testing). When the initial approximate values are obtained from the result of inverse RFM reconstruction, two iterations are usually enough.
model. This is understandable because the collinearity equations model is a special form of the RFM. The rational polynomial form can produce better accuracy than the regular polynomial form, observed from Tables 1 and 2. It is also found that the use of the coordinate normalization technique can achieve results with much better accuracy than ones without normalization. Applying normalization, for the cases with different denominators, the maximum absolute errors at the check grid points are $6.55 \times 10^{-8}$ meters for the inverse RFM and $5.52 \times 10^{-8}$ pixels for the forward RFM, respectively.

### Accuracy of 3D Reconstruction

The 7499 conjugate points in the left and right images were input to both the inverse RFM and the forward RFM reconstructions. All the 7499 3D object points were used to check the accuracy of the reconstructed object points. The results in Tables 3a and 3b are computed using Equations 4 and 5, respectively. The results show that use of the correction Equation 4 can produce higher accuracy to some extent. In Table 4, the median values of the three object coordinate ranges were used to start the reconstruction for the $p_2 = p_4$ and $p_2 = p_4 = 1$ cases, while the initial values solved for with Equation 7 were used for the case of $p_2 = p_4 = 1$. Again, use of normalization can obtain much better results than un-normalization for both the inverse and the forward RFM reconstruction. The results also show that no significant differences are found between the use of different denominators and the same denominator for the frame sensor. Applying normalization, for the cases

### Table 2. RMS (Max.) Errors at the Image Check Grid with Forward RFM Fitting (Unit: Pixels)

<table>
<thead>
<tr>
<th>Case</th>
<th>Column</th>
<th>Row</th>
<th>Un-normalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left Image</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p_2 \neq p_4$</td>
<td>1.16 $\times 10^{-11}$</td>
<td>4.40 $\times 10^{-12}$</td>
<td>3.64 $\times 10^{-04}$</td>
</tr>
<tr>
<td>$p_2 = p_4$</td>
<td>1.02 $\times 10^{-09}$</td>
<td>2.44 $\times 10^{-09}$</td>
<td>1.98 $\times 10^{-09}$</td>
</tr>
<tr>
<td>$p_2 = p_4 = 1$</td>
<td>9.44 $\times 10^{-04}$</td>
<td>1.39 $\times 10^{-04}$</td>
<td>1.01 $\times 10^{-04}$</td>
</tr>
</tbody>
</table>

| Right Image |        |           |                  |
| $p_2 \neq p_4$ | 3.34 $\times 10^{-10}$ | 1.03 $\times 10^{-09}$ | 2.25 $\times 10^{-04}$ |
| $p_2 = p_4$  | 4.07 $\times 10^{-11}$ | 1.92 $\times 10^{-10}$ | 4.96 $\times 10^{-04}$ |
| $p_2 = p_4 = 1$ | 1.01 $\times 10^{-05}$ | 1.48 $\times 10^{-05}$ | 1.03 $\times 10^{-03}$ |

### Table 3a. RMS (Max.) Errors at 7499 Checkpoints with Inverse RFM Reconstruction (1) (Unit: Meters)

<table>
<thead>
<tr>
<th>Case</th>
<th>Normalization</th>
<th>Un-normalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_6 \neq p_8$</td>
<td>$4.05 \times 10^{-06}$</td>
<td>$1.96 \times 10^{-06}$</td>
</tr>
<tr>
<td>$p_6 = p_8$</td>
<td>$3.31 \times 10^{-08}$</td>
<td>$1.53 \times 10^{-08}$</td>
</tr>
<tr>
<td>$p_6 = p_8 = 1$</td>
<td>$9.20 \times 10^{-06}$</td>
<td>$6.23 \times 10^{-05}$</td>
</tr>
</tbody>
</table>

### Table 3b. RMS (Max.) Errors at 7499 Checkpoints with Inverse RFM Reconstruction (2) (Unit: Meters)

<table>
<thead>
<tr>
<th>Case</th>
<th>Normalization</th>
<th>Un-normalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_6 \neq p_8$</td>
<td>$4.37 \times 10^{-06}$</td>
<td>$2.96 \times 10^{-06}$</td>
</tr>
<tr>
<td>$p_6 = p_8$</td>
<td>$3.46 \times 10^{-06}$</td>
<td>$1.65 \times 10^{-06}$</td>
</tr>
<tr>
<td>$p_6 = p_8 = 1$</td>
<td>$7.94 \times 10^{-05}$</td>
<td>$5.37 \times 10^{-05}$</td>
</tr>
</tbody>
</table>

### Table 4. RMS (Max.) Errors at 7499 Checkpoints with Forward RFM Reconstruction (Unit: Meters)

<table>
<thead>
<tr>
<th>Case</th>
<th>Normalization</th>
<th>Un-normalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_2 \neq p_4$</td>
<td>$2.47 \times 10^{-06}$</td>
<td>$1.89 \times 10^{-06}$</td>
</tr>
<tr>
<td>$p_2 = p_4$</td>
<td>$3.41 \times 10^{-07}$</td>
<td>$2.35 \times 10^{-07}$</td>
</tr>
<tr>
<td>$p_2 = p_4 = 1$</td>
<td>$6.60 \times 10^{-07}$</td>
<td>$4.50 \times 10^{-07}$</td>
</tr>
</tbody>
</table>
The stereo accuracy for these images is 25 m. The images were calibrated to the standard level without ground control. Two panchromatic 11-bit images that had been geometrically corrected by Space Imaging were determined by the terrain-independent approach. The physical Ikonos sensor model was derived from the satellite ephemeris and attitude without using ground control points. The regular polynomial based reconstruction method was used for 3D reconstruction, a vector data set was obtained from the CDAL (Canadian Geospatial Data Infrastructure) Data Alignment Layer. The CDAL Data Alignment Layer (CDAL) data set consists of intersection points that were derived from the National Topographic Database (NTDB) toposheets. The absolute accuracy analysis, the CDAL intersection points were used because the horizontal accuracy of the CDAL intersection points is specified as 10 m (CDAL, 1999) which is roughly equivalent to a 4.1-m RMSE.

The line (row) and sample (column) coordinates for 28 conjugate points were acquired in the overlapped area of the left and right images in Scene 1. The conjugate points were carefully selected so that they would correspond to the CDAL road intersection points. Any road intersections that seemed ambiguous were discarded. The line and sample coordinates were passed into the RFM 3D reconstruction software to calculate their latitude, longitude, and height coordinates. To facilitate accurate distance measurements, the CDAL points and the RFM output points were both projected into the Universal Transverse Mercator (UTM) projection. The horizontal distances between the two data sets were calculated and the statistics are summarized in Table 6. The result derived by the forward RFM method was more accurate than that with normalization applied. The only exception is that the regular polynomial form may obtain better or the same reconstruction accuracy if the reconstruction method is reliable. The physical Ikonos sensor model was derived from the satellite ephemeris and attitude without using GCPs. The line and sample coordinates were passed into the RFM 3D reconstruction software to calculate their latitude, longitude, and height coordinates. To facilitate accurate distance measurements, the CDAL points and the RFM output points were both projected into the Universal Transverse Mercator (UTM) projection. The horizontal distances between the two data sets were calculated and the statistics are summarized in Table 6. The result derived by the forward RFM method was more accurate than that with normalization applied.

### Table 5. A Comparison of Two Reconstruction Methods

<table>
<thead>
<tr>
<th>Comparison Items</th>
<th>Inverse RFM</th>
<th>Forward RFM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fitting Accuracy (RMSE in Pixels)*</td>
<td>$10^{-3}$ ~ $10^{-7}$</td>
<td>$10^{-12}$ ~ $10^{-9}$</td>
</tr>
<tr>
<td>Reconstruction Accuracy (RMSE in Meters)*</td>
<td>$10^{-6}$ ~ $10^{-7}$</td>
<td>$10^{-9}$</td>
</tr>
<tr>
<td>Initial Values Required</td>
<td>Z</td>
<td>X, Y, Z</td>
</tr>
<tr>
<td>Method to Obtain Initial Values</td>
<td>Set a priori</td>
<td>Linear solution or set a priori</td>
</tr>
<tr>
<td>Sensitivity to Initial Values</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Corrections Computed</td>
<td>$\Delta Z$</td>
<td>$\Delta X, \Delta Y, \Delta Z$</td>
</tr>
<tr>
<td>Need Matrix Inversion</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Need Iteration</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Convergence Speed</td>
<td>Fast, $\leq 3$</td>
<td>Slow, eight or more iterations</td>
</tr>
</tbody>
</table>

*for the rational function case with normalization

### Table 6. Absolute Accuracy Assessment of 3D Reconstruction (in Meters)

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Horizontal Accuracy</th>
<th>Vertical Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Error</td>
<td>6.83</td>
<td>-3.79</td>
</tr>
<tr>
<td>RMS Error</td>
<td>7.15</td>
<td>4.23</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>2.15</td>
<td>1.91</td>
</tr>
<tr>
<td>Minimum Absolute Error</td>
<td>2.37</td>
<td>0.97</td>
</tr>
<tr>
<td>Maximum Absolute Error</td>
<td>10.78</td>
<td>6.36</td>
</tr>
</tbody>
</table>

**Absolute Accuracy Test**

To measure the horizontal absolute accuracy of the RFM 3D reconstruction, a vector data set was obtained from the CDAL. The CDAL data set consists of intersection points that were derived from the National Topographic Database (NTDB) toposheets. For the absolute accuracy analysis, the CDAL intersection points were used because it was hoped that the road intersections could be identified more easily than other CDAL features on the Ikonos imagery. The horizontal accuracy of the CDAL intersection points is specified as 10 m (CDAL, 1999) which is roughly equivalent to a 4.1-m RMSE.

**IKONOS Scene 2**

The line (row) and sample (column) coordinates for 28 conjugate points were acquired in the overlapped area of the left and right images in Scene 1. The conjugate points were carefully selected so that they would correspond to the CDAL road intersection points. Any road intersections that seemed ambiguous (i.e., an accurate position could not be established) were discarded. The line and sample coordinates were passed into the RFM 3D reconstruction software to calculate their latitude, longitude, and height coordinates. To facilitate accurate distance measurements, the CDAL points and the RFM output points were both projected into the Universal Transverse Mercator (UTM) projection. The horizontal distances between the two data sets were calculated and the statistics are summarized in Table 6. The result derived by the forward RFM method was more accurate than that with normalization applied.
demonstrates the consistency to the specification from Space Imaging (Grodecki, 2001b).

For the vertical absolute accuracy analysis, a DEM from the Canadian Digital Elevation Data (CDED) was obtained for the test area. The product specification states that the vertical accuracy is dependent on the accuracy of the original NTDB topographic map contours that were scanned to produce the DEM. The CDAL intersection points were overlaid on the DEM to obtain orthometric heights (relative to the Canadian Geodetic Vertical Datum of 1928) at each of these points. Using software (i.e., GPS-HT) supplied by the Canadian Geodetic Survey Division, the heights were converted into ellipsoidal heights relative to the WGS84 ellipsoid. A comparison was done between these heights and those computed by the RFM 3D reconstruction software, and the statistics are listed in Table 6 (under the vertical accuracy column).

Relative Accuracy Test

The relative horizontal accuracy assessment was done by computing the dimensions of several buildings (width and length) using the RFM calculated corner coordinates and comparing them to the known distances. The RMS error of the building dimension is 0.52 m. More errors occur along the line direction.

To obtain a measure of the relative vertical accuracy of the RFM 3D reconstruction, conjugate points were collected on the rooftops of two buildings in Scene 1 and Scene 2. The 41 points on the rooftop of the building in Scene 2 are shown in Figure 5. Under the assumption that the building roof tops were level, a plane function was fit to each set of building points using a least-squares method. The RMS errors of the surface-fitting residuals for both buildings are 0.68 m. The tests show that the relative accuracy in both the horizontal and vertical directions is at a sub-meter level for Ikonos standard stereo products.

The above test validates that the reconstruction method described is correct and applicable to Ikonos imagery with the supplied RFCs. For a complete assessment of the Ikonos mapping capability, more comprehensive tests are required. Interested readers may refer to Fraser et al. (2001) and Baltavias et al. (2001).

Concluding Remarks

This study compared two methods for RFM-based 3D reconstruction, the inverse RFM method and the forward RFM method, in terms of the fitting accuracy, reconstruction accuracy, initial value determination, sensitivity to the initial values, iterations, normalization effects, etc. We have also compared them in different cases (with and without denominators, as well as with a regular polynomial form). The technique to determine the initial values for the RFM solution is also investigated.

Based on the tests, both methods converge well and are not sensitive to their initial approximate values. The forward RFM can produce better fitting and reconstruction results. However, the inverse RFM method is faster with fewer iterations. An important experience is that the normalization technique can improve the reconstruction accuracy for both the inverse and the forward RFM methods. In order to speed the convergence of the forward RFM method and also to improve the final reconstruction accuracy, one can use the reconstruction result from the inverse RFM as an initial value input to the forward RFM.

The RFM has been considered by the OpenGIS Consortium (OGC, 1999) as a part of the standard image transfer format due to its characteristic of sensor independence. We feel that use of the RFM would be a key driving force towards its interoperability with the image exploitation software (Tao and Hu, 2001a). In fact, one can develop a software package with a generic interface to handle the RFM for various sensors, provided that the values of the RFCs and normalization parameters are supplied. This is very beneficial in terms of making photogrammetric processing interoperable. If each sensor image comes with a set of RFCs (solved or supplied by the data vendor), end users will be able to perform the subsequent photogrammetric processing without knowing the sophisticated physical sensor model. Driven by this fact, we have developed a software package, Rational Mapper, which utilizes the RFM for image exploitation including ortho-rectification and 3D reconstruction. Figure 6 shows a web-based interface for 3D reconstruction with the Rational Mapper.

With additional control points available, one can improve the RFC accuracy provided. Hu and Tao (2001) have proposed an incremental method based on Kalman filtering to improve the solution of the RFCs. Thus, the 3D reconstruction result can be further improved using the updated values of the RFCs.

Acknowledgments

We would like to acknowledge that Mr. Steve Schnick performed the Ikonos data tests for this paper. We sincerely thank Dr. Xinghe Yang, ERDAS, Inc. and Dr. Clive Fraser, University...
of Melbourne, for the valuable information exchange and constructive comments. Special thanks go to Dr. Bob Truong, the Canadian Nuclear Safety Commission, Ottawa, for providing the Ikonos data set and Dr. Younian Wang, ERDAS, Inc. for providing the aerial stereo pair.

References


Jeanne Jack J.R. Arie Croitoru and DongMei Tianen Yerach Calemie.


Zhilin Li, Xiuqiao Yuan, and Kent W.K. Lam. Effects of JPEG Compression on the Accuracy of Photogrammetric Point Determination.


Fabio Maselli. Improved Estimation of Environmental Parameters through Locally Calibrated Multivariate Regression Analysis in Cloud-Shadow Discrimination from Landsat Images.


Jia Zong, Roger Davies, and Jan-Peter Muller. Photogrammetric Retrieval of Cloud Advection and Top Height from the Multi-Angle Imaging Spectroradiometer (MISR).


(Received 14 September 2001; revised and accepted 05 February 2002)