RFM-Based Image Matching Technique Using Object-Space Approach

**Abstract:** The Rational Function Coefficients (RFCs) provided with the high-resolution satellite imagery show an accuracy that is relatively low and inconsistent with their original image quality. Also, there have only been a few detailed researches about 3-D terrain data generation using Rational Function Model (RFM). In this study, we suggest an improved matching algorithm which uses RFM as an image to object transformation model and introduce the detailed procedure of generating 3-D surface models from the RFM for non-experts in photogrammetry and remote sensing. The proposed matching scheme constructs the Piecewise Matching Line (PML) by using elevation range in the object space. The constructed PML determines the initial location and confines the search space in image matching. The proposed scheme is applied to SPOT-3 stereo images which supply the detailed satellite orbital information. The RFCs of SPOT-3 are obtained using the terrain-independent approach. The 3-D positional accuracy of the calculated RFCs is about RMSE of 2 meter compared with that of the physical sensor model. Finally, DEM over test area is generated using the proposed matching scheme. DEMs generated from the physical sensor model and from 3 arc-second Digital Terrain Elevation Data (DTED) are used to evaluate the relative and absolute accuracies of the generated DEM. The result shows RMSE of 7.51 meter in relative accuracy and 16.50 meter in absolute accuracy.

**Keywords:** Rational Function Coefficients (RFCs), SPOT-3, Digital Elevation Model (DEM), Rational Function Model (RFM), 3 arc-second DTED.
ond DTED とを用いて制作された既製の DEM との比較を通じて相対的精度と絶対的精度を調べる事にした。そしてその結果、それぞれ7.51m の標準偏差値、16.50m の標準偏差値を示した。

1. Introduction

As the first commercial high-resolution satellite IKONOS-2 has been launched in September 1999, the field of photogrammetry and remote sensing faced a new era. The ground resolution of satellite imagery was remarkably improved and various users are able to acquire more accurate 3-D terrain data. However, in order to protect confidential information about sensor, the vendors of the high-resolution satellite imagery have not supplied the ephemeris data or have supplied the ephemeris data that has a lower accuracy than the original acquired level (QuickBird Product Guide, 2003). As a result, replace sensor model, which enables the general users to process the satellite imagery without having to expose the detailed satellite information, has been proposed (OGC, 1999; Tao and Hu, 2000). Rational Function Model (RFM), one of the replace sensor models, has been widely used in IKONOS-2 satellite. RFM defines the geometric relationship between the image space and object space in the form of the polynomial ratios.

RFM has the advantages to perform a series of photogrammetric processes without a complicated algorithm. It can also calculate the 3-D position of the surface, create a Digital Elevation Model (DEM), and generate ortho-images much faster than the conventional sensor models. After RFM has been verified as a replace sensor model (Madani, 1999; Dowman and Dolloff, 2000; Tao and Hu, 2001(a)), the accuracy of 3-D positioning using RFM has been studied by several researchers (Yang, 2000; Grodecki and Dial, 2001; Tao and Hu, 2001(b); Sohn et al., 2004). Also the 3-D positioning accuracy of inverse RFM and forward RFM have been compared (Tao and Hu, 2002). The above mentioned studies have proved that RFM can satisfy several important standards of the replace sensor model such as general tools for mapping, accurate 3-D positioning, and real-time processing.

In order to acquire high quality terrain data, the precise Rational Function Coefficients (RFCs) must be supplied and a solution for DEM generation using the RFM has to be developed. However, there have only been a few detailed publications about DEM generation using RFM. Especially, it is difficult to find studies related to image matching using RFM, even though it is a critical step in generating DEM.

Image matching for satellite imagery is not an easy work in case that the geometric differences between the stereo images are quite large. An epipolar solution, which is popular in aerial photograph, has some difficulties to be applied to satellite imagery due to different geometric characteristics (Otto, 1988; Baltsavias, 1993; Kim, 2000). To resolve this problem, an alternative method using the geometric model between image space and object space is developed (Baltsavias, 1993). It puts constraint condition on search space for image matching. The matching accuracy from SPOT imagery can be upgraded to a pixel level by the proposed method. However, this technique needs a physical sensor model and ephemeris data about the sensor that the current high-resolution satellite imagery does not provide to the general users.

This paper offers an efficient method to resolve the matching problems for the high-resolution satellites that adopt RFM as a basic sensor model. Especially, in order to improve the efficiency and quickness in image matching, the Piecewise Matching Line (PML) is proposed using object-space approach. PML helps to determine initial position and reduce the search space for image matching. In the following sections, the detailed processes for the proposed matching scheme are described. Also, the overall procedure for DEM generation using the RFCs is presented. Finally, the accuracy of generated DEM using the proposed scheme is presented.
2. Rational Function Model

RFM is usually classified into terrain-independent and terrain-dependent cases, depending on whether a physical sensor model is known or not. In addition, we can further classify RFM into the image-space oriented or the object-space oriented model so called the forward RFM and the backward RFM, respectively (Yang, 2000; Tao and Hu, 2002).

In RFM, image coordinates are expressed as the ratios of polynomials of object coordinates. The image and object coordinates are both offset and scaled to fit the range from -1.0 to +1.0 over the entire image in order to minimize computational errors during matrix operations. From the object space to image space transformation, the equation represented ratios of polynomials can be defined and is called the forward RFM

\[ r = \frac{p_r(X, Y, Z)}{p_d(X, Y, Z)}, \quad c = \frac{p_c(X, Y, Z)}{p_d(X, Y, Z)} \]  

(1)

where \( r \) and \( c \) are the normalized row and column index of pixels in image space. \( X, Y, \) and \( Z \) are the normalized coordinates in object space.

In Equation 1, polynomials \( p_i \) have a general form as shown in Equation 2.

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

(2)

where the \( a_{ijk} \) are polynomial coefficients called the RFCs and \( m_1, m_2, \) and \( m_3 \) are the maximum power of polynomial typically limited to 3.

The forward form, Equation 1, expresses the transformation from the object space to the image space. For the image space to object space transformation, an inverse form of Equation 1 can be used as written in Equation 3 (Yang, 2000).

\[ X = \frac{p_r(r, c, Z)}{p_d(r, c, Z)}, \quad Y = \frac{p_c(r, c, Z)}{p_d(r, c, Z)} \]  

(3)

Equation 3 is so called the backward RFM and expresses the horizontal object coordinates as the ratios of image coordinates \((r, c)\) and the vertical object coordinates \((Z)\).

Once the RFCs were calculated or supplied, 3-D object coordinates of the conjugate points in stereo images can be computed from the space intersection. Space intersection can be done by either the forward RFM or the backward RFM. Both methods provide good results, but it is reported that the forward RFM shows better accuracy in most cases (Tao and Hu, 2002).

3. RFM-Based Image Matching Using Object-Space Approach

For keeping the good matching performance and extracting proper seed points, a two-step preprocessing is applied for the original images. First, the periodic and pattern noise is eliminated. Second, the Wallis filter, which is an adaptive nonlinear local contrast enhancement filter, is applied for effective extraction of the seed points. In extracting seed points, the Canny operator is used. The Canny operator is the most widely used as an edge detector. It has the ability to mark all real edges and has only one response per edge (Canny, 1986).

The extracted seed points on the left image can be re-projected to the right image through the two-step transformations, image-to-object and object-to-image transformation, and the search space on the right image is confined according to the elevation range of the target area. The following section describes the image-to-object and the object-to-
image transformation based on the forward RFM. Figure 1 depicts the overall procedures of the proposed RFM-based image matching scheme.

3.1 Image-to-Object Transformation

Forward RFM represent the geometric relationship between the image space and object space. If the RFCs are supplied for the stereo images and the conjugate points on image are known, the corresponding object point can be uniquely defined. However, if only one image point is known, the object point does not have a unique solution. In other words, in order to set up the image-to-object transformation using one image point, one unknown variable must be eliminated. Assuming that the elevation $Z$ is a known value, an approximate transformation from reference image space to object space can be constructed. The assumed elevation range is divided into several levels. For each elevation level, two error equations, such as Equation 4, can be obtained.

$$f(r)_{\text{left}} = p(X, Y, Z) - r_{\text{left}}$$

$$f(c)_{\text{left}} = p(X, Y, Z) - c_{\text{left}}$$

where $Z_i$ represents the elevation value of the target area or the approximate value when the elevation information is not available.

Equation 4 can be written as a matrix form as shown in the Equation 5. The object coordinates at various elevation levels are calculated from the least squares solution.

$$\begin{align*}
\begin{bmatrix}
\frac{\partial f(r)_{\text{left}}}{\partial X} & \frac{1}{X_{\text{scale}}} & \frac{1}{Y_{\text{scale}}} \\
\frac{\partial f(c)_{\text{left}}}{\partial X} & \frac{1}{X_{\text{scale}}} & \frac{1}{Y_{\text{scale}}} \\
\end{bmatrix}
\begin{bmatrix}
\Delta X_o \\
\Delta Y_o \\
\end{bmatrix}
\end{align*}
$$

$$= 
\begin{bmatrix}
\frac{\partial f(r)_{\text{left}}}{\partial X} & \frac{1}{X_{\text{scale}}} & \frac{1}{Y_{\text{scale}}} \\
\frac{\partial f(c)_{\text{left}}}{\partial X} & \frac{1}{X_{\text{scale}}} & \frac{1}{Y_{\text{scale}}} \\
\end{bmatrix}
\begin{bmatrix}
r_{\text{left}} - r_{\text{left}} \\
\hat{c}_{\text{left}} - \hat{c}_{\text{left}} \\
\end{bmatrix}$$

where $\partial$ means partial derivatives of the given function. $X_{\text{left}}, Y_{\text{left}}$ is the scale factor of left image. $\Delta X_o, \Delta Y_o$ is correction for the estimated value.

Since Equation 4 is non-linear, the initial values for unknown variables are required and the final solutions are obtained from the iterative method. The initial value of object points $(X_0, Y_0)$ is computed from first-order RFM. The iterative solution is stopped when $\Delta X_o$ and $\Delta Y_o$ is smaller than 2 meter which is based on the 3-D positional accuracy of SPOT-3 using RFM. The final calculated 3-D object coordinates become as follows.

$$X_o = X_0^l + \Delta X, \ Y_o = Y_0^l + \Delta Y, \ Z = Z_i$$

3.2 Object-to-Image Transformation

3-D object coordinates calculated in the previous section can be re-projected to the right image by forward RFM using the RFCs. This process is particularly important in that the search space on the right image is confined within the linear-shaped areas. The re-projected image coordinates on right image construct PML. Actual image matching is performed along the PML. This proposed scheme makes it simpler and faster to perform image matching than the hyperbolic epipolar solution for satellite imagery. Figure 2 shows PML on the right image at each elevation level.

To determine the PML, the linear equation in each elevation level, Equation 7, is used.

$$i = \frac{i_{\text{max}} - i_{\text{min}}}{j_{\text{max}} - j_{\text{min}}} (j_i - j_{\text{max}}) + i_{\text{max}}$$

where $i_{\text{max}}$ and $j_{\text{max}}$ are the right image coordinates corresponding to the highest elevation in each level, $i_{\text{min}}$ and $j_{\text{min}}$ is the right image coordinates corresponding to the lowest elevation in each level.

![Figure 2. The PML depending on the height level of object space](image-url)
4. Experimental Results and Analysis

4.1 Dataset

Dataset used for testing the proposed algorithm is SPOT-3 panchromatic stereo images (Figure 3). The study area covers low relief terrain with the mountains and plains around the Daejeon Metropolitan City, located in the middle of South Korea. The test area (white rectangular area in Figure 3), about 15km × 15km, is selected for DEM generation using the proposed scheme. The elevation of test area ranges from 20 meter in the plains to 390 meter in the mountainous areas.

4.2 Generation of RFCs

The RFCs are calculated through the terrain-independent algorithm. A total of 36 GPS GCPs distributed over the study area construct the physical sensor model for SPOT-3 stereo images. The object grid consists with 11 by 11 on the overlapped areas. The overlap is about 90% as shown in Figure 3. The image grid for the corresponding object grid is calculated using the orientation parameters of the physical sensor model. The elevation of the study area is stratified into 6 elevation layers from −50m to 450m with interval of 100m. 726 GCPs (grid: 11 × 11, layer: 6) are contributed to set up the RFCs for the stereo images. In order to check the accuracy of the acquired RFCs, the 3-D object coordinates by the physical sensor model and those by forward RFM are compared for 1,543 check points. All check points are evenly distributed over the overlapped area. The 3-D positional differences between two sensor models are summarized in Table 1. In Table 1, case 1 means $p_2 = p_4$ on Equation 1, case 2 $p_2 = p_4$, and case 3 $p_2 = p_4 = 1$.

As shown in Table 1, three cases for $2^{nd}$ and two cases for $3^{rd}$ order polynomials are tested. The RMSE of X, Y, and Z coordinates are about 2m except for the case 1 of $3^{rd}$ order. In case 1 of $3^{rd}$, the least squares solution does not reach the convergence and the RFCs are not obtained. This may be due to high correlation among the parameters. We confirmed that the five types of RFMs show the acceptable accuracy because the error in all cases does not exceed 10 meter (1 pixel). This result is similar to the previous researches (Yang, 2000; Table 1. Accuracy of RFM compared with the physical sensor model in SPOT-3 stereo images

<table>
<thead>
<tr>
<th>ORDER</th>
<th>CASE</th>
<th>RMSE (m)</th>
<th>Mean Error (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>X</td>
<td>Y</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2.00</td>
<td>1.56</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.79</td>
<td>6.92</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2.02</td>
<td>1.54</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.74</td>
<td>7.25</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1.98</td>
<td>0.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.83</td>
<td>3.80</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>2.00</td>
<td>1.56</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.76</td>
<td>6.91</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>2.00</td>
<td>1.56</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.79</td>
<td>6.92</td>
</tr>
</tbody>
</table>
4.3 RFM-Based Image Matching

The RFM-based image matching scheme explained in section 3 was applied to SPOT-3 images of test area. In order to determine the PML, the assumed elevation range (0 to 400 meter) is divided into 4 different levels. Figure 4 (a) to (c) illustrates the PML and (d) shows the location of image point corresponding to (a), (b), and (c). The selected 3 points are located along the center of left image. The PML is calculated from the RFCs for case 2 of 3rd order.

The PML shows the similar patterns over the other parts of image as described in Figure 4 (a) to (c). Also, the same trends appear for all the other RFM cases. The maximum distance between the PML and the real conjugate point shows about 2 pixels. It verified that our image matching algorithm through object-space approach works well for the entire image and all RFM cases. The PML makes it help to find the initial position for image matching. In this experiment to guarantee a stable matching, the additional search space is extended to ±3 pixels around the PML.

A total of 222,445 seed points are extracted from the left image by the Canny operator. Four different elevation levels are applied to the selected test area as expressed in Figure 3. The final matching results for each elevation level are summarized in Table 2. In Table 2, MPs represents the total number of matched point in each elevation level and RMP represents the ratio of matched point which can be written as (the number of MPs)/(the number of Seed Points). Image matching for each level is begun with the same number of seed points (222,445 points).

About 1,000 final matched points are duplicated among elevation levels for all RFM cases. This is due to the fact that the actual conjugate point is located between the extended search space of the PML obtained from one elevation level and that of the PML obtained from the other elevation level.

The selection criteria for the final matched point
Table 2. Final matched points using the proposed algorithm

<table>
<thead>
<tr>
<th>Part</th>
<th>RFM Type</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>case1 for 2nd order</td>
<td>case2 for 2nd order</td>
<td>case3 for 2nd order</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>MPs</td>
<td>RMP</td>
<td>MPs</td>
<td>RMP</td>
<td>MPs</td>
<td>RMP</td>
</tr>
<tr>
<td>Seed Point</td>
<td>222445</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LVL. 1</td>
<td>193058</td>
<td>0.868</td>
<td>193096</td>
<td>0.868</td>
<td>193047</td>
<td>0.868</td>
</tr>
<tr>
<td>LVL. 2</td>
<td>20020</td>
<td>0.091</td>
<td>20029</td>
<td>0.090</td>
<td>20146</td>
<td>0.090</td>
</tr>
<tr>
<td>LVL. 3</td>
<td>5820</td>
<td>0.026</td>
<td>5797</td>
<td>0.026</td>
<td>5790</td>
<td>0.026</td>
</tr>
<tr>
<td>LVL. 4</td>
<td>4406</td>
<td>0.020</td>
<td>4402</td>
<td>0.020</td>
<td>4346</td>
<td>0.020</td>
</tr>
</tbody>
</table>

4.4 Generation of DEM

One of main purposes of image matching is to create various 3-D surface models. All matched points in Table 2 are transformed to 3-D object coordinates through the space intersection using Forward RFM. Final DEM over the study area is produced through the Kriging interpolation. Figure 5 shows the generated DEM from the RFM and that from the physical sensor model.

In order to assess the accuracy and check the performance of the proposed scheme, DEM generated from the RFM is compared with that from two different sources, one is DEM acquired from physical sensor model and the other is 3 arc-second Digital Terrain Elevation Data (DTED), which is 3-D terrain data format developed by National Geospatial-Intelligence Agency (NGA). Table 3 summarizes the elevation differences between DEM
Table 3. Accuracy of generated DEM (unit: m)

<table>
<thead>
<tr>
<th>Model</th>
<th>RMSE</th>
<th>Mean Error</th>
<th>Max. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical sensor model</td>
<td>7.51</td>
<td>5.25</td>
<td>143.55</td>
</tr>
<tr>
<td>DTED</td>
<td>16.50</td>
<td>15.13</td>
<td>229.04</td>
</tr>
</tbody>
</table>

from the RFM and the other two DEMs. DEM from the RFM shows a mean error of 5.25 meter and RMSE of 7.51 meter compared with that from the physical sensor model. This result sustains the pixel level accuracy checked in Table 1.

The comparison between DEM from the RFM and 3 arc-second DTED is to check the absolute accuracy. SPOT-3 stereo images are widely used to generate 1:50,000 topographic maps in general. Therefore, we compared with the similar contour interval data in a 1:50,000 map, that is, 3 arc-second DTED. This comparison shows a mean error of 15.13 meter and RMSE of 16.50 meter. It is lower accuracy than the study of Baltsavias (1993) using the physical sensor model, but can be recognized as a good outcome considering the fact that the full grid points are compared and only basic blunder detection is performed.

5. Conclusions

In this study, we proposed RFM-based image matching scheme using SPOT-3 stereo images. The RFCs for the proposed scheme are acquired from the terrain-independent method. The accuracy of acquired RFCs is checked for 1,543 points evenly distributed over the entire image. For a total of 5 cases RFCs, 3-D positional accuracy shows about RMSE of 2 meter in X, Y, and Z direction. It is confirmed that the result is suitable for generating DEM considering that of conventional method.

The generated PML using object-space approach is used to determine initial position and reduces the search space for image matching. This scheme leads the distance between the PML and the actual conjugate point to within 2 pixels. The RMP percentage based on seed points is over 85%. Therefore, the proposed algorithm is useful for high-resolution satellite image, especially when the epipolar solution is not impracticable or difficult to be applied.

DEM generated from the final matched points shows RMSE of 7.51 meter compared with that of the physical sensor model and RMSE of 16.50 meter compared with 3 arc-second DTED. As a result, the RFM-based image matching algorithm facilitates DEM generation for high-resolution satellite, which adopts the RFM as the basic sensor model, and is superior to the conventional matching algorithm considering the accuracy and real-time processing of RFM. Even though our scheme is only applied to the SPOT-3 stereo images, the most current high-resolution satellite images can utilize the scheme for generating 3-D terrain data.

References


(受付日2004.3.9, 受理日2004.7.9)
Methods of Sensor Orientation, Barcelona, Spain, pp.64–75.


