AUTOMATIC REGISTRATION OF SATELLITE IMAGE TO MAP

Kensaku FUJII
NTT Cyber Space Laboratories
1-1 Hikarinooka Yokosuka-shi, Kanagawa 239-0847, Japan
Tel:+81-468-59-2431, Fax:+81-468-59-2829, E-mail: fujii@marsh.hil.ntt.co.jp

Tomohiko ARIKAWA
NTT Cyber Space Laboratories

KEY WORDS: Registration, Satellite image, Vector map, Normalized Difference Vegetation Index, Generalized Hough Transform

ABSTRACT

This paper introduces a fully automatic method for registering satellite images to vector maps. The key idea is to automatically match the man-made objects extracted from satellite images against the corresponding objects in maps. The proposed method consists of two procedures: feature extraction for matching, and mismatch determination.

Matching features are extracted using indicators derived from multi-spectral data of satellite images. The normalized difference vegetation index is used as a spectral indicator. It is well known to provide the accurate differentiation of vegetation from man-made objects.

In order to minimize registration mismatch, we apply the voting technique based on the generalized Hough transform. Voting is used to judge the matching levels of displacement candidates. Mismatch is determined by voting using the positions of the features and those of topographical objects projected onto satellite images.

Our approach is validated in experiments using actual multi-spectral data of satellite images and standard topological maps. The experimental results show that the registration is extremely accurate; the registration mismatch is just a few pixels. This level of performance confirms that our approach can automatically register satellite images against maps. To quantify the average performance of the proposed approach, we also tested the approach of using image edges as matching features. These results reveal that our approach yields high accuracy and small mean square error. It is clear that our approach can extract effective features from satellite images because it makes good use of topographical reference objects through voting.

1. INTRODUCTION

The registration of airborne and spaceborne images plays a key role in the field of photogrammetry. The demand to automate such registration to maps is rapidly increasing in many applications such as agriculture, town planning, map making and updating, and disaster prevention.

Various sources can be used for GIS (Geographical Information System) applications. A very recent source for these applications is to use satellite images acquired by commercial satellite systems (Space Imaging), (ORBIMAGE). Recent progress in spaceborne sensing methods has made it possible to be easily acquired large amounts of data that offer excellent resolution (approximately a few meters). Therefore, satellite images are now one of the most important sources in GIS. Many approaches have been presented for the semi-automatic or automatic interpretation of satellite images (Vogtne & Steinel, 2000). (Klang, 1997). With regard to registration to maps, however, several problems remain.

The problems fall into two classes: problems with geometry mismatch and problems with image processing, i.e. feature extraction from satellite images. The former is mainly due to the fact that satellite images have some errors in geometry caused by mechanical device limitations or measurement error. With the rapid development in remote sensing, highly accurate geometry can be achieved by direct geo-referencing. However, the best level of position alignment accuracy equals that of 1/25,000 scale maps without Ground Control Point (GCP) correction. Since 1/25,000 scale maps are now being used as GIS sources, the use of GCP will always be needed for registration.

Due to the characteristic of satellite images, it is difficult to process images i.e. feature extraction (Vohra & Dowman, 2000); (Hill et al., 2000). Digital image processing becomes an important tool for the quantitative and statistical analysis of remotely sensed images (Ehner et al., 1999), (Cirinc & Baltsavias, 1997). However, such images often contain complex natural scenes and their resolution is at most a few meters. This seriously impacts feature extraction, and erroneous extraction seriously degrades registration accuracy.

It is necessary to achieve the robust interpretation of such images.

This paper describes a fully automatic method for the registration of satellite images to vector maps. The method provides the automatic analysis of satellite images and objects in maps, and completely eliminates the need for a priori information. The general idea is to automatically match man-made objects extracted from satellite images against the corresponding objects in maps. The proposed method consists of two procedures: feature extraction for matching, and mismatch determination. Our approach to identifying man-made objects is based on an indicator analysis of multi-spectral data of satellite images. The normalized difference vegetation index is used as a spectral indicator. It allows vegetation areas to be differentiated from man-made objects (Schilling & Vogtne, 1996). Additionally, in order to minimize registration mismatch, we apply the voting technique based on the generalized Hough transform. We also present the results of automatic registration experiments using actual satellite images and vector maps.

This paper is arranged as follows. Section 2 briefly explains the procedure of matching feature extraction. Section 3 details the process of determining registration mismatch. The results of some experiments are presented in section 4. Finally, we conclude this paper with future directions of our research.

2. MATCHING FEATURE EXTRACTION

This section describes our approach to extract matching features from satellite images. The general idea is that the features need to be correspond to map objects, which are usually man-made objects, for example building, house and so on. In detail, we use the NDVI (Normalized Difference Vegetation Index), which is well known to accurately differentiate
vegetation from man-made objects. Many proposed solutions use image edges as matching features. The problem is that these edges are usually detected without reference to the objects themselves. In addition, most satellite images often contain complex natural scenes with a resolution of at most a few meters. For these reasons, edges do not provide the discrimination performance needed. Our approach, therefore, is to use the NDVI for feature extraction. NDVI allows us to identify the most effective matching features. The index can reliably detect man-made objects in satellite images.

Our experiments used satellite images acquired by a commercial satellite. With the rapid development in remote sensing, we are easy to get these images. Current satellite imaging systems can take multi-spectral data over wide areas at a time with a resolution of a few meters. However, the accuracy of original position alignment is on the order of 125,000 scale map, which usually means a numerical position of 10 meters.

Reference and feature pixels are used for matching. The former are calculated by projecting map objects onto satellite images with their coordinates. The latter are calculated from satellite images as follows. The NDVI of an N x M satellite image is given by the discrete function (i, j, y), x ∈ {1, 2, 3, ..., N}, y ∈ {1, 2, 3, ..., M}, where (i, j) is the index value. (i, y) is calculated using the following arithmetic operation:

\[
(I(x, y) = R(x, y) - R(x, y)) / (R(x, y) + R(x, y))
\]  (1)

\(R(R)\) is the reflectance value in the near infrared channel (visible red channel) region. \(I(x, y)\) are normally used in identifying vegetation. Therefore, the binarization of \(I(x, y)\) leads to an image containing only vegetation. We use \(I(x, y)\) in the reverse manner to identify man-made objects. The effective threshold for binarization needs to be variable within the same image, because it is uneven to the index distribution. N x M satellite image is divided through n x m window. In this case, the image contains \((N/n) \times (M/m)\) windows. The window size \((n x m)\) depends on the characteristics of the image. \(I(x, y)\) is binary using the threshold for each window. Each threshold is calculated by averaging filtering toward \(I(x, y)\), where \(x\) or \(y\) is equal to \(x\) included in the window region. \(y\) or \(y\) is also included in the window region. A pixel \((x, y)\) is identified as a matching feature, if \((x, y)\) is below the adaptive threshold. This binarization is performed to all windows. In addition, isolated pixels are eliminated as features to suppress the obvious errors caused by shadow effects and spectral deviations. The remaining pixels are identified as feature pixels.

3. MISMATCH DETERMINATION

This section describes our approach to determine the mismatch between satellite images and vector maps. The general idea is to use the most reliable correspondence between projected map objects and real objects shown in satellite images by voting. Voting is based on the Generalized Hough Transform (GHT), which is often used to estimate geometry conversion parameters. GHT can be used when a known figure (i.e. template) exists in an arbitrary background and can account for unknown parallel displacement, rotation, and expansion (Duda & Hart, 1972). Here, it is assumed that only position shift (which means parallel displacement conversion) need be considered, because it is necessary in GIS to use satellite images and vector maps as they are. The parallel displacements in the x-direction and y-direction are estimated separately by one-dimensional GHT.

The displacements are determined as follows. Reference and feature pixels are extracted as mentioned above. The scan area is assumed to be bigger than the area that holds the map object area projected in the satellite image. The differences between the positions of all these pixels on each scan line are calculated. Each calculated value is taken as one vote. The peak corresponds to the number of votes obtained in the scan area. Displacement candidates are identified as those with high voting frequency on all scan lines. High vote frequency means that the voting score exceeds a threshold.

The final displacement is selected from among the candidates. The selection should pay attention to the consistency of pixel matching after displacement. Displacement candidates are estimated using the mean square error of all differences between corresponding pairs of feature and reference pixels (displaced). Here, it is assumed that the displacement is correct if feature pixels are sufficiently consistent with reference pixels. The mean square error is calculated for each candidate. The determined displacement is the one with the lowest mean square error value. Thus, the displacements in the x- and y-directions are estimated separately.

4. EXPERIMENTAL RESULTS

We performed experiments to verify our approach; actual satellite images and corresponding vector maps were used. Fig. 1 shows one part of a typical satellite image (RGB format) as acquired by IKONOS (Space Imaging). These images have multi-spectral data (red channel, green channel, blue channel and near infrared channel) with 11 bit steps and 4 meter (per pixel) resolution.

![Figure 1](image.png)

Commercial 1/2,500 scale maps of the same test area were used. The maps include topographical data identifying several types of man-made objects, such as buildings. An example is shown in Fig. 2. This figure contains some layers with regard to buildings, houses, and roads (which are man-made objects) to extract reference pixels. Following the procedure described in section 2, feature pixels regarded as man-made objects were extracted from the satellite image. Fig. 3 shows the red channel of the satellite image in Fig. 1. Fig. 4 shows the near infrared image. NDVI \(I(x, y)\) was calculated by Eq. 1 using the spectral reflectance values. The result is shown in Fig. 5, where \(I(x, y)\) values have been converted into gray scale values. The binarization and elimination of \(I(x, y)\) were performed using the 20 pixel x 20 pixel window. The resulting extracted features are shown in Fig. 6. The results in Fig. 6 demonstrate that the feature pixels were extracted comparatively well. Extracted pixels were also quite accurate, the main errors were excessive extraction and extraction mistakes. The results show that the NDVI approach is stable and reliable enough to make classification of man-
made objects possible. Note that there is a limit as a size of object can be discriminated, which depends on a resolution of the satellite image.

Figure 2: An example of 1/2,500 scale map

Figure 3: An example of red channel image

Figure 4: An example of near infrared channel image

Figure 5: The results of NDVI calculation

Figure 6: The extracted feature pixels

Following the procedures described in section 3, mismatch was determined using the resulting feature pixels. To simplify the discussion, only the results of y-direction displacement are shown below. Fig. 7 shows the resulting reference pixels in the satellite image; this was realized by projecting map objects onto the image using known coordinates. This figure shows that the projected positions are somewhat shifted against the real object positions shown in the satellite image. Fig. 8 shows that the voting scores exhibit one clear peak; 64.25% votes were cast in this case. The displacement with the lowest mean square error is -13 pixels in this case. This value is equivalent to about 52 meter mismatch. The mean square error is also 1.07 pixels.

Similarly, the registration results are compared in Fig. 9 to the original projection shown in Fig. 7. The figure shows that the registration of this image has slight errors. The results show that our approach is very effective in suppressing many of the errors common in mismatch determination.

To quantify the average performance of the proposed approach, the results of 5 regions (including the above area) are shown in Table 1. The size of each region is approximately 500 meters x 500 meters. In this case, the number of feature pixels was about 93,600 per region on average. The average error was about 1.87 pixels. We confirmed by projecting the maps over the satellite images that the registration has only slight error.
The results were extremely accurate; the registration mismatch was just a few pixels on average.

We also tested the approach of using image edges as matching features to quantify the advantage of the NDVI approach. The edges were extracted using the Robinson operator. Fig. 10 shows the edges extracted from the satellite image in Fig. 1. Fig. 11 shows the voting scores using these edges. No clear peak was detected in this case. Additionally, the mean square errors are calculated using resulting displacements of our approach. Table 2 shows the results of the 5 regions. The number of feature pixels was about 910,500 per region on average, and the average error was about 19.53 pixels. There is a large number of feature pixels, which include many pixels unsuitable for voting. Therefore, the error is about ten times worse. This fails to well determine mismatch. This is also the reason for the large computation time. These results reveal that the NDVI approach yields higher accuracy and smaller mean square error than the approach of using image edges. It is clear that our approach can extract effective features from satellite images because it makes good use of topographical reference objects in voting.

<table>
<thead>
<tr>
<th>Region</th>
<th>MSE (pixels)</th>
<th>Feature Pixels</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.17</td>
<td>193.582</td>
</tr>
<tr>
<td>2</td>
<td>1.91</td>
<td>80,864</td>
</tr>
<tr>
<td>3</td>
<td>2.23</td>
<td>92,069</td>
</tr>
<tr>
<td>4</td>
<td>1.57</td>
<td>60,461</td>
</tr>
<tr>
<td>5</td>
<td>1.48</td>
<td>48,343</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Region</th>
<th>MSE (pixels)</th>
<th>Feature Pixels</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16.26</td>
<td>1354399</td>
</tr>
<tr>
<td>2</td>
<td>16.32</td>
<td>632913</td>
</tr>
<tr>
<td>3</td>
<td>29.49</td>
<td>1218203</td>
</tr>
<tr>
<td>4</td>
<td>18.84</td>
<td>608366</td>
</tr>
<tr>
<td>5</td>
<td>18.77</td>
<td>612846</td>
</tr>
</tbody>
</table>
5. SUMMARY AND CONCLUSIONS

This paper introduced a fully automatic approach to the registration of satellite images against vector maps. The proposed approach consists of two procedures: feature extraction for matching, and mismatch determination. Our approach for the extraction of matching features uses the indicators based on the NDVI derived from multi-spectral satellite images. Mismatch is determined through the use of voting displacements between the positions of the features and those of topographical objects projected onto satellite images. Experimental results show that our method can register satellite images against maps automatically and accurately.

More experiments are needed including a wide variety of test areas. Several problems remain. As regards feature extraction, more studies on the threshold values are needed. Additionally, the choice of window size should also be examined. With regard to mismatch determination, more studies on the geometry conversion of rotation and expansion are needed. In particular, a method that handles these conversions is also examined.

ACKNOWLEDGMENTS

The authors are thankful to our research coworkers for valuable discussions.

REFERENCES


Klang, D., "Automatic ground control point extraction for satellite scene orientation", ISPRS Workshop "From producer to user"., Boulder USA, October 1997.

ORBIMAGE, http://www.orbimage.com/


Space Imaging, http://www.spaceimaging.com/
