Autonomous Conflation of Vector-to-Image

Gamal Seedahmed & Lou Martucci
Battelle/Pacific Northwest National Laboratory
902 Battelle Blvd. Richland, WA, 99354 USA
Gamal.Seedahmed@pnl.gov
Lou.Martucci@pnl.gov

Abstract

Autonomous, accurate, and robust vector-to-image conflation or alignment is a critical task in many geospatial applications, such as map update, change detection, navigation, smart image annotation, and even object recognition. This paper presents a novel approach for autonomous vector-to-image image conflation. This approach obviates the needs for the classical feature matching and searches for the conflation parameters in the parameter-space that characterizes the geometric transformation function between the vector and the image datasets. Point features extracted from the two datasets are compared through a mathematical transformation that describes the geometrical relationship between the two datasets. Point features extracted from the vector dataset are used to guide the extraction of point features in the image. The guided point extraction in the image alleviates the problem of global threshold selection for point feature labeling. Point features from the two sets are paired for comparison. The results of different comparisons are encoded and exploited in a histogram-like representation as clusters of votes. The most consistent cluster generates a hypothesis for the most likely conflation parameters. Matched point features are recovered as a by-product and used in a classical least squares fitting to further optimize the conflation parameters. The developed approach is applied successfully to conflate a GIS transportation layer to Landsat images with different spatial resolutions.

1. Introduction

Conflation is typically defined as the process of combining data from two sources or more. Several objectives can be achieved by performing data conflation such as accuracy enhancement, database update, change detection, etc. Alignment is one of the core aspects of conflation, which mainly concerned with bringing different data sets to a common coordinate space. Automated techniques for conflation are highly desirable. This paper presents an automated technique to conflate satellite imagery to vector data set, which could revolutionize the current practice of GIS in many respects such as: (1) GIS can be utilized to offer the prior knowledge for image analysis. (2) Images can be used to update the GIS database. The presented technique is named Geometrically Invariant-information Parameter Space Clustering (GIPSC), which will be explained in the sequel of this paper.
2. The Conflation Problem between Images and Vector Data

Miss-registration between images and GIS vector layers happens for several reasons. For example, when vector layers or images are georeferenced poorly and uncorrected for systematic errors in either or both of the data sets; see Fig. 1.

![Fig. 1: A typical miss-registration problem between the image data and vector information.](image)

3. Methodology

The basic idea underpinning GIPSC is to compare the elements of two data sets according to a mathematical transformation that describes the geometrical relationship between them. The current development considers two basic assumptions. First, the characteristics of the two data sets give rise to detectable features such as points, and at least part of these features are common to both of them. Second, the two data can be aligned by a 2-D transformation. The basic process starts with geometric invariant feature extraction, followed by parameter space clustering. For the interest of developing an intuitive understanding of the basic process, each step is highlighted briefly. First, in the presented study point features are dealt with. Second, the basic idea of parameter space clustering is to compare the data element gathered from two sets according to a pre-specified observation equation (voting function). The results of comparison will point to different locations in the parameter space. The pointing is achieved by incrementing each admissible location by one during the voting process. A coexisting location in the parameter space, defined by the data elements that satisfy the observation equation, will be incremented several times forming a global maximum in the parameter space. This maximum will be evaluated as a consistency measure between the two data sets.
GIPSC methodology as outlined above, is used to perform an autonomous image-to-vector conflation; see Fig. 2. In a nutshell, GIPSC obviates the needs for the classical feature matching and searches for the registration parameters in the parameter-space that characterizes the geometric transformation function between the image and the vector datasets. Point features extracted from the vector dataset are used to guide the extraction of point features in the image. The guided point extraction strategy alleviates the problem of global threshold selection for point feature labeling. The Moravec operator is used to extract point features from the images.

![Workflow Diagram](https://via.placeholder.com/150)

**Fig. 2:** General workflow of GIPSC for autonomous image-to-vector conflation.

### 4. Experimental Results

The GIPSC algorithms were tested on two Landsat images of different spatial resolutions. In particular, Landsat 30m and Landsat 15m images were used in this test. The Landsat 30m test will be presented to demonstrate the performance of GIPSC. Image and vector chips that have a size of $3K \times 3K$ were used; see Fig. 3. Table 1 shows the number of the extracted points from the image and vector layer and the number of matched points. By comparing the number of matched points to the extracted ones we can infer that GIPSC is very robust in the presence of outliers and very small percentage of common points ($<10\%$) are enough to obtain a unique peak in the parameter space; see Fig. 5. Matched points are used in a classical least squares adjustment to estimate the miss-registration function between the image and vector information; see Table 2. Fig. 6 shows the final registration results.
Fig. 3: A Landsat 30m and a transportation vector layer.

Fig. 4: The parameter space solutions obtained from GIPSC. The left Fig. Shows the rotation space and the right one shows the translation space.

<table>
<thead>
<tr>
<th>Point Description</th>
<th>Number of Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image points</td>
<td>1796</td>
</tr>
<tr>
<td>Vector points</td>
<td>1286</td>
</tr>
<tr>
<td>Matched points</td>
<td>51</td>
</tr>
</tbody>
</table>

Table 1: The number of the extracted points from the image and vector layer and the number conjugate or matched points.
<table>
<thead>
<tr>
<th>PAR.</th>
<th>SOLUTION</th>
<th>STANDARD DEV</th>
</tr>
</thead>
<tbody>
<tr>
<td>XT</td>
<td>-4.989 (-155.3 m)</td>
<td>±0.15291m</td>
</tr>
<tr>
<td>YT</td>
<td>10.818 (321.0 m)</td>
<td>±0.15217m</td>
</tr>
<tr>
<td>Scale</td>
<td>0.99986</td>
<td></td>
</tr>
<tr>
<td>Rotation Angle</td>
<td>0.0073712</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: The estimated parameters of the similarity transformation obtained from the matched points.

Fig. 5: An image and vector patches show some of the matched points.
Fig. 6: The final registration results of the image and vector information obtained by GIPSC.

5. Conclusions

a. GIPSC provides an autonomous and robust solution for the image-to-vector registration.
b. GIPSC delivers accuracy statement, which can be used to judge the final results of the registration.
c. GIPSC offers a flexible mathematical framework to integrate the available prior knowledge into the registration process.

6. References are available upon request.