Jiří Cajthaml, Tomáš Janata

Georeferencing of First Military Mapping survey maps in the area of Bohemia using polynomial method

Keywords: georeferencing, transformation, robust adjustment method, First Military Survey, multiple sheet map series

Summary: This paper describes an implementation and comparison of classical and robust methods of adjustment of multiple sheet map series on example of the First Military Survey maps originating in the Habsburg monarchy of the second half of 18th century, using own application for georeferencing of map series. The application computes adjustment of map sheets using least squares method with affine or second order polynomial transformation. It also applies a robust adjustment method (IRLS) when there are some outliers expected. The results of all mentioned methods are compared and the algorithms are tested for resistance to a presence of gross errors in collected control points.

Introduction

Old map works often involve not only individual maps, but also multi-sheet sets that could be used to cover a larger area with maps of a uniform layout, appearance and sign key. Typical representatives of these map series are cadastral maps or state map works.

For Central Europe, namely the Habsburg monarchy of the late 18th and entire 19th century, apart from cadastral maps, the three nation-widely accomplished military mappings are typical. They brought detailed topographic maps into being, which find wide application still in the present. These maps have roughly rectangular map sheets of unique dimensions. When making the Second and Third Military Survey, a surveying network formed by before created military triangulation has been used, but the First Survey has been created entirely without any geometric frame and thus no spatial coordinates of the sheet corners in the specified map layout are known. A large issue of this mapping is especially the impossibility to combine multiple adjacent map sheets into one compact area.

In digital form, the mentioned imperfection could be adjusted using an appropriate transformation method and thus the whole map work reconstructed in the form in which it was originally intended. This article provides a treatise on how to adjust the first military mapping using two approaches: the classical least squares method to the terms of follow-up edges of the individual map sheets and robust method of Reweighted Iterative Least Squares, which represents one of the possibilities of robust application of adjustment methods. Both these methods are combined in a software tool designed to adjust maps of the 1st Military Survey, but it can generally be used for any data consisting of a larger amount of vaguely defined individual sheets.
The state of georeferencing of old map series in the Czech lands

Map sheets of military mappings, which were based on mathematically defined grid and have a consistently designed map layout with known coordinates of the sheet corners, might adjust into a seamless maps directly, just using the coordinates of the corners, as it was performed in the case of the Second Military Survey (Bohemia mapped during 1806–73); this map series is currently available as the seamless digital data layer for the entire territory of the former Habsburg monarchy (Čada and Vichrová, 2009).

The First Military Survey cannot be so easily adjusted (for a lack of geometric basis) and therefore it was necessary to use a more complex procedure based on the conditional adjustment of the least squares method. Since the coordinates of corners could not be used for the adjustment of the map sheets, the fact that the sheets are interconnected and represent a perfect filling of the space without gaps and overlaps, still could; it was necessary to collect a large number of control points on whose basis the transformation was made.

Already during an experimental alignment of individual map sheets, it emerged that, as a result of mapping based on estimation and rough sketching with the minimum measured quantities, the sheets evince a complicated waveform of distortion in the map image, difficult to describe with a discrete set of several control points. Therefore, efforts of authors of the procedure were made to collect large amounts of usable control points (ideally, at least 30 per map sheet, in a suitable configuration within the entire surface of the sheet). During the collection of control points, maps of the 2nd Military Survey, in terms of time a few decades far, were primarily used. As objects, churches, manors, courtyards and other farming buildings, sacral monuments, mills and other objects have been selected.

For the whole historical territory of Bohemia thus roughly 7,000 control points were collected that represent an input into a global adjustment that allowed the creation of a seamless map layer of the 1st Military Survey for Bohemian area.

The adjustment technique

The least squares method (LSM) is used to adjust primarily for its simplicity and optimum results. The normal distribution of errors is a prerequisite for its proper functioning. The only gross error can significantly affect and deteriorate results of the whole calculation. In the real situation, the ensuring of the prerequisites is difficult; it is almost impossible to avoid outliers. It is therefore important to check the data before the actual computation. In case of collection of identical points, it is applicable to render the pairs of points and to make at least a visual inspection. Neither a careful check of the data has to reveal all the gross errors (outliers).

One of the ways to eliminate the effect of outliers is the use of robust statistical methods. A significant advantage of robust methods is their resistance to outliers; even in case of a larger number of outliers they can be used to obtain satisfactory results. The disadvantage can be a loss of a certain portion of optimality when the assumption of normal distribution of errors is met. Unlike the least squares method, the robust methods may not give the best unbiased estimate (also called Best Linear Unbiased Estimator, BLUE). Robust methods are numerous; they are characterized by a varying degree of resistance and optimizing criteria. The robustness of a method is usually evaluated using the
so-called point of failure. This expresses the smallest ratio of outlier observations, which can cause an erroneous estimate. For the least squares method, the point of failure is equal to 0 %, which means that a single outlier can lead to an incorrect result.

The best-known methods used are the Reweighted Iterative Least Squares (IRLS) in combination with different types of estimates, the method Least Median of Squares (LMS), which replaces the sum of squares with a median, the method of trimming or the iterative algorithm RANSAC. A more detailed overview and description of the methods mentioned is bringing Třasák (2013). Overviews of robust adjustment methods for the needs of cartography describe Beineke and Caspary (2003).

For georeferencing, the IRLS method appears appropriate for its simplicity, closeness to the least squares method and robustness. The method minimizes the expression

$$\sum w_i v_i^2 \to \min$$

where $v_i$ is a residual and $w_i$ a weight of the $i$-th measurement. This method is based on iterative solution of the weighted least squares method, where during (each) iteration all measurements are allocated their respective weights based on estimates of their remoteness. The more remote the measurement is, the lower is its weight. Most widely, the so-called M-estimators are used to determine the weights. In the application, known Huber’s M-estimator was used. When selecting, it proceeds from the normal distribution of a random variable, for which it replaces the outskirts of the distribution with the Laplace distribution. This way a bordered influential function is obtained, in which it is possible to influence the behaviour of the adjustment using a constant based on an estimate of contamination by outliers (gross errors). Huber’s estimator gives satisfactory results even when the ratio of contamination varies around 5 %.

During the calculation of IRLS method, the weights in the weight matrix are determined by Huber's estimate in each step of the iteration, while the normalization takes place through a robust statistical characteristic of a degree of dispersion called the median absolute deviation (MAD; a robust alternative to the standard deviation), which is calculated from the equation

$$MAD = med(|v - med(v)|)$$

where $v$ is a vector of residuals. As a condition for ending the cycle, an experimentally determined minimum percentage of the a posteriori standard deviation improvement compared to the previous iteration is used.

The disadvantage of the robust method is a manifold greater processing time, which is proportional to the number of iterations. Nevertheless, sometimes its use is necessary. An example may be the processing of data partially contaminated by outliers using polynomial transformation. As shown in Figure 1, due to a gross error in an identical point in a map sheet the result of using the classical least squares method is significantly distorted. The robust method, to the contrary, suppresses the impact of this error and the resulting frames of map sheets are shaped significantly more real.

It is also important to evaluate the results visually. Despite a small standard deviation the results may not be entirely correct. This occurs primarily in case of border map sheets as there it is not possible to collect control points complying the principle of even distribution throughout the sheet. This fact
often causes an unnatural shape of the resulting sheet and with regard to further planned use of the resulting map it is necessary to assess whether the outcome achieved is acceptable.

![Figure 1: Comparison of an adjusted pair of marginal map sheets insufficiently covered by control points using LSM (left image) and IRLS (right) algorithm.](image)

**The MultiGeoref tool**

The MultiGeoref software uses an affine transformation and 2nd order polynomial transformation, whereby it is possible to choose between ordinary least squares adjustment, which is suitable for the most common map series, and adjustment using IRLS, which is appropriate in the case of maps of the 1st Military Survey especially for its difficulty expressible inaccuracy with many local anomalies, moreover loaded with a difficult identification of control points on some map sheets.

The control points can be imported from tools of ESRI, QGIS and GRASS or other applications using the appropriate file formats. In each map sheet further 4 points representing corners of it have to be collected. Files of corner coordinates represent a second input for the MultiGeoref software. The last input is a matrix of mutual neighborhoods of the map sheets. From this matrix (which can be loaded from a file or interactively created by a user in the application’s GUI) the conditions used during adjustment are calculated.

To avoid singularity of the conditional matrix during the adjustment, the application selects only linearly independent conditions resulting from the identity of edges or corners of particular map sheets. For the affine transformation, only corners of map sheets are equated, while when using 2nd order polynomial transformation, a condition of identity of center point on an edge between two sheets is added, which is sufficient for the identity of the entire edge (after the adjustment generally expressed by a part of a second order parabola).

**Results. Testing the algorithm**

Using tool MultiGeoref tool, the map sheets of the First Military Survey within the entire territory of the historical land of Bohemia have been tentatively adjusted, except for marginal map sheets that are
inadequately covered by the map drawing upon which it was failed to identify more than 8 control points. Hence an area of 252 map sheets has arisen, which are shown for clarity in Figure 2.

In Table 1 the achieved results are summarized. The prerequisite for a proper functioning of the LSM is data without gross errors, which has been tried to achieve by the authors using visual inspections and a multiple verification of control points using different cartographic sources at several time levels.

Figure 2: Map sheets used for the global adjustment.

<table>
<thead>
<tr>
<th>Transformation</th>
<th>Method</th>
<th>Time</th>
<th>Accuracy [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Affine</td>
<td>LSM</td>
<td>t</td>
<td>1168</td>
</tr>
<tr>
<td></td>
<td>IRLS</td>
<td>9.4t</td>
<td>664</td>
</tr>
<tr>
<td>2nd Order</td>
<td>LSM</td>
<td>3.0t</td>
<td>776</td>
</tr>
<tr>
<td>Polynomial</td>
<td>IRLS</td>
<td>28.4t</td>
<td>304</td>
</tr>
</tbody>
</table>

Table 1: Results of the global adjustment of the entire Bohemia using various options of the software.

Notice that the time of calculation shown in Table 1 is significantly dependent on the hardware of the computer used. It therefore shows only relative approximate ratios of times for each calculation method. The accuracy is expressed using experimental a posteriori root mean coordinate error.
Figure 3: A detailed comparison of the adjusted map sheets: (left) affine transformation; (right) 2nd order polynomial transformation using IRLS method.

Figure 4: Overview of a part of the adjusted area (map sheets projected using affine transformation in red, using polynomial transformation in green; blue dots represent control points)

Within the adjustment also robustness of the IRLS algorithm against intentional gross errors was tested. Results of the calculation to the test area of 32 map sheets in Central Bohemia and various levels of contamination of data by outliers is summarized in Table 2.
In Table 2, several facts are apparent. Provided that only the final a posteriori coordinate error was selected for considering the most appropriate method, the polynomial transformation would seem to be better option for the given data set. The greatest accuracy evinces the IRLS method in combination with the polynomial transformation; it, however, represents a very time consuming method. As expected, the classic LSM and affine transformation is the fastest to use.

<table>
<thead>
<tr>
<th>Contamination</th>
<th>Transformation</th>
<th>Method</th>
<th>Time</th>
<th>Accuracy [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 %</td>
<td>Affine</td>
<td>LSM</td>
<td>1</td>
<td>1055</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IRLS</td>
<td>8.5t</td>
<td>628</td>
</tr>
<tr>
<td></td>
<td>2nd order polynomial</td>
<td>LSM</td>
<td>3.3t</td>
<td>738</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IRLS</td>
<td>32t</td>
<td>255</td>
</tr>
<tr>
<td>1 %</td>
<td>Affine</td>
<td>LSM</td>
<td>1.1t</td>
<td>1623</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IRLS</td>
<td>8.8t</td>
<td>654</td>
</tr>
<tr>
<td></td>
<td>2nd order polynomial</td>
<td>LSM</td>
<td>3.0t</td>
<td>1416</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IRLS</td>
<td>30.9t</td>
<td>283</td>
</tr>
<tr>
<td>5 %</td>
<td>Affine</td>
<td>LSM</td>
<td>1.2t</td>
<td>2154</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IRLS</td>
<td>6.7t</td>
<td>735</td>
</tr>
<tr>
<td></td>
<td>2nd order polynomial</td>
<td>LSM</td>
<td>3.0t</td>
<td>1961</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IRLS</td>
<td>32.4t</td>
<td>348</td>
</tr>
</tbody>
</table>

Table 2: Correlation of contamination of data and achieved accuracy.

The second interesting finding out is a comparison of coordinate errors of the resulting adjustment when using the classical LSM vs. robust methods. The greater is contamination of data by outliers, the more noticeable difference is in accuracy. While, in case of growing contamination of data by outliers, the accuracy of classic LSM deteriorates quite significantly (at 5% contamination about twice compared to the original data) for iterative methods to be reflected at least. As the main disadvantage of the robust method, primarily far greater time demands remain.

**Conclusion**

The article presents new ways of adjusting multiple-sheet map works together with software for this purpose, being created at the authors’ department and which is particularly suitable for georeferencing map works of unknown geometric parameters. As shown in the case of maps of the First Military Survey, it is possible, using robust statistical methods involved in the adjustment, to achieve significantly better outcomes (measured by means of the resulting a posteriori mean coordinate error) and to eliminate the influence of blunders and outliers, which cannot be easily avoided when collecting identical points on old map data.
References


