Georeferencing of Place Markers in Digitized Early Maps by Using Similar Maps as Data Source

Winfried Höhn
winfried.hoehn@uni.lu
ILIAS Lab, University of Luxembourg, Luxembourg

Christoph Schommer
christoph.schommer@uni.lu
ILIAS Lab, University of Luxembourg, Luxembourg

Early maps are usually only accessible for a small group of researchers and librarians because they are precious and fragile. In the age of Digital Humanities, online access and search in digitized historical documents and early maps allows people from all over the world to work with such artefacts of cultural heritage. However, the digitization solely generates images of the artefacts without any access to the semantics of the documents. For most digital libraries of early maps (e.g. Old Maps Online) the available metadata include only information about the map, e.g. author, title, size, creation date, covered region. Unfortunately, there is only little information about the data contained in the map. Thus, even if data about place development or toponym changes is present in the maps it is not easily accessible. Since a single map can easily contain many thousands of place markers, proper tool support and automation of the annotation and georeferencing of each single place marker are of interest.

For modern maps or aerial photos it is possible to use GIS software to georectify the images by specifying a few control points, thus this problem is seen as solved. But early maps contain many sources of distortion, for example inaccuracies during surveying, combining data from different sources, focusing on creating a visually pleasing map instead of an accurate one. So there is in general no simple mapping between modern geocoordinates and an early map. Our existing Referencing and Annotation Tool (RAT) (Höhn et al., 2013) already simplifies the annotation and georeferencing of place markers. RAT supports the annotation and georeferencing by using template matching to identify place markers and by suggesting the most likely modern places based on an estimated mapping between the pixel-coordinates and geocoordinates of the already georeferenced place markers. To further refine the suggestions a phonetic search can be used, where the historic spelling can be used to restrict the results to similarly sounding place names.

Even with tool support like provided by RAT the georeferencing and annotation process starts from scratch for each map. Despite the automation there is still manual effort needed for place marker annotation. Since early maps have often been copied from each other or share some underlying survey data, there should be some regularity between maps that we can exploit. To take advantage of the possible similarities in early maps we present an algorithm to identify similar maps and create a link between the place markers of these maps. This results in georeferencing an early map in relation to another early map, which can be much simpler than georeferencing in respect to modern data. When the maps are based on the same data, they share some of their distortions and so the transformation between them gets simpler. They will also more likely contain a similar set of places. This reduces the problem of identifying a matching place compared to a modern database containing all known places, even the smallest ones which will not be shown in medium or small scale maps.

Before we can apply the algorithm to a pair of maps we need to identify suitable maps. These are maps that already have some georeferenced place markers and share at least four mappings to modern places. Also the place markers in these maps must be already recognized, but not necessarily georeferenced.

The algorithm for linking corresponding place markers of two maps A and B can be split in two steps:

1. Estimation of a transformation between the maps.The coordinate mappings in $\mathbf{M}$, a bidirectional mapping containing the coordinates of the matching place markers in the two maps, are used to calculate the projective transformation between map A and map B.
2. Extending the linked place markers. Using the projective transformation calculated in step 1, map B is transformed into the coordinate system of map A. We will refer to the transformed map B as map B'. For each place marker in A the nearest place marker from B' is located. If for the place marker from B' also the place marker from A is the nearest, following checks are done:

- The second nearest place markers have to be at least two times further away than the distance of the two place markers under consideration.
- Both place markers must be connected to some place marker contained in **M** through edges in a Delaunay-Triangulation (Lee and Schachter, 1980) of map *A* and map *B*.

If both previous conditions are true, add the place markers to **M'**.

If **M'** has more elements than **M**, then set **M** to **M'** and continue with step 1.

M' is the resulting correspondence between the two maps. The steps of the algorithm are visualized in Fig. 3. The right column corresponds to step 1 and the left one to step 2. The rows show the different iterations of the algorithm. For all examples, the following maps are used: "*Nova Franconiae descriptio/Sculptum apud Abrahamum Goos. - Amsterdam: Joannes Janßonius, 1626*" referred to as Goos and "*Franckenlandt = Francia orientalis/Per Gerardum Mercatorem – o.O., ca. 1600*" referred to as Mercator.

These example maps both contain about 900 place markers and an overlapping area with about 800 place markers. For Goos, all place marker locations have been manually verified and for Mercator, the result of the template matching was kept. This resulted in an automatic detection of the correspondence of 755 place markers between the maps.

Another use case of this mapping is, that we can compare the automatically found place markers from two maps. We can highlight the differences between the sets of automatically detected place markers from two maps. This allows easily investigating the differences in the two sets of manually or automatically identified place markers. The identified differences highlight specific areas in these maps for further investigation. Two examples for detected differences between similar maps are shown in Fig. 1 and 2.

Figure 1. Corresponding map sections from Goos (left) and Mercator (right), where Goos has one place marker for Hoeltriech and Mercator one for Fuechstat which are both not in the other map.

Figure 2. Corresponding map sections from Goos (left) and Mercator (right), where Heibach is in Mercator located at the river and in Goos far away from the river.

There can be various reasons for differences between two maps. First, there could have been an error in the detection of the place markers, which then helps to spot such problems. Second, it is a genuine difference between the two maps, which itself can have many reasons, e.g. different decisions which places should be included on a map or errors while placing the places on a map.

This work shows that it's possible to create a correspondence between place markers in different maps with not more effort than for georeferencing a map, which then only provides the region covered by the map. A similar map can be identified, if one exists in the database, and the place markers between the maps can be connected. This then allows reusing the georeferencing of single place markers from one map in the other map and identifying differences in the sets of place markers.

Future Work

This method can also be used to quickly identify similar maps and the differences in them. In this way it could be useful for researchers who want to find out which sources were used to create a map or who copied from whom. MapAnalyst (Jenny and Hurni, 2011) is an already existing tool for this purpose. If one map is considered as a possible copy of another map, MapAnalyst is a tool used by researchers to explore if this is true. The method proposed in this work would allow doing this kind of analysis on a larger scale while also highlighting the differences between the maps. Although this area was not our primary focus, we plan to evaluate the usefulness of our method on this task.

The information from linked place markers could help in further analysis of other metadata items, such as place type or place name. The linked place markers already make this information available from the other maps and it could for example be used to improve the OCR process of place names.

Appendix

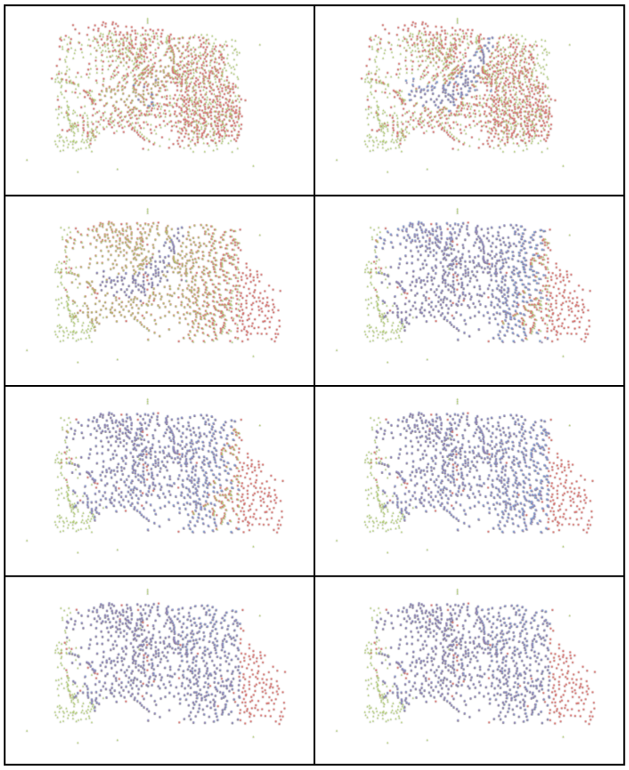

Figure 3: Visualization of the different steps in the matching procedure for Goos and Mercator, where the Mercator place markers are transformed into the Goos coordinate system. Triangles represent Goos place markers and squares Mercator ones. Green and red points don't have a mapping to a place marker in the other map, blue and purple ones have mappings. For further explanations see algorithm in main text.

Bibliography

Höhn, W., Schmidt, H.-G. and Schöneberg, H. (2013). "Semiautomatic Recognition and Georeferencing of Places in Early Maps." *Proceedings of the 13th ACM/IEEE-CS joint conference on Digital libraries*, pp. 335–38.

Lee, D. T. and Schachter, B. J. (1980). "Two algorithms for constructing a Delaunay triangulation." *International Journal of Computer & Information Sciences*, 9(3), pp. 219–42.

Jenny, B., and Hurni, L. (2011). "Studying cartographic heritage: Analysis and visualization of geometric distortions." *Computers & Graphics*, 35(2), pp. 402–11.