# Content-based Image Retrieval for Map Georeferencing

Jonas Luft [a]*, Jochen Schiewe [a]

[a] *Lab for Geoinformatics and Geovisualization, Hafencity University Hamburg, jonas.luft@hcu-hamburg.de*

* Corresponding Author

**Abstract:** In recent years, libraries have made great progress in digitising troves of historical maps with high-resolution scanners. Providing user-friendly information access for cultural heritage through spatial search and webGIS requires georeferencing of the hundreds of thousands of digitised maps.

Georeferencing is usually done manually by finding "ground control points", locations in the digital map image, whose identity is unambiguous and can easily be found in modern-day reference geodata/mapping data. To decide whether two symbols from different maps describe the same object, their semantic and spatial relations need to be matched. Automating this process is the only feasible way to georeference the immense quantities of maps in conceivable time. However, automated solutions for spatial matching quickly fail when faced with incomplete data – which is the greatest challenge when comparing maps of different ages or scales.

These problems can be overcome by computing map similarity in the image domain. Treating maps as a special case of image processing allows efficient and robust matching and thus identification of geographical regions without the need to explicitly model semantics. We propose a method to encode worldwide reference VGI mapping data as image features, allowing the construction of an efficient lookup index. With this index, content-based image retrieval can be used for both geolocating a given map for georeferencing with high accuracy. We demonstrate our approach on hundreds of map sheets of different historical topographical survey map series, successfully georeferencing most of them within mere seconds.

**Keywords:** image analysis, automatic georeferencing, historical maps, VGI

## 1. Introduction

Research in digital map processing has picked up the pace in recent years (Chiang et al., 2014, 2020, Jiao et al., 2021), mostly focusing on digitising content with vectorisation and attribution. Previous research often assumes already georeferenced maps (Iosifescu et al., 2016, Heitzler and Hurni, 2019) or leaves the tedious manual georeferencing task to crowdsourcing (Fleet et al., 2012, Tavakkol et al., 2019).

Automatic georeferencing can close the gap to a fully automatic map-to-geodata pipeline. Aligning maps makes comparison of map content across different series and designs easier (Schlegel, 2019). Comparability of old and present-day maps, in turn, leads to many new research possibilities and unleashes the full potential for automated spatial analyses of historical data, e. g. for long-term monitoring of coastlines (Fabris, 2021), finding potential archaeological sites (White, 2013) and many more. Furthermore, georeferencing can significantly increase the accessibility of the important cultural heritage that are historical maps (Crom, 2016, Buckley, 2019).

Some work has been done on the constrained task of aligning two maps which are roughly covering the same area (e. g. Howe et al., 2019, Duan et al., 2020). However, to solve the general task of geolocating a map with little or no prior information on location, maps have to be compared against large volumes of comparable reference geodata, possibly across the globe. Luckily, large amounts of reference data are available in the form of volunteered geographical information (VGI), for example OpenStreetMap (OSM)[1]. The challenge is to find a way to efficiently search through it to find the most relevant part. Here image retrieval, particularly content-based image retrieval (CBIR) when metadata of maps is not sufficiently available, comes into play.

CBIR is used to find images in a reference database that are most similar to a given query image. In the context of maps, this requires a spatial similarity function, which are often solved with graph representations Li and Fonseca (2006). However, graph representations require complex definitions (Janowicz et al., 2011). Solving similarity graphs algorithmically is computationally expensive and has not yet been proven to provide unambiguous results. Furthermore, getting from a map to semantic representation requires full vectorisation (Szendrei et al., 2011, Iosifescu et al., 2016) and attribution (Sun et al., 2020) first, which are intense fields of study and still not completely solved. Developing a general method that works for any and all maps is particularly challenging. A more robust and easily scalable approach to map similarity is needed.

### 1.1 Contributions

This work expands on our own method presented in Luft and Schiewe (2021) by developing a more general defini-

---

[1] https://www.openstreetmap.org/

tion for spatial similarity using robust image features (Section 2.3). This overcomes the need for fine tuning of parameters to reach high accuracy and consequently allows to transfer our method more easily to new map series. We argue that exploiting the image domain instead of complex semantic representations enables the development of efficient and robust algorithms for calculating the similarity of map content. This implementation of map similarity allows exploitation of a reverse index for fast querying in large reference databases (Section 2.5).

Assuming a mostly successful segmentation, retrieval of similar maps provides a general solution for automated georeferencing. In several experiments in Section 3, we will address the following challenges for georeferencing with our proposed CBIR approach:

- Erroneous symbols and noise introduced by segmentation.
- Occlusion: Missing or extra symbols due to historical change or ill-fitting reference data resulting from a different level of generalisation between query and reference maps.
- Ill-aligned reference data, because of datum shift or imprecise reference quadrangles.
- Maps have different scales.

## 1.2 Related Work

Howe et al. (2019) and Tavakkol et al. (2019) have proven that toponyms allow geolocating a diverse range of maps. However, it is challenging to reliably recognise toponyms on old prints with unusual type. Furthermore, cluttered topographic maps show many overlapping features, which make optical character recognition unreliable (Luft, 2020). Burt et al. (2019) and Heitzler et al. (2018) instead exploit the graticule visible on more modern topographic maps, which allows very precise rectification and alignment. However, the graticule is not visible on all maps and assigning graticule corners to geographic coordinates relies on marginal information and prior knowledge about the map.

Content-based georeferencing can overcome the challenges of both of these approaches (Luft and Schiewe, 2021). Using content requires a robust framework of spatial similarity, especially when using only the geometry of map symbols without further attribution (which would need to be inferred from symbol classification or text understanding).

A proven method for matching image content is the use of image features. Image features are frequently used to automatically extract tie-points for alignment of photogrammetric images. More recently, there have been proposals for cross-modal matching between different imaging systems (Ye and Shen, 2016, Zhuo et al., 2017, Li et al., 2020) or across time-series with changes in illumination and land cover (Liu et al., 2008).

In contrast to photogrammetry, not much research has gone into matching topographic maps by image features. The

reason may be that image features are designed to work on photos but can not immediately be applied to printed or manuscript maps. Without preprocessing, the generalised content and heterogeneous style of maps usually leads to low repeatability of feature detection.

## 2. CBIR on maps

Generally speaking, CBIR is the task of finding images in a huge database that are most similar to a given query image, specifically similar in content. The main challenge lies in finding a suitable definition of similarity that closes the "semantic gap", i. e. returns images that are not just similar in abstract mathematical properties but also aligned with the expectation of similarity of a human user.

The concept of CBIR can be applied to georeferencing of a map as follows: To find the geographic area shown in a map (the query map), we search through a comprehensive set of already georeferenced maps to find those with the most similar content (reference maps). The discovery of a reference map with identical content allows to use its geographical information to geolocate the query map. Solving georeferencing with this approach requires us to solve some tasks, our solution to which we will elaborate step-by-step in the following paragraphs (refer to Figure 1 for an overview):

- Obtaining georeferenced reference maps of possible map locations (Section 2.2).
- Definition of a similarity function for map content (Section 2.4).
- Since there can be a lot of reference maps, they need to be queried efficiently (Section 2.5).

## 2.1 Prerequisites

In order to calculate similarity between the query map and reference maps, we need to remove differences in cartographic design and interfering information. We ensure this by segmenting the query map to extract a single class of symbols. Segmentation eliminates clutter and makes maps comparable regardless of their design, but introduces noise and other errors. The key is to segment object classes which we can match reasonably well to the reference VGI data. Objects with complex geometry are most suitable, since they carry a lot of spatial information.

Segmentation and map symbol extraction is a challenging task that deserves a dedicated study. For this paper we assume a more or less reliable segmentation as a prerequisite.

## 2.2 Reference data

OSM is a valuable source of spatial information, because its increasing global availability in consistent format and it contains many different feature types in great detail. The selection of object classes happens by filtering for so called "tags". Tags need to be selected to produce reference data that is as close as possible to the symbols that have been segmented from the query map.
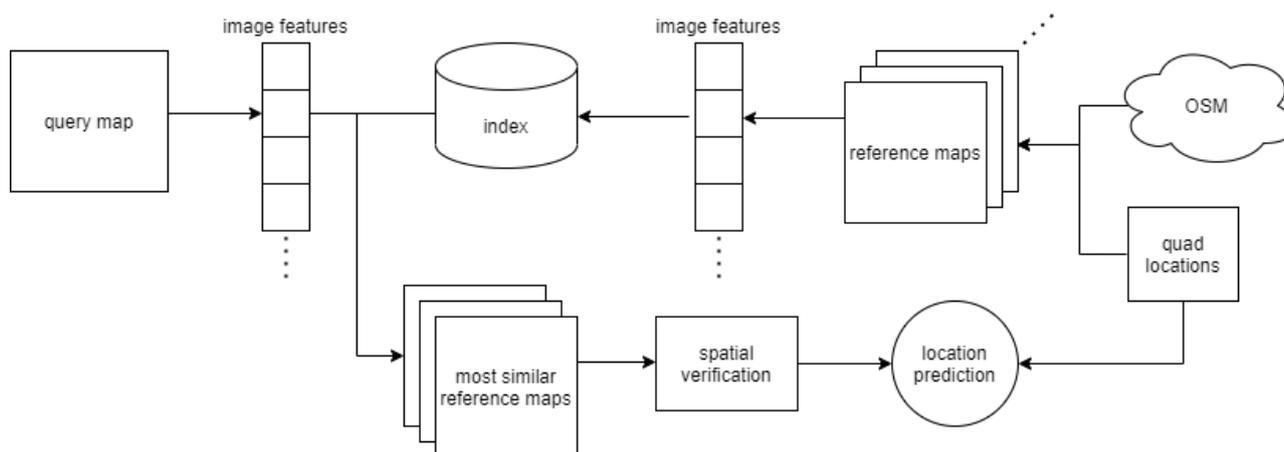
Figure 1. Overview of the CBIR process for georeferencing. To geolocate a given query map, first image features are extracted. With those image features, the index is queried, which was previously constructed from a large number of reference maps, created from OSM data at known quad locations. The index returns a set of reference maps, sorted by similarity. The features are matched between the query map and the reference maps and subsequently refined by spatial verification. The reference map with the most remaining matches is returned. Together with the known location of the reference map, we thus obtain a location prediction for the query map.

OSM reference data are vector based, but query maps are raster images of diverse style. The filtered OSM vector data are converted to binary raster images and subsequently processed analogously to the query maps. Reference images are created from quadrangles of all possible map locations which need to be known in advance. Alternatively, quadrangles can be automatically generated if they are arranged in a regular grid with known spacing.

## 2.3 Image features

In modern CBIR, image content is described by the image features that can be extracted from the image (Smeulders et al., 2000). Image features have been intensely investigated by the computer vision community. Image features provide a solid methodological basis for description and matching of image content with robust and efficient algorithms.

Image features are extracted at interest points in the image. This leads to a reduction of information, focusing on the most descriptive regions in the image. Matching a set of a few hundred image descriptors is significantly more efficient and robust than comparing full high resolution images.

A major consideration for the choice of feature descriptors is their ability to capture structure instead of texture. Structure is very important to identify geographical features. Conversely, there is no texture at all in the segmented maps and reference maps. Features should be extracted predominantly at informative points, such as bends of a river as in Figure 2. Furthermore, features should be scale-invariant, because input images come in different resolutions. Scale-invariant features are more likely to capture large-scale interest points and are less susceptible to small irregularities due to noise, artefacts of rasterisation or minuscule symbols with little information.

Specifically, we extract KAZE descriptors (Alcantarilla et al., 2012), since they showed the best indexing results in



Figure 2. KAZE interest points (coloured circles) detected on a KDR100 reference map. It is apparent, that the KAZE feature detector is able to detect structure-rich areas of the depicted rivers. Data © OpenStreetMap contributors.

comparative experiments. Apparently, they are more repeatable in face of the high contrast and low texture in segmented maps than e. g. the well-established SIFT or SURF descriptors. Furthermore, KAZE descriptors have a rotation-free variant, which reduces ambiguity on maps where the direction of features is relevant and we can expect roughly north-up oriented maps. The downside of KAZE descriptors is that they take about four times longer to compute than SURF.

## 2.4 Similarity

If an image feature in the query image is very similar to an image feature in a reference image (called a match), it is likely to be a sensible tie-point. Thus, a reference image is

similar to the query image, if a high number of tie points could be found.

When we compare a set of features, uncertainty is introduced by incorrect matches. Perhaps more limiting, the spatial relationship between image features is not taken into account. But for geographical objects, it is very relevant in what orientation their parts are (e. g. whether a river bends to the left or to the right after a divide).

Therefore, we apply a second step of spatial verification: the descriptor matches for each image pair are filtered by random sample consensus (RANSAC, Fischler and Bolles, 1981), significantly increasing the confidence of the matches. With a suitable choice of a transform model, RANSAC is able to compensate for anticipated deformations (introduced e. g. through different datum, projection or map extent). For the topographic maps in this study the similarity transform is sufficient to compensate for differently cropped map margins and slight rotation.

Basing similarity on a discrete set of tie points (in contrast to image-level metrics, such as the structural similarity-index or histograms) makes it robust to missing map symbols in either map (e. g. due to cartographic generalisation). When there are few symbols on one map there is, of course, a lower possible similarity score. It has to be noted that similarity scores do not adhere to an absolute scale and cannot be compared between image pairs. However, the number of tie-points can be used to order the reference maps by similarity to a single query image and find the most similar.

## 2.5 Efficient lookup

An index of all reference maps has to be constructed once to enable the efficient search for similar maps in two stages: index query and spatial verification.

### 2.5.1 Index construction

1. Query VGI database for each of the expected map locations while filtering the vector data for relevant object classes, using Overpass[2].

2. Paint binary raster maps from vector data of each map location.

3. Detect interest points and extract image descriptors in the raster maps.

4. Populate an approximate nearest neighbours tree with all descriptors, storing an identifier of the map they were extracted from. We implemented the tree with Annoy[3].

### 2.5.2 Index query

1. From a given segmented query map, extract image features in the same way as before.

2. Query the index to get the k nearest neighbours for each descriptor.

---

[2]https://overpass-api.de/
[3]https://github.com/spotify/annoy

3. Each of the k descriptor belongs to a reference map. To penalise bad matches, vote for each of the k reference maps with decreasing weight: nearest neighbour with a single vote, second-nearest neighbour with a half vote, etc. The votes for each reference map are added for every descriptor in the query map.

4. Sort all reference maps by their number of votes.

5. Return the N most similar reference maps as location hypotheses.

### 2.5.3 Spatial verification

1. The descriptors for each reference map that had been used for building the index have been stored alongside the index to save computation time. Get the descriptors of the previously determined location hypotheses.

2. Brute-force match the descriptors of interest points in the segmented query map to find all correspondences in the descriptors for each reference map.

3. Use RANSAC to fit a transform model from the query descriptors to the reference descriptors.

4. The reference map with the transform model that best explains the matches and has the highest number of inliers after RANSAC is returned as the location prediction of the query map.

## 3. Evaluation

For all experiments, we use hydrology features (rivers, lakes, coastlines) because their distinct colour makes them easy to segment for most maps and they have very distinct geometry (Wolter et al., 2017). Furthermore, hydrology is expected to have little historical change and can be filtered reasonably well with Overpass. Other features, such as roads or railroad tracks, are thinkable as well and can be used just as easily with our method, but they are harder to tell apart from other symbols on the map and they are expected to have changed significantly since printing of the investigated maps.

### 3.1 Baseline/proof-of-method

First, we demonstrate that the image domain is in fact suitable to calculate map content similarity by matching image features. We can demonstrate the applicability of our method and provide a baseline for the maximal reachable accuracy by using the artificial black-and-white reference map images used to build the index (with some padding to simulate the map margins). This baseline experiment shows, that the similarity function is able to find the correct maps if it is not impaired by erroneous segmentation or historical changes in topography.

### 3.1.1 Data

The query maps for this experiment are the same as the reference maps for building the index (see Section 2.5.1 above) and can thus be used to determine baseline accuracy (see Section 3.1 below). An example of a query map can be seen in Figure 3 on the left. The index is constructed from 911 sheets and thus contains more sheets than the 657 maps to be queried.

| map series | #maps | mean (median) index rank | maps w/ rank=0 (%) | maps w/ rank<N (%) | #correct (%) |
|---|---|---|---|---|---|
| KDR100 baseline | 657 | 0.9 (0) | 656 (99.8) | 656 (99.8) | 656 (99.8) |
| KDR100 degraded | 657 | 23 (0) | 377 (57.4) | 677 (87.8) | 529 (80.5) |
| KDR100 actual | 657 | 37.4 (0) | 399 (60.7) | 569 (86.5) | 539 (82.0) |
| KDR500 actual | 27 | 4.7 (3) | 9 (33.3) | 27 (100) | 15 (55.6) |
| USGS100 baseline | 207 | 511.4 (0) | 122 (59.5) | 125 (61.0) | 124 (60.5) |
| USGS100 actual | 207 | 430.7 (52) | 51 (24.6) | 107 (51.7) | 92 (43.6) |

Table 1. Comparison of results for all experiments. The fifth column shows the maximum reachable accuracy of spatial verification: for the KDR100 experiments only the N=30 most similar reference maps were verfied. For USGS100, N is 100. For KDR500 all sheets were verified.



Figure 3. OSM baseline image (left) and an artificially degraded image (right) for the same map (KDR100 sheet 29). The occluded map still provided enough information to be successfully georeferenced. Data © OpenStreetMap contributors.

### 3.1.2 Result

When querying the index with the OSM maps themselves, the correct reference map is returned as the most similar map for almost all maps, both by the index as well as by the subsequent spatial verification (compare Table 1). This is expected and proves that the structure of the map symbols and their representation as image features is not ambiguous. If a map cannot be matched to itself, it probably does not contain any discernible features and thus can not be localised.

## 3.2 Robustness to occlusion

### 3.2.1 Data

This dataset aims to simulate historical changes and bad segmentation quality with degraded OSM reference maps. 500 empty circles of varying diameter have been painted randomly over baseline images of KDR100 quadrangles resulting in 73% occlusion on average (compare Figure 3). The experiment uses the same index as for the baseline.

### 3.2.2 Result

The resulting performance is still quite good: 80% of maps could be located correctly (as can be seen in Table 1). Most maps with reduced image content could be matched unambiguously. This indicates, that our method is very robust to occlusion, which might occur because of historical changes to topography or because of incomplete segmentation.

## 3.3 Real-world application

### 3.3.1 Data

This experiment uses the actual map sheets of the survey map series *Karte des Deutschen Reiches* in scale

1 : 100 000 (KDR100). The maps were made at the turn of the 20th century and use a trapezoid projection with the Rauenberg datum. 657 sheets are aligned in regular grid. Again, we use the same index as for the baseline. The map are either one or three colour prints with hydrology in blue. Because the colours are almost identical across the series, we could use colour thresholding for segmentation. Using a single colour range across the whole series, however, introduces a lot of speckle noise and sometimes imprecise outlines (compare Figure 4).

### 3.3.2 Result

When working with real maps, the main issue are errors and noise introduced by segmentation which leads to uncertainty during matching. For the KDR100 experiments, we decided to only consider the 30 most similar reference maps for spatial verification (dashed red line in Figure 5), allowing a maximal accuracy of 86.45%. After spatial verification 82% of the 657 sheets were correctly predicted.

Respecting the datum shift, the projection introduces no apparent deformations with respect to WGS84. The coverage of hydrology features in OSM is similar to the symbols visible on the maps. Therefore, the maps can be matched to the reference data with high certainty in most cases. The results are quite similar to the experiment with degraded OSM data, which suggests, that the method is mostly robust to historical changes and segmentation noise. Incorrect matches are probably owed to segmentation errors.

## 3.4 Different scales and projections

### 3.4.1 Data

*Carl Vogel's Karte des Deutschen Reiches* consists of 27 sheets in scale 1 : 500 000 (KDR500) from the turn of the 20th century. It uses the Bonne projection with unknown central meridian and standard parallel. The maps are multi-colour prints, which we segmented by colour thresholding. The index is constructed from 36 quadrangles in a regular grid that have been reprojected to WGS84.

### 3.4.2 Result

The map projection is significantly different from the reference data. In WGS84, the maps look tilted and deformed. In this small scale, the coastline becomes the most descriptive symbol in a map. On the other hand, some inland sheets show very few rivers because of strong generalisation. This generalisation makes it hard to match some inland sheets to the reference data, which contains many
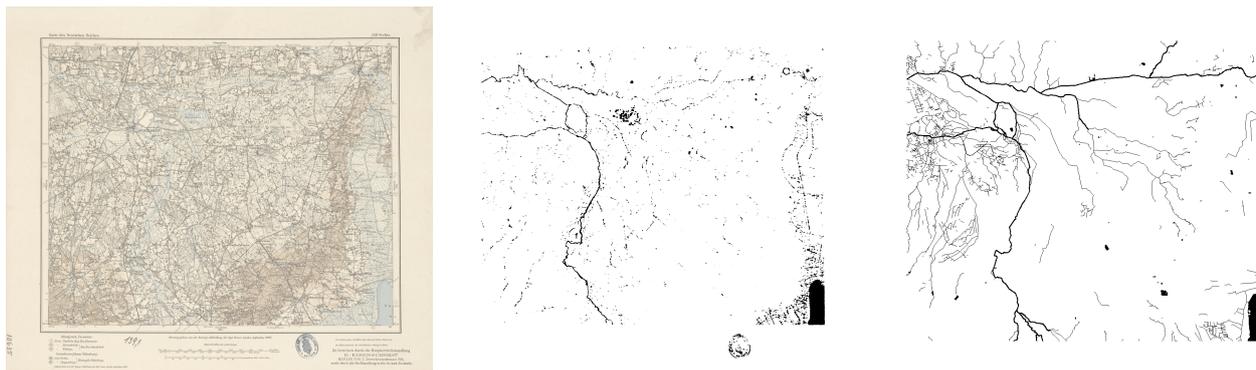
Figure 4. Segmentation errors of an actual KDR100 map (sheet 258) compared to OSM reference. Left: the original query map. Middle: segmentation of hydrology by colour thresholding. Right: the corresponding reference map from OSM. The dominant symbols in this map provided the information to geolocate this map despite the noise and missing symbols. Map © Staatsbibliothek zu Berlin, reference data © OpenStreetMap contributors.

more rivers, as we could not find OSM tags to filter them more precisely. Subsequently those inland sheets were not successfully located, resulting in the low accuracy of 55.6%, as seen in Table 1.

### 3.5 Different designs across the world

#### 3.5.1 Data

A set of 197 quadrangles of historical topographic maps across the United States in 1 : 100 000 scale from the years 1972–1980 in NAD23-projection (USGS100). Segmentation is a bit more difficult than for the other maps because the USGS maps use more colours. Therefore, we use a convolutional neural network for segmentation, as has been proposed by Jiao et al. (2020). The index covers the United States completely with 1772 reference quadrangles.

#### 3.5.2 Result

The maps show many hydrology symbols which are not mapped in OSM at all. Particularly maps washes in desert areas and small creeks in mountain areas could not be matched to the corresponding reference maps, which are often almost empty. Areas with big rivers and coastlines, on the other hand, are matched to OSM better. The result is an overall subpar performance of only 43.6% correct predictions. The comparison with the OSM baseline for the USGS100 quadrangles (see Table 1) suggests that a better source of reference data[4] or different symbols altogether (e. g. roads) are necessary to increase the accuracy for USGS maps.

### 3.6 Runtime performance

On a single CPU core with 2.9 GHz and 24 GB of RAM, constructing the index takes about 0.7 s per reference quadrangle excluding the download of reference data. Querying the index takes approximately 6.7 s per uncompressed query map, plus the variable time needed for spatial verification. Each iteration of spatial verification takes 0.3 s,

---

[4]such as the National Hydrography Database: https://www.usgs.gov/core-science-systems/ngp/national-hydrography/national-hydrography-dataset
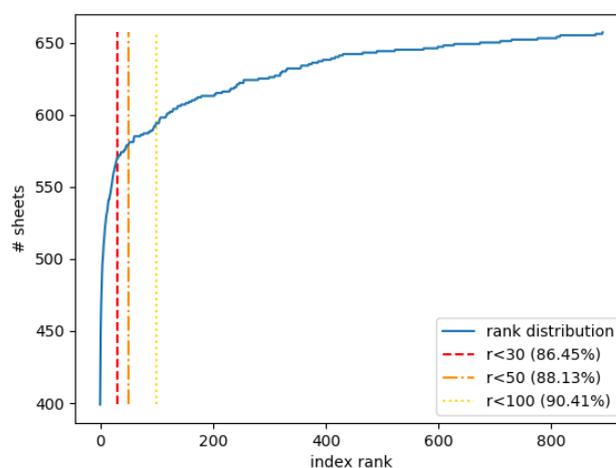


Figure 5. Index rank distribution and effect on the maximal reachable accuracy in relation to the cutoff value for KDR100 maps. Increasing the number of location hypotheses to verify increases computation time linearly, but has diminishing returns for accuracy.

leading to approximately 9 s when verifying the 30 most similar reference maps.

This clearly illustrates the importance of a sensible choice for the number of location hypotheses to return from the index. We want to save as many iterations of spatial verification as possible, to save computation time. However, this comes as a trade-off, since when the correct match is not among the location hypotheses returned from the index, it cannot be located correctly. Nonetheless, a very high number of hypotheses does not always turn out to be helpful, since maps which rank low in the index often do not carry enough information to be successfully predicted by spatial verification either. Figure 5 illustrates, how the number of location hypotheses to verify influences the maximum accuracy of KDR100 predictions.

We have experimented with an early termination heuristic, which aims to save some spatial verification steps for the maps that have already been correctly predicted by the index alone. It works by investigating the index votes of the index response. When the number of votes for the n most

similar reference image is two times larger than for the n+1 most similar image, spatial verification stops after n steps. For the KDR100, this heuristic enables 270 correct skips, usually after the first reference image, thus saving almost a third of the total computation time when processing the complete series, at the cost of failed predictions for only 18 additional sheets.

## 4. Conclusion

### 4.1 Summary

We have shown that CBIR with sparse image features-based similarity can successfully geolocate most sheets of different topographic map series. With our method, large volumes of maps can be georeferenced in a short time with significantly reduced manual labour. The algorithm scales very well and can be parallelised to be incorporated into an interactive web application such as the one proposed by Tavakkol et al. (2019).

Many maps are located with high confidence, but others might not have been successfully located with increased uncertainty (e. g. through noise). We can provide a measure of confidence with the number and distribution of RANSAC inliers. This allows to reduce manual quality control and error correction by only investigating the low confidence predictions.

RANSAC not only outputs a score, but also a transform model, which can be used for alignment of the geolocalised maps. However, these transformations did not consistently produce acceptable registration accuracy on their own and should be supplemented with one of the proposed methods of alignment in order to exploit more available spatial information (Howe et al., 2019, Duan et al., 2020, Luft and Schiewe, 2021).

Comparing with previous studies, our numbers suggest a worse performance than the 96% accuracy of Luft and Schiewe (2021). However, we have substantially increased the size of the test data set. Furthermore, the improved method presented here significantly improves query time and is better suited to apply to various input maps without much parameter tuning.

Calculating the total accuracy of all three map series, our method slightly outperforms the method proposed by Tavakkol et al. (2019). They have found location predictions for 68% of 500 very diverse maps of different type and scale. However, the quality of their solutions, given by error distance, is not particularly meaningful without taking the scales of the respective maps into consideration. Their work indicates that toponym-based geolocation transfers easily to a variety of maps but is error-prone and challenging to implement robustly.

### 4.2 Future work

The KDR500 reference maps were shifted and skewed slightly because of reprojection. The experiment shows that the method is to some extent robust to inexact reference locations in the index. Pushing the boundaries of the index robustness to displacement, in the hopes of being able to do without the prior knowledge of quadrangle locations altogether, is subject to further research.

Complex or unusual map projections make direct comparison to the reference OSM data difficult. Reprojection of either the reference data or the input maps (compare our experiments on KDR500) avoids any issues. However, reprojection requires prior knowledge of the used map projections. Sometimes the exact spatial reference system might be unknown and in order to be able to deal with deformations in the input maps, more complex transform models for RANSAC have to be evaluated. This also raises the question of how robust image feature descriptors are to projection deformations. They are usually designed to be invariant to geometric deformations, like rotation and scaling, but their performance under e. g. polynomial deformations has to be evaluated.

Segmentation is the crux for generalisability. The method here works on an abstracted image and is universal, but getting there by extracting the "right" symbols is different for each map series and does not easily transfer without manual parameter tuning. Deep learning segmentation models which are able to transfer between different map styles could be the solution. High performance with little training data is paramount, because collecting and annotating large amounts of training data quickly gets out of proportion to the expected time savings of an automated georeferencing system. We are currently working on a deep learning segmentation model using synthetic training data, which is the preferred solution when there is limited availability of training data, as is the case with historical maps.

Our research continues with the goal of improving accuracy, making our method more general by allowing the use of multiple different symbols at once and increasing ease of use by reducing metadata requirements.

## References

Alcantarilla, P. F., Bartoli, A. and Davison, A. J., 2012. KAZE features. In: A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato and C. Schmid (eds), *Proceedings of the 12th European Conference on Computer Vision*, Lecture Notes in Computer Science, Vol. 4, Springer, p. 214–227.

Buckley, A. R., 2019. Sharing collections of historical maps online. *Abstracts of the ICA*.

Burt, J., White, J., Allord, G., Then, K. and Zhu, A.-X., 2019. Automated and semi-automated map georeferencing. *Cartography and Geographic Information Science* 47, pp. 40–66.

Chiang, Y.-Y., Duan, W., Leyk, S., Uhl, J. H. and Knoblock, C. A., 2020. *Using Historical Maps in Scientific Studies: Applications, Challenges, and Best Practices*. SpringerBriefs in Geography, Springer International Publishing, Cham, Switzerland.

Chiang, Y.-Y., Leyk, S. and Knoblock, C. A., 2014. A survey of digital map processing techniques. *ACM Computing Surveys* 47(1), pp. 1–44.

Crom, W., 2016. Kartendigitalisierung – buntes Bild oder Mehrwert. *Kartographische Nachrichten* 66(5), pp. 243–248.

Duan, W., Chiang, Y.-Y., Leyk, S., Uhl, J. H. and Knoblock, C. A., 2020. Automatic alignment of contemporary vector data and georeferenced historical maps using reinforcement learning. *International Journal of Geographical Information Science* 34(4), pp. 824–849.

Fabris, M., 2021. Monitoring the coastal changes of the Po river delta (northern italy) since 1911 using archival cartography, multi-temporal aerial photogrammetry and lidar data: Implications for coastline changes in 2100 a.d. *Remote Sensing* 13(3), pp. 1–23.

Fischler, M. A. and Bolles, R. C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM* 24(6), pp. 381–395.

Fleet, C., Kowal, K. C. and Pridal, P., 2012. Georeferencer: Crowdsourced georeferencing for map library collections. *D-Lib Magazine* 18(11/12), pp. 1–11.

Heitzler, M. and Hurni, L., 2019. Unlocking the geospatial past with deep learning – establishing a hub for historical map data in switzerland. *Abstracts of the ICA* 1, pp. 1–2.

Heitzler, M., Gkonos, C., Tsorlini, A. and Hurni, L., 2018. A modular process to improve the georeferencing of the siegfried map. *e-Perimetron* 13(2), pp. 85–100.

Howe, N. R., Weinman, J., Gouwar, J. and Shamji, A., 2019. Deformable part models for automatically georeferencing historical map images. In: *Proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, SIGSPATIAL '19, Association for Computing Machinery, Chicago, IL, USA, p. 540–543.

Iosifescu, I., Tsorlini, A. and Hurni, L., 2016. Towards a comprehensive methodology for automatic vectorization of raster historical maps. *e-Perimetron* 11(2), pp. 57–76.

Janowicz, K., Raubal, M. and Kuhn, W., 2011. The semantics of similarity in geographic information retrieval. *Journal of Spatial Information Science* 2011(2), pp. 29–57.

Jiao, C., Heitzler, M. and Hurni, L., 2020. Extracting wetlands from swiss historical maps with convolutional neural networks. In: *Automatic Vectorisation of Historical Maps: International workshop organized by the ICA Commission on Cartographic Heritage into the Digital*, Department of Cartography and Geoinformatics ELTE, Budapest, Hungary, p. 33–38.

Jiao, C., Heitzler, M. and Hurni, L., 2021. A survey of road feature extraction methods from raster maps. *to appear in Transactions in GIS*.

Li, B. and Fonseca, F., 2006. TDD: A comprehensive model for qualitative spatial similarity assessment. *Spatial Cognition & Computation* 6(1), pp. 31–62.

Li, J., Hu, Q. and Ai, M., 2020. RIFT: Multi-modal image matching based on radiation-variation insensitive feature transform. *IEEE Transactions on Image Processing* 29, pp. 3296–3310.

Liu, L., Wang, Y. and Wang, Y., 2008. SIFT based automatic tie-point extraction for multitemporal SAR images. In: *2008 International Workshop on Education Technology and Training 2008 International Workshop on Geoscience and Remote Sensing*, Vol. 1, Shanghai, China, p. 499–503.

Luft, J., 2020. Automatic georeferencing of historical maps by geocoding. In: *Automatic Vectorisation of Historical Maps: International workshop organized by the ICA Commission on Cartographic Heritage into the Digital*, Department of Cartography and Geoinformatics ELTE, Budapest, Hungary, p. 77–89.

Luft, J. and Schiewe, J., 2021. Automatic content-based georeferencing of historical topographic maps. *to appear in Transactions in GIS*.

Schlegel, I., 2019. Empirical study for a deployment of a methodology for improving the comparability between historical and current maps. *KN - Journal of Cartography and Geographic Information* 69(2), pp. 121–130.

Smeulders, A., Worring, M., Santini, S., Gupta, A. and Jain, R., 2000. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(12), pp. 1349–1380.

Sun, K., Hu, Y., Song, J. and Zhu, Y., 2020. Aligning geographic entities from historical maps for building knowledge graphs. *to appear in International Journal of Geographical Information Science*.

Szendrei, R., Elek, I. and Márton, M., 2011. A knowledge-based approach to raster-vector conversion of large scale topographic maps. *Acta Cybernetica* 20, pp. 145–165.

Tavakkol, S., Chiang, Y.-Y., Waters, T., Han, F., Prasad, K. and Kiveris, R., 2019. Kartta labs: Unrendering historical maps. In: *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on AI for Geographic Knowledge Discovery*, GeoAI 2019, Association for Computing Machinery, New York, NY, USA, p. 48–51.

White, A. P., 2013. X marks the spot: Extracting data from historical maps to locate archaeological sites. *Journal of Map & Geography Libraries* 9(1–2), pp. 140–156.

Wolter, D., Blank, D. and Henrich, A., 2017. Georeferencing river networks using spatial reasoning. In: *Proceedings of the 11th Workshop on Geographic Information Retrieval - GIR'17*, ACM Press, Heidelberg, Germany.

Ye, Y. and Shen, L., 2016. HOPC: a novel similarity metric based on geometric structural properties for multi-modal remote sensing image matching. III–1, pp. 9–16.

Zhuo, X., Koch, T., Kurz, F., Fraundorfer, F. and Reinartz, P., 2017. Automatic UAV image geo-registration by matching UAV images to georeferenced image data. *Remote Sensing* 9(44), pp. 376.