ROBUST GRID DETECTION IN HISTORICAL MAP IMAGES

Josef Baloun[1,2], *Ladislav Lenc*[1,2], *Pavel Král*[1,2]

[1] Dept. of Computer Science & Engineering
University of West Bohemia
Plzeň, Czech Republic

[2] NTIS - New Technologies for the Information Society
University of West Bohemia
Plzeň, Czech Republic

{balounj, llenc, pkral}@kiv.zcu.cz

ABSTRACT

This paper presents a novel method for grid detection in historical maps. The approach is based on Hough transform accompanied with a sophisticated post-processing. They are applied to detect the grid that consists of graticule lines. It works without any training and does not require any annotated data. The proposed approach is very efficient in detecting the rectangular grid and the intersection points as shown in the international "MapSeg" segmentation competition, where it won the Task 3 with a significant margin. The robustness of the proposed method has been demonstrated by evaluating on another dataset composed of significantly different cadastral map images with excellent results.

***Index Terms*—** Grid Detection, Historical Map Segmentation, Hough Transform, Geo-reference

1. INTRODUCTION

Digitization of historical maps is an ongoing process and a large number of such maps are already scanned and accessible. However, to be able to efficiently utilize the maps, scanning and storing it in a digital form is not enough. It is beneficial to perform several additional tasks which brings many research problems in the image processing field.

A crucial task, which is often performed is geo-referencing. It allows a proper positioning of the map images and aligning them with recent maps. Important features that can be used for geo-referencing the maps are graticule lines indicating the North/South/East/West major coordinates and their intersections. [1] Another use case is assembling the single map sheets into a larger seamless map. This task was also the part of the ICDAR 2021 Competition on Historical Map Segmentation[1] [1] (*MapSeg*). In this work we concentrate on the graticule line intersections detection according to the *Task 3* of the competition and present the winning *UWB* method.

The task offers several solutions like point or line detection using conventional computer vision (CV) and also deep learning techniques. Especially with little training data, deep learning approaches contain noise in predictions and require additional post-processing in order to provide good results. This can be the reason, that traditional CV methods can still compete with deep learning approaches and perform well as stated in [2]. Traditional CV approaches usually start with the detection of significant parts of the grid (e.g. longest line) and then build the grid. This can produce fault results, especially if this part is out of the grid.

Therefore, we propose a method that detects the grid directly as a set of lines. The result is thus much more robust. Further we propose an approach for intersections refinement to provide the precise outcome that can be directly used for geo-referencing the map images.

In order to show the robustness of the proposed method we evaluate it on another dataset composed of cadastral map images which significantly differ to the "MapSeg" competition data.

[1] https://icdar21-mapseg.github.io/

2. RELATED WORK

For the solved task, methods for object detection, line detection or segmentation are potentially useful. We can mention template matching approaches [3], Hough transform [4], Convolutional neural networks [5] and more recently fully convolutional networks [6, 7, 8]. Further we report approaches that were submitted to the MapSeg competition and are designed specifically for this task.

2.1. CMM

The CMM [1] approach uses morphological processing and the Radon transform to detect the graticule lines. The lines correspond to the Radon transform maxima at corresponding and orthogonal angle. The angle is selected based on the longest detected line. Further it uses auto-correlation to obtain the period of the grid and finally it uses morphological closing with two perpendicular lines as a refinement.

2.2. IRISA

The IRISA [2] method consists of a line segment detector and grammar rules. Firstly, a large amount of line segments are detected. These are filtered using a set of rules that constraint spacing or perpendicularity for example. The grid is composed of a longest segment cross and the additional parallel rulings. Finally, the line intersections are computed.

2.3. L3IRIS

The L3IRIS [1] approach uses the U-Net [6] as intersection point detector. There are several noisy point predictions in the output so

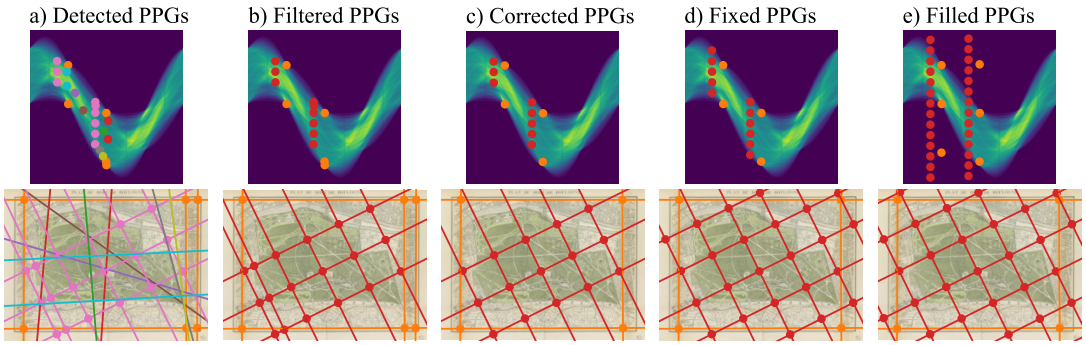

Fig. 1. Graticule candidate generation process: Hough accumulator (top) peaks from the same perpendicular peak group (PPG) and their visualization (bottom) have the same color. Note that image origin is placed at top left corner and y-axis is inverted so the angle goes clockwise.

the Hough transform is used to detect candidate lines. These lines are filtered using a proposed clustering. The intersections of resulting parallel and orthogonal lines are finally filtered using the map content area mask.

3. GRATICULE DETECTION

Graticule lines cover the whole map content area overlapping many elements. The lines form practically a grid and they can be of any angle. Usually, they are not strictly straight and can be suppressed or degraded.

To detect graticule line intersections, we first locate the grid-forming graticule lines using Hough transform [4]. The Hough transform produces an accumulator (Fig. 1) containing more or less noise depending on the input image. Since the graticule lines in the input image are thin and not easily distinguishable, the input image is firstly binarized and pre-processed. Then, the grid can be detected under the following presumptions:

1. graticule lines are straight;
2. graticule lines are equidistant;
3. graticule lines are parallel or perpendicular to each other;
4. there are at least three detectable graticule lines, of which at least two lines are perpendicular.

It is expected that graticule lines will locally over-vote other lines and noise. This expectation depends on the first presumption. Generally, the presumption does not have to be fully satisfied and allows some tolerance but it affects the local maxima positions and values in the Hough accumulator. Tolerance is also allowed in the second and third presumptions, which are important for the target peak pattern. The last presumption is strict and allows determining the distance interval between lines, filtering and rating of the candidates. The better fulfilled the presumptions, the better results can be expected.

Using parameters t_a (line angle difference tolerance in grid), d_m (minimal allowed distance of parallel lines) and t_{d_r} (parallel line relative distance tolerance in grid), the proposed method firstly detects peaks in the Hough accumulator and selects angle candidates. Based on the angle candidates, peaks are grouped into perpendicular peak groups (PPGs). Further steps include filtering, correcting, fixing, rating and filling the PPG. As a result, we have several graticule candidates that are represented by the PPGs and contain information about their rating, angle and distance between the lines. Based on these, we can select the best candidate, detect intersection points and also refine them.

3.1. Pre-processing

Bigger components and mainly thin lines are crucial for the task. Therefore, a recursive Otsu binarization method [9] is used. It preserves well thin lines, treats noise, brightness inconsistency and other degradation in the map frame. It also deals in an adaptive manner with large homogeneous areas that could be disruptive for Hough transform. Optionally, the result is masked with map content area from the Task 2 of the MapSeg competition.

3.2. Peak detection

Peaks in the Hough accumulator (Fig. 1.a) are detected as local maxima in a certain area. That area should not be neither too small (can result in noisy peaks) nor too big (can result in missed graticule peaks). The height of an area is given by $h = 2d_m + 1$ where d_m is a minimal distance parameter which determines how close the lines can be so that they are still detectable. The width of the area is set experimentally to $\frac{\pi}{5}$. This setup results in suitable amount of peaks even in noisy inputs. Since the graticule lines usually overvote "noise" lines in the image, the width settings from $\frac{\pi}{10}$ to $\frac{\pi}{2}$ yields comparable results on both examined datasets.

3.3. Angle candidate selection

The angle error of 2 parallel lines can be approximated with normal distribution. The angle candidate is then selected as the local maxima of weighted arithmetic mean. As weights, we use Gaussian window given by $2\sigma = t_a$. It is used for convolution with angle histogram. The histogram is obtained summing the peak values according to the angle coordinate of the peak. Taking into account perpendicularity, the perpendicular angle values are summed. Since it is periodic, only the peaks between $\frac{\pi}{4}$ and $\frac{3\pi}{4}$ are considered.

3.4. Peak grouping

The grouping of peaks utilizes the 3. presumption saying that the peaks should have corresponding angle-axis value. The perpendicular peak group (PPG) (Fig. 1) is composed of two peak groups. Peaks in the PG have the same angle $\pm$ tolerance (t_a) and represent

parallel lines. For a given angle candidate a, first and second PG of the PPG contain peaks whose angle-axis values match $a \pm t_a$ and $a + \frac{\pi}{2} \pm t_a$ respectively.

3.5. PPG filtering

As can be seen in Fig. 1.a, there are several detected noisy PPGs that do not satisfy the presumption no. 4, which can be rewritten into the following conditions:

1. the PPG has to contain at least three peaks;
2. the PPG has to have at least one peak in each PG.

Then the filtration is straightforward, the PPG is filtered out if it does not satisfy all conditions.

3.6. Distance interval selection

According to 2. presumption, graticule line intervals should be the same, but usually there is an error. This error can be approximated by normal distribution and candidates selected as in Sec. 3.3.

The histogram is created from distance-axis differences between two consecutive peaks within each PG. Gaussian window is given by Eq. 1, where t_{d_r} is relative distance tolerance parameter and i_e is median interval as initial estimate.

$$2\sigma = t_{d_r} \cdot i_e \tag{1}$$

According to Fig. 2, the distance interval is identified either as the maximal value or additional interval candidates can be identified from the peaks.

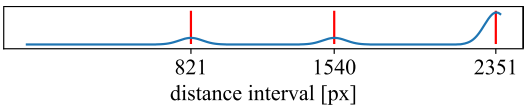

Fig. 2. Distance interval probabilities and detected peaks of red PPG from Fig. 1.b.

3.7. PPG correcting

As can be seen in Fig. 1.b and 2, there are several intervals due to noisy peaks. For every possible interval, we try to find the biggest group of peaks that corresponds to that interval within both PGs. This way, several PPGs can be made of a single PPG. Finally, filtering according to Sec. 3.5 is repeated.

3.8. PPG fixing

The purpose of this step is to detect previously undetected insignificant peaks that match the expected position. The position is given by the same angle, corresponding distance interval and tolerance (t_{d_r} and t_a). This way, the peaks are not over-voted by noisy peaks and even insignificant lines can be detected as can be seen in Fig. 1.d.

3.9. PPG rating

Since we have detected several grid candidates, we need to distinguish the good ones. Therefore we propose heuristic rating inspired by heart rate variability analysis [10]. According to Eq. 2, we consider the set of intervals i as a signal s, where the mean value $\bar{i}$ is subtracted because of 0 Hz component and the result is divided by probable interval i_p to consider relative differences among intervals.

$$s = (i - \bar{i})/i_p \tag{2}$$

The rating r is computed according to Eq. 3, where n stands for the number of intervals. $\overline{FFT_{as}(s)}$ represents the mean value of amplitude spectrum and penalize the interval differences. The bigger amount and more stable intervals results in lower rating. The lower the rating, the better it is.

$$r = \overline{FFT_{as}(s)}/n^2 \tag{3}$$

3.10. PPG filling

Last but not least, PPG is filled with missing but expected peaks using probable distance interval and corresponding angle coordinate according to Fig. 1.e.

3.11. Graticule line intersection detection

Intersections are determined using homogeneous coordinates [11]. It is straightforward if we represent lines in normal form and detect peaks with angle and distance coordinates.

3.12. Graticule line intersection refinement

The 1. presumption on straight lines is not usually satisfied so detected intersections are rather estimates of intersections and their localization can be improved as can be seen in Fig. 3.

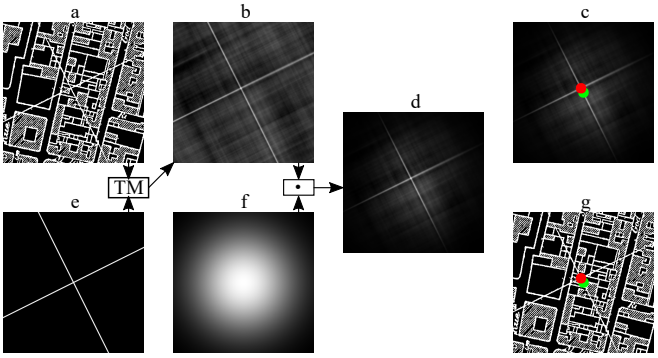

Fig. 3. Graticule lines intersection refinement process: *(a)* cropped binarized input image, *(b)* cropped result of template matching, *(c)* red refinement of green estimated point using maximal value of *d*, *(d)* pixel-wise product of *b* and *f*, *(e)* template for template matching, *(f)* Gaussian hill with estimated point as center, *(g)* red refinement of green estimated point visualization.

To do that we use normalized correlation coefficient template matching [3]. The template is a cross rotated accordingly to the grid (Fig. 3.e). The result of the template matching is multiplied with Gaussian hill to limit the refinement distance and also to approximate error distribution. The center of the Gaussian hill is positioned at the intersection estimate and its size is given by the maximal refinement distance parameter (r_d). Finally, the maximal value determines position of refined intersection.

4. EXPERIMENTAL SET-UP

We use the same scenario, provided evaluation tools and dataset [12] as in MapSeg competition [1]. The source codes and other related materials are freely available for non-commercial purposes at `https://gitlab.kiv.zcu.cz/balounj/21_icdar_mapseg_competition`.

For grid detection, the tolerance parameters $t_a = \frac{\pi}{360}$ and $t_{d_r} = 0.05$ are used. Since there are two possible distance intervals between lines ($d = \pm 1180$ or ± 2360 pixels), we use two minimal distance parameters as $d_m = 0.6 \cdot d$ for each. The best PPG is selected from the candidates with regard to its rating and distance interval. PPGs with intervals close to d are favored. For the intersections refinement (Sec. 3.12), r_d is set to 200 pixels. Finally, the map content area mask from Task 2 of the MapSeg competition is used to filter the intersection points. For more details, visit the provided website.

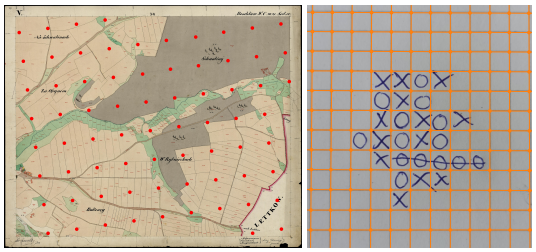

Fig. 4. Grid detection example in different images.

In order to show the robustness of the algorithm, we further annotated 21 cadastral Austro-Hungarian maps from the 19th century[2] (Fig. 4 left). Compared to MapSeg dataset, these sheets contain indistinct lines, thus the binarization is not optimal and many graticule line segments are missing. On the other hand, there are more lines and they are straighter. We use the same parameters except the distance interval $d = 826$ pixels and thus $d_m = 0.6 \cdot d$.

Generally, the parameters are tolerant to improper setting. Changes in t_a ($[\frac{0.5\pi}{180}, \frac{4\pi}{180}]$), t_{d_r} ($[0.05, 0.2]$) and d_m ($[0.2, 0.8]$ of expected distance interval) produced comparable results on both datasets. Except the r_d for refinement that depends on the binarization quality as can be seen in Table 1.

Dataset \ r_d	0	20	50	200	400
MapSeg	77.7	82.4	87.1	93.6	91.4
Ours	91.9	97.1	96.6	89.7	84.3

Table 1. PDS [%] for different maximal refinement distances (r_d) on different datasets.

4.1. Evaluation criteria

The provided evaluation tools and the same scenario as in [1] are used for evaluation. Points detection score (PDS) represents the area under $F_{\beta=0.5}$ vs distance threshold curve. The predicted point is considered correct if it is closer than threshold. The threshold ranges from 0 to 50 pixels. The final measure is computed as an average of the all test image measures.

[2] `https://corpora.kiv.zcu.cz/map_border/`

Method	CMM [1]	IRISA [2]	L3IRIS [1]	Ours[3]
PDS [%]	86.6	89.2	73.6	**93.6**

Table 2. Final PDS for locating graticule line intersections task.

5. RESULTS

The proposed method shows excellent results even on noisy historical map images. As can be seen in Table 2 and Fig. 5, our method surpassed the other methods by a significant margin.

There may appear a small amount of fault refinements that are caused usually by noisy lines. Some extra predictions are also present in the map content area exceeding map frame or at map borders. It is hard for us to decide their correctness, since they correspond to the expected position in the grid. Those points could be probably filtered with improved map content area segmentation for example. The correction is also very beneficial as could be seen in Table 1. In total, more than 90 % of all predicted points are closer than 3 pixels from provided GT. This can be considered perfect taking into account high image resolution and possible GT error.

Based on the $F_{\beta=0.5}$ from Fig. 5 and taking into account the approach order, we deduce that focusing on bigger part of the grid (whole grid, cross, longest line etc.) brings more correct detections and is thus more robust.

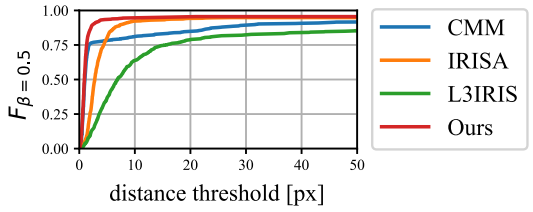

Fig. 5. Locate graticule lines intersections results: $F_{\beta=0.5}$ vs distance threshold across test images.

6. CONCLUSIONS

In this paper we have proposed an efficient and robust approach for map grid (graticule) detection with three intuitive parameters. Hough transform, several structure constraints and sophisticated post-processing are used for determining a set of graticule lines and their intersections. In order to allow the best possible georeferencing, we have further proposed the refinement approach to maximize the localization accuracy of the intersections.

The proposed approach is very efficient as shown in the international "MapSeg" segmentation competition, where it won the Task 3 with a significant margin. We have also shown the robustness of the proposed approach by evaluating it on another significantly different dataset with excellent results.

7. ACKNOWLEDGEMENTS

This work has been partly supported by Grant No. SGS-2022-016 Advanced methods of data processing and analysis.

[3] Compared to PDS = 92.5 % in [1], the improvement is due to minor changes in the algorithm.

8. REFERENCES

[1] Joseph Chazalon, Edwin Carlinet, Yizi Chen, Julien Perret, Bertrand Dumenieu, Clement Mallet, Thierry Geraud, Vincent Nguyen, Nam Nguyen, Josef Baloun, Ladislav Lenc, and Pavel Kral, "ICDAR 2021 Competition on Historical Map Segmentation," in *DOCUMENT ANALYSIS AND RECOGNITION, ICDAR 2021, PT IV*, Llados, J and Lopresti, D and Uchida, S, Ed. IAPR, 2021, vol. 12824 of *Lecture Notes in Computer Science*, pp. 693–707, 16th IAPR International Conference on Document Analysis and Recognition (ICDAR), ELECTR NETWORK, SEP 05-10, 2021.

[2] Lemaitre Aurelie and Camillerapp Jean, "Segmentation of historical maps without annotated data," in *The 6th International Workshop on Historical Document Imaging and Processing*, 2021, pp. 19–24.

[3] Roberto Brunelli, *Template matching techniques in computer vision: theory and practice*, John Wiley & Sons, 2009.

[4] Richard O Duda and Peter E Hart, "Use of the hough transformation to detect lines and curves in pictures," *Communications of the ACM*, vol. 15, no. 1, pp. 11–15, 1972.

[5] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik, "Region-based convolutional networks for accurate object detection and segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 1, pp. 142–158, 2015.

[6] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, Eds., Cham, 2015, pp. 234–241, Springer International Publishing.

[7] Josef Baloun, Pavel Král, and Ladislav Lenc, "Chronseg: Novel dataset for segmentation of handwritten historical chronicles," in *Proceedings of the 13th International Conference on Agents and Artificial Intelligence - Volume 2: ICAART,*. INSTICC, 2021, pp. 314–322, SciTePress.

[8] Farzin Foroughi, Jikai Wang, Alireza Nemati, Zonghai Chen, and Haoyuan Pei, "Mapsegnet: A fully automated model based on the encoder-decoder architecture for indoor map segmentation," *IEEE Access*, 2021.

[9] Oliver Nina, Bryan Morse, and William Barrett, "A recursive otsu thresholding method for scanned document binarization," in *2011 IEEE Workshop on Applications of Computer Vision (WACV)*. IEEE, 2011, pp. 307–314.

[10] A John Camm, Marek Malik, J Thomas Bigger, Günter Breithardt, Sergio Cerutti, Richard J Cohen, Philippe Coumel, Ernest L Fallen, Harold L Kennedy, Robert E Kleiger, et al., "Heart rate variability: standards of measurement, physiological interpretation and clinical use. task force of the european society of cardiology and the north american society of pacing and electrophysiology," 1996.

[11] Václav Skala, "Intersection computation in projective space using homogeneous coordinates," *International Journal of Image and Graphics*, vol. 8, no. 04, pp. 615–628, 2008.

[12] Joseph Chazalon, Edwin Carlinet, Yizi Chen, Julien Perret, Clément Mallet, Bertrand Duménieu, and Thierry Géraud, "Icdar 2021 competition on historical map segmentation dataset," online dataset, 2021.