Detection of incomplete rectangles with an implicit fragment-based model

Extended abstract*

Igor Zingman[1], Dietmar Saupe[1], Karsten Lambers[2]

[1] Department of Computer and Information Science, University of Konstanz, Germany.

[2] Institute of Archaeology, Heritage Sciences and Art History, University of Bamberg, Germany

1 Introduction

During recent years substantial progress has been made in object detection in images. Most approaches were focused on development of efficient algorithms for detecting complex patterns, based on learning from a large number of available examples. Prominent examples of such complex patterns are faces, pedestrians, or vehicles, [4]. Some approaches use local low-level features and their more global summaries to capture these patterns. Other approaches use higher-level descriptions that capture spatial relations among object parts identified by dedicated detectors. Potentially, such descriptions are more discriminative, but may also suffer from lower robustness to object variability.

Detection of simple geometric shapes such as rectangles, seems to be much easier to deal with. The detection task is indeed much easier when these shapes are perfect and located on simple background. The Hough transform, for example, can be used to detect a rectangle based on detection of its parts, i.e. lines, [2]. Moreover, a generalized Hough transform can be used to detect any shape analytically described by an equation. However, in real images the shapes might be distorted, only approximately resembling an exemplary shape. Complex backgrounds with clutter and with irrelevant structures make the problem even worse. The task becomes very challenging when the contrast of the target shape is comparable or even lower than of the clutter or irrelevant structures. The parts of the geometric shapes, e.g. line segments, are simple enough to be falsely detected with a high rate even in relatively simple backgrounds. To prevent resulting false detections of the target shapes, a suitable description of relations among their parts must be highly distinctive. In this respect, the detection of simple shapes in complex backgrounds is even more difficult than detection of complex shapes. To derive distinctive descriptions, the most powerful approaches involve learning from a large set of examples. In some applications, however, only a few positives are available. These are representative enough to allow humans

* The corresponding full paper was recently submitted to Computer Vision Image Understanding journal (the draft is available online).

to reliably detect the target shapes, but are hardly enough for computer vision classification algorithms with dozens of features involved.

In our work, we address the problems mentioned above developing a mid-level description of a set of extracted line segments summarized in new rectangularity feature. The rectangularity feature quantifies the degree of consistency of an optimal subset of line segments with a contour of rectangular shape, arbitrary size, and aspect ratio. We show that the rectangularity feature is effective for detection of ruins of rectangular livestock enclosures in high-resolution remotely sensed images. We also show that the detection performance can be improved using an additional feature, and learning feature combination from a small number of representative examples of the livestock enclosures and a large number of available negatives. In the presence of the difficulties described above, the obtained performance is hardly achievable with other approaches for detection of rectangles or with related approaches used, e.g., for detection of buildings, [3].

2 Approach

In our approach, a binary map of edges accompanied by angle information is computed first. Candidate points with the suitable sizes of an analysis window are determined as in [6]. Line segments are then found and modeled by a triple of parameters (θ, r, l), the orientation, position, and size, with the use of a variant of a local Hough transform. Our strategy for grouping rectangle fragments is based on an undirected graph. It is constructed with nodes corresponding to line segments and edges encoding spatial relations between line segments. Each node of the graph is attributed by the triple (θ, r, l) of the corresponding line segment. An edge (k, j) of the graph is attributed with the mutual angle $\beta_{k,j}$ and the pairwise convexity $\tau_{k,j}$ of the corresponding pair of line segments. An edge (k, j) is included in the graph if $\beta_{k,j}$ and $\tau_{k,j}$ satisfy particular angle and convexity constraints. This attributed graph encodes properties of line segments and their spatial relationships. Due to the construction of the graph, its maximal cliques correspond to valid configurations $\mathbf{C}$ of line segments matching a hypothesized rectangle. These configurations are then ranked by a rectangularity measure ρ that encodes the goodness of grouping the segments into a rectangular structure

$$\rho^4(\mathbf{G}^c) = \left(\sum_{\{k,j\} \in E^c} l_k l_j \mathrm{f}_{90}(\beta_{k,j}) \mathrm{f}_{cv}(\tau_{k,j}) \right) \left(\sum_{\{k,j\} \in E^c} l_k l_j \mathrm{f}_{180}(\beta_{k,j}) \mathrm{f}_{cv}(\tau_{k,j}) \right) \tag{1}$$

where $\mathbf{G}^c$ denotes a graph corresponding to valid configuration $\mathbf{C}$ of lines segments, E^c is a set of graph edges, and $\mathrm{f}_{90}, \mathrm{f}_{180}, \mathrm{f}_{cv}$ are weighting functions with peaks at 90 and 180 degrees, respectively, for mutual angles, and at 0 value for convexity measurements. We define the rectangularity feature as the largest rectangularity measure among the maximal cliques $\mathcal{M}$ of the graph $\mathbf{G}^w$

$$f_R(\mathbf{G}^w) = \max_{\mathbf{G}^c \in \mathcal{M}(\mathbf{G}^w)} \rho(\mathbf{G}^c) \tag{2}$$

The rectangularity feature is a function of an attributed graph $\mathbf{G}^w$ or a corresponding set of line segments $\mathbf{W}$ within an analysis window. To allow fast processing, an analysis window is placed at carefully chosen candidate points [6]. The rectangularity feature effectively quantifies the distinctive alignment of a group of line segments that originated from an approximately rectangular contour. The feature is sensitive to incomplete and fragmented rectangular structures. However, it has zero values at junctions, corners, lines or other configurations of line segments that do not form at least three sides of a rectangle.

A detector of rectangular structures can be built by setting an appropriate threshold on the rectangularity feature. It can be either automatically learned from the data or manually set based on experiments. We also introduced an adjusted rectangularity feature to reduce the false positives rate. The adjusted rectangularity feature is a weighted sum of the rectangularity feature and an additional feature that measures the structure size. To find optimal weights, we developed the Fisher Linear Discriminant - like approach that learns from the large number of negative examples and adapts itself to a few positives, which are the settings in our particular problem of detection of livestock enclosures [5].

3 Experiments

The livestock enclosures are manmade structures that sparsely appear in alpine environments (see Fig. 1 on the left). They usually resemble rectangular contours of various sizes and aspect ratios with nearly linear walls that may be heavily ruined. We use satellite and aerial images of 0.5m resolution where the wall width does not exceed two pixels, see Fig. 1 on the right. The ruined walls are usually of a low height, which results in low contrast linear features in images. Nearby non-relevant structures, such as rivers, trails, and rocks are often of higher contrast due to larger size or distinctive spectral properties. Detection of such faint enclosure structures in a complex terrain is a very challenging task. Because of the difficulties described above, commonly used methods for rectangle detection are not applicable. Unfortunately, only nine examples of well-preserved enclosures taken from aerial and satellite images were available to us in this study. On the other hand, a large number of negative examples can easily be generated from the available satellite imagery of the large area of the Silvretta mountains. The data stems from the recent Silvretta Historica project [5]. In our experiments we used an image of 19000×10000 pixel size, which corresponds to the area of about 48km^2, in order to generate negative examples.

We compared the performance of the rectangularity feature with the performance of the feature recently proposed in [1] for building detection. It is based on the estimate of the gradient orientation density function (GODF) that captures the distribution of orientations of intensity gradients. To compare discrimination ability of the features we tested several suitable measures. The most intuitive among them is the number of false detections obtained for the sensitivity allowing detection of all available positive examples. Using our data, we obtained $170, 120, 6522$ false detections for the rectangularity feature, the adjusted rectangularity feature, and the GODF-based feature, respectively. The false structures were detected out of 403716 candidate positions extracted from the image.

4 Conclusions

We proposed a rectangularity feature that predicts the locations of approximately rectangular contours that may be incomplete and fragmented. These structures were modeled by convex configurations of line segments with orientation angles constrained to be close to zero or ninety degrees. The model was implicitly defined by pairwise segment constraints with evidence for line segments obtained by means of a variant of a local Hough Transform. The rectangularity feature is not sensitive to structures composed of less than three sides, which allows us to avoid a large number of random configurations of extracted lines occurring in images with complex or cluttered background.

We suggested an FLD-like classifier that may incorporate additional features and can learn from a large number of readily available negatives and just a few representative positives. We showed that adding a feature that quantifies the size of the structure improves the detection performance on our data.

The introduced rectangularity and the adjusted (learnt) rectangularity features have shown good performance discriminating ruined enclosures from irrelevant structures and clutter in remotely sensed images. Due to the inherent difficulties of the problem, such a performance is hardly achievable with other approaches for detection of rectangles. Experiments showed that the rectangularity feature is far more discriminative in the detection of the enclosures in comparison with the GODF-based feature recently proposed for detection of buildings in remotely sensed images.

Fig. 1. Left: Example of well-preserved livestock enclosure of rectangular shape. Right: Aerial image with the enclosure structure corresponding to the left picture.

References

1. Benedek, C., Descombes, X., Zerubia, J.: Building development monitoring in multitemporal remotely sensed image pairs with stochastic birth-death dynamics. IEEE Trans. Pattern Anal. Mach. Intell. 34(1), 33–50 (2012)
2. Jung, C.R., Schramm, R.: Rectangle detection based on a windowed Hough transform. In: Proceedings of the Computer Graphics and Image Processing (SIBGRAPI), XVII Brazilian Symposium. pp. 113–120 (2004)
3. Mayer, H.: Automatic object extraction from aerial imagery - a survey focusing on buildings. Computer Vision and Image Understanding 74(2), 138–149 (1999)
4. Szeliski, R.: Computer vision: algorithms and applications. Springer (2010)
5. Zingman, I., Lambers, K., Saupe, D.: Detection of archaeological sites in high resolution remotely sensed imagery (Jun 2014), https://www.informatik.uni-konstanz.de/en/saupe/research/ongoing-projects/detection-of-archaeological-sites-in-high-resolution-remotely-sensed-imagery/
6. Zingman, I., Saupe, D., Lambers, K.: Automated search for livestock enclosures of rectangular shape in remotely sensed imagery. In: Proc. SPIE, Image and Signal Processing for Remote Sensing (2013)