Detection of incomplete rectangular contours with application in archaeology

Igor Zingman\textsuperscript{a,}, Dietmar Saupe\textsuperscript{a}, Karsten Lambers\textsuperscript{b}

\textsuperscript{a}Department of Computer and Information Science, University of Konstanz, Germany
\textsuperscript{b}Institute of Archaeology, Heritage Sciences and Art History, University of Bamberg, Germany

Abstract

We introduce a feature useful for detection of structures in images that are perceived as approximate rectangular contours. Given a set of extracted line segments, this rectangularity feature quantifies the degree of alignment of an optimal subset with a contour of rectangular shape, arbitrary size, and aspect ratio. The rectangularity feature has high values not only for perfect rectangles, but also for rectangles with distorted angles, fragmented, or with a completely missing side. However, it has zero response for configurations of lines that do not form at least three sides of a rectangular contour. This feature is based on a graph with nodes corresponding to line segments in a particular neighborhood and with edges encoding spatial relations between the lines. It is defined as the maximum of a target function of node and edge attributes of the graph cliques. The target function assigns high values to nearly convex line configurations with angles close to zero or right angles. 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 how to improve the detection performance using an additional feature that quantifies the size of a structure, and by learning a linear feature combination from a small number of representative examples of the livestock enclosures and a large number of available negatives.

1. Introduction

In this paper, we address the problem of detecting manmade rectangular patterns that can be roughly represented by rectangular contours. Since in practical applications it is important to be robust to deviations from a perfect rectangular shape, we allow rectangle distortions. Targeted rectangular contours are at times incomplete such that even an entire side is missing. Rectangle angles may deviate from right angles, and rectangle sides may be fragmented. The angle between adjacent fragments of the same (broken) side may deviate from 180 degrees. To be robust to such irregularities we have developed an algorithm with dedicated parameters that control the maximum tolerance to rectangle distortions. This algorithm relies on a new powerful descriptor that discriminates rectangular patterns from other structures in complex cluttered background. We call this descriptor the rectangularity feature.

It is known that local low-level features alone have certain limitations in the task of shape recognition. Spatial relations need to be quantified in order to find appropriate groupings of low-level primitives allowing us to capture higher level concepts that have a much stronger discrimination power. We have therefore developed a mid-level feature that robustly captures the concept of rectangularity for a set of line segments.

In our approach, a binary map of edges accompanied by angle information is computed first. Line segments are then found and modeled by a few parameters with the use of a variant of a local Hough transform. In the following stage an undirected graph is constructed, nodes of which correspond to line segments and graph edges encode spatial relations between line segments. Particularly, we use angle and convexity properties to encode appropriate spatial relations. Due to the construction of the graph, its maximal cliques correspond to appropriate configurations of line segments. These configurations are then ranked by a new rectangularity measure that encodes the goodness of grouping the segments into a rectangular structure. We call the highest rectangularity of an optimal configuration the rectangularity feature that is a function of an attributed graph or a corresponding set of line segments within an analysis window. To allow fast processing, an analysis window is placed at carefully chosen candidate points. This mid-level feature effectively quantifies the distinctive alignment of a group of line segments that originated from an approximately rectangular contour. Such alignment of line segments rarely happens randomly. 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 rectangular structure. Distorted rectangular structures result in low values of the rectangularity feature.

We recently presented in [1] the basic ideas underlying our approach for the detection of incomplete contours. Here we substantially extend this approach by designing a new more robust rectangularity feature, which is computed at candidate points extracted in accordance with [1]. Hard thresholds are avoided, such that the new feature encodes deviation of structure from a rectangular form. In contrast to our previous work, the feature does not rely on a heuristic partitioning of the set of lines into four subsets. We show the discriminative ability of the rectangularity feature for the problem of detecting livestock enclosures. These are ruins of manmade structures that sparsely

\textsuperscript{a}Corresponding author.

Email address: igor.zingman@uni-konstanz.de (Igor Zingman)
appear in alpine environments (example is shown in Fig. 1). The livestock enclosures usually resemble a rectangular contour with nearly linear walls that may be heavily ruined. These man-made structures offer important insights into historical development of alpine pasture economy and their automated detection was addressed in a recent archaeological project [2, 5, 8].

The rectangularity feature is based on a prior model of a fragmented rectangle, which is a convex polygon with constrained angles. We show that we can improve the detection based solely on the rectangularity feature by introducing an additional feature proportional to enclosure size and learning from the available data. Particularly, we design a linear classifier that learns from the large number of negative examples and adapts itself to a few available positive examples of livestock enclosures.

In the next section we review previous approaches for detection of rectangular structures and their applications. In Sec. 2 we introduce the new rectangularity feature. In Sec. 3 we show the use of this feature for detection of livestock enclosures in remotely sensed imagery. Using this data, we provide a comparison of the introduced rectangularity feature and a feature based on a gradient orientation density function recently proposed for building detection in [5]. In Sec. 4 we introduce an adjusted rectangularity feature in order to reduce the number of false detections. We show how this feature can be learned from a small number of positive examples and a large number of negatives.

2. Related work

Detection of rectangular structures was previously addressed in different contexts. Examples are detection of buildings in remotely sensed images [6, 7, 8, 9, 10, 11, 12, 13], traffic signs from ground level images [12, 15], and particles of a rectangular shape in cryo-electron microscopy images [16, 17]. Here we will not give an exhaustive review, but only mention a few representative examples, some of which are related to our approach. Most techniques for detection of rectangular structures dealt with buildings rooftops in remotely sensed images. In [7] the authors used geometric and projective constraints for generating building hypotheses from lines. The lines were extracted by means of an edge detector followed by linking and line approximation. The process of hypotheses generation involved time expensive reasoning. The hypotheses were further verified using 3D cues, which are not available in our case. Heuristic decisions during edge linking and line approximations with identification of the end points may reduce the robustness to clutter and to small changes in appearance of linear features. In our opinion this is a weakness of many algorithms developed for detection of building rooftops.

Markov Random Fields (MRF) were used in [10] to delineate buildings. Spatial dependencies were specified in a probabilistic framework using the MRF model. An energy function associated with the MRF was minimized resulting in an appropriate grouping of lines. A similar approach was used in [14] for detection of traffic signs from color images. The approach is sensitive to inaccuracy of extracted edge information and cannot detect incomplete rectangles, as it requires the presence of all four sides of a rectangular structure. Parameter tuning is difficult and an optimal solution of the MRF optimization problem is not guaranteed.

In [8], a search on a directed graph was used as follows. Line approximations of linked edges served as nodes of a directed graph while values of graph edges encoded line relations. A search for closed loops in the graph accompanied by an additional extensive set of rules and thresholds was used to generate building hypotheses. An extensive set of rules and hard thresholds present limitations on the robustness of the approach. A softer encoding of candidate line configurations is desirable. A subsequent hypothesis validation was based on 3D cues and existence of shadows. These, however, are very weak or absent for the rectangular enclosures.

The Hough transform is known as a robust technique for detection of lines (or other parameterized shapes) in cluttered environments. In [9] the Hough transform was proposed for the detection of two sets of perpendicular lines in the context of rooftop polygon extraction in urban areas. A general approach for detection of rectangular contours based on the Hough transform was developed in [18]. In this work, it was suggested to compute a local Hough transform using a local analysis window. A set of simple constraints on peaks in the Hough plane was designed to search for two orthogonal pairs of parallel lines. Unfortunately, this approach may result in detection of rectilinear configurations that cannot form a rectangular contour. Such configurations are prevented in our approach by adding a convexity constraint. The constraints used in [18] restrict opposite lines to be symmetric relative to the central point of a local window and to be of similar length. In many practical applications, however, including the detection of livestock enclosures, such an approach with very strict rules fails to detect fragmented or incomplete structures.

Some approaches avoid explicit detection of edges or lines, which potentially may avoid information loss during edge binarization. However, the time complexity of such methods is usually large. In [19] a template-based approach was proposed for detection of arbitrary, but a priori known shapes. A shape template was used to aggregate edge intensity information along the template border. This approach, however, requires a large number of templates when object sizes or shapes vary. This makes a template-based approach computationally not feasible. In [11] a set of local features was extracted from gray-tone images. These features, which carried local edge or corner information, were jointly used in order to produce a probability map of building rooftops. In many cases, including our case of spatially extended rectangular contours that appear in cluttered background with a variety of irrelevant structures, local features do not suffice. A more global description that takes into account spatial relations between local features is necessary. For example, in [5, 13] gradient orientation density function (GODF) was computed from image gradients and used for building detection. This function captures the distribution of orientation of gradient vectors in a neighborhood region. A correlation of this function with a mixture of two Gaussians having mean values separated by ninety degrees served as a GODF-based feature indicating the presence of rectilinear structures. Though effec-
tive for building detection, this feature loses its discrimination power when the target structure is a contour of a low contrast on cluttered background, which is the case we deal with in this paper. In Sec. 5 we compare the rectangularity feature introduced in this paper with the GODF-based feature applied to the task of detection of livestock enclosures [3, 11].

Some methods use analysis windows of a size much smaller than the size of the target object, meaning that the size of elementary patches is set such that only a group of patches covers an object or a region of interest. In this case, the fact that the group of elementary patches describes the object or a region of interest allows more reliable detection. For example, in [12] multiple detections of SIFT keypoints (with corresponding SIFT features) in urban areas and their relations allowed for reliable detection of buildings. In our case, objects of interest appear separately from similar instances and therefore detection of a single instance does not influence the likelihood of nearby detections. In [20], groups of elementary patches capture manmade structures located at medium to long distances from the camera. Each small patch was characterized by a set of features having distinctive values at rectilinear structures. A MRF was used in order to model spatial consistency of patch detections within a texture built up from manmade structures. This approach is not suitable for our target objects, which are spatially extended contours described by a sparse set of linear fragments rather than by texture built of small rectilinear elements. Moreover, a few positive examples do not allow usage of techniques that need extensive learning of statistical dependencies between neighboring elements.

The building detection methods described above are hardly applicable to our task because buildings are very distinct structures. In contrast, walls of livestock enclosures are very low (which corresponds to low contrast features in an image), may be highly fragmented, or even completely ruined. Irrelevant structures of higher contrast may appear inside or outside of rectangular structure in an immediate neighborhood. 3D cues are not available and shadows frequently do not exist or are very weak. None of the approaches mentioned above can cope with such a set of difficulties.

3. Measuring structure rectangularity

We introduce a rectangularity feature \( f_R \) computed from a set of linear segments \( W = \{S_i, i = 1, ..., m\} \) that were extracted from a gray-scale image. Linear segments that are approximately aligned in a rectangular structure result in a high value of the rectangularity feature. Computing the rectangularity feature at each point in the image is computationally expensive. Therefore, localization of sparse candidate points is necessary. In Sec. 6 we give an example of an approach we designed to extract ridges and valleys (bar edges) and localize candidate points for the task of detection of remains of livestock enclosures of rectangular shape in alpine regions, see Fig. 8.

3.1. Grouping edge points into line segments

Given a candidate location and edge points accompanied by estimated orientation we extract and parameterize linear segments, each of which is a group of aligned edge points. Linear segments are represented by a triple of parameters \((\theta, r, l)\) found by means of a local Hough transform centered at candidate points. We use the Hough transform in the form introduced in [21], where a line is defined by the orientation \(\theta\) of the normal and a distance \(r\) from the origin

\[
 r = x \cos \theta + y \sin \theta,
\]

where \(x, y\) are spatial coordinates of edge feature. A peak in the \((\theta, r)\) plane, also called Hough plane, corresponds to a linear segment \(S\), which can also be fragmented. The parameter \(l\) in the triple \((\theta, r, l)\) is the number of points that belong to the linear segment and is computed as the height of the corresponding peak in the Hough plane. The peaks are found as regional maxima in the Hough plane that was discretized with \(\Delta \theta = 3\) degrees and \(\Delta r = 1\) pixel. We use the parametrization \(\theta \in [0, 360)\), \(r \in (0, \infty)\) of a Hough plane that defines any line in a spatial plane. To better relate the parameter \(l\) to the length and avoid its dependence on the width of the extracted edges, we perform their thinning [22] prior to clustering in a Hough plane. The thinning reduces the width of all edges to one pixel.

Since we require edges to be extracted together with their orientations, \(r\) can be directly computed for each edge feature \((x, y)\) using Eq. (1). Thus, each edge feature votes for a single point in the \((\theta, r)\) plane instead of voting for a curve as suggested in [21]. This idea, which was already used in [23] for clustering of short ridge features and in [24] for clustering of point features, considerably eases extraction of meaningful peaks in the Hough plane. This clustering technique computed in a local window can actually be considered an extension of edge orientation histograms, which are at the core of most successful feature sets used for detection of certain object classes, [25, 26]. In contrast to edge orientation histograms, the \((\theta, r)\) plane based technique allows detection not only of dominant orientations of local features, but also their spatial alignment.

3.2. Valid configurations of line segments

Below we define a valid configuration of line segments \(C \subseteq W\) such that it can be a part of rectangular structure. We require all mutual angles \(\beta_{k,j}\) between line segments \(S_k, S_j \in C\) of the valid configuration to be close to either zero or right angles. An angle tolerance \(\alpha\) will be set to control the strictness of the angle constraint. We define mutual angles as

\[
 \beta_{k,j} = \min(\mid\theta_{S_k} - \theta_{S_j}\mid, 360 - \mid\theta_{S_k} - \theta_{S_j}\mid).
\]

Note that \(\beta_{j,k} = \beta_{k,j}\) and \(\beta \in [0, 180]\), since \(\theta \in [0, 360)\).

The angle constraint alone does not suffice to restrict configurations to be perceptually close to rectangles or parts of rect-angles. We define a second constraint that requires the valid

\[1\text{See section Sec. 6 for details on methods we used in our experiments to extract candidate locations and edges.} \]
configuration be approximately convex in the sense that extension of all linear segments of the configuration can form an approximately convex contour. The convexity tolerance $t$ will be defined to control the strictness of the convexity constraint. To ensure the convexity of a configuration of linear segments, it is sufficient to require that a half plane generated by each segment includes all other segments of the configuration. Additionally, we require that all these half planes contain the candidate point around which we search for a rectangular structure. Convexity constraints that are pairwise suffice to verify the convexity of a configuration containing the given candidate point. We define the pair-wise convexity measure $\tau$ for a pair of line segments $S_k, S_j$, each of which with corresponding attributes of size $l_S$, orientation $\theta_S$, and distance $r_S$ to the candidate point $p_0$, as
\[
\tau_{k,j} = \max(\tilde{\tau}_{k,j}, \tilde{\tau}_{j,k}),
\]
\[
\tilde{\tau}_{k,j} = \frac{1}{I_j} \sum_{p \in S_j} H((p-p_0)\cdot n_k - r_k),
\]
where $n_k = (\cos \theta_k, \sin \theta_k)^T$ is the unit normal of the segment $S_k$ and
\[
H(u) = \begin{cases} 1, & u > 0 \\ 0, & u \leq 0. \end{cases}
\]
$\tilde{\tau}_{k,j}$ measures the relative number of points in segment $S_j$ that are behind the segment $S_k$, relative to the given candidate point $p_0$ as illustrated in Fig. 1. Note that $\tilde{\tau}_{k,j} \neq \tilde{\tau}_{j,k}$, while $\tau_{k,j} = \tau_{j,k}$. Note also that $\tau \in [0, 1]$.

Definition 1. Let $\alpha \in [0, 45], \beta \in [0, 1], \gamma \in (0, 1]$, a candidate point $p_0$, and a configuration $C$ of linear segments be given. If for all pairs $S_k, S_j \in C, j \neq k$, one of the inequalities of the angle constraint
\[
\beta_{k,j} \leq \alpha \text{ or } |90 - \beta_{k,j}| \leq \alpha \text{ or } 180 - \beta_{k,j} \leq \alpha
\]
and the convexity constraint
\[
\tau_{k,j} \leq t
\]
both hold, then $C$ is a $(t, \alpha)$-valid configuration located around $p_0$, and denoted by $C_{p_0}^{t, \alpha}$.

For shortness, we sometimes omit the indices $t, \alpha$ and the reference point $p_0$, only mentioning that $C$ is a valid configuration. Valid configurations include not only perfect rectangles or their parts, but also convex polygons (or their parts) with angles around either 90 or 180 degrees. This is important in practical applications where approximately rectangular structures are better modeled by such polygons rather than by perfect rectangles.

3.3. Rectangularity of a valid configuration

A couple of poorly aligned short segments (as far as tolerances $t$ and $\alpha$ allow) or a rectangle are valid configurations $C_{p_0}^{t, \alpha}$. There is a need to rank valid configurations according to their similarity to a canonical rectangle. To find and rank valid configurations we construct an undirected graph from the given set $W$ of line segments in an analysis window centered at a candidate point $p_0$, and denote it as $G^w$. The graph $G^w$ has nodes $j = 1, \ldots, m$ corresponding to the segments $S_1, \ldots, S_m \in W$. Each node $j$ is attributed by a triple of parameters $(\theta_j, r_j, l_j)$, i.e. orientation, distance to the reference point $p_0$, and size of the line segment. An edge $[k, j]$ is attributed with the mutual angle $\beta_{k,j}$ and the pair-wise convexity $\tau_{k,j}$ of the corresponding pair of line segments $S_k, S_j$. An edge $[k, j]$ is included in the graph $G^w$ if $\beta_{k,j}$ and $\tau_{k,j}$ satisfy the constraints given in Eq. (5) and in Eq. (6). This attributed graph encodes properties of line segments and their spatial relationships. Due to the graph construction and Definition 1, valid configurations correspond to fully connected subgraphs, also called cliques, of the graph $G^w$. A clique of a graph $G^w$ that corresponds to a line configuration $C \subseteq W$ will be denoted by $G^C$.

Below we introduce the rectangularity measure $\rho(G^C)$ that ranks a clique $G^C$ corresponding to a particular valid configuration $C$. We define the measure with the following properties in mind. The rectangularity measure shall yield higher values for configurations with

1. higher degree of convexity given by lower values of the convexity measure $\tau$
2. higher degree of angle alignments given by mutual angles $\beta$
3. longer line segments given by larger $l$.

In addition, the proposed rectangularity measure shall

4. have the increasing property $\rho(G^C_1) \leq \rho(G^C_2)$ for $G^C_1 \subseteq G^C_2$. Thus, the rectangularity measure of a larger encompassing clique has a higher value
5. yield a zero value for configurations of line segments with less than three sides of a matched rectangle. Thus, a non-zero rectangularity indicates existence of a three or four-sided structure.

We define the rectangularity measure of a graph clique $G^C$ in terms of sums over its undirected edges $[k, j] \in E^C$
\[
\rho(G^C) = \left( \sum_{[k, j] \in E^C} l_k l_j f_{90}(\beta_{k,j}) f_{180}(\tau_{k,j}) \right)^{\frac{1}{2}},
\]
\[
\left( \sum_{[k, j] \in E^C} l_k l_j f_{90}(\beta_{k,j}) f_{180}(\tau_{k,j}) \right)^{\frac{1}{4}},
\]
\[
\left( \sum_{[k, j] \in E^C} l_k l_j f_{90}(\beta_{k,j}) f_{180}(\tau_{k,j}) \right)^{\frac{1}{4}},
\]
where \( f_{90}, f_{180}, \) and \( f_{cv} \) (see Fig. 2) are defined with the use of mode-functions \( m_{\mu,\delta} \) with peak location \( \mu \) and spread \( \delta \) as follows

\[
\begin{align*}
f_{90}(\beta) &= m_{90,\delta}(\beta), \quad \delta < 45, \\
f_{180}(\beta) &= m_{180,\delta}(\beta), \quad \delta < 45, \\
f_{cv}(\tau) &= m_{0,\delta}(\tau),
\end{align*}
\]

where

\[
m_{\mu,\delta}(u) = \begin{cases} 
\frac{1}{\sigma^2} (g_{\mu,\sigma}(u) - a), & \mu > 0 \\
0, & \text{otherwise}
\end{cases} ,
\]

and \( a \) is determined such that

\[
m_{\mu,\delta}(a) = 0, \quad \text{iff} \quad |\mu - \mu| > \delta.
\]

Note that only half of the second mode is present in \( f_{180} \), which is shown in Fig. 4. For \( f_{90} \) and \( f_{180} \) the peaks are at \( \mu = 90 \) and \( \mu = 180 \) degrees, respectively. The spread \( \delta = \alpha \) for both \( f_{90} \) and \( f_{180} \), where \( \alpha \) is the angle tolerance used in constraint Eq. (3). For \( f_{cv} \), the parameters are \( \mu = 0 \) and \( \delta = \tau \), where \( \tau \) is the convexity tolerance used in constraint Eq. (3). The parameter \( \sigma \) controls the peakiness of the function shape and is less influential in practice. In our experiments described in the following sections, we set the angle tolerance \( \alpha = 35 \) degrees, convexity tolerance \( \tau = 0.3 \), and \( \sigma = \frac{\alpha}{2} \) for \( f_{180} \) and \( f_{90} \) functions and \( \sigma = \frac{\tau}{2} \) for \( f_{cv} \).

![Image showing functions](image)

Figure 2: Functions \( f_{90} \) (left figure, solid blue curve), \( f_{180} \) (left figure, dashed red curve), and \( f_{cv} \) (right figure) used in definition of the rectangularity measure in Eq. (4).

The first factor of \( \rho(G^w) \) in Eq. (10) yields a non-zero value only if the valid configuration \( C \) contains at least one pair of approximately perpendicular line segments. The second factor is non-zero only if the valid configuration \( C \) contains at least one pair of approximately parallel line segments. The product of these two factors is non-zero only if the valid configuration \( C \) contains at least one pair of parallel and at least one pair of perpendicular line segments. The angles between line segments of these parallel and perpendicular pairs are restricted to be approximately 0, 180, or 90 degrees since \( G^w \) is a clique corresponding to a valid configuration with line segments constrained by Eq. (3). Thus, non-zero rectangularity measure insures a valid configuration \( C \) containing at least one triple of linear segments arranged in a \( II \)-like structure, as was stated in property 5 above. This property allows suppression of a large number of line configurations originating from irrelevant structures or clutter. It is easy to verify that the other four properties are also satisfied by the rectangularity measure in Eq. (4).

It follows directly from the definition in Eq. (4) that scaling a configuration of line segments of a corresponding graph clique scales its rectangularity measure by the same factor. Therefore, the rectangularity measure scales linearly with the spatial size of rectangles. In fact, it can be shown that for a rectangular structure with perfectly aligned linear segments (the structure can be fragmented), i.e. for the case of all angles \( \beta \) being equal to either 90 or 180 degrees and all \( \tau \) being equal to 0, the rectangularity measure reduces to \( ((L_1 + L_3)(L_2 + L_4)(L_1 + L_3 + L_2 L_4))^2 \), where \( L_i, i = 1, ..., 4 \) are the sums of segment sizes of four sides of a rectangle, such that index pairs 1, 3 and 2, 4 correspond to parallel sides. If more sides than one are missing, the expression above equals zero. Note that the rectangularity measure is a function of graph nodes and edge attributes, it does not require explicit partitioning of a valid configuration of line segments into four subsets corresponding to four sides of hypothesized rectangle as was required in our earlier work [1].

3.4. Rectangularity feature

Given a set of line segments \( W \) in an analysis window and the corresponding graph \( G^w \), we define the rectangularity feature \( f_\rho(G^w) \) of the graph \( G^w \) based on the rectangularity measure of its cliques \( G^c \). Let us denote the set of cliques as \( \mathcal{K}(G^w) \). The rectangularity feature of \( G^w \) is defined as the maximal rectangularity measure of the cliques in \( \mathcal{K}(G^w) \)

\[
f_\rho(G^w) = \max_{G^c \in \mathcal{K}(G^w)} \rho(G^c). \tag{10}
\]

The corresponding optimal clique is

\[
G^c_{opt} = \arg \max_{G^c \in \mathcal{K}(G^w)} \rho(G^c). \tag{11}
\]

Due to the increasing property of \( \rho \) (the third property of the rectangularity measure stated above), the maximum can be searched over the set of maximal cliques only, denoted here by \( \mathcal{M}(G^w) \)

\[
f_\rho(G^w) = \rho(G^c_{opt}) = \max_{G^c \in \mathcal{M}(G^w)} \rho(G^c). \tag{12}
\]

Since the set of maximal cliques \( \mathcal{M}(G^w) \subseteq \mathcal{K}(G^w) \) is much smaller than the set of graph cliques \( \mathcal{K}(G^w) \), the number of times the rectangularity measure \( \rho \) needs to be evaluated in Eq. (11) is considerably reduced in comparison to Eq. (3). Since, in addition, there are efficient algorithms for the search of maximal cliques [17], computing the rectangularity feature is not computationally demanding.

Fig. 3 (left) shows an example of a given set \( W \) of linear segments and the optimal configuration of line segments \( C_{opt} = \{S_1, S_2, S_3, S_4\} \) in red, while Fig. 3 (right) shows the corresponding graph \( G^w \) and the optimal maximal clique.

---

2Maximal cliques are cliques that are not contained in larger cliques.
There are two additional maximal cliques $G_1^c$ and $G_3^c$ and corresponding valid configurations $C_1 = \{S_2, S_5, S_6, S_3\}$, $C_2 = \{S_1, S_5, S_3, S_4\}$ for for this particular case of the graph. They, however, have lower rectangularity values $\rho(G_1^c) < \rho(G_2^c), \rho(G_1^c) < \rho(G_3^c)$.

Figure 3: Left: A given set $W = \{S_1, S_2, \ldots, S_6\}$ of line segments around a candidate point $p_0$. Right: A graph $G^w$ for the set of linear segments. We assume an angle tolerance $\alpha$ such that all mutual angle constraints are satisfied. Several node pairs of the graph are not connected by an edge due to the convexity constraint, which is not satisfied for some assumed convexity tolerance $t$. The red nodes of the graph are the nodes of the optimal maximal clique $G_{\text{opt}}^c$. The corresponding valid configuration $C_{\text{opt}}$ of line segments is marked in red on the left figure.

3.5. Normalized rectangularity feature

The rectangularity feature in Eq. (12) scales with the size of the structure to be detected. Since the rectangularity feature has low values for small structures, a detector based on such a feature is prone to dismiss small rectangles. Therefore, we normalize the rectangularity feature by an estimated size of the structure, represented by the optimal clique $G_{\text{opt}}^c$, which is a subgraph of $G^w$, as follows

$$f_S(G^w) = \frac{\sum_j l_j f_j}{\sum_j l_j},$$

where the sums are over all nodes of the optimal clique $G_{\text{opt}}^c$. $f_S$ is the weighted distance of the line segments of $C_{\text{opt}}$ from the corresponding candidate point, where the weights are line segment sizes. We call it the size feature. The normalized rectangularity feature is defined as

$$f_R(G^w) = \frac{f_R(G^w)}{f_S(G^w)}.$$  

The normalized rectangularity feature is not sensitive to the size of the structure to be detected.

4. A comparative study of the rectangularity feature applied to detection of livestock enclosures

4.1. Problem description

Automated visual analysis has substantially advanced in recent years, allowing a variety of targets to be automatically detected. Remarkably successful algorithms and technologies have been developed for face detection and for object detection for autonomous car navigation [28]. In archaeology, on the other hand, the interpretation of remotely sensed images is still the domain of archaeologists visually inspecting the images. Aerial or satellite images may be visually inspected prior to field survey with the goal to identify potential sites to guide fieldwork. Such an approach considerably saves fieldwork time. A large amount of high-resolution image data has become available during recent years due to developments of satellite technology. However, a visual inspection of the image data over vast unexplored areas is not feasible. Although, human vision is clearly superior to the current state of computer vision, in contrast to a human expert, machine vision algorithms are capable of screening a large amount of imagery, generating candidate locations, which can be verified easily and timely by the archaeologist. Such a semi-automated approach can guide archeological survey and increase its efficiency in vast unexplored areas [13]. It should be noted that the detection of archaeological structures in vast unknown areas differs essentially from contrast enhancement and automated mapping of known sites and structures as undertaken in many archaeological projects (e.g. [22]).

In this section, we show the advantages of the rectangularity feature applied to the detection of remains of livestock enclosures. These manmade structures sparsely appear in grassland of mountainous regions. They are of a special archaeological interest because they offer important insights into historical development of alpine pasture economy. Their automated spotting was addressed in a recent archaeological project [2, 3]. Examples of the well-preserved enclosures are shown in Fig. 4. These structures are usually composed of linear walls that may be heavily ruined. Though the most common shape of livestock enclosures resembles a rectangular contour, its size and aspect ratio may greatly vary. We use satellite and aerial images of 0.5m resolution where the width of linear walls does not exceed two pixels. The walls are usually of a low height, which results in low contrast linear features in images. The first row of Fig. 4 shows a satellite (on the left) and an aerial (on the right) image with structures corresponding the livestock enclosures shown in Fig. 4. Nearby non-relevant structures, such as rivers, trails, rocks are often of higher contrast due to larger size or distinctive spectral properties. Detection of such faint enclosure structures in a complex terrain, which is characterized by various irrelevant structures and textures having much higher contrasts, is a very challenging task. Even the detection of easily modeled circular soil structures [31] had very limited success due to low contrast and complex terrain. Because of these difficulties described above, commonly used methods for rectangle detection are hardly applicable.

Figure 4: Livestock enclosures in alpine environment
The spectral properties of livestock structures are very similar to the spectral properties of the surrounding terrain, rocks, and other irrelevant objects. Though the internal area of enclosures sometimes exhibit distinctive spectral signatures, they are not consistent from site to site and depend on time of the year the image was captured and on the type of imaging modality. On the other hand, geometrical properties appear to be more distinctive and do not depend on such factors. We therefore investigate the use of the rectangularity feature that can capture the distinctive geometry of enclosures.

4.2. The data used

In our experiments we used panchromatic images captured by the GeoEye1 satellite and the red channel of Swiss Topo aerial images. Both types of images are at 0.5m resolution. We used images from different sources in order to emphasize that the rectangularity feature does not depend on illumination conditions and contrast of enclosures (provided that contrast is high enough allowing detection of edges), but only on geometry and completeness of enclosures.

Unfortunately, only nine examples of well-preserved enclosures taken from aerial and satellite images were available for us in this study. On the other hand, a large number of negative examples (examples of rectangularity features at points with no real livestock enclosure) 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 [2, 3, 4]. In our experiments we used an image of 19000×10000 pixel size, which corresponds to an area of about 48km², in order to generate negative examples and estimate the type and the rate of false detections. The false positives rate (FPR) should not be too large, otherwise the detector applied to large areas becomes useless in practice. The examples generated from this large image are actually unlabeled and may contain a few unknown enclosures, but since these structures are very rare, they practically do not influence the estimation of the FPR and other performance measures.

4.3. Preprocessing

We used the same approaches as in [11] to generate a binary map of bar edges (both ridge and valley line features, which are required for computation of the rectangularity feature as described in Sec. 3.1) and initial candidate points. Bar edges were extracted using the Morphological Feature Contrast based line detector from [31, 32]. This technique is capable of extracting line features, while suppressing texture elements of cluttered background. Extraction of candidate points was carried out by sampling the skeleton points of a complementary binary map of detected bar edges [11]. The second row of Fig. 3 shows two examples of maps of bar edges (black) and candidate points (red) for the two corresponding images shown in the first row. Along with the sampled skeleton, we computed the distance transform of the set of detected bar edges. The values of the distance transform at candidate points were used to adaptively define the sizes of an analysis window for computation of the rectangularity feature at the corresponding points. Note that the skeleton

Figure 5: First row: 600 × 600 satellite (©GeoEye 2011) and aerial (SWISSTOPO) images of 0.5m resolution with structures corresponding to livestock enclosures shown in Fig. 4. Second row: Bar edges (black) and candidate points (red) generated from the images shown in the first row as suggested in [11]. Third row: The rectangularity feature that was raised to the 4th power to better visualize the range of obtained values. The feature was computed at each candidate point shown in the second row and visualized by a transparent colored disk. Color saturation increases and hue is changing from yellow to red for growing values of the feature in accordance with the color bar shown in the bottom of the figure. Fourth row: The GODF-based feature raised to the 8th power. The 4th power used for visualization of the feature did not suffice to visually distinguish values of \( f_g \). Note that though we have used the same color bar for several examples, the feature values corresponding to the same color are not identical for different figures.
could also be directly computed from the distance transform by means of localization of its ridges [53, 54]. However, we have found that such an approach may miss some skeleton points because these points are not always the local maxima of the distance transform and are not always the local maxima in the direction of the gradient of the distance transform. As in [41] we discarded all candidate points having a distance smaller than 10 or greater than 90 pixels, which bounds the distances between parallel walls of the structures.

Rectangularity features may result in a large number of false detections in textured regions such as urban areas or forests. Since we are only interested in livestock enclosures that sparsely appear in grassland areas, we filtered out high contrast texture regions using the morphological texture contrast descriptor [55, 51, 52] thresholded at the level obtained with the Otsu method [55]. This effectively filters out urban areas, forests, rocky mountains, and other high contrast texture regions, but preserves isolated or individual structures.

### 4.4. The gradient orientation density function based feature

We compared the performance of the rectangularity feature with the performance of the feature recently used in [8, 73] for detection of buildings. It is based on the estimate \( \lambda(\theta) \) of the gradient orientation density function (GODF) that captures the distribution of orientations of image intensity gradients. The correlation of this function with a mixture of two Gaussians having mean values separated by 90 degrees served as a GODF-based feature \( f_g \) indicating the presence of rectilinear structures.

Let \( A \) be the neighborhood around a candidate point and let us denote by \( g(p) \) the intensity gradient at point \( p \in A \) (the Prewitt operator was used). \( \lambda(\theta) \) is computed in the neighborhood \( A \) as a weighted gradient orientation histogram with gradient magnitudes \( ||g(p)|| \) as weights, and discrete orientation \( \theta \in [0, 180) \), \( \theta = k\Delta \theta \), where \( k = 0, 1, 2, \ldots \),

\[
\lambda(\theta) = \frac{1}{B} \sum_{p \in A} ||g(p)|| I(\theta, \varphi = \theta - \varphi(g(p)), \theta - \varphi(g(p)) + \Delta \theta),
\]

(15)

\( B \) is a normalizing constant such that \( \lambda(\theta) \) is a unit vector, and \( I \) is the indicator function

\[
I(\theta, \varphi) = \begin{cases} 
1, & \varphi \in [\theta, \theta + \Delta \theta) \\
0, & \text{otherwise}.
\end{cases}
\]

The GODF-based feature \( f_g \) at the candidate point is then defined as a circular correlation of the orientation histogram \( \lambda(\theta) \) with the function \( f_{\lambda_0} \) with modes separated by 90 degrees

\[
f_g = \max_{\theta \in [0, 180)} \sum_{\varphi = 0}^{\Delta \theta} \lambda(\theta) f_{\lambda_0}(\theta - \varphi) \text{ modulo } 180),
\]

(16)

The function \( f_{\lambda_0} \) is defined in the interval \([0, 180)\), shown in Fig. 6, and computed as a sum of the three mode functions \( m_{0, \delta}, m_{90, \delta}, \) and \( m_{180, \delta} \) given in Eq. (6).

In \( f_{\lambda_0} \) we used \( \delta = 35 \) degrees, and \( \sigma = \frac{90}{\delta} \) (see Eq. (22)) as for the rectangularity feature. Discrimination step \( \Delta \theta \) was set to one. Note that the correlation above can be computed with the circular convolution [67].

![Figure 6](image)

The constant \( B \) in Eq. (15) was set such that \( ||A||^2 = \sum_{\theta} \lambda^2(\theta) = 1 \). Such normalization insures maximal values of the GODF-based feature \( f_g \) at candidate points with \( \lambda \) optimally matching \( f_{\lambda_0} \) (up to the circular shift), i.e. with small angles between these vectors. We obtained considerably better results for such a choice of \( B \) than for \( B = \sum_{p \in A} ||g(p)|| \) that was used in [55].

### 4.5. Measuring discrimination power

To detect livestock enclosures, an appropriate threshold on the value of the rectangularity features should be set. The appropriate value of the threshold can be chosen based on the available positive and negative examples. Setting a particular threshold defines the true positive rate (TPR) or sensitivity of the detector and the false positives rate (FPR). The effectiveness of rectangularity features is in our case their ability to discriminate livestock enclosures from irrelevant structures and clutter. A possible measure of this ability is the minimal achievable rate of false detections generated with the threshold that insures \( TPR \geq \xi \), where \( \xi \) is the predefined rate of true positives. This corresponds to the so-called Neyman-Pearson task [88]. In our experiment we computed \( FPR \) for \( \xi = 1 \), denoted in the following by \( FPR_{100} \). This was done by setting the detection threshold to the minimum value of the rectangularity features for nine available positives and counted the relative number of false detections \( F_{100} \). The false detections were obtained in about 48km² area covered by our 19000 x 10000 satellite image. Obviously, the threshold used to obtain the detection rate \( TPR = 100\% \) on a very small number of available examples may not be optimal and it does not insure a detector with the true detection rate equal 100%. This threshold is only used in order to generate a measure of discrimination ability of compared features.

An alternative measure of the discrimination ability is an area under receiver operating characteristic (ROC) curve, which is especially useful in the presence of unbalanced classes [89, 90]. In contrast to the \( FPR_{100} \), the area under receiver operating characteristic (AUC) does not rely on a particular threshold and a corresponding operating point on the ROC curve, but instead summarizes the detection performance for different values of the threshold. In fact, it is an average of true positive rates estimated for all possible false positive rates. The AUC has an important statistical property. It equals the probability that a ran-
domly chosen sample \( y \) from the population of positives \( P \) has a higher score \( f(y) \) (the rectangularity feature in our case) than the score \( f(x) \) for randomly chosen sample \( x \) from the population of negatives \( N \), i.e. \( \text{AUC}(f) = P(f(y) \in P) > f(x \in N) \). This probability of a correct ranking can be estimated by means of the Wilcoxon-Mann-Whitney statistic \([21, 41]\) as

\[
\frac{1}{npnN} \sum_{i=1}^{np} \sum_{j=1}^{nN} I(f(y_i) \in P, f(x_j) \in N),
\]

where \( np \) and \( nN \) denote the number positive and negative samples, respectively, and \( I(u, v) \) is the indicator function defined as

\[
I(u, v) = \begin{cases} 
1, & u > v \\
0, & u < v \\
0.5, & u = v.
\end{cases}
\]

It should be noted that for the case of normally distributed features \( f(x), f(y) \), the AUC has a very simple relation to the Fisher criterion \([22, 23]\), which is also frequently used as a separability measure between distributions. Namely, \( \text{AUC} = \Phi \left( \frac{\mu_p - \mu_n}{\sqrt{\sigma_p^2 + \sigma_n^2}} \right) \), where \( \Phi \) is the normal cumulative distribution function, evaluated for the Fisher criterion with distribution means \( \mu_p, \mu_n \) and standard deviations \( \sigma_p, \sigma_n \) for positive and negative populations, respectively.

### 4.6. Comparative experiments

The rectangularity feature \( f_R^k \) computed at candidate points is visualized in the third row of Fig. 8. The rectangularity feature was raised by power four in order to visually better distinguish its low and high values. The value of the rectangularity feature at each candidate point is visualized by a transparent colored disk. Color saturation increases and hue is changing from yellow to red for growing values of the feature in accordance with the color bar shown in the bottom of the figure. As expected, high values were obtained at positions of livestock enclosures while zero or low values were obtained at most other candidate positions. Less convincing results are obtained for the GODF-based feature shown in Fig. 8 (fourth row), where \( f_g \) is computed in the same regions. The GODF-based feature was raised by power eight in order to visually better distinguish its low and high values. The 4\(^{th}\) power used for visualization of the feature \( f_R \) did not suffice to visually distinguish values of \( f_g \).

The discrimination measures (Sec. 43) confirm the considerable difference in the performance of the features \( f_R \) and \( f_g \). Using our data (Sec. 42), for the rectangularity feature we obtained 170 false positives, \( FP_{100}(f_R) = 170 \). These false structures were detected out of 403716 candidate positions in approximately 48\( \text{km}^2 \) in the 19000 \times 10000 pixel image. The corresponding false positives rate \( FPR_{100}(f_R) \) and the area under the curve measure \( \text{AUC}(f_R) \) are given in Table 1 in the first column. The normalized rectangularity feature \( \hat{f}_R = f_R/f_s \) yields inferior results: 279 false structures (\( FP_{100} \)) were detected in the same 403716 candidate positions. The corresponding false positives rate \( FPR_{100}(\hat{f}_R) \) and the area under the curve measure \( \text{AUC}(\hat{f}_R) \) are given in Table 1 in the second column. The

<table>
<thead>
<tr>
<th>Feature</th>
<th>( FP_{100} )</th>
<th>( FPR_{100} )</th>
<th>( \text{AUC} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f_R )</td>
<td>170</td>
<td>( 4.2 \times 10^{-4} )</td>
<td>0.99993</td>
</tr>
<tr>
<td>( \hat{f}_R )</td>
<td>279</td>
<td>( 6.9 \times 10^{-4} )</td>
<td>0.99973</td>
</tr>
<tr>
<td>( f_g )</td>
<td>6522</td>
<td>( 162 \times 10^{-4} )</td>
<td>0.99683</td>
</tr>
<tr>
<td>( \hat{f}_g )</td>
<td>120</td>
<td>( 3.0 \times 10^{-4} )</td>
<td>0.99995</td>
</tr>
</tbody>
</table>

Table 1: Discrimination measures for the rectangularity feature \( f_R \), the normalized rectangularity \( \hat{f}_R \), the GODF-based feature \( f_g \), and the adjusted rectangularity feature \( \hat{f}_g \) (Sec. 44).

The possible reason of inferior performance of the normalized rectangularity feature, which is insensitive to the size of the structures estimated with \( f_g \), is that random configurations of larger size are more rare. In section Sec. 4.7 we use both \( f_R \) and \( \hat{f}_R \) features to build a detector with improved performance. For the GODF-based feature we obtained 6522 false detections (\( FP_{100} \)) out of 403716 candidates. The corresponding false positives rate \( FPR_{100}(f_g) \) and the area under the curve measure \( \text{AUC}(f_g) \) are summarized in Table 1 in the third column. The experiments were carried out using Matlab on a machine with an Intel Core 2 2.83 GHz processor. The generation of all 403716 candidate locations and computation of the rectangularity features \( f_R \) or \( \hat{f}_R \) took about three hours.

In summary, we have shown that the rectangularity feature is far more discriminative in detection of the enclosure. Though effective for building detection, the GODF-based feature loses its discrimination power when the target structure is a contour of low contrast on cluttered background. In general, this feature is effective only when computed within relatively small analysis windows. For large windows its effectiveness drops since there is only a small number of points that belongs to the structure (a contour) of interest relative to the large number of points that belong to irrelevant structures and background clutter.

### 4.7. Learning from the data

As we have mentioned in Sec. 4.3, the rectangularity measure scales linearly with the size (perimeter) of the enclosures and so does the rectangularity feature \( f_R \). In the previous section we have seen that the normalized rectangularity feature \( \hat{f}_R = f_R/f_s \) is not as effective as \( f_R \). Though we normalized the rectangularity feature in order to be independent of the structure size, it is actually biased towards small enclosures because small enclosures are more likely to have a higher ratio of length of preserved walls to length of the structure perimeter. On the other hand, small false structures due to random configurations of irrelevant fragments or clutter are more frequent. This may explain the higher rates of false positives for the normalized rectangularity feature.

The dependency of the rectangularity feature \( f_R \) of false structures on their size is visualized in the two dimensional histogram in Fig. 8 (left). This figure shows the frequency of

---

3 The experimental code along with examples of the livestock enclosures shown in the first row of Fig. 8 is available online under the following link: [https://www.informatik.uni-konstanz.de/en/sauge/research/on-going-projectedetection-of-archaeological-sites-in-high-resolution-remotely-sensed-imagery](https://www.informatik.uni-konstanz.de/en/sauge/research/on-going-projectedetection-of-archaeological-sites-in-high-resolution-remotely-sensed-imagery).
structures (candidate points extracted as mentioned in Sec. 5) as a function of the rectangularity \( f_R \neq 0 \) and the size \( f_S \). The structures are limited in their size by setting two thresholds on the minimal and maximal values of the distance transform, see Sec. 4.3. Real livestock enclosures are very rare in the field and though a few points of the distribution may correspond to unknown livestock enclosures, the vast majority of the candidates correspond to false structures. Relying on this fact, will refer to this distribution as distribution of negatives and denote it by \( X \). The curved shape of the upper margin of the distribution \( X \) indeed shows that the frequency of candidate structures with similar ratio \( f_R/f_S \) decreases with structure size.

Thresholding the feature \( f_R \) corresponds to classifiers with decision boundaries in the feature space \((f_R, f_S)\) that are horizontal lines, while thresholding the normalized feature \( f_R^* = f_R/f_S \) corresponds to straight lines passing through the origin of feature space with slope given by the threshold. Instead of normalizing the rectangularity feature \( f_R \) by \( f_S \), we can use both features independently in order to learn a decision boundary from the available data in the two-dimensional feature space. This may improve the trade-off between the sensitivity and the number of false detections in comparison to one-dimensional case. However, only a few positive examples are available in our case, see Fig. 4 on the right. Thus, a classification approach should be carefully chosen.

Figure 7: Rectangularity-size feature space. Left: Distribution of candidates (with non-zero rectangularity feature) extracted from the satellite image of 19000 × 10000 pixel size as described in Sec. 4.3. Right: Scatter plot of the candidates (cyan) corresponding to the distribution on the left and examples of livestock enclosures (red).

### 4.7.1. Detection of rare events

The Neyman-Pearson approach, commonly used for detection tasks, is a non-Bayesian decision making that is especially useful when priors are not available or misclassification risks are not comparable \(^4\). The Neyman-Pearson classification method maximizes the sensitivity of the classifier given an upper bound for the rate of false detections \(^4\). This strategy is directly applicable to our problem of enclosure detection. It can be interpreted as setting the maximal number of false detections that can be visually verified by an expert, while maximizing the sensitivity of the detection. As for the case of Bayesian classifiers, the solution is based on the ratio of class-conditional
distributions. Unfortunately, we have a very small number of positive examples, which makes a reliable estimation of the distribution of the target class (positives) impossible.

One-class classifiers are usually employed in situations with available samples from a single class only \[^{[43]}\]. Samples from the other class are either not available, difficult to obtain, or very rare. The instances from the second class, which is poorly or not at all represented, are called novelties, outliers, or anomalies. Several approaches were developed in order to approach the one-class classification task. The distribution of the well-described dominant class can represented by a model of choice. The samples that are very distant from the modeled distribution in accordance with the chosen metric are then assigned to the second class of novelties. Alternatively, the reconstruction error of representing the sample by the chosen model can be used as the measure of novelty \[^{[44]}\]. In \[^{[45]}\] a support vector data description was developed, where a decision boundary separating the dominant class from novelties is a hypersphere of minimum volume containing samples of the dominant class. An important advantage of this method is its ability to incorporate examples from the class of novelties (if a small number of such examples is available) while learning the decision boundary.

One-class classifiers usually tend to produce a decision boundary that compactly encloses the samples of the well-represented dominant class \( X \). All the other samples are assigned to the second class of novelties. In fact, such classifiers imply a uniform distribution for the class of novelties \[^{[46]}\]. In our case and in many other applications, where novelties are positives of the target class that describes a particular category of objects to be detected, the distribution of novelties is far from being uniform. For example, due to construction of the rectangularity feature \( f_R \), the samples with lower value of \( f_R \) and for the same value of the \( f_S \), have lower probability of being positive examples. The one-class classifiers may yield erroneous results assigning samples with very low rectangularity values outside of the enclosed distribution \( X \) to the class of enclosures (novelties). A possible solution is to construct a decision boundary that cannot fold up. The simplest choice is a hyperplane, which is beneficial for our case with very limited number of positives, preventing the classifier from overfitting.

The optimal direction \( w \) of the separating hyperplane (i.e. the direction of the normal to the hyperplane) can be found by means of the Fisher linear discriminant analysis (FLD). In this approach, the optimal direction is determined such that the data from two classes projected on \( w \) is maximally separated. The separation is measured by the squared distance between class means normalized by the sum of their variances \[^{[42]}\]. This approach results in a simple solution represented in terms of class means and covariance matrices. In our case, however, the number of positive examples is very limited and the covariance matrix cannot reliably be estimated.

#### 4.7.2. Proposed linear classifier

We propose a direction \( w \) of the separating hyperplane based on the large number of available examples from one class and just a few examples from the other. Let us define the expected
signed distance between an arbitrary point \( y \) and the distribution \( X \), both projected to the direction \( w \) and normalized by the standard deviation of the projected distribution by

\[
D_w(y, X) = \frac{E_x (w^T x - w^T y)}{\sqrt{E_x (w^T x - w^T \mu_x)^2}}.
\]  

(18)

where \( \mu_x \) denotes the mean of the distribution \( X \). It can also be expressed as

\[
D_w(y, X) = \frac{w^T (y - \mu_x)}{\sqrt{w^T C_x w}},
\]

(19)

where \( C_x \) is the covariance matrix of the distribution \( X \), which is a positive semidefinite matrix. Next, we define the average signed distance between set of points \( \{y_i, i = 1, \ldots, n\} \) and the distribution \( X \)

\[
D_w(\{y_i\}, X) \equiv \frac{1}{n} \sum_{i=1}^{n} D_w(y_i, X).
\]

(20)

It follows that due to linearity of \( D_w(y, X) \) with respect to \( y \) we have

\[
\bar{D}_w(\{y_i\}, X) = \frac{w^T (\bar{y} - \mu_x)}{\sqrt{w^T C_x w}}.
\]

(21)

where \( \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i \). We now define the optimal direction of \( w \) as the direction that maximizes the absolute value of the average signed distance between a small set of points corresponding to positive examples and the distribution of the dominant class of negatives \( X \), i.e.

\[
w_{opt} \equiv \arg \max_w |\bar{D}_w(\{y_i\}, X)|.
\]

(22)

From Eqs. (21, 22) we obtain

\[
w_{opt} = \arg \max_w \frac{|w^T (\bar{y} - \mu_x)|}{\sqrt{w^T C_x w}}.
\]

(23)

The objective function in Eq. (23) is invariant with respect to scale and sign of the vector \( w \). We, therefore, can add constraints \( w^T C_x w = 1 \) and \( w^T (\bar{y} - \mu_x) \geq 0 \) that uniquely define the norm and the orientation of \( w \), and solve the constrained optimization problem

\[
\bar{w}_{opt} = \arg \max_w \frac{w^T (\bar{y} - \mu_x)}{w^T C_x w = 1 \quad w^T (\bar{y} - \mu_x) \geq 0}.
\]

(24)

Using Lagrange multipliers it can be shown that maximum is achieved for

\[
\bar{w}_{opt} = \frac{C_x^{-1}(\bar{y} - \mu_x)}{\sqrt{(\bar{y} - \mu_x)^T C_x^{-1}(\bar{y} - \mu_x)}}.
\]

(25)

Since, the length of the optimal \( w \) is not important we will further use the scaled simpler version of \( \bar{w}_{opt} \)

\[
w_{opt} = C_x^{-1}(\bar{y} - \mu_x)
\]

(26)

that also solves Eq. (23).

The direction of \( w_{opt} \) is similar to the one obtained in the Fisher linear discriminant analysis [21]. In contrast to the FLD solution, Eq. (14) includes the covariance matrix of the class of negatives only, preferring the solution in the direction of the small variance of negatives. Negatives are well sampled in our problem and their covariance matrix can be robustly estimated. The positives (novelties) are treated as deterministic points in the feature space and influence the solution only via their average. Literally, the average only weekly guides the solution showing were the novelties of our interest may reside. In case of unreliable examples of positives, one may consider using the median or a weighted average. Note also that the signed distance in Eq. (21) for \( w_{opt} \) given in Eq. (20) yields a positive value equal to the Mahalanobis distance \( D_{opt}(\{y_i\}, X) = \sqrt{(\bar{y} - \mu_x)^T C_x^{-1}(\bar{y} - \mu_x)} \) with the metric \( C_x \).

Our approach conceptually relates to the work in [H7], where a few examples of a poorly represented class (called novelties here) weakly guided a non-linear SVM classifier. For some choices of the kernels the decision boundary, which corresponds to the hyperplane in a high dimensional feature space associated with the kernel, may enclose the samples of the well-represented class in the original feature space. As we have discussed in Sec. [H7], this may cause erroneous detection results in cases with category-specific features and with target novelties belonging to a particular category. We carefully constructed such category-specific features, i.e. the rectangularity and the size features. As was noted in [H8], in contrast to many generic features, a small number of category-specific features allows use of simple linear classifiers. These classifiers are especially favorable when there is a danger of overfitting the data due to limited number of available examples. Additionally, in comparison to our simple FLD-like solution, support vector based classifiers are more demanding for memory and computational effort.

The samples of \( X \) may include outliers. Though the estimate of the mean given a small percentage of outliers might still be robust, the estimate of the covariance matrix may be sensitive to the outliers. Therefore, we use the robust Multivariate Trimming (MVT) estimates of the mean and the covariance matrix [H9]. The MVT is an iterative technique with mean and covariance matrices recomputed at each iteration. Given the current estimates of \( \mu_x \) and \( C_x \) the Mahalanobis distance is computed for all the data points. A specified percentage of the observations with the largest Mahalanobis distance is discarded and the remaining data is used to recompute the estimates of \( \mu_x \) and \( C_x \). The technique is initialized with the sample mean and covariance matrix computed from the whole data. In our experiments we used three iterations and discarded 10% of observations in each iteration.

Given the optimal hyperplane defined by \( w_{opt} \), samples with coordinates \( (f_S, f_R) \) in feature space can be characterized by the adjusted (scalar) rectangularity feature

\[
f_R = (f_S, f_R) w_{opt}.
\]

(27)

Thereby, the adjusted rectangularity feature is an optimal linear combination of the rectangularity and size features. Note that
our approach is not limited to two dimensional feature spaces, but also applies to higher dimensions.

4.7.3. Results

Using the same data as in the experiments in Sec. 4.6, we evaluated the discrimination ability of the adjusted rectangularity feature $\hat{f}_R$, which gave us 120 falsely detected structures ($FP_{100}$) out of 403716 candidate positions. The corresponding false positives rate $FPR_{100}(\hat{f}_R)$ and the area under the curve measure $AUC(\hat{f}_R)$ are summarized in Table 1 in the fourth column. Comparing with the performance of $f_R$ given in the table and in Sec. 4.6, the experimental results confirm that the adjusted rectangularity feature $\hat{f}_R$ has better discrimination ability.

We also verified the performance of the adjusted feature by learning the hyperplane $w$ from candidate structures (negatives) of a different large image, but using the same set of nine positives. For our test image of a 19000x10000 pixel size, we obtained 126 false detections (versus 120 previously) out of 403716 candidates. Since we had a very well-represented class of negative candidates this gave us practically the same discrimination results.

The threshold value for the adjusted (learnt) rectangularity feature defines the particular position of the separating hyperplane in the feature space. In our application this threshold can be interactively set by the user in accordance with the maximally allowable number of false detections. Fig. 8 shows an example of classification with the optimal $w_{opt}$ in Eq. (26) when the threshold was set to allow 200 detections. With the threshold used, all positive examples were detected. Fig. 8 shows a couple of examples of false structures detected in this case. In the upper-left figure barriers used for preventing avalanches and in the upper-right figure a group of streams were falsely detected. Additional features are needed to discriminate such cases from livestock enclosures. For example, spectral features could be useful to discard false candidates caused by streams. Geo-referenced digital elevation models can be used to prevent detections in regions with steep slopes. In the bottom-left figure, several manmade structures were detected at the edge of the urban area. Though we segmented out high contrast texture regions, suburban sparsely built-up areas still remained within our search area. This is because the used texture segmentation is accurate at texture borders and avoids segmenting out individual structures, which potentially may be structures of our interest. Manmade structures in the immediate vicinity of urban areas are, however, not of our interest and considered false detections. Such false detections can be easily avoided by extension of segmented regions in the vicinity of urban areas. This requires discrimination of urban texture from other texture types. Bottom right figure shows another example of false detection caused by roads. Note that many false positives appear in groups in the image due to the dense candidate positions. It implies that visual validation of all detected sites is faster than individual inspection of the detected candidates.

5. Discussion and Conclusion

In this paper, we proposed a rectangularity feature that has large values in regions containing approximately rectangular contours, which may be incomplete and fragmented. These structures were modeled by convex configurations of line fragments with orientation angles constrained to be close to zero or ninety degrees. The 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.

A common measure of contour quality (also called saliency criterion) that was used for detection of convex contours in [50, 51] is the ratio of the length of detected fragments and the length of enclosing contour (including gaps). The candidate contour with the highest saliency needs to be determined. In case of incomplete and fragmented target contours, such a criterion is not suitable. This is because detection of an additional fragment may yield an enclosing contour with large gaps decreasing the saliency criterion.

Our approach is similar to [8] in the sense that it uses graph nodes to represent lines and graph edges to represent line relations. However, our undirected graph is built in a way that maximal cliques correspond to appropriate candidate groupings of line segments. To determine appropriateness of the groupings we used a very simple reasoning based on only two rules related to angle and convexity constraints. The strictness of these rules is controlled by parameters. In contrast to [8] we avoid hard thresholds in the geometric rules. Line configurations that closer match the rules result in a higher rectangularity feature.

Some approaches for the detection of manmade structures involved learning model parameters from the data (e.g. [21]). Potentially, such approaches are more robust to changes in images or target objects and are easier adapted to new data than approaches that require complex manual setup of parameters. In our case, we developed a feature that ranks the rectangularity of a structure within a candidate neighborhood. An appropriate detection threshold can be learned from the data, for example by constraining the number of false detections. We suggested learning an FLD-like classifier that may incorporate additional features and showed that adding a feature that quantifies the size of the structure improves the detection performance on our data. The obtained detector can also be seen as thresholding the adjusted rectangularity feature. The major difficulty that arose in our case and may arise in other methods that learn from the data is the small number of representative examples. We proposed a methodology to learn from a large number of readily available negatives and just a few representative positives, which are in our case target enclosures rarely occurring in alpine regions.

The introduced rectangularity and the adjusted (learned) rectangularity features have shown good performance discriminating exemplary ruined enclosures from irrelevant structures and clutter in remotely sensed images. Due to the inherent difficulties in our problem, such a performance is hardly achievable with other approaches for detection of rectangles or with related approaches used, e.g., for detection of buildings.

In general, methods for buildings detection are not suitable for our case because of very low contrasts of linear features, which might be fragmented or absent, and complex background (not only around the contour, but also inside it). Low contrasts are due to low height of the preserved walls of enclosures. Some parts of the walls may also be missed in the edge extraction stage (the width of linear features usually does not exceed two pixels in images of 0.5m resolution we are working with). Complexity of the background stems from various irrelevant structures (trails, streams, rocks), which may have spectral properties similar to those for rectangular structures and may occasionally form rectilinear configurations.

It is also interesting to investigate the usefulness of the rectangularity features for detection of targets other than livestock enclosures. We believe, for example, that our approach may have comparative performance for tasks such as detection of single abandoned buildings or other architectural structures in rural or mountainous areas. In the future, we plan to incorporate additional informative features and apply the FLD-based detector to vast alpine areas captured by remotely sensed images in order to spot unknown livestock enclosures.

Acknowledgments

This work was partially supported by the Interreg IV Program “Alpenhein - Bodensee - Hochrhein”, the Zukunftskolleg University of Konstanz, and by the DFG Research Training Group GK-1042 “Explorative Analysis and Visualization of Large Information Spaces.”

References

References


13


