A Topological Sampling Theorem for Robust Boundary Reconstruction and Image Segmentation

Hans Meine [a,*] Ullrich Köthe [b] Peer Stelldinger [a]

[a] *University of Hamburg, 22527 Hamburg, Germany*

[b] *University of Heidelberg, 69120 Heidelberg, Germany*

Abstract

Existing theories on shape digitization impose strong constraints on admissible shapes, and require error-free data. Consequently, these theories are not applicable to most real-world situations. In this paper, we propose a new approach that overcomes many of these limitations. It assumes that segmentation algorithms represent the detected boundary by a set of points whose deviation from the true contours is bounded. Given these error bounds, we reconstruct boundary connectivity by means of Delaunay triangulation and α-shapes. We prove that this procedure is guaranteed to result in topologically correct image segmentations under certain realistic conditions. Experiments on real and synthetic images demonstrate the good performance of the new method and confirm the predictions of our theory.

Key words: alpha-shapes, geometric sampling theorem, Delaunay triangulation, edgel linking, topology preservation

1 Introduction

Image segmentation is an important component of many image analysis systems. Since the performance of subsequent analysis steps depends on the quality of the segmentation, it is important to understand to which degree a computed image segmentation corresponds to some underlying real-world

* Corresponding Author. Address: Vogt-Kölln-Str. 30, 22527 Hamburg, Germany
Email addresses: meine@informatik.uni-hamburg.de (Hans Meine), ullrich.koethe@iwr.uni-heidelberg.de (Ullrich Köthe), stelldinger@informatik.uni-hamburg.de (Peer Stelldinger).

partitioning. A number of partial answers to this question have been obtained in the past, but they are not sufficiently realistic to model many actual imaging situations (see below). In this paper, we present a new approach that overcomes many of these limitations. In particular, it explicitly takes into account that real-world data is never free of measurement errors.

The analysis we are going to present is based on a clear distinction between the ideal geometric image, which cannot be observed in practice, and the actually available digital image. The geometric image is defined as a perfect geometric projection of the scene onto a plane and has infinite resolution (i.e. is an analog function). However, we do not consider the details of the projection in this work, but consider the geometric image as a given geometric partitioning of the plane into distinct regions. The interior of each region is described by some simple function (possibly even by a constant), while the transitions between regions are discontinuous. Boundaries in the ideal image mostly correspond to object boundaries in the real world, but may also include other visible boundaries such as shadow edges. In our model, this ideal analog image is then transformed into a digital image by a real camera. Besides geometric projection, real cameras are characterized by their point spread function and by the sensor's sampling grid, quantization method, and noise properties. Other effects, such as chromatic aberrations and defocus blur, are not explicitly considered in this work, but are easily included when their consequences can be quantified in terms of known error bounds.

Low-level image segmentation is now defined as the task of reconstructing, as well as possible, the ideal geometric partition from the limited and distorted information in the digital image. In this sense, the partition of the ideal geometric image plays the role of ground-truth, i.e. defines the desired result of low-level segmentation. We ask how accurate the ground truth can be reproduced by a digital reconstruction. In particular, we investigate conditions which ensure that the regions of the reconstruction correspond to the true regions of the ideal geometric partition. That is, we are interested in the question whether and when low-level segmentation methods are able to correctly infer important properties such as the number of regions and their neighborhood and inclusion relations. Since our answers to this question are closely related to the sampling properties of digital images and digital boundaries, we refer to our main results as "geometric sampling theorems".

So far, a number of geometric sampling theorems have been developed that are restricted to binary partitionings, i.e. the plane is split into (not necessarily connected) fore- and background components. Under these conditions, it has been proven that the topology of the partition is preserved under various discretization schemes when the original regions are r-regular (see definition below) and the sampling grid has a maximum pixel radius of at most $r' < r$ [1, 2]. By making slightly stronger assumptions ($r' + p < r$), this property is

preserved when the shapes are blurred by a disc or square of radius p prior to discretization [3, 4]. It is even possible to relax the requirement of r-regularity somewhat to r-halfregularity [5], when the size of the regions is sufficiently large.

However, the above approaches have two important limitations. First, they are not applicable to images where regions cannot be labeled as either fore- or background, e.g. when three distinct regions meet at a junction. Second, they do not predict what happens when the segmentation contains a certain measurement error, i.e. when the available image data is distorted and corrupted by noise. These limitations are partly caused by the fact that the theorems are based on the assumption of a regular (or fixed irregular) sampling grid. In this paper, we are going to drop this assumption in favor of *adaptive sampling* where the location of the sampling points is adapted to the data. In particular, we assume that sampling points are placed roughly along the contour of the regions to be segmented.

We can obtain an adaptive contour sampling in two ways: First, we can keep the restriction to points defined by a regular grid (i.e. to points at pre-defined coordinates), but select suitable subsets of these points, namely points located near the contours of interest. Two variants of this approach are common: we can either retain points in the grid itself, thus arriving at a so-called pixel edges, or we can retain points in the dual grid, thus arriving at the crack edge or inter-pixel boundary located between digitized regions (cf. section 4.1, Fig. 4 and 5 respectively).

Alternatively, we can allow sampling points to be placed on arbitrary coordinates in the plane. These points may, for example, result from a geometric smoothing of the inter-pixel boundary (Euclidean paths [6]), a sub-pixel accurate version of Canny's algorithm [7], or from exact contour following in a smoothly interpolated image, e.g. by means of the predictor-corrector method or the sub-pixel watershed transform [8–10] (Fig. 8). These approaches may be computationally more expensive, but will allow for significantly higher accuracy than the grid-based approaches.

Our analysis of adaptively placed sampling points is inspired by research on laser range scanning. Here, a number of isolated points is scattered over the surface of the object of interest, and the task is to reconstruct the surface from the set of points. A successful solution of this problem is the concept of α*-shapes* [11, 12]. The α-shape is essentially defined as the subset of the Delaunay triangulation of the points where the Delaunay cells' radius is below $\alpha \in \mathbb{R}^+$. Under certain conditions, an α-shape is homeomorphic or at least homotopy equivalent to the desired object surface.

By applying this idea to the problem of image segmentation, we are able to

derive a new condition on object shape which ensures homotopy equivalence of the computed segmentation with the plane partition of the ideal geometric image. In particular, this means that there is a one-to-one mapping between the computed and the ground-truth regions. By imposing slightly stronger requirements on region shape, these properties can even be guaranteed when the segmentation is subject to measurement errors. Our theoretical framework further allows us to give sufficient conditions for preserving neighborhood relations.

After giving preliminary definitions in the following section, our main contribution is made in section 3 in the form of our reconstruction algorithm and corresponding boundary sampling theorem. Subsequently, we apply our method in section 4 to derive both theoretical properties of existing sampling and segmentation methods and experimental results on synthetic and real image data. Finally, the topic of neighborhood relations and a method for thinning our reconstructed boundaries are considered in section 5.

2 Preliminaries

We consider the task of reconstructing the plane partition of an ideal geometric image from a given digital image. The plane partitions to be recovered are defined as follows (see for example [13, 14]):

Definition 1 *A* partition of the plane $\mathbb{R}^2$ *is defined by a finite set of* points $P = \{p_i \in \mathbb{R}^2\}$ *and a set of pairwise disjoint* arcs $A = \{a_i \subset \mathbb{R}^2\}$ *such that every arc is a mapping of the open interval* $(0, 1)$ *into the plane, the start and end points* $a_i(0)$ *and* $a_i(1)$ *are in* P *(but not in* $\cup_A a_i$*). The union of the points and arcs is the* boundary *of the partition* $B = P \cup A$*, and the* regions $R = \{r_i\}$ *are the connected components (maximal connected sets) of the complement of* B*.*

A partition is called *binary* when we can assign two labels (foreground and background) to its regions such that every arc is in the closure of exactly one foreground and one background region.

A binary partition is called *r-regular*, when at every boundary point there exist two osculating discs of radius r which are entirely in the foreground and background respectively. This implies that the regions are morphologically open and closed with respect to discs of radius $\leq r$, and that the curvature of the boundary cannot exceed $1/r$. This, in turn, means that regions cannot have corners, and that it is impossible to represent junctions between three or more regions.

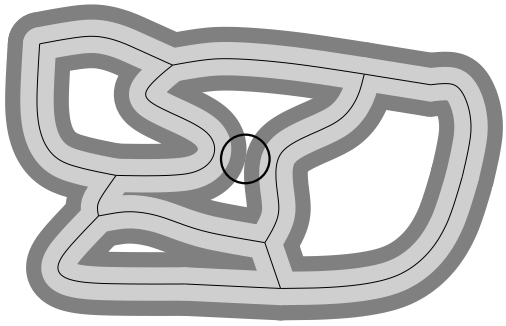


Figure 1. An r-stable plane partition does not change the homotopy type when dilated with a disc of radius of at most r (light gray), while dilations with bigger radius (dark gray) may connect different arcs at waists as marked by the circle.

These restrictions are somewhat relaxed by the notion of *r-halfregular partitions*, where an osculating r-disc must exist at least in the foreground *or* the background, and the number of regions (connected components) must not change under either morphological opening or closing with discs of radius $\leq r$. This definition makes corners possible, but still requires a binary partition that cannot have junctions. The two notions of r-regularity and r-halfregularity have been central to all existing geometric sampling theorems.

In this work, our only requirement is that the boundary of a plane partition does not to change topologically when we thicken it to a certain amount. Therefore we need the topological concept of *homotopy type* [15]. The general definition of homotopy type for spaces of any dimension is rather complicated. In the 2D case however, it has been shown that two bounded sets in $\mathbb{R}^2$ have the same homotopy type if their enclosure trees are isomorphic [2]. The enclosure tree (or enclosure hierarchy) of a bounded set A is defined as follows: The tree's root is given by the unique infinite component of A^c, and its children are the components of A that are adjacent to the root. Each node has all adjacent regions as children that are not already its parent. Thus, any region of A or A^c is enclosed by its parent in the enclosure tree. In our context, the role of A will be played by the boundary set B of the partition, whereas A^c will be the union of all regions.

Now the class of admissible plane partitions used in this work is defined as follows:

Definition 2 *A plane partition is called* r-stable *when its boundary B can be dilated with a closed disc of radius s without changing its homotopy type for any $s \leq r$.*

In other words, we can replace an infinitely thin boundary with a strip of width $2r$ such that the number and enclosure hierarchy of the resulting regions is preserved. In particular, "waists" are forbidden, whereas junctions

are allowed, see Fig. 1. The set of r-stable partitions includes r-regular and r-halfregular partitions, but also allows for non-binary partitions and partitions with junctions. In particular, *polygonal partitions* (all arcs are straight lines) are always r-stable for some sufficiently small r. Thanks to these generalizations, the notion of r-stability is a much more realistic model for the images occurring in practice than r-regularity and r-halfregularity. Unfortunately, the traditional method of proving geometric sampling theorems – which is to establish a topological equivalence between the original and its reconstruction on a fixed sampling grid – cannot be applied to r-stable partitions: since region adjacency is not in general preserved (cf. theorem 18), the reconstruction of an r-stable partition is usually not homeomorphic to the original.

Therefore, we consider another approach to digitization: We approximate the *boundary* of the partition with a finite set of *adaptively placed* sampling points. Such sampling points are selected somehow "near" the boundary. We formalize this as follows:

Definition 3 *A finite set of sampling points $S = \{s_i \in \mathbb{R}^2\}$ is called a* (p, q)-sampling of the boundary B *when the distance of every boundary point $b \in B$ to the nearest point in S is at most p, and the distance of every sampling point $s \in S$ to the nearest point in B is at most q. The elements of S are called* edgels.[1] *The sampling is said to be* strict *when all sampling points are exactly on the boundary, i.e. $q = 0$.*

The Hausdorff distance between the boundary and its sampling is $d_H(S, B) \leq \max(p, q)$ (it is equal to $\max(p, q)$, if both p and q are chosen to be minimal). Non-zero edgel shifts $q > 0$ can be caused by systematic or statistical measurement errors.

2.1 Possible Sources for Boundary Sampling Points

Edgels may be determined in various ways. We distinguish between boundary digitization and boundary detection. In boundary digitization, the ideal geometric partition is given, but has to be encoded in digital form. Thus, the transition to a finite representation (including round-off of real values to integers) is the only source of error in digitization. Our results can be used in the context of digitization to predict the fidelity of a rendering method, e.g. how

[1] Various definitions of the term edgel exist in the literature. In some cases, an edgel is a pixel that has been marked as belonging to an edge. Other authors use the term for a very short 1-dimensional piece of edge, equipped with an orientation. In the present paper, it is a 0-dimensional point located near the true edge. Considering that edgel is an abbreviation for "edge element", all three usages are plausible in their respective contexts.

accurate a letter will be reproduced on a digital display. In contrast, boundary detection is concerned with the reconstruction of real world shapes from their digital images. Here, the correct partition is generally unknown (except in specific experimental settings for algorithm verification), and additional real camera effects such as blurring and noise have to be taken into account.

In boundary digitization, we can start from a set of candidate points at predefined positions (e.g. on some regular grid) and select a subset of these points representing the boundary of interest. If grid points themselves are retained in the subset, the representation may be referred to as a boundary's grid intersection or supercover digitization (see Fig. 4 and formal definitions in section 4). These representations result, for example, from the pixel-accurate version of Canny's algorithm and from many variants of region growing and the watershed transform [7, 16].

Instead, one can also select points of the pixels' dual grid by using crack edges, i.e. edgels are located at the vertices of the inter-pixel boundary (Fig. 5 and formal definition in section 4). These representations typically result from thresholding, but also from grid-based energy minimization approaches such as level-set methods or graph cuts. Alternatively, we can represent the boundary by points at arbitrary plane positions, independent of a grid. For example, we can improve grid-based methods by shifting initial edgels to sub-pixel positions that are supposedly closer to the true boundary, as done in the subpixel-accurate version of Canny's algorithm (Fig. 8a) or when defining Euclidean paths based on discrete tangents [6]. Grid-independent edgels can also be obtained from sub-pixel accurate edge following algorithms [8–10] (Fig. 8b) and active contours [17], to mention just a few possibilities.

In other words, our results can be applied to any segmentation method that represents the computed boundary by a set of points. It is only required that the accuracy of the representation, i.e. the maximum errors p and q, can be estimated when the algorithm is applied to the image class of interest. We then reconstruct a connected boundary from the points by means of the algorithm introduced in section 3.

2.2 Triangulations and α-Shapes

Our new algorithm is based on the Delaunay triangulation:

Definition 4 *The* Delaunay triangulation D *of a set of points* S *is the set of all triangles formed by triples* $t \subset S$ *such that the open circumdisc (i.e. the interior of the circumcircle) of every triangle does not contain any point of*

S.[2] *If the points are in general position (i.e. no four of them are on a common circle), the Delaunay triangles, their edges and corners (also denoted as 2-, 1- and 0-cells in this context) form a uniquely defined, connected simplicial complex. The union of all cells $|D| = \bigcup_{c \in D} c$ is called the* polytope of D.

The Delaunay triangulation completely partitions the plane into edges and triangles. However, in the context of segmentation we are only interested in those edges and triangles of the Delaunay triangulation that are related to the boundary of the desired segmentation. Edges and triangles lying within the regions of the desired segmentation should be removed. A suitable subset of the Delaunay triangulation is defined by the α-complex introduced in [11]:

Definition 5 *The α-complex D_α of a set of points S is defined as the subcomplex of the Delaunay triangulation D of S which contains all cells c such that*

- *the radius of the smallest circumcircle of c is smaller than α, and the interior of this circle contains no point of S, or*
- *an incident cell c' with higher dimension is in D_α.*

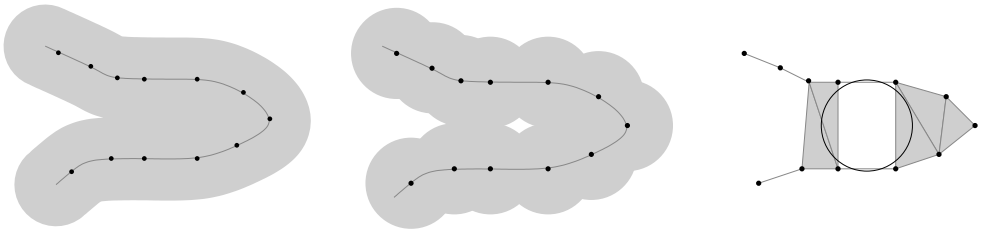
The polytope $|D_\alpha|$ is called *α-shape*. Since cells are removed from the Delaunay triangulation, the α-complex has holes which are closely related to the regions of the desired segmentation. In order to make this relationship precise, the following theorem is of fundamental importance (the proof can be found in [12]):

Theorem 6 (Edelsbrunner) *The union of closed discs of radius α centered at the points $s_i \in S$ covers the polytope $|D_\alpha|$, and the two sets are homotopy equivalent.*

The intuitive meaning of the theorem is as follows: Suppose the edgels $s_i \in S$ are located near the boundary of a given plane partition. Then, the polytope $|D_\alpha|$ is homotopy equivalent to this plane partition if and only if the dilation of the edgels with α-discs is homotopy equivalent to the boundary of the partition. The following theorem shows that this requirement is indeed fulfilled under certain conditions:

Theorem 7 (Bernardini & Bajaj) *Suppose the plane partition is r-regular, and S is a strict sampling of its boundary B such that $p \leq r$, $q = 0$. Then the polytope $|D_\alpha|$ is homotopy equivalent and even homeomorphic to the boundary B for all $p < \alpha < r$.*

[2] Note that an *open* circumdisc does not contain the corners of the constituting triangles since it does not contain its own boundary.



(a) (b) (c)

Figure 2. It may happen that the α-dilation (a) of the boundary of an α-stable plane partition is not homotopy equivalent to the union (b) of the α-discs centered at the edgels. Thus the α-shape (c), which is always homotopy equivalent to the union of discs (b), may contain unwanted holes. These holes consist of Delaunay triangles of radius greater than α, thus there exists an α-disc centered in the hole which does not cover any edgel, as shown in (c).

The proof of this theorem is given in [18]. Under these particular conditions, D_α does not contain any triangles – it only consists of edges and points and thus defines a plane partition in itself. According to the theorem, this plane partition is topologically equivalent to the original partition whose boundary was sampled by S. In other words, the α-complex completely defines the correct linking of edgels into edge chains.

Unfortunately, this no longer applies when the original partition is not r-regular and/or the edgels are not exactly on the original boundary. Fig. 2 shows an example where the r-dilation of the boundary is homotopy equivalent to the boundary (i.e. the partition is r-stable), but the dilation of the edgels is not.

3 Segmentation with Alpha-Shapes

Since the holes of an α-complex do not necessarily correspond to regions of the original plane partition, we must characterize these holes in more detail.

This is facilitated by the following definition:

Definition 8 *Consider the Delaunay triangulation D of a point set S and the complement $D_\alpha^C = \mathbb{R}^2 \setminus |D_\alpha|$ of the corresponding α-polytope with $\alpha > 0$. A connected component of D_α^C is called an* α-hole of $|D_\alpha|$. *When the radius of the circumcircle of the largest Delaunay triangle in an α-hole's closure is at least $\beta \geq \alpha$, we speak of an (α, β)-hole.*

For simplicity, we also use the term "hole" for the component which contains the infinite region. It is an (α, β)-hole for arbitrary large β.

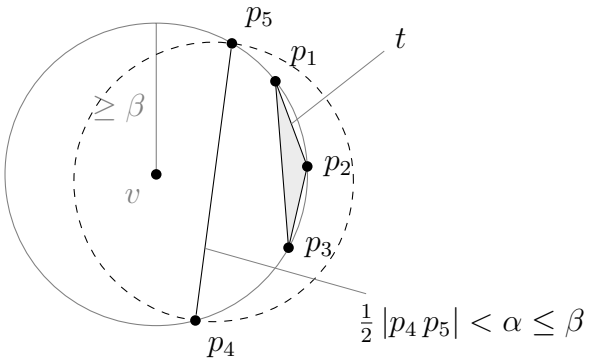


Figure 3. Any circumcircle around p_4 and p_5 contains p_1, p_2, and p_3 (see lemma 9).

It follows from theorem 6 that there is a 1-to-1 relation between α-holes and the holes in the union of α-discs around the edgels. The following lemma establishes that a similar relationship exists for (α, β)-holes:

Lemma 9 *An α-hole h is an (α, β)-hole if and only if it contains a point v whose distance from the nearest edgel is at least β.*

PROOF. I ($d_H(v \in h, S) \geq \beta \Rightarrow h$ is an (α, β)-hole): when v is in the infinite region, the claim follows immediately. Otherwise, v is contained in some Delaunay triangle. By assumption, the corners of this triangle must have distance $\geq \beta$ from v. Therefore, the radius of the triangle's circumcircle must be at least β, and the claim follows.

II (h is an (α, β)-hole $\Rightarrow \exists\, v \in h$ with $d_H(v, S) \geq \beta$): by assumption, the closure of h contains a Delaunay triangle t with circumradius $\beta' \geq \beta \geq \alpha$. Consider the center v of its circumcircle. If v is within the triangle t, it is also in h and the claim follows. Otherwise, v is at least in *some* (α, β)-hole (since the open β'-disc contains no edgel), and we must prove that t is in the same hole. Suppose the contrary, and consider Fig. 3, which shows the triangle t and its circumcircle (gray) with center v. For v and t to be in different α-holes, there needs to exist a Delaunay triangle t' or a single edge e between t and v whose smallest circumcircle is smaller than α. The corners of t' or e cannot be inside the circumcircle of t because otherwise t would not be a Delaunay triangle. Neither t' nor e can contain v because their circumcircle radius would then be at least β. Let the points p_4 and p_5 be the end points of e or of one side of t'. Their distance $|p_4\,p_5|$ must be greater than $|p_1\,p_3|$. Consequently, any circumcircle with radius $\leq \alpha$ (dashed) around p_4 and p_5 contains t, contrary to the condition that it must not contain any other edgel (according to the definition of the α-complex). The claim follows from the contradiction. $\square$

The reason for defining (α, β)-holes is that even for optimally chosen α, the

α-complex does not necessarily have the same homotopy type as the ground-truth boundary, since it may contain too many holes, as can be seen in Fig. 2. We solve this problem by introducing a second parameter β for the size of such holes (we will soon determine a proper value for β), i.e. we use the notion of (α, β)-holes to "repair" α-complexes that contain too many holes:

Definition 10 *An* (α, β)-boundary reconstruction *from a set of edgels* S *is defined as the union of the polytope* $|D_\alpha|$ *with all* α*-holes of* D_α *that are not* (α, β)*-holes.*

In other words, holes smaller than a certain size are simply "painted over", and (α, β)-boundary reconstruction essentially amounts to edgel linking by hysteresis thresholding on the triangle size of a Delaunay triangulation. Our complete boundary reconstruction algorithm is thus summarized as follows:

(1) Compute the Delaunay triangulation D of the edgels S.
(2) Mark all triangles in D (including their edges) with a circumradius $< \alpha$.
(3) Additionally mark Delaunay edges whose circumcircle contains no edgel and has a radius smaller than α.
(4) Find connected components of unmarked triangles and edges.
(5) For each component from step 4 which does *not* contain any triangle with a circumradius of at least β, mark all its triangles and edges.

A key feature of the algorithm is that it does not enforce the boundary to be thin. This is important because the accuracy of the edgels may be insufficient to decide unambiguously and precisely where the boundary runs. Instead of forcing possibly wrong decisions at such locations, our algorithm keeps triangles in the boundary which cover the correct boundary without deciding on their precise location. It is even possible that a piece of expanded boundary covers several pieces of the correct boundary running in close vicinity to each other, when the accuracy of the edgels does not allow to tell whether or not these pieces meet, and if they meet, exactly where that happens.

Thanks to this relaxed definition of the boundary, the (α, β)-boundary reconstruction algorithm manages to preserve exactly those holes that are in one-to-one correspondence to the regions of the ground-truth, as long as the values of α and β are properly chosen. This is demonstrated in the following theorem which can be interpreted as a new sampling theorem for region boundaries.

Theorem 11 (Boundary sampling theorem) *Let* $\mathcal{P}$ *be an* r*-stable plane partition, and* S *be a* (p, q)*-sampling of* $\mathcal{P}$*'s boundary* B*. Then the* (α, β)*-boundary reconstruction* $\mathcal{R}$ *defined by* S *is homotopy equivalent to* B*, and the* (α, β)*-holes of* $\mathcal{R}$ *are topologically equivalent to the regions* r_i *of* $\mathcal{P}$*, provided the following conditions are met:*

(1) $p, q < \alpha \leq r - q$
(2) $\beta = \alpha + p + q$
(3) every region r_i *contains an open* γ*-disc with* $\gamma \geq \beta + q > 2(p + q)$.

PROOF. Let $U = S \oplus \mathcal{B}^o_\alpha$ be the union of open α-discs centered at the points of S. Furthermore, let $B^\oplus_q = B \oplus \mathcal{B}_q$ be the dilation of B with a closed q-disc, $B^\oplus_{\alpha+q} = B \oplus \mathcal{B}^o_{\alpha+q}$ the dilation of B with an open $(\alpha + q)$-disc, and $r_i^\ominus = r_i \ominus \mathcal{B}_{\alpha+q}$ the erosion of region $r_i \in \mathcal{P}$ with a closed $(\alpha + q)$-disc.

(Same number of components in B and $|D_\alpha|$)[3] Due to r-stability of the partition $\mathcal{P}$ and $\alpha + q \leq r$ (which follows from condition 1), both $B^\oplus_q$ and $B^\oplus_{\alpha+q}$ have as many connected components as B. According to the definition of a (p, q)-sampling, $B^\oplus_q$ covers S. Moreover, along with condition $p < r$ this definition ensures that every component of $B^\oplus_q$ covers at least one edgel from S. Thus, the number of edgels is at least as big as the number of components in B. Since $B^\oplus_{\alpha+q}$ and U are obtained from $B^\oplus_q$ and S by dilation with an open α-disc, $B^\oplus_{\alpha+q}$ covers U, and every component of $B^\oplus_{\alpha+q}$ contains at least one component of U. It follows that the number of components in U is at least as big as the number of components in B.

Conversely, since $\alpha > q$ and $\alpha > p$, every open α-disc around a point of S intersects B, and the union U of these discs covers the entire boundary B. It follows that U cannot have more components than B, because otherwise there would be an edgel in S with distance of at least α from B, in contradiction to the definition of a (p, q)-sampling. The number of components of B and U is thus equal. Due to homotopy equivalence of U and $|D_\alpha|$ (theorem 6), this also holds for the components of $|D_\alpha|$.

(Injective mapping from regions to α-holes) Since $\mathcal{P}$ is r-stable with $r \geq \alpha + q$, each $r_i^\ominus$ is a connected set with the same topology as r_i. The intersection $r_i^\ominus \cap B^\oplus_{\alpha+q}$ is empty, and $r_i^\ominus$ cannot intersect $U \subset B^\oplus_{\alpha+q}$ and $|D_\alpha| \subset U$. Hence, $r_i^\ominus$ is completely contained in a single α-hole of $|D_\alpha|$, and a function which maps every region r_i onto the α-hole containing $r_i^\ominus$ is injective.

(Injective mapping from regions to (α, β)-holes) Due to condition 3, r_i contains a point whose distance from B is at least $\gamma = \beta + q$. Its distance from S is therefore at least $\gamma - q = \beta$. Due to lemma 9, the α-hole which contains $r_i^\ominus$ is therefore also an (α, β)-hole. Thus, the function already defined for α-holes is an injective mapping from regions to (α, β)-holes.

(Bijective mapping from regions to (α, β)-holes) An α-hole that does not intersect any region $r_i^\ominus$ must be completely contained within $B^\oplus_{\alpha+q}$. Every point $v \in B^\oplus_{\alpha+q}$ has a distance $d < \alpha + q$ to the nearest point of B. In turn,

[3] This part of the proof is based on the following general property of sets: Let $U \subset V$ be two sets, and $\mathrm{cc}(U)$, $\mathrm{cc}(V)$ the numbers of their (connected) components. Then, if every component of V contains at least one point of U, it follows that $\mathrm{cc}(U) \geq \mathrm{cc}(V)$.

every point in B has a distance of at most p to the nearest point in S. Hence, the distance from v to the nearest point of S is $d' < \alpha + p + q = \beta$. According to lemma 9, this means that an α-hole contained in $B^{\oplus}_{\alpha+q}$ cannot contain a triangle with circumradius β and cannot be an (α, β)-hole. This implies that the already defined mapping is also surjective.

(Bijective mapping between the components of B and $|\mathcal{R}|$) The previous observation has two consequences: (i) All holes h_i remaining in $\mathcal{R}$ intersect a region $r_i^{\ominus}$. Therefore, the correspondence between r_i and h_i is 1-to-1, and B and $|\mathcal{R}|$ enclose the same number of regions. (ii) All differences between $\mathcal{R}$ and D_α (i.e. all Delaunay cells re-inserted into $\mathcal{R}$) are confined within $B^{\oplus}_{\alpha+q}$. This implies that $|\mathcal{R}|$ cannot have fewer components than $B^{\oplus}_{\alpha+q}$ and B. Since all re-inserted cells are incident to D_α, $|\mathcal{R}|$ cannot have more components than $|D_\alpha|$, which has as many components as B (see first observation). Hence, B and $|\mathcal{R}|$ have the same number of components.

(Homotopy equivalence of B and $|\mathcal{R}|$) Consider the components of the complement $(r_i^{\ominus})^C$ and recall that $r_i^{\ominus}$ is a subset of both r_i and h_i for any i. Since B and $|\mathcal{R}|$ have the same number of components, it is impossible for h_i^C to contain a cell that connects two components of $(r_i^{\ominus})^C$. This means that the sets r_i^C and h_i^C have the same number of components. This finally proves the topological equivalence of r_i and h_i, and implies homotopy equivalence of B and $|\mathcal{R}|$. $\square$

If there exists no r such that all conditions of theorem 11 are fulfilled for a given plane partition (or if the chosen α is too large), it cannot be guaranteed that the regions of the (α, β)-boundary reconstruction have the same topology as the ground-truth ones. E.g. if a ground-truth region is too small, it may happen that it is lost in the reconstruction, because the resulting edgels are so close to each other that the configuration could also have resulted from measurement errors along a single boundary. On the other hand, if a region has an s-waist for $s \leq 2\alpha$, i.e. if the s-dilation of the boundary cuts the region into two or more components, edgel localization errors may cause a single true region to be split into two or more holes of the (α, β)-boundary reconstruction. In case of very narrow waists, i.e. when $s+2p+2q \leq \alpha$, this is even guaranteed to happen, i.e. the two sides of a narrow waist are always connected by at least one line segment in the reconstruction (and if the different parts of the original region are big enough, every part occurs in the reconstruction as a separate region).

Even if a plane partition does not fulfill the requirements of our sampling theorem, we may still derive useful properties of its reconstruction. Suppose that the original plane partition has waists whose width is below some desired distance $2r$. Now construct a new plane partition where these waists are closed by drawing a new arc between the two sides of every narrow waist. When the modified plane partition fulfills the requirements (in particular, when it is

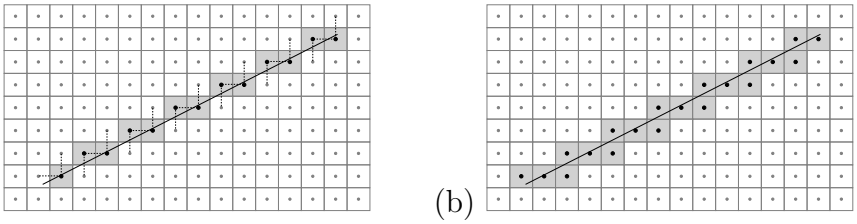


Figure 4. (a) Where the boundary intersects the dual grid, the nearest sampling points form the *grid intersection digitization*. (b) The *supercover digitization* contains all sampling points whose pixel facets intersect the arc.

now r-stable with the desired r), the theorem now guarantees preservation of the modified topology. Thus, (α, β)-reconstruction handles narrow waists in a well-defined, predictable way, namely as if they had been connected. This fact is illustrated by the example shown in the second row of Fig. 13.

4 Application to Common Sampling and Segmentation Schemes

In theorem 11, the parameters p and q are assumed to be given. In order to make their meaning and effects more intuitive, we compute or estimate these numbers for common sampling and segmentation schemes in the following.

4.1 Sampling Schemes

Let's first look at digitization schemes, i.e. at the case where the ground-truth partition is known and has to be represented by means of a regular or irregular grid. This situation is common in the context of rendering, and the sampling theorem may be used to compute at which resolution a given shape should be rendered in order to ensure topological correctness of the rendered shape.

4.1.1 Grid Intersection Digitization

One important digitization scheme called grid intersection digitization is defined on regular grids:

Definition 12 *Consider a plane partition $\mathcal{P}$ with boundary B and a square grid. Compute all intersection points of B with the grid lines (i.e. with the lines connecting 4-adjacent grid points) and round their coordinates to the nearest grid point coordinate. The set of edgels thus defined is called* grid intersection digitization *of B, see Fig. 4a.*

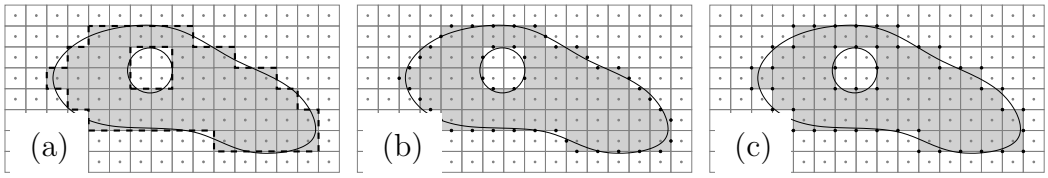


Figure 5. The *inter-pixel boundary* (dashed) can be extracted from the subset digitization (a). It includes both the *midcrack digitization* (b) and the *endcrack digitization* (c).

For simplicity, let the grid size (i.e. the smallest distance from one sampling point to another) be unity. When each component of B crosses at least one grid line, the distance p of any point of B to the nearest selected grid point is less than $\sqrt{2}$, and the distance q of any grid intersection to its rounded coordinate cannot exceed $1/2$. Inserting this into the conditions of theorem 11, we get $\alpha \geq \sqrt{2}$, $r \geq \sqrt{2} + \frac{1}{2}$, $\beta \geq 2\sqrt{2} + \frac{1}{2} \approx 3.3$, and $\gamma \geq 2\sqrt{2} + 1 \approx 3.8$. However, the worst case configurations giving rise to the values of β and γ in the theorem cannot actually occur in a square grid because Delaunay edges between grid points cannot have arbitrary length. It can be shown that the largest circumradius in an undesirable α-hole is below $\sqrt{34} \approx 2.9$, so that $\gamma \approx 3.4$ (circle area 37 pixels) is sufficient.

Generally the grid intersection digitization of a connected curve is an 8-connected digital curve. It is identical to Bresenham's digital straight line in case of a straight arc. Moreover the grid intersection digitization is a subset of the supercover digitization on a square grid, which produces a 4-connected digital curve for any connected curve:

Definition 13 *Let $\mathcal{P}$ be a plane partition with boundary B and $\mathcal{G}$ a finite set of sampling points such that the Voronoi cells of $\mathcal{G}$ have a radius of at most g. The* supercover digitization *of B is the set of all sampling points whose Voronoi cell intersects B, see Fig. 4b for the common case that $\mathcal{G}$ is a square grid.*

The constraint on the size of the Voronoi cells implies that $p = g$ and $q < g$. Hence, $\alpha > g$, $r > 2g$, $\beta > 3g$ and $\gamma > 4g$ are required. For example, in a unit square grid we have $q < p = \sqrt{2}/2$ and $\gamma > 2\sqrt{2} \approx 2.8$. Thus, the supercover digitization imposes weaker constraints on the original plane partition $\mathcal{P}$ than the grid-intersection digitization. This is mainly due to the denser sampling of the boundary (smaller spacing of the edgels) in the former. As stated in [19], the supercover digitization is a Hausdorff discretization, i.e. a set of sampling points which minimizes the Hausdorff distance to the boundary B. Since this Hausdorff distance is equal to $\max(p, q)$, the given bounds for α, β and γ are sufficient for all Hausdorff discretizations.

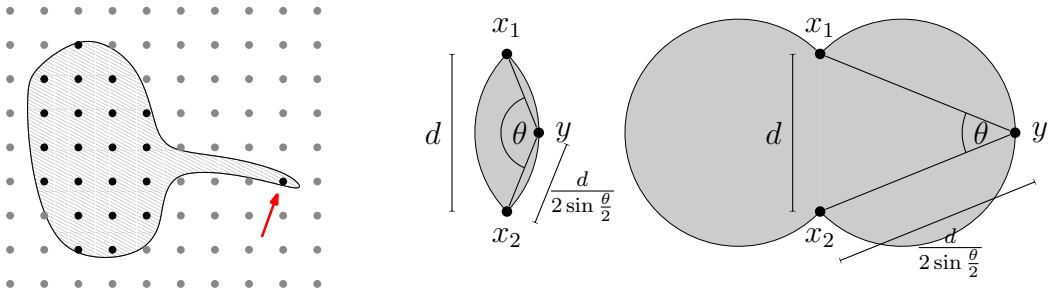


Figure 6. *left:* An r-stable region (hatched) whose subset digitization $\hat{r}_i$ (black circles) has a different topology (note the isolated point indicated by the arrow). *center and right:* In order to preclude (θ, d)-spikes, the contour must contain a path between its two points x_1 and x_2 which remains entirely within the shaded region. (*center*: $\theta > 90°$, *right*: $\theta < 90°$)

4.1.2 Inter-pixel Boundaries from Region-Based Methods

Another interesting question is what can be said about region based digitization methods, in particular the subset digitization:

Definition 14 *Let $\mathcal{P}$ be a plane partition with regions $R = \{r_i\}$ and $\mathcal{G}$ a finite set of sampling points such that the Voronoi cells (i.e. the pixels) of $\mathcal{G}$ have a radius of at most g. The* subset digitization *$\hat{r}_i$ of region r_i is the union of all Voronoi cells whose sampling point is in r_i, see Fig. 5a. The union of the boundaries of all $\hat{r}_i$ is called the* inter-pixel boundary*. A boundary digitization scheme where all edgels are on the inter-pixel boundary $\mathbb{B}$ is an* inter-pixel digitization*. Two examples are the the* midcrack digitization *(Fig. 5b) where the center points of all pixel edges inside the inter-pixel boundary $\mathbb{B}$ are chosen as edgels, and the* endcrack digitization *(Fig. 5c) where all pixel corner points lying on the inter-pixel boundary $\mathbb{B}$ are used.*

Thus, *boundary-based* digitizations like endcrack and midcrack digitization can be derived from the *region-based* subset digitization. While the maximal distance q of any edgel to the nearest boundary point cannot exceed g, the distance p from any boundary point to the nearest edgel can be arbitrary large, as the following considerations illustrate:

An r-stable region is not necessarily r-regular. In particular, the region may have *spikes* as shown in Fig. 6 (left). Consequently, there is no guarantee that the subset digitization $\hat{r}_i$ of the region r_i is topologically equivalent to r_i. As illustrated in that figure, $\hat{r}_i$ may even be disconnected (despite r_i being connected), and the distance between the components of $\hat{r}_i$ can become arbitrarily large.

Obviously, this prevents us from defining a useful upper bound for the value of p. Thus, we need a condition which is stronger than r-stability, but weaker than r-regularity and which preclude these undesirable spikes:

Definition 15 *Let $\mathcal{P}$ be a plane partition with boundary B. We say two points $x_1, x_2 \in B$* delimit a (θ, d)-spike*, if the distance from x_1 to x_2 is at most d and if there exists no path on B from x_1 to x_2 which contains only points y with $\angle(x_1\, y\, x_2) \geq \theta$.*

Conversely, we say that $\mathcal{P}$ has no (θ, d)-spikes *if for any pair of boundary points $x_1, x_2 \in B$ with distance of at most d, there exists a path $Y \subset B$ between x_1 and x_2 such that $\angle(x_1\, y\, x_2) \geq \theta$ for all points $y \in Y$.*

This definition is illustrated by Fig. 6 (center and right): Let x_1 and x_2 be two points on B with distance d. Then, the points which enclose angles of at least θ with x_1 and x_2 are necessarily located in the shaded region (center: $\theta < 90°$, right: $\theta > 90°$). Among these points, the indicated point y has the maximal distance from x_1 and x_2, namely $\frac{d}{2\sin\frac{\theta}{2}}$. By requiring a path on the contour between x_1 and x_2 which remains entirely within the shaded region, the size of admissible spikes is restricted with Def. 15.

Intuitively, two points delimit a (θ, d)-spike, if the shortest boundary path between them deviates too much from a straight line, i.e. it leaves the shaded region in Fig. 6. But this intuitive description cannot be used for the definition, since we want to apply it to fractal arcs as well. A fractal arc has infinite length, so the notion of shortest path is not applicable, but the arc may nevertheless be free of θ-spikes (see below).

It should be noted that the definition also covers the unintuitive situation where the two points x_1 and x_2 are located on different components of the boundary B. Since no path exists between these points at all, the partition $\mathcal{P}$ can only be free of (θ, d)-spikes if the distance between x_1 and x_2 is larger than d. This requirement is closely related to the absence of waists in $\mathcal{P}$. Thus, this counter-intuitive kind of "spikes" cannot occur in practice if $\mathcal{P}$ is r-stable with $r > d/2$.

In case of r-regular sets we can give a precise bound for the possible angle θ given two boundary points x_1, x_2 with some distance d: Due to r-regularity, there exists a path from x_1 to x_2 on B, which lies inside the intersection of the two r-discs that have x_1 and x_2 on their exterior. By simple geometric construction it follows that for any point x inside this intersection, the angle $\angle(x_1\, x\, x_2)$ is at most equal to $2\arctan\left(\frac{d}{2r-\sqrt{4r^2-d^2}}\right)$. Thus, an r-regular partition has no (θ, d)-spikes for all $d \leq r$ and $\theta = 2\arctan\left(\frac{d}{2r-\sqrt{4r^2-d^2}}\right)$ (e.g. for $\theta = 90°, 60°$ we get $d = r$ and $d = \sqrt{3}r$ respectively). By sampling the boundary of an r-regular partition densely enough, we can make d arbitrary small, which implies that the angles $\theta = \angle(x_1\, y\, x_2)$ become arbitrarily flat, i.e. tend toward 180°. In other words, we can always enforce absence of (θ, d)-spikes by making d small. In general, absence of (θ, d)-spikes will not even imply r-stability (let alone r-stability), so we have to require both.

Absence of spikes does not only restrict the angle between points of the original ground-truth contour. It also implies a restriction on the angles between pair of adjacent line segments in the boundary reconstruction, provided the edgels are placed with sufficient density (small distance along the contour). Consider Fig. 6 and let y be a point on the contour path Y connecting the contour points x_1 and x_2. It is easily seen that the distance from y to the nearer point among x_1 and x_2 is at most $\frac{d}{2\sin\frac{\theta}{2}}$. When all edgels lie on the contour (i.e. $q = 0$), this immediately implies that adjacent line segments of the reconstruction cannot enclose an angle smaller than θ. If $q > 0$ (e.g. if the digitization suffers from round-off error), the conclusion is not so simple, but the following theorem shows that the absence of (θ, d)-spikes can still be used to compute an upper bound for p:

Theorem 16 *Let G be a square grid with sampling distance h (pixel radius $g = \frac{h}{\sqrt{2}}$). Further, let $\mathcal{P}$ be a plane partition such that every region $r_i \in \mathcal{P}$ contains a closed g-disc and the boundary B has no (θ, d)-spikes. Then the endcrack digitization of B is a (p, q)-boundary sampling with $q = \frac{h}{\sqrt{2}}$ and $p = q + \left(\frac{h}{2} + q\right) / \sin\frac{\theta}{2}$, provided that $h \leq \frac{d}{1+\sqrt{2}}$. Likewise, the midcrack digitization is a (p, q)-boundary sampling with $q = \frac{h}{2}$ and $p = q + \left(\frac{h}{2} + q\right) / \sin\frac{\theta}{2}$, provided that $h \leq \frac{d}{2}$.*

PROOF. First, we prove the bounds on q. Let x, y be two 4-adjacent square grid points. Their common pixel edge is in the inter-pixel boundary if and only if x and y lie in different regions r_i and r_j, i.e. the grid line between x and y intersects the boundary B in at least one point v. The endcrack edgels are exactly the end points of these pixel edges, and their distance to v is at most $\frac{h}{\sqrt{2}}$. It follows that $q = \frac{h}{\sqrt{2}}$ for the endcrack digitization. The midcrack edgels are the center points between x and y, so their maximum distance to v is $\frac{h}{2}$. Hence, $q = \frac{h}{2}$ for the midcrack digitization. The maximum distance between neighboring edgels on the inter-pixel boundary is h in both cases.

Now, we prove the bound on p given q. By definition $B = \bigcup \partial r_i$, where ∂r_i is the boundary of region r_i. Since every region contains a closed disc of radius $g = \frac{h}{\sqrt{2}}$, and every such disc contains at least one grid point, every region r_i contains a grid point, i.e. $\hat{r}_i$ is not empty, and there exist at least four edgels near ∂r_i. Due to the nonexistence of (θ, d)-spikes any two boundary points with distance of at most d must belong to the same component. This implies that any two components $(\partial r_i)_j$ and $(\partial r_i)_l$ of the boundary ∂r_i must have a minimum distance of more than d. Note that $d \geq 2h = 4q$ holds. So, for every component there exists a set of edgels which are closer to $(\partial r_i)_j$ than to any other component.

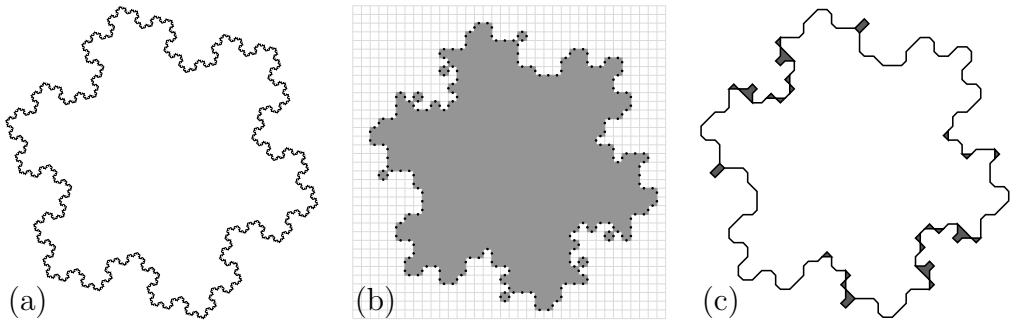


Figure 7. (a) Koch Snowflake; (b) subset digitization of (a) with midcrack edgels marked (note the topology violations); (c) (α, β)-boundary reconstruction from midcrack edgels. Uncertain areas where the edgels do not unambiguously determine the shape of the original boundary pop out by being thick.

Obviously, every contour $(\partial r_i)_j$ constitutes a closed curve. Thus by mapping every edgel to the nearest point of B, one gets a cyclic list of points $[b_k]^{(ij)}$ for every component $(\partial r_i)_j$, and each point b_k has a distance of at most $h+2q$ to its successor b_{k+1} in the list. For endcrack edgels, we have $h+2q = (1+\sqrt{2})h \leq d$, and for midcrack edgels $h+2q = 2h \leq d$. Thus, the boundary part between b_k and b_{k+1} includes no point with an angle smaller than θ. As shown in Fig. 6, this implies that the distance from any boundary point between b_k and b_{k+1} to the nearer one of these two points is at most $\left(\frac{h}{2}+q\right) / \sin\frac{\theta}{2}$. Thus, the maximum distance to the nearest of the two edgels which are mapped onto b_k and b_{k+1} is $p = q + \left(\frac{h}{2}+q\right) / \sin\frac{\theta}{2}$. □

For instance, when $h = 1$ and the plane partition has no $(60°, d)$-spikes with $d > 2.4$, we get $p \approx 3.12$, $q \approx 0.71$ for endcrack and $p = 2.5$, $q = 0.5$ for midcrack digitization. It follows that midcrack digitization should be favored over endcrack digitization.

The nonexistence of spikes allows us to even digitize objects topologically correctly that have a fractal boundary, like the Koch Snowflake (see Fig. 7): Let K be the object bounded by the Koch Snowflake based on a triangle of side length 1. It follows that K is r-stable for all $r < \frac{1}{\sqrt{3}}$, it has no $(60°, d)$-spikes for $d < \frac{1}{\sqrt{3}}$, and it contains a γ-disc for any $\gamma \leq \frac{1}{\sqrt{3}}$. Thus, the (α, β)-boundary reconstruction based on the midcrack digitization with a square grid of grid size h is correct for all $h < \frac{1}{\sqrt{27}} \approx 0.192$.

4.2 Segmentation Schemes

Let us now turn our attention from digitization to segmentation, where the distortions of real cameras have to be taken into account. Many segmentation

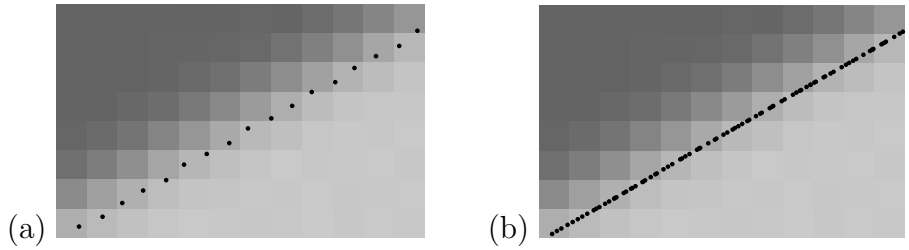


Figure 8. Subpixel-accurate edgels from Canny's algorithm (a) and the subpixel watershed algorithm (b). Note the lower density and higher displacement of the edgels in (a).

algorithms (e.g. zero-crossing-based edge detectors and the watershed algorithm) compute image labellings similar to subset digitization, which can be used to define endcrack and midcrack edgels. However, their error bounds differ from the ideal ones obtained above. To quantify these differences, we need a model of the transformation from analog to digital images in real cameras:

$$f_{ij} = (\mathrm{PSF} \star f(x,y))_{ij} + n_{ij} \tag{1}$$

where $f(x,y)$ is the ideal geometric image, PSF is the point spread function, subscripts denote sampling, and n_{ij} is additive Gaussian white noise. (Lens effects not captured by the PSF, such as vignetting and coma, and gray-level quantization are neglected). We assume that the ideal geometric image is composed from regions r_i whose interior can be described by a set of smooth functions f_i, but the transition between two such functions f_i and f_j is almost everywhere discontinuous along the common boundary between regions r_i and r_j. The ideal image is thus defined as

$$f(x,y) = \sum_i \rho_i(x,y) f_i(x,y) \tag{2}$$

where ρ_i is the indicator function of region r_i. The discontinuities between the regions define a plane partition which we regard as the ground truth of the segmentation problem. Convolution of $f(x,y)$ with PSF (which shall be band-limited) suppresses high spatial frequencies, and the resulting smooth transitions between regions allow for sub-pixel accurate edge localization. On the other hand, systematic localization errors are introduced because blurring distorts curved edges and boundaries near corners and junctions. Noise causes additional statistical errors in p and q.

We have estimated these errors for a number of exemplary edge detectors. In this work, we consider two variants of the watershed transform, Canny's algorithm, and the Haralick detector respectively. The latter serves as an example of zero-crossing-based algorithms, while the former methods represent gradient-based edge detection.

The watershed transform and Canny's edge detector look for relative maxima

of the gradient magnitude $g = \sqrt{f_x^2 + f_y^2}$ which indicates discontinuities. The methods differ in the definition of (1-dimensional) maxima: When interpreting the values of g as the height of a terrain, watersheds are those ridges which shed raining water into separate catchment basins. A common method for finding these ridges is a simulated flooding of the terrain [16], which produces a pixel-accurate segmentation with inter-pixel boundaries defining end-crack or mid-crack edgels (in our experiments, we use end-crack edgels). Higher accuracy is possible by means of the *subpixel watershed algorithm*, which first locates the saddle points in the continuously interpolated gradient magnitude image g and then traces edges by upwards path-following from these points [9, 20], see examples in the left column of Fig. 9.

Canny's algorithm [7] looks for relative maxima along the gradient direction (center column of Fig. 9). In its pixel-accurate variant, the resulting edgels are similar to a grid-intersection digitization, i.e. we get 8-connected edgel chains. Better localization (significantly smaller q) is achieved by placing the edgels at the subpixel maxima of the gradient g, either by means of an approximating parabola across the edge, or by means of Newton iterations on a continuously interpolated version of the gradient image. Fig. 8 shows that the results of the subpixel watershed algorithm are slightly better than those of the subpixel Canny algorithm, mostly because Canny's algorithm can find at most one edgel per pixel, resulting in rather large values for p.

In contrast, Haralick [21] defines edgels at the zero-crossing of the second derivative along the gradient direction:

$$b = f_x^2 f_{xx} + 2 f_x f_y f_{xy} + f_y^2 f_{yy} = 0 \tag{3}$$

provided that the third derivative along the same direction is negative (indicating a local gradient maximum), and the gradient magnitude is above a threshold. Crack edges between positive and negative pixels of b where the constraints are fulfilled define a set of end-crack or midcrack edgels (in the experiments, we use midcrack edgels). The limited accuracy of this grid-based representation can be improved when a continuous function $\tilde{b}$ is computed by means of spline interpolation of b, and edgels are located in $\tilde{b}$ by means of Newton iteration along the gradient direction. In our implementation of this variant, edgels are placed roughly at a distance of 0.1 pixels along the edge, Fig. 9 right column.

We estimate p and q on a large set of synthesized images. The original geometric images of the test set consisted of straight lines, curved lines, corners, and junctions at various angles, translations and contrasts. The actual test images were generated from the ground truth by numerical solution of the convolution integral (1) with a Gaussian PSF at scale $\sigma_{\mathrm{PSF}} = 1$ and additive white noise with various signal-to-noise ratios, see Fig. 9. Derivatives are computed by

Gaussian filters at scale σ_E. To avoid aliasing in the sampled filter coefficients we use $\sigma_E = 1$ (cf. [22]).

First, consider low-noise straight edges (SNR = 100). In this case, a radially symmetric PSF does not cause the edge detector to exhibit localization bias, and q should be close to zero. Subpixel methods achieve $q \approx 0.05$ pixels. With the exception of the subpixel-accurate watershed algorithm and Haralick operators (which places edgels at a maximal distance of 0.2 and 0.1 pixels respectively), p roughly equals the pixel radius. Row 1 in Table 1 lists the maximum errors we found.

The effect of higher noise levels on straight edge localization was analyzed by Canny [7]. When the noise is Gaussian distributed with zero mean and standard deviation s_N, the expectation of the squared localization error (in units of pixel^2) is

$$E[\xi^2] = \frac{3}{8}\left(\frac{s_N}{a}\right)^2\left(1 + \frac{\sigma_{PSF}^2}{\sigma_E^2}\right)^3 \tag{4}$$

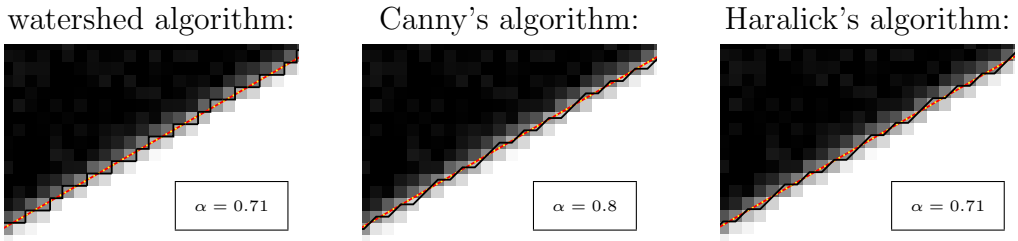
where a is the height of the step, and a/s_N is the signal-to-noise ratio (SNR). When $\sigma_{PSF} \approx \sigma_E$, we get $\sqrt{E[\xi^2]} \approx 1.7\frac{s_N}{a}$. For $\sigma_E \to \infty$, the error approaches $0.6\frac{s_N}{a}$ (the common belief that the error increases with σ_E is only justified in 1D). In typical images $\frac{a}{s_N}$ is between 5 and 100. The expected statistical error is then below 0.2 pixels, and the maximum error does not exceed $3\sqrt{E[\xi^2]} =$ 0.6 pixels with probability 0.997. Rows 1 and 2 of Table 1 confirm these predictions.

Smoothing of curved boundaries with the PSF results in biased edgel positions. When the ground-truth contour is a disc with radius ρ and contrast a, the expected value of the gradient magnitude in the digital image is [23]

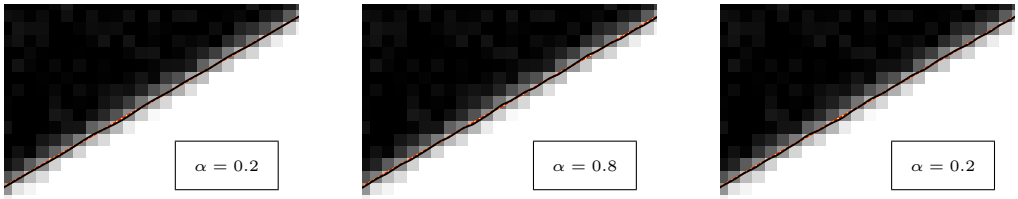
$$g(r) = |a|\frac{\rho}{\sigma^2}e^{-\frac{r^2+\rho^2}{2\sigma^2}}I_1\left(\frac{r\rho}{\sigma^2}\right) \tag{5}$$

where r is the distance from the center of the disc, I_1 is the modified Bessel function of order 1, and $\sigma^2 = \sigma_{PSF}^2 + \sigma_e^2$ is the combined scale of the PSF and edge operator. The edge is located at the maximum of this function. The bias of this position relative to the ground-truth edge position depends on the curvature radius ρ and the scale σ. The displacement is directed toward the concave side of the curve when $\sigma < 0.8\rho$ (which is true in most practical situations). Row 3 of Table 1 compares theoretical predictions and experimental estimates for $\rho = 4$. It can be seen that subpixel-accurate methods are very close to the theoretical limit.

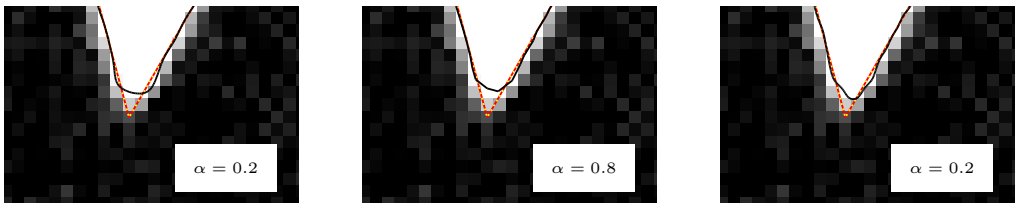
A bias toward the concave side of the contour is also observed at corners. Its magnitude depends on σ and the corner angle φ and is maximal along the bisector of the corner. The location of the gradient maximum along the



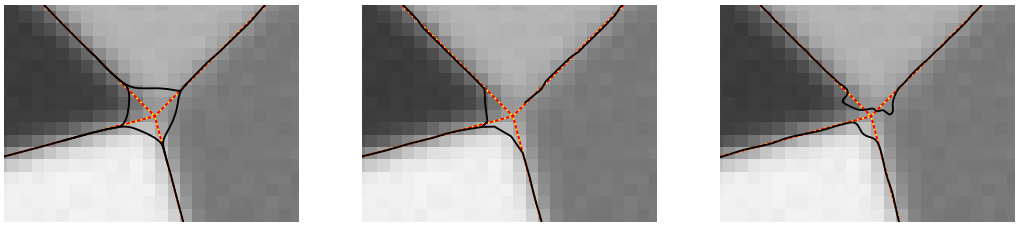
The errors of pixel-accurate algorithms are dominated by the round-off to grid.



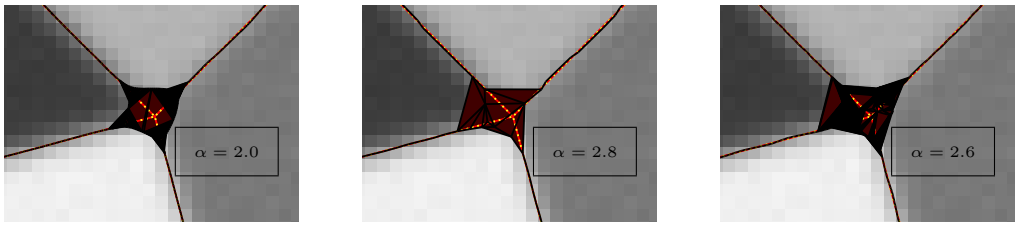
The subpixel variants of the same algorithms are much more accurate, and their errors are determined by the signal-to-noise ratio.



Systematic distortions occur at corners, but the topology is not violated if the regions are large enough.



Edge detector responses near a junction are often topologically incorrect (left: phantom region, center/right: gap between two regions), which requires much larger α.



The corresponding (α, β)-reconstructions recover the correct region topology by creating a thick boundary where the edge detector responses are uncertain.

Figure 9. (α, β)-reconstruction (black) on generated test images with $\sigma_{\mathrm{PSF}} = \sigma_E = 1$ and SNR $= 30$. We set $\alpha = p$ according to the errors reported in Table 1 and $\beta = \alpha$. The ground truth is marked in red/yellow (dotted). (In rows 1 to 3, the (α, β)-reconstruction is identical to the original edge detector response).

	Watersheds (crack edge)		Watersheds (sub-pixel)		Canny (pixel-based)		Canny (sub-pixel)		Haralick (midcrack)		Haralick (sub-pixel)	
	p	q	p	q	p	q	p	q	p	q	p	q
straight line	0.71	0.71	0.13	0.06	0.79	0.52	0.73	0.09	0.71	0.53	0.11	0.09
SNR = 100	[0.71]	[0.71]	[0.11]	[0.05]	[0.79]	[0.5]	[0.72]	[0.05]	[0.71]	[0.5]	[0.07]	[0.05]
straight line	1.0	2.4	0.61	0.58	1.21	1.01	1.14	0.48	0.99	0.96	0.62	0.57
SNR = 10		[1.4]		[0.52]		[1.0]		[0.52]		[1.0]		[0.52]
disc, radius = 4	0.78	1.0	0.30	0.29	0.96	0.74	0.79	0.34	0.87	0.72	0.34	0.33
SNR = 100		[0.96]		[0.25]		[0.75]		[0.25]		[0.75]		[0.25]
corner 90°	1.34	1.55	1.06	0.74	1.17	0.75	1.29	0.73	1.43	0.96	1.06	0.74
SNR = 100			[1.0]	[0.71]			[1.0]	[0.71]			[1.0]	[0.71]
corner 30°	2.84	1.84	3.00	0.94	2.69	0.99	2.85	0.60	2.34	0.84	2.42	0.55
SNR = 100			[2.4]	[0.62]			[2.4]	[0.62]			[2.4]	[0.62]
T-junction $\geq 15°$	3.19	4.54	2.89	3.81	3.46	1.68	3.40	1.40	3.26	3.88	3.21	3.40
SNR = 100												
T-junction $\geq 30°$	2.30	4.54	2.00	3.81	2.89	1.32	2.80	1.40	2.61	3.88	2.60	3.40
SNR = 100												
X-junction	2.65	4.53	3.87	2.86	3.01	2.2	3.07	1.86	3.07	3.31	2.78	3.82
SNR = 100												

Table 1
Experimental estimates of the maximum errors p and q (pixels). Theoretical predictions (if available) are given in brackets.

bisector (i.e. the estimated edge location) has been computed by [24]. Let r_0 be the solution of the implicit equation

$$\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{r_0^2}{2}\right) - \left(\tan\left(\frac{\varphi}{2}\right)\right)^2 \frac{r_0}{2}\left(1 + \operatorname{erf}\left(\frac{r_0}{\sqrt{2}}\right)\right) = 0 \tag{6}$$

where erf is the error function. Then the bias is

$$r = \sigma\, r_0 \sqrt{1 + \left(\tan\left(\frac{\varphi}{2}\right)\right)^2} \tag{7}$$

The sharper the corner, the higher the bias. E.g. for $\varphi = 90°, 45°, 15°$ we get $r \approx 0.71\sigma$, 1.33σ, 2.20σ respectively. Rows 4 and 5 in Table 1 show that actual errors are even slightly bigger than theory predicts.

The situation at junctions is even more complicated. The large number of degrees of freedom (number of constituent regions, angles, intensities) does not allow the error to be described in a simple formula. In general, the error becomes quite large when two contours enclose a very small angle. Moreover, many edge detectors are unable to maintain closed contours near a junction. The resulting gaps contribute to large values of p in the neighborhood of junctions, as rows 6 to 8 of Table 1 show. The watershed algorithms do not have this problem, so their p-error is comparatively smaller. However, they tend to produce oversegmentations, i.e. false positive edges, which leads to an increase in q (because false positive edgels tend to be far from the true edge).

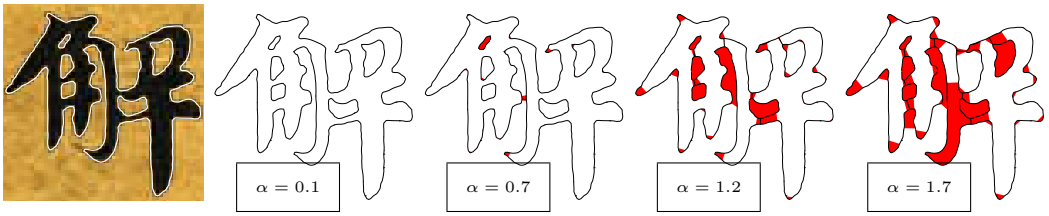


Figure 10. Chinese character (*white*: contours extracted by level-contour tracing [8]), (α, β)-boundary reconstructions with increasing values of α (*red*: before thinning, *black*: minimal boundary reconstruction)

Fig. 10 and Fig. 14 show results of α, β-reconstruction in real images. Region topology is correctly recovered when α and β are properly chosen. Since edgels are considered as isolated points, our new algorithm also facilitates the combination of edgels from different sources, cf. Fig. 11: The edgels computed by Canny's algorithm are not very accurate near corners and junctions, and this requires large α and β causing the reconstruction to be thick in problematic areas (gray). In a second step, a maximum likelihood junction position is computed from the gradient magnitudes and directions at the edgels in a neighborhood of each thick area, resulting in the red points. These points are simply added to the set of edgels, and the reconstruction from the new set is much more accurate than the original one.

Taking everything together, we arrive at the following approximate bounds: suppose the original partition is r-stable and free of $(60°, 2r)$-spikes (i.e. corners enclose at least 60°, curved arcs have at least curvature radius $\rho = 2r/\sqrt{3}$), and the combined PSF and edge detector scale is at most $\sigma = 0.8r$. Moreover, σ should not be smaller than 0.9 pixels in order to avoid aliasing [22], so the pixel distance must be $h \approx r$. Then q does not exceed $0.9\sigma + 0.3 \approx 1.1$ pixels when the boundary contains corners or junctions and SNR = 10 (this is quite visible noise), and $q \approx 0.2$ pixels when the partition is (4-pixel)-regular and SNR = 30. Note that these bounds are maximum errors, the average error is much lower and approaches zero along straight edges. When the edgels are not represented with subpixel accuracy, a round-off error of $h/\sqrt{2}$ must be added, and the average error cannot fall below 0.4 pixels (the standard deviation of a uniform distribution in the unit square) even in case of straight edges.

5 Boundary Thinning and Neighborhood Relations

The boundary sampling theorem presented in section 3 tells us how to reconstruct all regions of a plane partition with correct topology and how to reconstruct the boundary of the partition with correct homotopy type. However, it does not consider the preservation of neighborhood relations for the

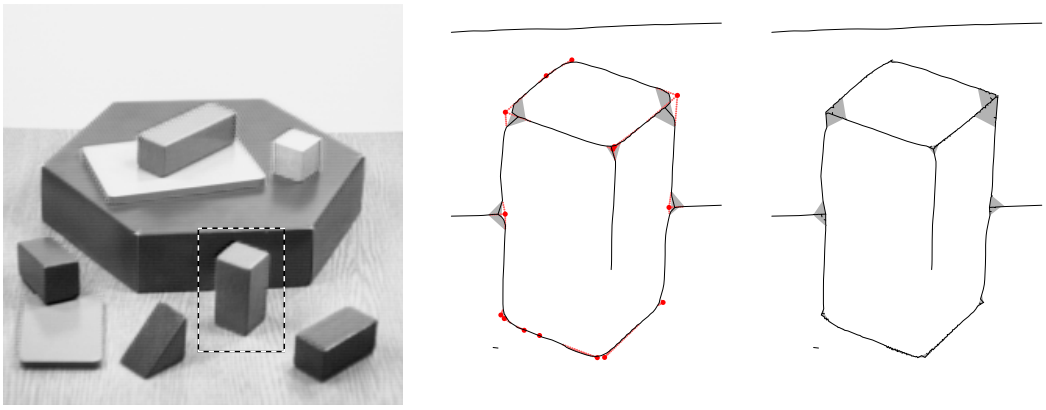


Figure 11. *Left:* original image and ROI;*center:* (α, β)-boundary reconstruction from subpixel Canny edgels (black and gray), thinned reconstruction (black only) and additional edgels to be added (red);*right:* modified reconstruction including new edgels.

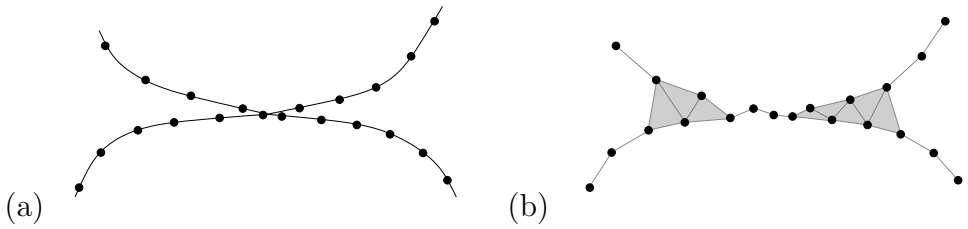


Figure 12. Narrow spikes can lead to a boundary reconstruction where originally unconnected regions (a) look like they had a common boundary edge (b).

following two reasons: First, it is not straightforward to define adjacencies when the boundary representation can be thick (i.e. may contain triangles). Second, two regions whose reconstructions are adjacent (i.e. have a common thin boundary) have indeed not necessarily been neighbored in the original partition, as can be seen in Fig. 12. In this section, we will discuss these two problems and show when and how they can be solved.

5.1 Recovering a Thin Boundary

Many algorithms that build upon segmentation results cannot handle partially thick boundary representations. We can recover a thin (i.e. locally 1-dimensional) boundary by *topology-preserving thinning*, which works similar to skeletonization in a pixel-based region representation.

An edge in the (α, β)-boundary reconstruction is called *simple* if its removal does not change the topology of the reconstructed regions. Simple edges can be easily recognized: they bound an (α, β)-hole on one side and a triangle in the boundary reconstruction on the other. Thinning removes simple edges until any further removal would change the topology (i.e. create an isolated

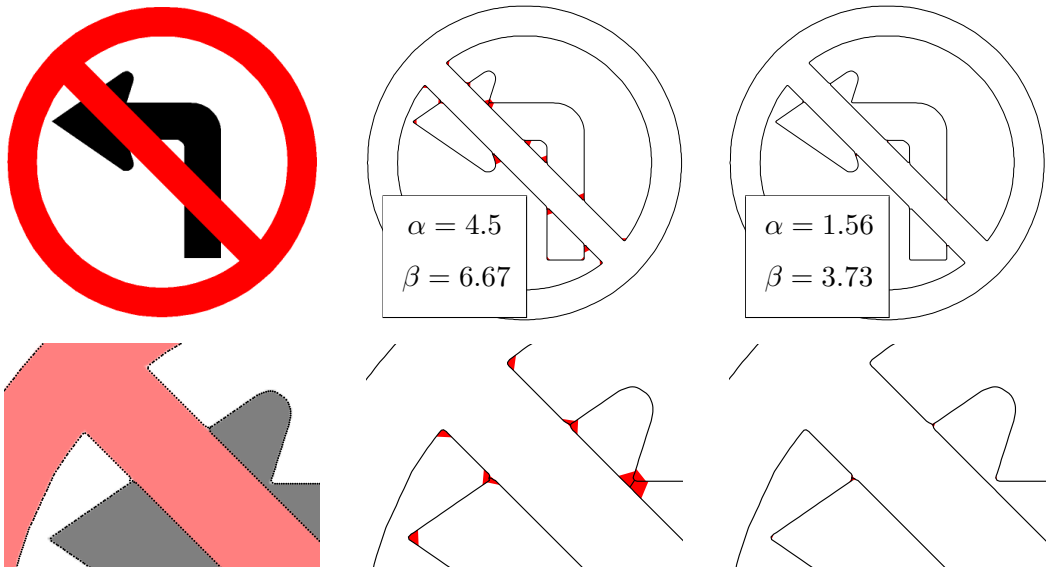


Figure 13. *Top row*: generated image, reconstructions before (red and black) and after (black only) boundary thinning. *Second row*: details (left: original with edgels). Note the connectivity error in the center image where α is too large.

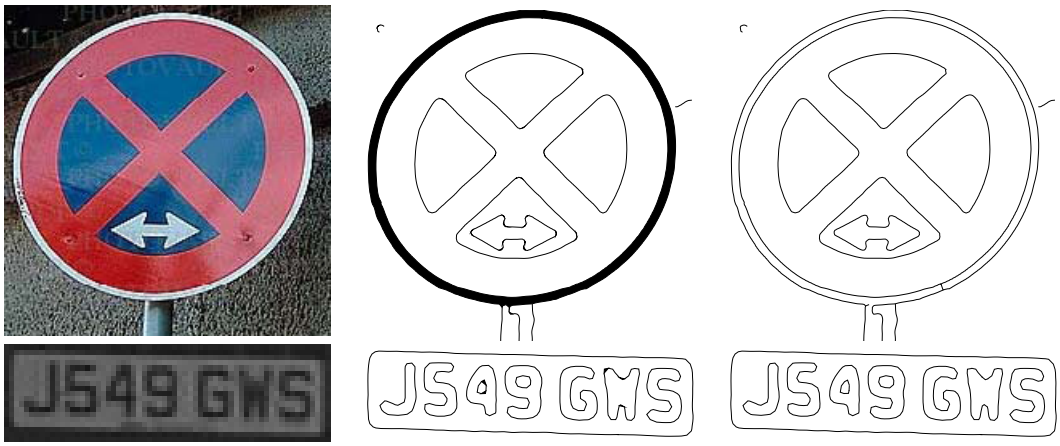
sampling point or merge two regions). The algorithm is as follows:

(1) Find all simple edges of the given (α, β)-boundary reconstruction and put them in a priority queue (the sorting will be discussed below).
(2) As long as the queue is not empty:
   (a) Get the edge e with the highest priority from the queue.
   (b) If e is not simple anymore, it has lost this property during the removal of other edges. Skip the following and recommence with step 2.
   (c) Otherwise, remove e and the adjacent triangle $t \in \mathcal{R}$ from the boundary reconstruction $\mathcal{R}$.
   (d) Check whether the other edges adjacent to t have now become simple and put them in the queue if this is the case.

As far as region topology is concerned, the ordering of the edgels in the priority queue is arbitrary. For example, we can measure the contrast (image gradient) along each edge and remove weak edges first. A particularly interesting ordering however is defined by the length of the edges:

Definition 17 *A (not necessarily unique)* minimal boundary reconstruction *is obtained from an (α, β)-boundary reconstruction by means of topology-preserving thinning where the longest edges are removed first.*

An example for resulting boundaries is given in Fig. 13. Since region topology is preserved, each minimal boundary reconstructions is homotopy equivalent to the boundary of the original plane partition $B = \partial\mathcal{P}$. The two boundaries do not in general have the same topology, because the adjacency relations



(a) original images

(b) (α, β)-boundary reconstruction

(c) minimal reconstruction after thinning

Figure 14. Boundary reconstruction examples on real images (Edgels have been computed by Canny's algorithm on a color and intensity gradient, respectively).

between regions may differ (see below for details), and the reconstruction may contain "dangling" edges, which end in the interior of a region. Since only the shortest edges survive, dangling edges cannot reach very far into a region.

Since a minimal boundary reconstruction can be shown to be a shortest possible one with correct topology, the surviving edges connect edgels closest to each other. Neighboring edgels therefore align in an optimal way on the thinned boundary. The length $d_{\max}$ of the longest surviving edge is a measure of the density of the boundary sampling. The maximum distance p between a true boundary point and the nearest edgel may be much larger than $d_{\max}/2$ if the displacement of neighboring edgels is highly correlated as is usually the case in practice. For example, edgels along a circular arc are consistently biased toward the concave side of the curve. When we set $\alpha' = d_{\max}/2 + \epsilon < p$ (with arbitrarily small ϵ), an (α', β) reconstruction of the edgel set is still correct in the sense of theorem 11: since a minimal reconstruction is a subset of the (α', β) reconstruction, no true regions can get merged. Since $\alpha' < \alpha$, no region can get lost, and since β remained unchanged, no additional holes can be created. In fact, $\beta' = \alpha' + p + q < 2p + q$ would have been sufficient.

We found experimentally that undesirable holes (α-holes that are not (α, β)-holes) are actually quite rare, and their largest triangles are hardly ever as large as the maximal possible circumradius β allows. Therefore, an (α', β')-boundary reconstruction with β' even smaller than $\alpha'+p+q$ often produces the correct region topology. We are currently investigating the conditions which permit weaker bounds. This is important, because a smaller β leads to a correspondingly reduced γ, i.e. the required size of the original regions is reduced, and more difficult segmentation problems can be solved correctly.

So far, we have only shown under which circumstances the topology of each single region of a plane partition can be correctly reconstructed. This does not imply that the region adjacencies are also correct. It may happen that some reconstructed regions become connected while the respective original regions are not, and vice versa. E.g. if two regions of a plane partition have a distance of at most $2p$, their reconstructions may appear as connected. Thus a minimal boundary reconstruction does not necessarily reproduce the original region adjacencies and it is interesting to ask whether some neighborhood relations can nevertheless be recovered from a (p, q)-boundary sampling and the associated (α, β)-boundary reconstruction.

We demonstrated with Fig. 12 that narrow spikes in the original boundaries may cause false adjacencies. When the boundary B is free of (θ, d)-spikes, thin parts of the boundary reconstruction that exceed a certain length can never arise from unfortunate spike configurations, but reflect the true adjacency of two original boundaries:

Theorem 18 *Let $\mathcal{P}$ be an r-stable plane partition with regions r_i and boundary B having no (θ, d)-spikes. Further, let S be a (p, q)-sampling of B and $\mathcal{R}$ the (α, β)-boundary reconstruction of S with regions h_i, such that all requirements of theorem 11 are fulfilled. $S_i = \partial h_i \cap S$ denotes the set of edgels on the boundary of h_i. When $d \geq 2(\alpha + q)$ the following holds with $p' := d/\left(2\sin\frac{\theta}{2}\right) + q$:*

(1) If the distance between the two nearest edgels of S_i and S_j exceeds $2p'$, the corresponding original regions r_i, r_j are not adjacent, i.e. $\partial r_i \cap \partial r_j = \emptyset$.
(2) When there exists a point x with $d_H(x, S_i) \leq p'$, $d_H(x, S_j) \leq p'$ and $d_H(x, S_k) > 2p'$ for all $k \neq i, j$, the original regions r_i, r_j are arc-adjacent.
(3) If two regions r_i, r_j have a distance greater than $2(p' + q)$, the conditions of item 1 are always fulfilled.
(4) If two regions r_i, r_j have a common boundary point x such that $d_H(x, r_k) > 3p'$ for all $k \neq i, j$, the conditions of item 2 are always fulfilled, i.e. adjacency of r_i and r_j can be detected in the boundary reconstruction.

PROOF. (1) For any $s_t \in S_i$ let $x_t \in \partial r_i$ be the nearest boundary point from s_t. Then for any two s_{t_1}, s_{t_2} being connected by a line segment of ∂h_i, the distance between x_{t_1} and x_{t_2} is smaller than $2(\alpha + q)$. Since (θ, d)-spikes do not exist, the distance of each point of ∂r_i to the nearest x_t cannot exceed $d/\left(2\sin\frac{\theta}{2}\right)$ and thus the distance of ∂r_i to ∂h_i is bounded by p'. The same holds for h_j. When the shortest distance between S_i and S_j is larger than $2p'$, ∂r_i and ∂r_j cannot intersect.

(2) Both S_i and S_j intersect the disc $\mathcal{B}_{p'}(x)$. Since $d_H(x, S_k) > 2p'$ for every $k \neq i, j$, no part of ∂r_k can intersect $\mathcal{B}_{p'}(x)$. Thus r_i and r_j are the only regions which intersect $\mathcal{B}_{p'}(x)$, which is only possible when they have a common edge.
(3) Since the distance between r_i and r_j exceeds $2\,(p' + q)$, S_i, S_j have to be more than $2p'$ away from each other.
(4) Due to the absence of (θ, d)-spikes, the distance $d_H(x, S_k), k \neq i, j$ must be greater than $2p'$. For the same reasons, $d_H(x, S_i) \leq p'$ and $d_H(x, S_j) \leq p'$. □

As Fig. 12 shows, junctions of $\mathcal{P}$ with degree 3 or higher are not automatically reconstructed topologically correctly. If however every junction of $\mathcal{P}$ has only degree 3, a sufficiently dense boundary sampling is enough to reconstruct not only the topology of every region of a plane partition, but also the complete neighborhood relations. We only have to ensure that p' and q are small enough such that $2p' + q$ does not exceed the smallest distance between two non-adjacent regions. Then it follows with theorem 18 that the complete topological structure of the original plane partition and of the reconstruction are identical for the right choice of α.

6 Conclusions

To our knowledge, this paper proposes the first geometric sampling theorem that explicitly considers measurement errors. Moreover, our new theorem applies to a much wider class of shapes (r-stable partitions) than existing theorems (r-regular partitions). The situation in real images is thus modeled much more faithfully because shapes may now have corners and junctions, and standard segmentation algorithms can be used. We carefully derive the theoretical properties of several well-known edge detectors in order to apply our new theorem and demonstrate theoretically correct edgel linking. The resulting segmentations are similar to what one gets from traditional heuristic edgel linking, but their properties can now be formally proven thanks to their theoretical foundation in Delaunay triangulation. The key to these advancements has been the shift of attention from region-based digitization models to boundary-based ones: The assumption that no sampling points are in the interior of any region (beyond the known error bound) allows us to reliably recover region and boundary connectivity. Our approach (including boundary thinning) provides a novel way for computing a combinatorial map representation [10, 14] of the boundaries in real images.

We demonstrated that many known digitization and segmentation methods can be analyzed and applied in the new framework by simply determining their error bounds. We can predict whether a given image will be handled properly by an algorithm with a certain error bound. When the error in-

creases, the performance degrades gracefully: First, the recovered boundary becomes thick when the detailed curve shape or junction connectivity can no longer be unambiguously determined. Then, regions get split at too narrow waists, and finally too small regions will be lost. When additional edgels are added within the thick part of the (α, β)-boundary reconstruction, the error bounds p and q will never increase. This opens up new possibilities for algorithm combination. For example, one could start with an edge detector, which produces thick boundaries near corners and junctions. Additional edgels can then be computed by a corner detector whose output is confined to these areas, so that it cannot produce false positives within regions. In future research we will investigate how false positives (large q) and false negatives (large p) can be recognized and removed.

Acknowledgements

We want to thank our anonymous reviewers; in particular, we are grateful for the exceptionally thorough comments of one reviewer, which led to significant improvements and clarifications of our proofs.

References

[1] T. Pavlidis, Algorithms for Graphics and Image Processing, Computer Science Press, 1982.

[2] J. Serra, Image Analysis and Mathematical Morphology, Academic Press, Inc., Orlando, FL, USA, 1983.

[3] L. J. Latecki, C. Conrad, A. Gross, Preserving topology by a digitization process, Journal of Mathematical Imaging and Vision 8 (2) (1998) 131–159.

[4] P. Stelldinger, U. Köthe, Towards a general sampling theory for shape preservation, Image and Vision Computing 23 (2) (2005) 237–248.

[5] P. Stelldinger, Digitization of non-regular shapes, in: C. Ronse, L. Najman, E. Decencière (Eds.), Mathematical Morphology: 40 Years On, Vol. 30 of Computational Imaging and Vision, Springer, Dordrecht, 2005, pp. 269–278.

[6] J.-P. Braquelaire, A. Vialard, Euclidean paths: a new representation of boundary of discrete regions, Graphical Models and Image Processing 61 (1) (1999) 16–43.

[7] J. F. Canny, A computational approach to edge detection, IEEE Trans. Pattern Analysis and Machine Intelligence 8 (6) (1986) 679–698.

[8] E. L. Allgower, K. Georg, Numerical path following, in: P. G. Ciarlet, J. L. Lions (Eds.), Techniques of Scientific Computing (Part 2), Vol. 5 of Handbook of Numerical Analysis, North-Holland, Amsterdam, 1997, pp. 3–207.

[9] H. Meine, U. Köthe, Image segmentation with the exact watershed transform, in: J. Villanueva (Ed.), Proceedings of the 5^{th} IASTED Intl. Conf. Visualization, Imaging, and Image Processing, IASTED, ACTA Press, 2005, pp. 400–405.

[10] H. Meine, U. Köthe, A new sub-pixel map for image analysis, in: R. Reulke, U. Eckardt, B. Flach, U. Knauer, K. Polthier (Eds.), Proc. 11^{th} Intl. Workshop on Combinatorial Image Analysis, Vol. 4040 of Lecture Notes in Computer Science, Springer, 2006, pp. 116–130.

[11] H. Edelsbrunner, E. P. Mücke, Three-dimensional alpha shapes, ACM Transactions on Graphics 13 (1) (1994) 43–72.

[12] H. Edelsbrunner, The union of balls and its dual shape, Discrete & Computational Geometry 13 (1995) 415–440.

[13] G. A. Jones, D. Singerman, Theory of maps on orientable surfaces, Proc. London Math. Soc. 3 (37) (1978) 273–307.

[14] A. Braquelaire, Representing and segmenting 2D images by means of planar maps with discrete embeddings: From model to applications, in: L. Brun, M. Vento (Eds.), Proc. 5^{th} Workshop on Graph-Based Representations in Pattern Recognition, Vol. 3434 of LNCS, Springer, 2005, pp. 92–121.

[15] K. Jänich, Topologie, Springer, 2005.

[16] L. Vincent, P. Soille, Watersheds in digital spaces: an efficient algorithm based on immersion simulations, IEEE Trans. Pattern Analysis and Machine Intelligence 13 (6) (1991) 583–598.

[17] M. Kass, A. Witkin, D. Terzopoulos, Snakes: active contour models, International Journal for Computer Vision 1 (4) (1988) 321–331.

[18] F. Bernardini, C. L. Bajaj, Sampling and reconstructing manifolds using alpha-shapes, in: Proc. 9^{th} Canadian Conf. Computational Geometry, 1997, pp. 193–198.

[19] C. Ronse, M. Tajine, Discretization in Hausdorff space, Journal of Mathematical Imaging and Vision 12 (3) (2000) 219–242.

[20] C. Steger, Subpixel-precise extraction of watersheds, in: ICCV '99, Proc. 7^{th} Intl. Conf. Computer Vision, Vol. 2, IEEE Computer Society Press, 1999, pp. 884–890.

[21] R. M. Haralick, L. G. Shapiro, Computer and Robot Vision, Vol. 1, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1992.

[22] E. B. Dam, B. ter Haar Romeny, Front end vision and multi-scale image analysis, in: Deep Structure I, II & III, no. 1-4020-1507-0 in Computational Imaging and Vision, Kluwer, Kluwer academic publishers, 2003, pp. 215–276.

[23] H. Bouma, A. Vilanova, L. J. van Vliet, F. A. Gerritsen, Correction for the dislocation of curved surfaces caused by the PSF in 2D and 3D CT images, IEEE Trans. Pattern Analysis and Machine Intelligence 27 (9) (2005) 1501–1507.

[24] K. Rohr, Localization properties of direct corner detectors, Journal of Mathematical Imaging and Vision 4 (2) (1994) 139–150.