

Digitization of non-regular shapes in arbitrary dimensions

Peer Steldinger *, Kasim Terzic

Cognitive Systems Group, University of Hamburg, Vogt-Kölln-Str. 30, D-22527 Hamburg, Germany

Received 15 September 2005; received in revised form 11 May 2007; accepted 26 July 2007

Abstract

The preservation of topological properties during digitization is a hard problem in 3 and higher dimensions. Only for the very restricted class of r -regular shapes it is known that the connectivity and inclusion properties of shape components do not change.

In a previous paper it was shown for the 2D case, how a much wider class of shapes, for which the morphological open–close and the close–open-operator with an r -disc lead to the same result, can be digitized correctly in this sense by using an additional repairing step.

This paper extends this to the arbitrary dimensions and analyses the difficulties which occur in 3 or higher dimensional spaces.

The repairing step is easy to compute, parallelizable and does not change as much hyper-voxels as a preprocessing regularization step. The results are applicable for arbitrary, even irregular, sampling grids in arbitrary dimensions.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Shape; Digitization; Repairing; Topology; Reconstruction; Irregular grid

1. Introduction

It is a well known problem that the topological properties of a given object are in general not the same as the topological properties of its digitized version. Since topology is a very fundamental and important aspect of objects, a lot of research has been done on how to preserve topology during digitization.

In 2D it is proven, that a very restricted class of binary objects, called r -regular sets, does not change topology under digitization with a sufficiently dense grid, see [3,5,4,6,7,9].

But most shapes are not r -regular, e.g. have corners. To solve this problem Pavlidis said [6]:

“Indeed suppose that we have a class of objects whose contours contain corners. We may choose a radius of

curvature r and replace each corner by a circular arc with radius r .”

This approach to make 2D shapes r -regular has two problems:

- (1) Pavlidis gives no algorithm how to do it exactly. He also does not say, for which shapes it is possible without changing the topology of the set.
- (2) It is a preprocessing step and thus cannot be computed by a computer, which only gets the digitized information.

In a previous paper [8] these problems have been solved by introducing a new shape class, called r -halfregular shapes, which is more general than the class of r -regular shapes, but also allows to reconstruct the correct topology given a digitized version with an extremely simple algorithm.

In higher dimensional spaces digitization is even more difficult than in 2D. In [10] the first author proved together

* Corresponding author. Tel.: +49 40 42883 2570; fax: +49 40 42883 2572.

E-mail address: steldinger@informatik.uni-hamburg.de (P. Steldinger).

with Ullrich Köthe that the digitized version of a 3D r -regular object cannot be guaranteed to be topologically equivalent even for arbitrary small sampling densities. But it was possible to prove that a weaker topological criterion, the identity of the homotopy trees, is preserved when digitizing r -regular objects with sampling grids of sufficient density. This means that connectivity, adjacency relations and inclusion properties between different foreground and background components remain the same.

In this paper the results for digitizing non-regular objects are generalized to higher dimensions. It is shown that the homotopy tree does not change, if one digitizes an n -dimensional r -halfregular object with a sufficiently dense sampling grid and applies a simple, parallelizable repairing algorithm.

After a short introduction in the definitions of r -regular images, sampling and reconstruction (Section 2), the class of r -halfregular sets is defined in Section 3, whose elements can be converted into r -regular sets by using a very simple morphological preprocessing step. Then it is shown how these shapes can be digitized without changing the homotopy tree by using a postprocessing algorithm instead of the preprocessing. These results are applicable for digitization of objects of any dimension with any type of sampling grid – only a certain density is needed.

In Section 4 it is shown that the concept of r -halfregular shapes includes several other shape descriptions, although it is not as general as the same concept in 2D. Finally in Section 5 the postprocessing step is even more simplified in case of certain regular sampling grids. This chapter also corrects a wrong postprocessing step given in [8]. Finally a discussion about the consequences of the results for applications is given in Section 6, followed by the conclusions (Figs. 1–3).

2. Regular images, sampling and reconstruction

If nothing else is said, all the statements in this paper are regard to the n -dimensional space \mathbb{R}^n with arbitrarily chosen finite $n \geq 2$. At first some basic notations are given: the Euclidean distance between two points, x and y , is noted as $d(x, y)$ and the Hausdorff distance $d_H(A, B)$ between two sets A, B is the maximal distance between one point of one set and the nearest point of the other, i.e. $d_H(A, B) := \max(\max_{x \in A} \min_{y \in B} d(x, y), \max_{y \in B} \min_{x \in A} d(x, y))$. The complement of a set A will be noted as A^c . The boundary ∂A is the set of all common accumulation points of A and A^c . A set A is open, if it does not intersect

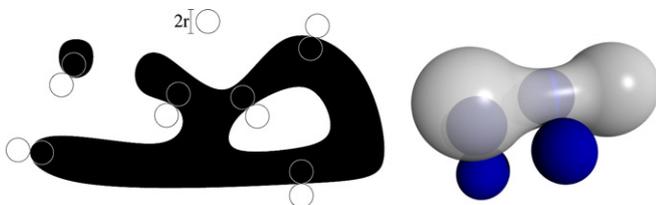


Fig. 1. For each boundary point of a 2D/3D r -regular set there exists an outside and an inside tangent open disc/ball of radius r .



Fig. 2. A lot of 2D and 3D objects can be seen as halfregular sets for some radius r . As illustrated by the 2D text object ‘IMAGE’ and the 3D chessmen object, halfregular objects can have sharp corners and edges which is impossible for regular objects. r -regular versions of these objects can be seen in the left part of Fig. 5.



Fig. 3. The areas which change during regularization can be classified into three types: r -tips (left), r -waists (center), in both cases the semitransparent parts show the remaining object after regularization, and finally r -spots (right), here the covering semitransparent r -regular object (a torus plus a ball) illustrates that the spots, which are independent components of the object, do not contain any r -ball.

its boundary and it is closed if it contains the boundary, $A^0 := A \cup \partial A$, $\bar{A} := A \cup \partial A$. $\mathcal{B}_r(c) := \{x \in \mathbb{R}^n | d(x, c) \leq r\}$ and $\mathcal{B}_r^0(c) := (\mathcal{B}_r(c))^0$ denote the n -dimensional closed and the open ball of radius r and center c . If $c = (0, 0)$, write \mathcal{B}_r and \mathcal{B}_r^0 . The r -dilation $A \oplus \mathcal{B}_r^0$ of a set A is the union of all open r -balls with center in A and the r -erosion $A \ominus \mathcal{B}_r^0$ is the union of all center points of open r -balls lying inside of A . The morphological opening with an open r -ball is defined as $A \circ \mathcal{B}_r^0 := (A \ominus \mathcal{B}_r^0) \oplus \mathcal{B}_r^0$ and the respective closing as $A \bullet \mathcal{B}_r^0 := (A \oplus \mathcal{B}_r^0) \ominus \mathcal{B}_r^0$. The concept of r -regular images was introduced independently by Serra [7] and Pavlidis [6] for the 2D case. These sets are extremely well behaved – they are smooth, round and do not have any cusps (e.g. see Fig. 1). Furthermore r -regular sets are invariant under morphological opening and closing with an open r -ball, as already stated by Serra [7] and generalized to higher dimensions in [10].

Definition 1. A set $A \subset \mathbb{R}^n$ is called r -regular if for each boundary point x of A it is possible to find two tangent open balls B_1, B_2 of radius r , lying entirely in A , respectively, A^c , i.e. $B_1 \subset A, B_2 \subset A^c$ and $x \in \partial B_1 \cap \partial B_2$.

Each outside or inside tangent ball at some boundary point x of a set A defines a tangent hyperplane through x , which is unique if there exists both an outside and an inside

tangent ball. The definitions of r -erosion and r -dilation imply that the boundary of a set does not change under opening or closing with an r -ball if and only if it is r -regular.

Lemma 2. *A set $A \subset \mathbb{R}^n$ is r -regular if and only if A does not change outside of its boundary during morphological opening or closing w.r.t. \mathcal{B}_r^0 : $A \setminus \partial A = A \circ \mathcal{B}_r^0 = (A \bullet \mathcal{B}_r^0)^0$.*

Proof. It only needs to be shown that the boundary does not change under opening or closing with \mathcal{B}_r^0 . This directly follows from the definitions of r -erosion and r -dilation. \square

In order to compare analog with digital images, a definition of the processes of sampling and reconstruction is needed. The most obvious model for sampling is to restrict the domain of the image function to a set of sampling points, called sampling grid. In most approaches only special 2D grids like square or hexagonal ones are taken into account [5–7]. A more general approach only needs a 2D grid to be a countable subset of \mathbb{R}^2 , with the sampling points being not too sparse or too dense anywhere and pixels being defined as Voronoi regions [9]. This has already been generalized to dimensions greater than 2 in [10]. Together with Köthe the first author proved a sampling theorem, saying that an n -dimensional closed r -regular object has the same homotopy tree as its reconstruction with an r' -grid if only $r' < r$ [10]. The homotopy tree of a sets describes the inclusion properties of the foreground and background components, i.e. its root stands for the infinite background component, its children are the neighboring foreground components, each of them having its remaining neighboring background components as children and so on. If two objects have the same homotopy tree, their number of foreground and background components are the same and the neighborhood and connectivity relations between different components are the same. Unfortunately this implies not topological equivalence if the dimension n of the space is at least 3, e.g. a sphere has the same homotopy tree as a torus.

Definition 3. A countable set $S \subset \mathbb{R}^n$ of *sampling points*, where the Euclidean distance from each point $x \in \mathbb{R}^n$ to the nearest sampling point is at most $r' \in \mathbb{R}$, is called an r' -*grid* if $S \cap A$ is finite for any bounded set $A \in \mathbb{R}^n$.

Definition 4. The voxel $\text{Voxel}_S(s)$ of a sampling point s is defined as the Voronoi region of the sampling point, i.e. the set of all points lying at least as near to this point than to any other sampling point. (For simplicity the term voxel is used in all dimensions instead of the theoretically better founded term hyper voxel. In 2D a voxel is called pixel.)

Definition 5. The union of the voxels with sampling points lying in A is called the reconstruction of A w.r.t. S : $\hat{A} := \bigcup_{s \in S \cap A} \text{Voxel}_S(s)$.

Definition 6. Two voxels are adjacent if their intersection has dimension $n - 1$ (thus two adjacent 3D voxels share a face and two 2D pixels share an edge). Two voxels of \hat{A} are connected if there exists a chain of adjacent voxels

in \hat{A} between them. A (*connected*) *component* of \hat{A} is a maximal set of connected voxels (this is equal to 6-connectivity in 3D and to 4-connectivity in 2D for the usual cubic respectively square grids).

3. Digitization of halfregular sets

Most shapes are not r -regular for any r . So if one wants to apply the above mentioned sampling theorem, an r -regular version of the shape has to be constructed before it can be sampled, as suggested by Pavlidis. The question is how to define such a preprocessing step. Following the method described in [10] for the 2D case, we apply a morphological opening and a morphological closing with an r -ball. If the result does not depend on the order of the operations we call the set regularizable.

Definition 7. A set A is called *r -regularizable* if the open-close and the close-open-operator of radius r lead to the same (except of the boundary), $(A \circ \mathcal{B}_r^0) \bullet \mathcal{B}_r^0 = (A \bullet \mathcal{B}_r^0) \circ \mathcal{B}_r^0$. The *r -regularization* of A is $(A \circ \mathcal{B}_r^0) \bullet \mathcal{B}_r^0$.

Lemma 8. *The r -regularization of an r -regularizable set A is r -regular.*

Proof. Since opening and closing are idempotent, $(A \bullet \mathcal{B}_r^0) \circ \mathcal{B}_r^0$ is open and $(A \circ \mathcal{B}_r^0) \bullet \mathcal{B}_r^0$ is closed w.r.t. \mathcal{B}_r^0 . This implies r -regularity of $(A \circ \mathcal{B}_r^0) \bullet \mathcal{B}_r^0$. \square

The disadvantage of the regularization step is that it has to be applied before sampling. Thus this method cannot be used for digitization of a real world object, since one does not know the exact object which has to be digitized. It is only applicable for discretization of an analytic object representation, i.e. if one has an exact description of the object, e.g. as a mathematical formula and wants to have a discrete voxel representation. But in practice one does not have this information. Then one needs to have a repairing step which solely relies on the sampled data. In the following we will derive such a postprocessing step which can always be applied if one can assume that the original object is r -halfregular for some known r . Thus one can derive the exact homotopy tree of the object by looking at its digitization. Note, that a shape and its regularization in general do not need to have the same homotopy tree, since the topology can be totally changed during the regularization step (e.g. a small region vanished under regularization). The changing areas can be classified into waists, tips and spots (see Fig. 3).

The waists cause the biggest problems, because even big and thus important connected components can change their topology under regularization if they have waists. So if one wants to regularize a set it should have no waists.

Definition 9. For some set A let A' be a connected component of $(A \setminus (A \circ \mathcal{B}_r^0))^0$. If no boundary point of A' is also boundary point of $A \circ \mathcal{B}_r^0$, A' is called *r -spot* of A . If

each point x in A' has exactly one point in $\partial(A \circ \mathcal{B}_r^0) \cap \partial A'$ with its normal going through x , A' is called r -tip of A and if at least one point x in A' has more than one point in $\partial(A \circ \mathcal{B}_r^0) \cap \partial A'$ with its normal going through x , A' is called r -waist of A (see Fig. 3).

Note, that this definition of spots, tips and waists seems to differ from the one given in [8], where the number of adjacent r -balls was used for classification. While these two definitions are equivalent in 2D, the old one could not directly be generalized to higher dimensions.

Definition 10. A set A is called r -halfregular if for each boundary point there exists an open inside or an open outside tangent r -ball, lying completely inside, respectively outside of A , and if neither A nor A^c has an s -waist for any $s \leq r$.

Obviously an r -halfregular set is also s -halfregular for any $s < r$ and the complement of an r -halfregular set is also r -halfregular. For the rest of this section let A be an r -halfregular set with $r > 0$. Further let $B := (A \circ \mathcal{B}_r^0) \bullet \mathcal{B}_r^0$ be its r -regularization. $C := A \setminus (A \circ \mathcal{B}_r^0)$ shall be the difference between A and its opening with \mathcal{B}_r^0 and $D := (A \bullet \mathcal{B}_r^0) \setminus A$ the difference between A and its closing with \mathcal{B}_r^0 (see Fig. 4). The connected components of C respectively D are the r -spots, r -tips and r -waists of the foreground respectively background, which change during the preprocessing regularization step. Note, that $D = (A^c \bullet \mathcal{B}_r^0) \setminus A^c$ and $C = A^c \setminus (A^c \circ \mathcal{B}_r^0)$, i.e. C and D are duals.

Lemma 11. For each boundary point of C or of D there exists an outside tangent open r -ball and no connected component of $C \oplus \mathcal{B}_r^0$ or of $D \oplus \mathcal{B}_r^0$ contains an open ball of radius $2r$ as subset.

Proof. C cannot have an open ball of radius r as subset, because due to the definition of C the center of such a ball is not in C . Now let x be a boundary point of C . Then x is boundary point of A or of $A \circ \mathcal{B}_r^0$. In the first case there exists an tangent open r -ball lying completely outside of C since A is r -halfregular and C cannot include an inside tangent r -ball. In the second case there also exists an outside tangent ball for C , since $A \circ \mathcal{B}_r^0$ is open w.r.t. \mathcal{B}_r^0 . Thus \bar{C} is closed w.r.t. \mathcal{B}_r^0 and $C \oplus \mathcal{B}_r^0$ does not contain any open ball of radius $2r$ as subset. The proof for D is analog. \square

Lemma 12. Let A be an r -halfregular set. Then every boundary point $y \in \partial A$ is also boundary point of $A \bullet \mathcal{B}_r^0$ or $A \circ \mathcal{B}_r^0$ and A is r -regularizable.



Fig. 4. From left to right: an r -halfregular set A , its r -regularization B , and the changes of the fore- and background due to regularization C and D .

Proof. Let y be some boundary point of A . If there exists an outside [inside] tangent r -ball, then y remains boundary point after r -closing [r -opening].

Now suppose A is not r -regularizable. Then $(A \circ \mathcal{B}_r^0) \bullet \mathcal{B}_r^0$ is not equal to $(A \bullet \mathcal{B}_r^0) \circ \mathcal{B}_r^0$ and there either exists a point $x \in (A \circ \mathcal{B}_r^0) \bullet \mathcal{B}_r^0$, which is not in $(A \bullet \mathcal{B}_r^0) \circ \mathcal{B}_r^0$ or there exists a point $x \in (A \bullet \mathcal{B}_r^0) \circ \mathcal{B}_r^0$, which is not in $(A \circ \mathcal{B}_r^0) \bullet \mathcal{B}_r^0$.

Such an x cannot lie inside or on the boundary of an r -ball being subset of A , because then x would be element of $A \circ \mathcal{B}_r^0$ and thus $x \in (A \circ \mathcal{B}_r^0) \bullet \mathcal{B}_r^0$, and – since closing is extensive and opening is increasing – $x \in (A \bullet \mathcal{B}_r^0) \circ \mathcal{B}_r^0$ is true.

If x lies inside or on the boundary of an r -ball in A^c , it cannot be in $(A \circ \mathcal{B}_r^0) \bullet \mathcal{B}_r^0$ or $(A \bullet \mathcal{B}_r^0) \circ \mathcal{B}_r^0$ for analog reasons.

Now suppose, x is in A , but not inside or on the boundary of some r -ball in A nor inside or on the boundary of some r -ball in A^c , thus $x \in C^0$. C has an outside tangent open r -ball at any boundary point y due to Lemma 11. Now let y be the boundary point of C being nearest to x . Then any y tangent has to be orthogonal to \overline{xy} . Thus there exists a unique tangent and also a unique outside tangent r -ball. Obviously, the distance between x and y is smaller than r . Since y remains boundary point after r -closing of C , x cannot be in $(A \bullet \mathcal{B}_r^0) \circ \mathcal{B}_r^0$.

Now let $A' := A \circ \mathcal{B}_r^0$ and let C' be the connected component of C containing x . Due to the absence of r -waists, there exists at most one point y in $\partial A' \cap \partial C'$ having a normal going through x . Since y lies on $\partial A'$ and since any point of C' lying on the normal though y lies on no other such normal, it is also boundary point of $A' \bullet \mathcal{B}_r^0$ and thus x cannot be element of $(A \circ \mathcal{B}_r^0) \bullet \mathcal{B}_r^0$.

Analogously any x in D^0 is element of both $(A \circ \mathcal{B}_r^0) \bullet \mathcal{B}_r^0$ and $(A \bullet \mathcal{B}_r^0) \circ \mathcal{B}_r^0$.

Thus any $x \in \mathbb{R}^n$ is element of $(A \circ \mathcal{B}_r^0) \bullet \mathcal{B}_r^0$ if and only if it is element of $(A \bullet \mathcal{B}_r^0) \circ \mathcal{B}_r^0$. \square

As a consequence of Lemma 12, A can be constructed (except of its boundary) as $A^0 = (B \cup C)^0 \setminus D$. In the following the sampling theorem for halfregular sets is developed. Therefore, one lemma needs to be proved before.

Lemma 13. No background [foreground] component in the reconstruction of $B \cup C$ [$B \setminus D$] w.r.t. an r' -grid, $r' < r$, is subset of $C \oplus \mathcal{B}_r^0$ [$D \oplus \mathcal{B}_r^0$].

Proof. Let $c \in A^c$ be a background sampling point in $C \oplus \mathcal{B}_r^0$. Due to Lemma 11 there exists an open r -ball in C^c such that c lies in the ball. This ball can be chosen such that it lies either completely in B or completely outside of B . The center m of the ball is not in $C \oplus \mathcal{B}_r^0$. The halfline starting at c and going through m crosses $\partial \text{Voxel}(c)$ at exactly one point c' . If $d(c, m) \leq d(c, c')$, m lies in $\text{Voxel}(c)$ and thus the voxel is connected to the area outside of $C \oplus \mathcal{B}_r^0$, which implies that c cannot be part of a separate background component covered by $C \oplus \mathcal{B}_r^0$. If $d(c, m) > d(c, c')$, let g be the hyperplane (i.e. a plane in

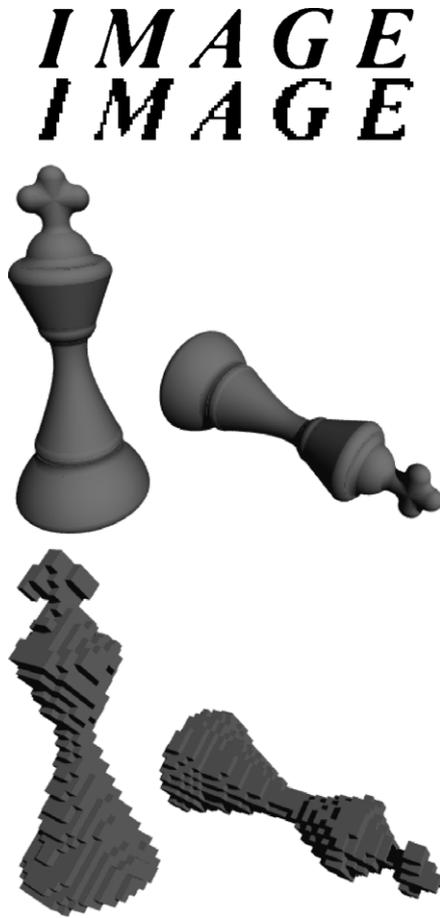


Fig. 5. The regularization of a halfregular set (first two rows) does not change the homotopy tree during digitization (last two rows).

3D and a line in 2D) defined by the hyperface of $\partial\text{Voxel}(c)$ containing c' . If there is more than one such hyperface (e.g. at a corner or edge point), one is chosen arbitrarily. The point c'' constructed by mirroring c on g is also a sampling point, and their voxels are adjacent. c'' lies on the hypersphere of radius $d(d', c) = d(c', c'')$ with center c' . Among all points on this hypersphere, c has the largest distance to m , and in particular $d(m, c'') < d(m, c)$. Thus, c'' lies outside of $C \oplus \mathcal{B}_r^0$, and is closer to m than c . By repeating this construction iteratively we obtain a chain of adjacent voxels whose sampling points successively get closer to m . Since $C \oplus \mathcal{B}_r^0$ contains only a finite number of sampling points, one such voxel will eventually not be covered by $C \oplus \mathcal{B}_r^0$. The constructed chain consists of voxels whose sampling points lie in a common r -ball outside of C . If this

ball lies in B , they are not in the background, in contradiction to the supposition. Otherwise the voxels cannot be part of a separate background component in $C \oplus \mathcal{B}_r^0$. Since the chain does not contain any sampling point lying in B , they also cannot be part of a separate background component in the reconstruction of $B \cup C$ which is subset of $C \oplus \mathcal{B}_r^0$.

Analogously there exists no foreground component in the reconstruction of $B \setminus D$, which is subset of $D \oplus \mathcal{B}_r^0$. \square

Theorem 14. *Let A be a closed r -halfregular set with no $3r$ -spot in A or in A^c , let \hat{A} be the reconstruction of A with an r' -grid, $r' < r$, and let \hat{A}' be the result of filling [deleting] all connected components of \hat{A}^c [\hat{A}], which do not contain an open $2r$ -ball. Then \hat{A}' has the same homotopy tree as A (see Fig. 7) and the number of different voxels from \hat{A} to \hat{A}' is as most as high as from A to the reconstruction of the r -regularization B (see Fig. 6).*

Proof. \hat{A} is equal to the union of the reconstructions of B and $C = A \setminus (A \circ \mathcal{B}_r^0)$ minus the reconstruction of $D = (A \bullet \mathcal{B}_r^0) \setminus A$. Due to Lemmas 8 and 12 B is an r -regular set and thus it has the same homotopy tree as its reconstruction (see [10]).

The connected components of \hat{C} are either also separated components in \hat{A} (which would change the homotopy tree) or connected with some component of \hat{B} (which does not affect the homotopy tree). Lemma 11 states that no connected component of $C \oplus \mathcal{B}_r^0$ contains an open ball of radius $2r$. Due to $r' < r$, \hat{C} is a subset of $C \oplus \mathcal{B}_r^0$. It follows that no connected component of \hat{C} can contain an open ball of radius $2r$. Analogously no connected component of \hat{D} can contain an open ball of radius $2r$.

Due to Lemma 13 there cannot exist any background component in the reconstruction of $B \cup C$, which is subset of $C \oplus \mathcal{B}_r^0$, and there cannot exist any foreground component in the reconstruction of $B \setminus D$, which is subset of $D \oplus \mathcal{B}_r^0$. This implies that any separate component of \hat{C} or of \hat{D} is surrounded by voxels belonging to connected components which do not vanish under dilation or erosion with an open $3r$ -ball. This also implies that the resulting image is independent of the order of the filling and deleting of connected components, which are subsets of $C \oplus \mathcal{B}_r^0$ or of $D \oplus \mathcal{B}_r^0$.

So by filling all connected components of \hat{A}^c and deleting all connected components of \hat{A} , which do not contain an open ball of radius $2r$, any component caused by C and D is affected, which is not part of bigger components in \hat{A} and \hat{A}^c , respectively. Any connected component of \hat{B} [\hat{B}^c] is not

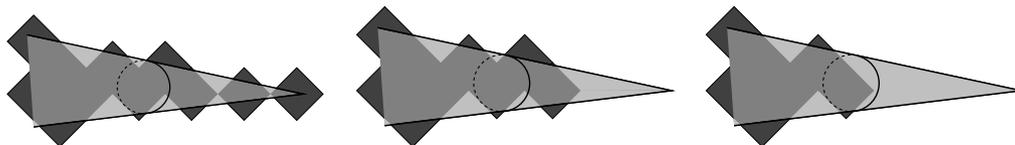


Fig. 6. The straightforward digitization may cause topological errors (i.e. new connected components) at r -tips (left) in contrast to the use of the regularization step (right) and the even better repairing step (center).

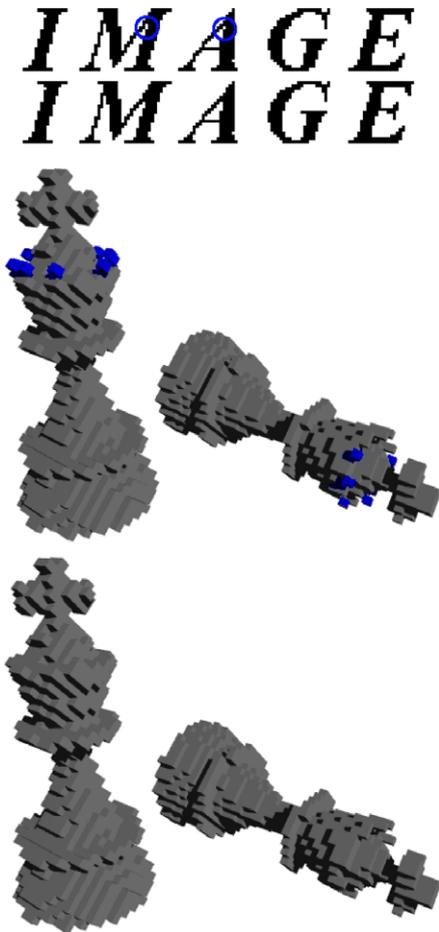


Fig. 7. The topological errors which may happen during digitization (first two rows) can be identified and corrected in order to get a result with correct homotopy tree (last two rows) which is as similar as possible to the straight forward digitization – in contrast to the digitization of the regularized version (see Fig. 5).

deleted [filled], because it contains an open ball of radius $2r$ due to the fact that the corresponding component in $B[B^c]$ contains a ball of radius $3r$.

It follows that the connected components of \hat{A}' and \hat{B} differ only in a way, which does not affect the neighborhood relations. Thus \hat{A}' and B have the same homotopy tree. Moreover they differ only in voxels lying in C or D . Since \hat{A} and \hat{A}' differ in all these voxels, the number of voxels, which changes due to the postprocessing repairing step is at most as big as the number of voxels, which changes due to the preprocessing regularization step.

Since B can be constructed from A by removing r -tips from A and A^c , which does not change the topology, \hat{A}' and A have the same homotopy tree. \square

In [8] it was shown that for the 2D case not only the homotopy trees are identical but the sets are even \mathbb{R}^2 -homeomorphic. This follows simply from the identity of the homotopy trees and the connectivity via direct pixel neighborhood.

Another important fact is that every connected component is connected via direct voxel adjacency. Since this is

the most restrictive way to define connectivity, any definition of connectivity leads to the same connected components. That is, in 2D point-, edge- and face-connectivity (and higher degree connectivities in higher dimensional spaces) all lead to the same connected components. If two voxels intersect in at least one point, they belong globally to the same connected component.

This is also true for a lot of other reconstruction methods. For example, if one uses the marching cubes algorithm for 3D cubic grids to generate the surfaces between the different components, the connected components are preserved since every two adjacent voxels are also connected in the marching cubes reconstruction if they are both in the foreground or both in the background.

4. Examples for r -halfregular sets

In the last section it was shown, that digitization with repairing is sufficient to get a correct digital version (in the sense of connectivity) of some r -halfregular set, if each of the connected components of the set and its complement has a certain size (no $3r$ -spots). Note that this is only a restriction to the sampling density and not a restriction to the class of correctly digitizable sets, since for each r -halfregular set there exists an $s \leq r$ such that the set has no $3s$ -spots and is s -halfregular. In [8] it was shown how general the concept of r -halfregular sets is and proved that it can be applied to polygonal objects, to convex and to locally convex/concave objects. Unfortunately it is not as general in three or more dimensions.

While two-dimensional polygonal shapes are always locally convex or concave (note, that due to [1] every two-dimensional shape can be approximated by a polygonal one), this are not the only possibilities in three or more dimensions. One can show that any n -dimensional polytope ($n \geq 3$), which is not globally convex (or the complement of a convex object), must have for any $r > 0$ at least one surface point where no outside and no inside tangent r -ball exist. Most of such polytopes are even not r -regularizable for any sufficiently small r .

These problems do not occur for convex or locally convex/concave objects in higher dimensions. Another example are shapes, which are bounded by spline hypersurfaces, if they are convex or concave at any point where two or more hypersurfaces meet. One has only to determine the minimal size of the spots and waists and the maximal curvature at inner hypersurface points, which can be used to compute the necessary minimal sampling density.

As the 3D chessmen example illustrates, rotationally symmetric objects, e.g. objects bounded by surfaces of revolution (SORs), are also r -halfregular if the defining set of lower dimension is also r -halfregular.

Theorem 15. (a) Each convex set is r -halfregular for any $r \in \mathbb{R}_+$. (b) Each set A , where for each boundary point $x \in \partial A$ either $A \cap \mathcal{B}_{2r}^0(x)$ or $A^c \cap \mathcal{B}_{2r}^0(x)$ is convex, is r -halfregular. (c) Let A be an n -dimensional r -halfregular set

which is invariant under mirroring at some hyperplane l (i.e. in 2D it is symmetric under mirroring at some line). Further let A' be the result of rotating A around l in the $(n + 1)$ -dimensional space (here l defines a rotation axis). Then A' is r -halfregular.

Proof. (a) A convex set can be described as the intersection of hyper-halfspaces. Any outside tangent r -ball of some hyper-halfspace is completely outside of this intersection. Thus for any boundary point there exists an outside tangent r -ball. Obviously convex sets cannot have waists, which implies r -halfregularity.

(b) First suppose there exists an s -waist A' of A with $s \leq r$. Now let x be a boundary point of A which is also boundary point of A' . If $A^c \cap \mathcal{B}_{2r}^0(x)$ is convex, then it is subset of one half of the $2r$ -ball and the other half contains an r -ball lying in A . Thus $A \cap \mathcal{B}_{2r}^0(x)$ has to be convex. Since this has to be true for any such boundary point, A must be convex in the area $A' \oplus \mathcal{B}_{2r}^0$ of radius $2r$ around the waist. This implies that the union of the waist and all tangent r -balls inside A also has to be convex. But this cannot be since the convex hull of this union is morphologically open w.r.t. \mathcal{B}_s^0 in contrast to the set itself. Thus there cannot exist any s -waist in A and analogously in A^c .

For each boundary point x there exists an outside tangent r -ball if $A \cap \mathcal{B}_{2r}^0(x)$ is convex and an inside tangent ball if $A^c \cap \mathcal{B}_{2r}^0(x)$ is convex. Thus A is r -halfregular.

(c) Any n -dimensional tangent r -ball of A results in a torus-like object due to the rotation. Since this object is equal to the trace of rotating an $(n + 1)$ -dimensional r -ball around l , every boundary point of A' has one inside or one outside tangent r -ball. Since any normal in a point $y \in \partial A'$ lies in the hyperplane containing y and l , every two intersecting normals lie in the same hyperplane, whose intersection with A' is simply a rotated version of A . Thus two normals of A' intersect if and only if the regarded normals intersect in A . Thus the absence of waists in A implies the absence of waists in A' and A' must therefore be r -halfregular. \square

5. Discrete repairing

Although the repairing process is very simple – you only have to find connected components which do not contain a ball of a certain size – an implementation into a discrete algorithm is not straightforward since subvoxel positions of such balls have to be considered. In this section it is shown that there are even better ways to find such connected components. The idea is that for any regular r' -grid there are only finitely many patterns which cover a $2r$ -ball ($r = r' + \varepsilon$) such that each voxel intersects the ball, and some of these patterns include others.

Definition 16. Let S be the set of all voxel patterns in an r' -grid, which cover a $2r$ -ball ($r = r' + \varepsilon$) for some $\varepsilon > 0$. Then any pattern of S , which itself does not include any other pattern of S , is called an *indicator pattern*.

Theorem 17. Let A be a digital component, i.e. a connected component of the digital reconstruction of some set. Then A contains a $2r$ -ball if and only if it contains an indicator pattern.

Proof. If a digital component contains a $2r$ -ball then the union of all voxels which intersect this $2r$ -ball includes an indicator pattern. If otherwise a digital component contains an indicator pattern, it must contain the $2r$ -ball which is covered by this pattern. \square

In the case of digitization of two-dimensional objects such a simplification is already mentioned in a previous paper [8] (see also [2]). Unfortunately there was given a wrong criterion which is only feasible for identifying connected components not including an r -ball instead of a $2r$ -ball. The following two theorems correct this mistake. The first theorem gives a criterion to detect connected components which have to be removed, and the second gives the indicator sets for square and hexagonal grids.

Theorem 18. Let S be an r' -grid and let A' be one connected component of the digital reconstruction \hat{A} of a set A . If there exists no voxel such that the voxel and all voxels sharing at least one boundary point with it, are inside of A' , then A' contains no $(2r' + \varepsilon)$ -ball.

Proof. Suppose to the contrary there exists such a $(2r' + \varepsilon)$ -ball B . Then let V be the voxel in A' containing the center of B . Since V can be covered by an r' -ball and since this r' -ball lies inside of B having a Hausdorff distance of at least ε to ∂B , every voxel sharing at least one boundary point with V intersects with B and thus lies inside of A' . \square

Generally it is possible that some of the connected components to be removed do not fulfill this criterion. This is always the case if the biggest ball which can be covered with the union of a voxel with all neighboring voxels has radius of at most $2r'$. Unfortunately this is true for all n -dimensional hypercubic sampling grids with $n \geq 3$. But in case of the mostly used 2D sampling grid, the square grid, the above criterion is both necessary and sufficient.

Theorem 19. Let A be a 2D closed r -halfregular set with no $3r$ -spot in A or in A^c , let \hat{A} be the reconstruction of A with a 2D square or hexagonal r' -grid with $r' < r$ and let \hat{A}' be the result of filling/deleting all connected components of \hat{A}^c and \hat{A} , which do not contain one of the patterns given in Fig. 8. Then \hat{A}' is \mathbb{R}^2 -homeomorphic to A and the number of different pixels between \hat{A} and \hat{A}' is as most as high as between \hat{A} and the reconstruction of the r -regularization B of A .

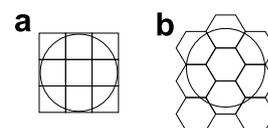


Fig. 8. There is only a finite number of indicator patterns, i.e. minimal pixel configurations which cover a $2r$ -ball. In case of a square (a) or hexagonal (b) grid there is even only one indicator pattern (disregarding of rotation).



Fig. 9. The shown figures are based on the 3D aneurysm dataset by Philips Research, Hamburg. (a) Original dataset. Note the disconnected regions on the top of the image due to insufficient sampling. (b) Dataset after regularization. Since the vessels are too thin regarding the sampling resolution, a lot of details have been removed and the object is splitted into several parts. (c) Reconstruction with the proposed postprocessing step. The result is connected, small vessel parts have been removed only if they are not connected to big ones. Thus as much information as possible is preserved.

The theorem can be verified by simply looking at every possible pattern supercovering a $2r$ -ball.

6. Applications

The presented work is theoretical and it has not been done for a specific application. But it is of interest for every application where one needs to measure certain topological or geometric properties. One example can be found in [11]. There the task was to segment 3D MRT scans of wooden medium density fibreboards (MDF) and one subtask was to measure the intra fibre volume, i.e. the amount of air between the fibres. This could not be done through direct voxel classification and counting, since the air between the fibres had to be distinguished from the air inside the fibres. Since the intra fibre volume is connected and it has large regions with big diameters while the diameter of the inter fibre regions cannot be bigger than the fibre diameter, the intra fibre volume could successfully be separated by using the repairing step mentioned above.

Another example is in medical image analysis, if one wants to measure the cortical surface area in 3D MRI images of the human brain (see [12]). Surface area estimation only makes sense if a surface with correct topology can be reconstructed. If one knows by experience a good value of r for modeling the cortical brain structure as r -halfregular object, one knows how dense one has to digitize. The proposed postprocessing filter not only reconstructs topological information, it also suppresses a large amount of noise, such that afterwards the surface area estimation is much more accurate.

A further application in medical image analysis is segmentation of blood vessels. Such structures are extremely thin and thus digitization problems arise. Since blood vessels can be interpreted as r -halfregular objects with r being the vessel radius, our results allow to compute a minimal vessel radius given some digitization, such that the network of all blood vessels with bigger radius can be found with correct topology, i.e. one has the guaranty that if two vessel

parts are connected in reality, their digital reconstructions are also connected. Fig. 9 shows an example.

Of course, if one knows the exact geometry of the objects to be digitized, e.g. in industrial production, then the exact value of r can be computed and thus one knows the necessary sampling density. Sometimes one can even manipulate the object, e.g. if the object is a mathematical deformable model. If one tries to fit such a deformable model to a real world object, but always takes care that the model is r -halfregular, one can derive a voxel-based representation with the correct connectivity, which can be advantageous for further processing steps. The postprocessing step allows to model objects much better than by modeling them as r -regular objects, which would be equivalent to the preprocessing step (regularization).

7. Conclusions

The new class of r -halfregular shapes was introduced for arbitrary dimensions and it was shown that this class can be digitized without changing connectivity and inclusion properties by using a simple postprocessing step. The main result simply says that the digitization of an r -halfregular shape with an arbitrary sampling grid of sampling density $r' < r$ does not change the homotopy tree after applying a postprocessing step which simply removes all connected components, which do not exceed a certain size. This is much more general than the restriction to r -regular shapes, which was used in the literature before. It was also shown that the postprocessing step leads to better results than a morphological preprocessing step, which makes an r -halfregular shape r -regular. Further on, it was proven that the class of r -halfregular shapes subsumes other shape classes like locally convex shapes. Finally, the postprocessing step was even more simplified in case of using regular sampling grids.

Thus a method is given which allows to guarantee the preservation of important topological properties during digitization in arbitrary dimensions, which is applicable

to a class of shapes being much more general than the one used in previous approaches.

References

- [1] R.H. Bing, Approximating surfaces with polyhedral ones, *Annals of Mathematics* 65 (1957) 456–483.
- [2] A. Giraldo, A. Gross, L.J. Latecki, Digitizations preserving shape, *Pattern Recognition* 32 (1999) 365–376.
- [3] U. Köthe, P. Steldinger, Shape preserving digitization of ideal and blurred binary shapes, in: I. Nyström, G. Sanniti di Baja, S. Svensson (Eds.), *Discrete Geometry for Computer Imagery, Proceedings of DGCI 2003, LNCS, vol. 2886, Springer, Berlin, 2003*, pp. 82–91.
- [4] L.J. Latecki, C. Conrad, A. Gross, Preserving topology by a digitization process, *Journal of Mathematical Imaging and Vision* 8 (1998) 131–159.
- [5] L.J. Latecki, *Discrete Representation of Spatial Objects in Computer Vision*, Kluwer, Dordrecht, Netherlands, 1998.
- [6] T. Pavlidis, *Algorithms for Graphics and Image Processing*, Computer Science Press, Rockville, Maryland, 1982.
- [7] J. Serra, *Image Analysis and Mathematical Morphology*, Academic Press, London, New York, 1982.
- [8] P. Steldinger, Digitization of non-regular shapes, in: C. Ronse, L. Najman, E. Decenciere (Eds.), *Mathematical Morphology, Proceedings of ISMM 2005, Springer, 2005*, pp. 269–278.
- [9] P. Steldinger, U. Köthe, Shape preservation during digitization tight bounds based on the morphing distance, in: B. Michaelis, G. Krell (Eds.), *Pattern Recognition, Proceedings of DAGM 2003, LNCS, vol. 2781, Springer, Berlin, 2003*, pp. 108–115.
- [10] P. Steldinger, U. Köthe, Towards a general sampling theory for shape preservation, in: *Image and Vision Computing, Special Issue on Discrete Geometry for Computer Imagery, vol. 23 (2), Elsevier, Amsterdam, Netherlands, 2005*, pp. 237–248.
- [11] T. Walther, K. Terzic, T. Donath, H. Meine, F. Beckmann, H. Thoemen, Microstructural analysis of lignocellulosic fiber networks, in: U. Bonse, (Ed.), *Developments in X-Ray Tomography, vol. 6318, CID No. 631812, SPIE, 2006*.
- [12] G. Windreich, N. Kiryati, G. Lohmann, Voxel-based surface area estimation: from theory to practice, in: *Pattern Recognition, vol. 36 (11), Elsevier, Amsterdam, Netherlands, 2003*, pp. 2531–2541.